# Mott-Hubbard Metal-Insulator Transition and Optical Conductivity in High Dimensions

Von der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Augsburg
zur Erlangung eines Doktorgrades der Naturwissenschaften
genehmigte Dissertation

von
Diplom–Physiker Nils Blümer
aus
München

# Contents

# Nomenclature

| | |
|---|---|
| $0^+$ | Positive infinitesimal |
| $\langle\langle \hat{B}, \hat{A} \rangle\rangle$ | Correlation function, p. 308 |
| $\langle f(\epsilon) \rangle$ | Thermal expectation value, p. 195 |
| $\langle f(\epsilon) \rangle_{\rho(\epsilon)}$ | Expectation value, p. 31 |
| $\langle f(\boldsymbol{k}) \rangle_{\boldsymbol{k}}$ | Expectation value, p. 31 |
| $\langle\langle i, j \rangle\rangle$ | NNN pairs (in summation), p. 31 |
| $\langle i, j \rangle$ | NN pairs (in summations), p. 9 |
| $\{s\}$ | Auxiliary Ising spin field, p. 20 |
| $\lvert \boldsymbol{v} \rvert$ | Euclidean metric, 2-norm, p. 29 |
| $\lVert \boldsymbol{v} \rVert$ | "Taxi cab" vector metric, 1-norm, p. 30 |
| $\alpha$, $\beta$, $\gamma$ | Cartesic indices |
| $\beta$ | Inverse temperature, p. 17 |
| $\epsilon(\omega)$ | Dielectric function, p. 177 |
| $\epsilon_{\boldsymbol{k}}$, $\epsilon(\boldsymbol{k})$ | Electronic dispersion (noninteracting energy), p. 14 + 47 |
| $\epsilon_D(\boldsymbol{k})$ | Contribution to dispersion from hopping to $D^{\text{th}}$-NN, p. 47 |
| $\eta_{\boldsymbol{q}}$ | DMFT momentum transfer parametrization, p. 14 |
| $\Lambda$ | Number of time slices, p. 19 |
| $\tilde{\Lambda}$ | NRG frequency discretization, p. 111 |
| $\mu$ | Chemical potential, p. 10 |
| $\mu$, $\nu$ | Band indices, p. 8 |
| $\rho(\epsilon)$ | Noninteracting DOS, p. 28 |
| $\tilde{\rho}(\epsilon)$ | Lattice input for $\sigma(\omega)$, p. 28 |
| $\tilde{\rho}'(\epsilon)$ | Derivative of $\tilde{\rho}(\epsilon)$, p. 195 |
| $\Sigma$ | Proper self-energy, p. 16 |
| $\sigma$ | Electronic spin, p. 8 |
| $\sigma_0$ | Prefactor for $\sigma(\omega)$, p. 193 |
| $\sigma(\omega)$ | Optical conductivity, p. 177 + 189 |
| $\tau$, $\Delta\tau$ | Imaginary time (p. 17), Trotter discretization (p. 19) |
| $\boldsymbol{\tau}$ | Hopping vector, p. 47 |
| $\boldsymbol{\tau}_l$ | Primitive lattice vector, p. 291 |
| $\hat{\psi}$ | Field operator, p. 8 |
| $\psi, \psi^*$ | Grassmann variables, p. 17 |
| $\omega_n$ | Matsubara frequency, p. 16 |
| $\omega_{\text{p}}$ | Plasma frequency, p. 180 |
| $a$ | Asymmetry of model DOS (p. 43), lattice spacing (p. 188) |
| $a^*$ | Ratio of scaled hopping amplitudes for hc lattice with NN and NNN hopping, p. 31 |

| | |
|---|---|
| $a_n$ | Weight of self-avoiding loops of length $n$, p. 44 |
| $a_{1g}$ | Atomic orbitals, p. 281 |
| $\mathcal{A}$ | Single-site action, p. 17 |
| $A(\omega)$ | Full single-particle spectral function, p. 23 |
| $A_\epsilon(\omega)$ | "$\boldsymbol{k}$-resolved" spectral function, p. 193 |
| $\boldsymbol{A}$ | Electromagnetic vector potential, p. 305 |
| $A$, $B$ | Sublattices of a bipartite lattice, p. 32 |
| $\hat{c}^\dagger_{i\sigma}$,$\hat{c}_{i\sigma}$ | Creation, annihilation operator, p. 9 |
| $\hat{c}^\dagger_{i\nu\sigma}$,$\hat{c}_{i\nu\sigma}$ | Creation, annihilation operator (multi-band), p. 8 |
| $c$ | Speed of light, p. 305 |
| $\boldsymbol{C}$ | General matrix or tensor (p. 293), covariance matrix (p. 157) |
| d | Differential operator |
| $\mathrm{d}\boldsymbol{x}$ | Infinitesimal volume $\mathrm{d}^d x$ |
| $d$ | Band character in atomic limit ($s$, $p$, $d$, $f$,...), p. 281 |
| $d$ | Dimension, p. 10 |
| $D_{\boldsymbol{q}}(\epsilon_1, \epsilon_2)$ | Two-particle density of states, p. 14 |
| $\mathcal{D}[\psi]$ | Functional differential, p. 17 |
| $\mathrm{erf}(x)$ | Error function (of $x$) |
| e | Euler's number ($\mathrm{e}^x \equiv \exp(x)$) |
| $\boldsymbol{e}_\alpha$ | Cartesic unit vector, p. 30 |
| $e$ | Electronic charge $e \approx -1.6 \times 10^{-19}$C, p. 6 |
| $e_g$, $e_g^\pi$ | Atomic orbitals, p. 281 |
| $\boldsymbol{E}$ | Electric field, p. 176 |
| $\mathcal{F}(x)$ | DOS transformation function, p. 49 |
| $F_0$ | Multi-band model parameter, p. 282 |
| $G$, $G^0$ | Full vs. noninteracting Green function, p. 16 |
| $G_\epsilon(i\omega_n)$ | "Momentum-dependent" Green function, p. 78 |
| $\mathcal{G}$ | Effective local bath propagator, p. 17 |
| $\mathrm{He}_n(x)$ | Hermite polynomial, p. 49 |
| $H$, $\hat{H}$ | Hamiltonian |
| $\hat{H}_{\mathrm{Hub}}$ | Hubbard Hamiltonian, p. 9 |
| $\mathrm{Im}\, x$ | Imaginary part (of $x$) |
| $\boldsymbol{j}$, $\hat{\boldsymbol{\jmath}}$ | Paramagnetic current density (operator), p. 186 |
| $\boldsymbol{J}$ | Total current, p. 176 |
| $\boldsymbol{k}$ | Lattice momentum, p. 8 |
| $K$ | Connectivity for Bethe pseudo-lattice, p. 40 |
| $K$, $K_0$ | Kinetic energy with (without) coupling to $\boldsymbol{A}$, p. 188 |
| $L$ | Number of ions (p. 6), number of lattice sites (p. 8) |
| $m$ | Electronic mass $m \approx 9.1 \times 10^{-31}$kg, p. 6 |
| $m^*$ | Effective (electron) mass, p. 76 |
| $\hat{n}_{i\sigma}$ | Occupation of lattice site $i$ for spin $\sigma$, p. 9 |
| $n$ | Average band filling, p. 10 |
| $n_{\epsilon,\sigma}$, $n_\epsilon$ | Momentum distribution function, p. 196 |

$n_\mathrm{f}(\omega)$ Fermi function, p. 193

$N_e$, $N_v$ Number of electrons (p. 6), number of valence electrons (p. 7)

$\mathcal{O}(x^n)$ Of the order $x^n$

$\hat{\boldsymbol{P}}$ Polarization operator, p. 190

$\boldsymbol{q}$ Lattice momentum, momentum transfer, p. 14

$\boldsymbol{Q}$ Antiferromagnetic wave vector, p. 14

$\mathrm{Re}\, x$ Real part (of $x$)

$r$ Distance, radius, p. 6

$\boldsymbol{R}$ Position of ion or lattice site, p. 6

$\hat{\boldsymbol{S}}_{i\nu}$, $\hat{\boldsymbol{s}}_{i\nu}$ Spin operator, p. 282

$t$ NN hopping matrix element, p. 9

$t_{ij}^{\nu}$ Hopping matrix element between sites $i$ and $j$ in band $\nu$, p. 8

$t_D$ Hopping matrix element for (taxi cab) distance $D$, p. 38 + 48

$t_{2g}$ Atomic orbitals, p. 281

$\mathrm{Tr}\{\}$ Trace

$T$ Temperature, p. 10

$U$ Hubbard on-site interaction, p. 9

$\boldsymbol{v}$ Generic vector

$\boldsymbol{v}_{\boldsymbol{k}}$, $\boldsymbol{v}_{\boldsymbol{k},\nu}$ Fermi velocity, p. 185

$V_0$ Multi-band model parameter, p. 282

$x$ Generic variable

$Z$ Lattice coordination number, p. 11

$Z$ Quasiparticle weight, p. 76

$\mathcal{Z}$ Partition function, p. 17

AF Antiferromagnet(ic), p. 10

ASA Atomic sphere approximation, p. 251

CDMFT Cellular (or cluster) DMFT, p. 286

DCA Dynamical cluster approximation, p. 286

DFT Density functional theory, p. 249

DMFT Dynamical mean-field theory, p. 11

DMRG Density matrix renormalization group, p. 289

DOS Noninteracting density of states, p. 14

ED Exact Diagonalization, p. 18

EELS Electronic energy loss spectroscopy, p. 183

FLEX Fluctuation-exchange approximation, p. 18

HF Hartree-Fock theory, p. 11

IPT Iterated perturbation theory, p. 18

L(S)DA Local (spin) density approximation, p. 250

LMTO Linear muffin-tin orbitals, p. 251

MEM Maximum entropy method, p. 25

NCA Non-crossing approximation, p. 18

| | |
|---|---|
| NN | Nearest neighbors, p. 9 |
| NNN | Next-nearest neighbors, p. 27 |
| NRG | Numerical renormalization group, p. 18 |
| PES | Photoemission spectroscopy, p. 258 |
| QMC | Quantum Monte Carlo, p. 19 |
| RDA | Random dispersion approximation, p. 288 |
| SIAM | Single impurity Anderson model, p. 11 |
| XAS | X-ray absorption spectroscopy, p. 268 |
| $f$-sum | Sum rule for $\sigma(\omega)$, p. 179 |
| fcc | (Hyper) face centered cubic (lattice), p. 32 |
| hc | Hypercubic (lattice), p. 27 |
| hd | Hyperdiamond (lattice), p. 291 |

# Introduction

Solid state theory aims at a description of intrinsic properties of solid materials by finding and extracting information from suitable models. A good model is complicated enough to capture interesting aspects of nature, but simple enough to be solvable (at least within some limits or within controlled approximations) and to provide insights about the associated mechanisms. Even if it was possible to construct a theory which exactly predicts all measurable properties of solids, more abstract models would still be necessary in order to classify the different properties and the different classes of physical systems and in order to identify which "ingredients" are important for some particular effect. In this sense, the presence of free parameters in *model Hamiltonian approaches* is not necessarily a shortcoming compared to *ab initio* methods, but a very useful handle for understanding phenomena in a general context. Often, different simplified models give reasonable descriptions for different properties of the same material. For a more realistic description one can then build up a hierarchy of more complex theories, thus broadening the range of validity and improving on the accuracy of agreement to experiment. The main advantage of a *microscopic model* over a phenomenological theory is that its inherent approximations are known so that it can, at least in principle, be made more realistic in a controlled way.

*Solids* are constituted of a macroscopic number of positively charged atomic nuclei and negatively charged electrons. In single crystals, the heavy nuclei can be assumed to form a rigid periodic lattice. Furthermore, due to the strong Coulomb force, a large fraction of the electrons is typically tightly bound in atomic-like shells around the ions. The resulting core ions then provide a periodic background potential for and shield the interaction between the remaining electrons. These *valence electrons* determine, to leading order, the electronic, magnetic, and thermal properties of solids. In this work, we will exclusively study the electronic system and mostly restrict our treatment to the valence electrons.

Even in the limit of vanishing interaction between the (valence) electrons, their dispersion is split by the lattice potential into an infinite number of energy bands (Bloch, 1928). Due to the Pauli exclusion principle each eigenstate, characterized by its momentum (within the Brillouin zone), band index, and spin, can only be occupied by one electron. Consequently, there is some trivial correlation between the electrons even at the noninteracting level. The (lattice-dependent) energy of noninteracting electrons is commonly referred to as *kinetic energy*. It is the *potential energy* associated with the (shielded) electron-electron *Coulomb interaction* which introduces genuine *correlations* between the valence electrons and makes the problem interesting and complicated.

The low-energy electronic properties of some materials can be reasonably well understood without explicitly taking electron-electron interactions into account. For good metals, this surprising fact is elucidated by the phenomenological Landau Fermi-liquid theory (Landau, 1957*a*; Landau, 1957*b*) which replaces electrons by noninteracting quasiparticles with a renormalized mass and a finite lifetime. Even when the noninteracting picture fails, phenomena such as long-range order can sometimes be described within a static mean-field theory; here one replaces the effect of the electron-electron interaction on each electron by that of a time averaged (and possibly spin-dependent) electronic charge density. In general however, in particular for transition metal compounds, such effective single-particle pictures fail to satisfactorily reproduce the interesting physical phenomena. These *strongly correlated electron systems*, for which the transition metal compounds are prototype examples due to the strong (and locally essentially unshielded) Coulomb interaction between their well-localized $3d$ and $4f$ valence orbitals, then call for theories which retain the full dynamics of the electronic correlation problem. Theories for such $d$ and $f$ systems typically take only few orbitals per lattice site into account and assume an effectively short-ranged Coulomb interaction. An extreme case is the *Hubbard model* where the Coulomb interaction is reduced to its on-site part (Hubbard, 1963; Gutzwiller, 1963; Kanamori, 1963), i.e., where electrons only interact with each other if they occupy the same lattice site (in the Wannier picture).

The *Dynamical Mean-Field Theory (DMFT)* is a nonperturbative approximation for strongly correlated electron models, namely variants of the Hubbard model, which becomes exact in the limit of infinite dimensionality or, equivalently, infinite lattice coordination number (Metzner and Vollhardt, 1989). It neglects (short-range) spatial correlations between the electrons while it retains their dynamical correlations which are crucial for the description of many interesting phenomena observed in correlated electron systems. The DMFT simplifies the lattice problem by mapping it onto a single-impurity model embedded in a medium that has to be determined self-consistently. Still, explicit solutions can only be obtained by application of either further approximations or of numerical techniques. The *quantum Monte Carlo* (QMC) method is numerically exact, i.e., its error can in principle be made arbitrarily small by increasing the numerical effort; even with today's high-performance computers the method is, however, restricted to not too low temperatures.

The low-temperature electronic properties of materials can be classified as insulating, metallic, or superconducting depending on whether the resistivity increases or decreases upon lowering the temperature or precisely vanishes (below some critical temperature), respectively. Changes in the resistivity and, in particular, phase transitions from a metallic to an insulating state can be induced by, e.g., change of stoichiometric composition, pressure, temperature, or magnetic field or by introducing disorder. In contrast to transitions to (possibly long-range ordered) band insulators or those induced by Anderson localization which can be understood on a noninteracting or static mean-field level, nonperturbative approaches are required for a quantitative theory of the correlation-induced *Mott metal-insulator transition (MIT)* which occurs at a point where potential and kinetic energy are of the same order.

The main focus of this work are studies of correlated electron systems near a Mott metal-insulator transition. In particular, we will present the first controlled DMFT calculation of the complete phase diagram of the fully frustrated single-band Hubbard model with semi-elliptic density of states at half filling using the QMC method. We will also perform the first calculations of the associated optical conductivity which do not depend on the assumption of anisotropy or disorder; these will be based on the general theory of densities of states and transport properties in high dimensions which is also developed in this thesis. In addition to these pure model studies, we will present results specific to the doped transition metal oxide $La_{1\text{-}x}Sr_xTiO_3$ which are obtained using the hybrid LDA+DMFT technique. This new method employs ab initio density functional theory in the local density approximation (LDA) for defining a general multi-band Anderson-Hubbard model which is then treated within the DMFT.

## Structure of this Thesis

In chapter 1, we introduce the general electronic Hamiltonian and its reduction to the Hubbard model. We characterize the DMFT and its relation with mean-field approximations to spin systems and present the DMFT mean-field equations as well as their solution using the auxiliary-field QMC method. Finally, we discuss the analytic continuation of imaginary-time Green functions to the real axis by the maximum entropy method (MEM).

In chapter 2, we study the relations between lattice types, frustration, and densities of states (DOS). On the basis of Monte Carlo computations of momentum sums, we also evaluate the convergence of the density of states to its infinite-dimensional limit. We present new insights on the fractal tree commonly referred to as Bethe lattice and on the impact of longer-range hopping for this model. We develop a new formalism that allows to construct models with hypercubic symmetry which reproduce an arbitrary target DOS in the limit of infinite dimensionality ($d = \infty$). Using this approach, we can for the first time define a regular and translationally invariant lattice with semi-elliptic DOS in $d = \infty$.

In the central chapter 3, we thoroughly explore the low-temperature properties of the fully frustrated Hubbard model with semi-elliptic DOS within the DMFT. We determine the boundaries of a coexistence region of metallic and insulating solutions with high accuracy, thereby resolving a controversy on the existence of a first-order transition within this model. In the course of these studies, we detect and correct deficiencies in previously used QMC schemes and develop an improved criterion for the detection of phase transitions. Going beyond previous work, we also develop a method for a reliable determination of the first-order transition line and perform various tests of the accuracy of the result. Finally, we compute local MEM spectra and suggest some methodological improvements to be used in future calculations.

Transport properties are discussed in chapter 4, where we carefully review the formalism and its range of applicability and develop new expressions for the optical $f$-sum rule which are valid in $d = \infty$. We show that the correct consideration of lattice properties is quantitatively important also in this limit and that simplifying assumptions made in a previous study on the impact of frustration lead to large errors. We point out the ambiguities associated with any DMFT calculation of transport properties for "the Bethe lattice" and review the possible concepts for making the problem well-defined; among these choices, the model defined in chapter 2 will be seen to have the most desirable properties. We then present accurate numerical results for the optical conductivity which are based on the MEM spectra computed in chapter 3.

In chapter 5, we give an introduction to the density functional theory and its local density approximation and introduce the recently developed hybrid LDA+DMFT scheme. We discuss the construction and simplification of the resulting multi-band Hubbard model and its solution within the DMFT as well as the extraction of photoemission and x-ray absorption spectra. In addition to numerical results obtained in a collaboration with Anisimov's group, we present new calculations and quantify the QMC discretization error. We also apply the formalism developed in chapter 2 in order to derive a definition of the optical conductivity which is compatible with the LDA DOS and present corresponding numerical results.

With the exception of the following introductory part, each chapter is supplemented by a conclusion. A brief overview over the results and new insights obtained in this thesis is given in the Summary. Finally, some aspects not directly within the main scope of this thesis are treated in the appendices.

# Chapter 1

# Models and Methods

In this chapter we set the framework for the main part of this thesis by introducing the models under investigation and the approximations and methods employed. Studies of abstract models represent an idealized and focused view on physics which should be complemented by a clear knowledge about the inherent limitations. Therefore, we will emphasize those steps of abstraction from a "complete model" to the models studied in this work which potentially cause properties of the resulting theory to differ qualitatively from seemingly corresponding properties of real materials. Later, we will make contact with these observations, e.g., in the discussion of the $f$-sum rule of the optical conductivity (see chapter 4) and discuss a more complete picture in the context of the LDA+DMFT method and its application to $La_{1\text{-}x}Sr_xTiO_3$ in chapter 5. Since the density functional theory (DFT) and its local density approximation (LDA) are not central to this work, these approaches are introduced directly in the context of their application in chapter 5. As far as practical, methods like the quantum Monte Carlo (QMC) algorithm or the maximum entropy method (MEM) are covered here on a descriptive level while original methodological contributions are mainly discussed in the following chapters.

In the following, we will first discuss the Hubbard model in a general context in Sec. 1.1, then its nonperturbative treatment within the dynamical mean-field theory (DMFT) in Sec. 1.2. We describe the quantum Monte Carlo method of solving the DMFT self-consistency equations in Sec. 1.3 and the maximum entropy method of obtaining real-time dynamical information by analytic continuation of QMC imaginary-time data in Sec. 1.4.

## 1.1 Hubbard Model

The one-band Hubbard model is *the* minimal lattice model for strongly correlated electron systems, i.e., for describing electronic properties of materials which cannot be treated without explicitly taking the Coulomb interaction between the (valence) electrons into account. Typical examples are the transition metals and their compounds. In subsections 1.1.1 – 1.1.4, we explore the microscopic foundation of the Hubbard model and review some of its basic properties and limitations. Extensions of the one-band Hubbard model which can potentially overcome some of these limitations are listed in App. A.1. While most of these extensions are beyond the scope of this thesis, the multi-band model (A.5) will be central to chapter 5.

### 1.1.1 Solid State Theory for Crystals

The typical condensed matter system contains a macroscopic number (usually $10^{20}$ or more) of atomic nuclei and electrons.[1] In this situation, a complete description of a particular system for given initial conditions is not possible. In fact, it might not be desirable since the truly interesting information would be hidden in an extensive amount of details. It is the concept of *statistical physics* to abstract from some specific system and view it as a realization of an ensemble of systems which share only some important features like the atomic constitution exactly but are allowed to differ in others. Among the observations that can be made for such ensembles are not only those that pertain to each individual realization (like energy conservation), but also others that involve averages over the ensemble, over time, and/or over space (like density-density pair correlation functions).[2] All results of this work will apply to the thermodynamic limit of a *(grand) canonical ensemble* of unbounded systems with no net charge.

Neglecting intra-nuclear effects and the gravitational, “weak”, and “strong” forces (which are extremely small on typical condensed matter length scales of $10^{-10}$ m to 1 m), we assume that condensed matter systems are composed of inert pointlike atomic nuclei and electrons which interact by the Coulomb $1/r$ potential (for a distance $r$) and have no internal degrees of freedom other than (possibly) spin. Although, in general, relativistic effects are important for the systems under consideration, they often contribute only a constant energy offset to the solid-state problem (e.g., for core states of heavy nuclei) or can be treated approximately in a nonrelativistic framework (e.g., spin-orbit coupling for valence electrons; see App. A.1). Thus, the usual starting point for condensed matter theory is *nonrelativistic quantum statistical mechanics* for charged electrons and ions with the Hamiltonian[3]

$$H = \sum_{i=1}^{N_e} \frac{\boldsymbol{p}_i^2}{2m} + \sum_{k=1}^{L} \frac{\boldsymbol{P}_k^2}{2M_k} + \sum_{i<j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \sum_{k<l} \frac{Z_k Z_l e^2}{|\boldsymbol{R}_k - \boldsymbol{R}_l|} - \sum_{i,k} \frac{Z_k e^2}{|\boldsymbol{r}_i - \boldsymbol{R}_k|} \tag{1.1}$$

Here, $\boldsymbol{r}_i$ ($\boldsymbol{R}_k$), $\boldsymbol{p}_i$ ($\boldsymbol{P}_i$), and $m$ ($M_k$) label the positions, momenta, and masses of electrons (ions), respectively; $e$ is the electronic charge, $N_e$ the total number of electrons, $L$ that of ions, and $Z_k$ the atomic number of the ion with index $k$. Charge neutrality requires that $N_e = \sum_k Z_k$. In the Gaussian system of units, $4\pi\varepsilon_0 = 1$ drops out of the expressions for the Coulomb energy.

[1] Exceptions to this rule are, e.g., mesoscopic systems where the physics may be dominated by finite-size effects like in the single-electron transistor.

[2] Note that for ergodic systems, the infinite-length limit of the time average by definition agrees with the ensemble average.

[3] The discussion of couplings to external electromagnetic fields is deferred to Sec. 4.2. We note at this point, however, that the assumption of instantaneous Coulomb interaction used here and in the following is exact within the Coulomb gauge.

Since the long-range mobility of ions is vanishingly small in the solid state (except for solid Helium), one can estimate the kinetic energy of the ions to be smaller than that of the electrons by a factor of $(m/M_k)^{1/2} \approx 10^{-2} - 10^{-3}$, which suggests using $m/M$ as a perturbation parameter where $M$ is, e.g., the smallest ion mass in the system. In the *adiabatic approximation* (Born and Oppenheimer, 1927) one treats the electronic problem for fixed ion coordinates; feeding back the electronic eigenenergies as a function of the ion coordinates then allows for an approximate treatment of the ionic contribution which reduces the error to $(m/M)^{3/4}$ (see, e.g., Czycholl, 2000). We will take the zeroth-order approximation, i.e., use immobile ions on a periodic lattice and view them as an external potential for the electrons. The effect of this considerable simplification should be kept in mind when later interpreting results, in particular with respect to transport properties.

### 1.1.2 Electronic Lattice Models

The resulting purely electronic Hamiltonian,

$$H = \sum_{i=1}^{N_e} \frac{\boldsymbol{p}_i^2}{2m} + \sum_i V(\boldsymbol{r}_i) + \sum_{i<j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}, \tag{1.2}$$

where the external potential $V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R}_\alpha)$ is periodic on the lattice (with primitive lattice vectors $\boldsymbol{R}_\alpha$ for $\alpha = 1, \ldots, d$, where $d$ is the dimension), defines the electronic lattice problem.

A full solution of this problem would have to include energy contributions (of positive and negative sign) of at least the highest x-ray absorption edge of the constituting atoms, which can be more than 6 orders of magnitude larger than energy differences between, e.g., metallic and insulating phases considered in this work. In this situation it is useful to distinguish between *core electrons* and *valence electrons* where the former are assumed to be tightly bound to the nuclei. The resulting new “elementary particles”, the *core ions*, then define a new, drastically weaker crystal potential. Since core ions are polarizable they also modify the infinite-range $1/r$ potential of the bare Coulomb electron-electron interaction to some shorter ranged form.[4] Provided that these effects can be quantified, we may restrict the treatment to few valence electrons per lattice site,

$$H = \sum_{i=1}^{N_v} \frac{\boldsymbol{p}_i^2}{2m} + \sum_{i=1}^{N_v} V^{\text{ion}}(\boldsymbol{r}_i) + \sum_{i=1}^{N_v-1} \sum_{j=i+1}^{N_v} V^{ee}(\boldsymbol{r}_i, \boldsymbol{r}_j), \tag{1.3}$$

where $N_v$ is the number of valence electrons and $V^{\text{ion}}$ the shielded lattice potential. Here, we have implicitly assumed an instantaneous effective interaction of the density-density type which is certainly not exact. We also point out that the (screened) effective interaction potential $V^{ee}$ between valence electrons is in general neither homogeneous nor isotropic since the shielding charges are localized. Now switching

[4]Valence electrons can also take part in shielding the electron-electron interaction. This effect should, however, not be built into a microscopic valence-electron model, but result from its solution.

to the occupation number formalism ("second quantization"; see, e.g., Fetter and Walecka, 1971) in real space, the Hamiltonian for electrons with spin $\sigma$ takes the form

$$\hat{H} = \hat{H}_0 + \hat{H}_{\mathrm{int}}, \tag{1.4}$$

where

$$\begin{aligned} \hat{H}_0 &= \sum_{\sigma} \int \mathrm{d}\boldsymbol{r}\, \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{r}) \left[ -\frac{\hbar^2}{2m}\Delta + V^{\mathrm{ion}}(\boldsymbol{r}) \right] \hat{\psi}_{\sigma}(\boldsymbol{r}) & (1.5) \\ \hat{H}_{\mathrm{int}} &= \frac{1}{2} \sum_{\sigma\sigma} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, V^{\mathrm{ee}}(\boldsymbol{r},\boldsymbol{r}')\, \hat{n}_{\sigma}(\boldsymbol{r})\, \hat{n}_{\sigma'}(\boldsymbol{r}')\,. & (1.6) \end{aligned}$$

Here, $\hat{\psi}_{\sigma}(\boldsymbol{r}), \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{r})$ are field operators which annihilate and create electrons of spin $\sigma$ at site $\boldsymbol{r}$, respectively; $\hbar = h/2\pi$ is Planck's constant, $\Delta$ the Laplace operator, and $\hat{n}_{\sigma}(\boldsymbol{r}) = \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{r})\hat{\psi}_{\sigma}(\boldsymbol{r})$ the operator measuring the local density of electrons with spin $\sigma$ at position $\boldsymbol{r}$.

### 1.1.3 Wannier Representation

In terms of the lattice momentum $\boldsymbol{k}$ and Bloch eigenfunctions $\phi_{\nu\boldsymbol{k}}(\boldsymbol{r})$ of the noninteracting[5] Hamiltonian (1.5), we may introduce Wannier functions predominantly localized at site $\boldsymbol{R}_i$ by

$$\chi_{i\nu}(\boldsymbol{r}) = \frac{1}{\sqrt{L}} \sum_{\boldsymbol{k}} \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{R}_i}\; \phi_{\boldsymbol{k}\nu}(\boldsymbol{r})\,, \tag{1.7}$$

where $L$ is the number of lattice sites, and thus construct creation and annihilation operators $\hat{c}^{\dagger}_{i\nu\sigma}$,$\hat{c}_{i\nu\sigma}$ for electrons with spin $\sigma \in \{\uparrow,\downarrow\}$ in the band $\nu$ at site $\boldsymbol{R}_i$ as

$$\hat{c}^{\dagger}_{i\nu\sigma} = \int \mathrm{d}\boldsymbol{r}\; \chi_{i\nu}(\boldsymbol{r})\; \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{r}) \quad \longleftrightarrow \quad \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{r}) = \sum_{i\nu}\; \chi^{*}_{i\nu}(\boldsymbol{r})\hat{c}^{\dagger}_{i\nu\sigma}\,. \tag{1.8}$$

Using this basis, the Hamiltonian may be written in the lattice representation as (Hubbard, 1963; Hubbard, 1964*a*)

$$\hat{H} = \sum_{i\nu j\sigma} t^{\nu}_{ij}\, \hat{c}^{\dagger}_{i\nu\sigma}\; \hat{c}_{j\nu\sigma} + \frac{1}{2} \sum_{\nu\nu'\mu\mu'} \sum_{ijmn} \sum_{\sigma\sigma'} \mathcal{V}^{\nu\nu'\mu\mu'}_{ijmn}\; \hat{c}^{\dagger}_{i\nu\sigma}\; \hat{c}^{\dagger}_{j\nu'\sigma'}\; \hat{c}_{n\mu'\sigma'}\; \hat{c}_{m\mu\sigma}\,, \tag{1.9}$$

where the matrix elements are given by

$$\begin{aligned} t^{\nu}_{ij} &= \int \mathrm{d}\boldsymbol{r}\; \chi^{*}_{i\nu}(\boldsymbol{r}) \left[ -\frac{\hbar^2}{2m}\Delta + V^{ion}(\boldsymbol{r}) \right] \chi_{j\nu}(\boldsymbol{r}) & (1.10) \\ \mathcal{V}^{\nu\nu'\mu\mu'}_{ijmn} &= \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\; V^{ee}(\boldsymbol{r},\boldsymbol{r}')\, \chi^{*}_{i\nu}(\boldsymbol{r})\; \chi^{*}_{j\nu'}(\boldsymbol{r}')\; \chi_{n\mu'}(\boldsymbol{r}')\; \chi_{m\mu}(\boldsymbol{r})\,. & (1.11) \end{aligned}$$

[5]In principle, we mean by "noninteracting" the absence of interaction between electrons. In practice, however, it is often necessary to assume that at least a long-range Hartree part of the interaction between the valence electrons is also included as a part of the kinetic energy and used for the definition of the Wannier orbitals (Gebhard, 1997) which then better justifies a very short-ranged form of the remaining explicit interaction. In any case, however, we want to assume that the external potential need not be determined self-consistently (which might affect conclusions about phenomena such as phase separation).

Several observations can be made at this point: First of all, in contrast to the field-operator representation defined in the continuum, in the Wannier representation even the bare Coulomb interaction not only depends on the densities $\hat{n}_{i\sigma} = \hat{c}^\dagger_{i\sigma}\hat{c}_{i\sigma}$ but also contains explicit off-diagonal contributions, for example, on-site Hund's rule couplings and Heisenberg nearest-neighbor exchange couplings (see App. A.1). Secondly, when the explicit treatment is restricted to a few valence electrons, the interaction $V^{ee}(\boldsymbol{r}, \boldsymbol{r}\,')$ cannot be expressed in the translationally invariant form $V^{ee}(\boldsymbol{r} - \boldsymbol{r}\,')$. This affects symmetry considerations for the matrix elements $\mathcal{V}^{\nu\nu'\mu\mu'}_{ijmn}$.[6] Finally, the matrix elements of the kinetic energy $t^\nu_{ij}$, called *hopping amplitudes*, can be expected to fall off rapidly with increasing $|\boldsymbol{R}_i - \boldsymbol{R}_j|$ since the associated atomic orbitals and, consequently, the corresponding Wannier functions hardly overlap, at least for $3d$ or $4f$ valence electrons (see, e.g., Ashcroft and Mermin, 1976).

### 1.1.4 One-band Hubbard Model

The simplified valence-electron model (1.9) still contains an infinite number of input parameters that cannot be reliably determined from first principles.[7] This situation calls for a minimalistic approach where the number of parameters is chosen to be just large enough to capture the interesting effects, at least qualitatively.

If we restrict the Hamiltonian (1.9) to one valence band ($\nu = \nu' = \mu = \mu' = 1$) with isotropic hopping to nearest neighbors (NN) only[8] and assume "perfect" screening, i.e., choose

$$t^\nu_{ij} \equiv t_{ij} = \begin{cases} -t & \text{if } i \text{ NN of } j \\ 0 & \text{otherwise,} \end{cases} \tag{1.12}$$

$$\mathcal{V}^{\nu\nu'\mu\mu'}_{ijmn} \equiv \mathcal{V}_{ijmn} = U\,\delta_{ij}\delta_{im}\delta_{in}\,, \tag{1.13}$$

we arrive at the one-band Hubbard model,

$$\hat{H}_{\text{Hub}} = -t \sum_{\langle i,j\rangle,\sigma} \left(\hat{c}^\dagger_{i\sigma}\hat{c}_{j\sigma} + \text{h.c.}\right) + U \sum_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\,. \tag{1.14}$$

Here, $t$ parameterizes the kinetic energy, $U$ the Coulomb interaction, and the bracket $\langle i, j\rangle$ restricts the sum to nearest-neighbor pairs $i, j$. The operator $\hat{n}_{i\sigma} = \hat{c}^\dagger_{i\sigma}\hat{c}_{i\sigma}$ measures the occupancy of the site $i$ with electrons of spin $\sigma$. Consequently, $\hat{n}_{i\uparrow}\hat{n}_{i\downarrow}$ is the double occupancy, i.e., its expectation value corresponds to the density of doubly occupied sites. Studies of this model, originally proposed for the description of

[6] Formally, higher symmetry can be retained by applying (1.10) and (1.11) to the full unshielded problem. A truncation of the inherent sums over nearest-neighbor shells, however, can only be justified after the divergencies have been removed by renormalization, i.e., shielding.

[7] For an approximate material specific determination of matrix elements using density functional theory, see Sec. 5.3.

[8] Since the methods introduced in the following do not rely on this restriction, we will later also consider extended models with hopping beyond the NN shell. In this context, we refer to the model with NN hopping only as the "pure" Hubbard model.

itinerant ferromagnetism (Hubbard, 1963; Gutzwiller, 1963; Kanamori, 1963), have become a major subfield of condensed matter theory. Since the hopping and interaction terms in (1.14) do not commute, they cannot be simultaneously diagonalized which leaves the treatment of the Hubbard model highly nontrivial. In the (grand) canonical ensemble, the model is fully specified by choosing $t$ and $U$ plus the underlying lattice, the temperature $T$, and the average band filling $n := \frac{1}{L}\sum_{i=1}^{L}\langle \hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}\rangle$ (or, equivalently, the chemical potential $\mu$).

For dimension[9] $d = 1$, many properties of the Hubbard model such as the ground state wave function, magnetic susceptibility, Drude weight, and excitation spectra can be exactly calculated using the Bethe ansatz (Bethe, 1931; Lieb and Wu, 1968) in the thermodynamic limit. For higher dimensions, little is known rigorously. The generic low-temperature phase for half-filled ($n = 1$) bipartite lattices (for a definition, see Sec. 2.1) and $d \geq 2$ (only at $T = 0$ for $d = 2$) is antiferromagnetic (AF), i.e., both sublattices have a finite (opposite) magnetization. This phase is found in weak coupling ($U \to 0$; qualitatively in Hartree-Fock, i.e., Slater mean-field theory) and strong coupling (Anderson's "super-exchange" mechanism) for $d > 2$ and can also be approached by renormalization group methods, at least for $d = 2$ (Halboth and Metzner, 2000; Honerkamp and Salmhofer, 2001). Ferromagnetism is rigorously established in the pure Hubbard model only in the thermodynamically irrelevant case of half filling minus one electron at $U = \infty$ for a variety of lattice types (in dimensions $d \geq 2$; in one dimension it is excluded by the Lieb-Mattis theorem) and also for so-called "flat band systems" (Mielke and Tasaki, 1993). For $n = 1$ and $U \to \infty$, the Hubbard model can be mapped to the spin-1/2 Heisenberg model for which a vast literature exists; less is known about the $t$-$J$ model to which the Hubbard model is closely related for $U \to \infty$ and $n \neq 1$. In $d = 2$, the Hubbard model is widely used for modeling high-$T_c$ superconductors. While the existence of long-range order in the $d$-wave pairing is rigorously excluded (Mermin and Wagner, 1966; Su and Suzuki, 1998) at finite temperatures for narrow-band models, instabilities towards such order can still be inferred from corresponding correlation functions or from calculations at $T = 0$. For a more comprehensive overview over the physics of the Hubbard model see, e.g., the proceedings edited by Baeriswyl, Campbell, Carmelo, Guinea, and Louis (1995).

The one-band Hubbard model represents a highly idealized view on strongly correlated materials. Its applicability to $d$ or $f$ electron systems is a priori questionable since the partially filled bands correspond to atomic orbitals which are 5-fold and 7-fold degenerate (for each spin direction), respectively. While this degeneracy is at least partially lifted by crystal field effects, a more realistic description can potentially be achieved using multi-band versions of the Hubbard model, possibly also

[9] Note that for a lattice model with local Coulomb interaction, the dimensionality of the problem is completely determined by the hopping matrix elements $t_{ij}$ in (1.9). The physics of the problem is of lower dimension $d$ than the host lattice if the sublattice formed by all sites connected to one given site via $t_{ij}$ is topologically equivalent to a *regular* lattice of dimension $d$ since then the Hamiltonian factorizes into identical contributions from each of the unconnected sublattices. This fact explains why even some (unisotropic) bulk systems in our three-dimensional world may be modeled as one-dimensional or two-dimensional. Deviations from this idealized behavior arise both from residual hopping matrix elements and from long-range Coulomb interactions.

taking intersite interactions into account. Apart from just keeping additional terms of (1.9), one may also choose a semi-effective approach including Kondo type interactions, e.g., for modeling manganites. Compounds with several inequivalent ions per unit cell may be modeled directly or by using orbitals on an effective Bravais lattice. Further possible extensions which introduce qualitatively new physics [compared to (1.14)] are disorder, phonons, and spin-orbit interaction. A more detailed discussion (including Hamiltonians and references) can be found in App. A.1. While most of the material provided there is intended to put our work into perspective, we will analyze the implications of (off-diagonal) disorder for magnetic order, the density of states, and transport properties in Sec. 2.2, chapter 3, and subsection 4.4.4 and study a multi-band model in chapter 5. For the moment, however, we will concentrate on the (clean) one-band case for which we can achieve the high accuracy (in the limit of infinite dimensionality, see below) needed for a reliable determination of the low-temperature phase diagram in chapter 3.

## 1.2 Dynamical Mean-Field Theory

Even though the Hubbard model (1.14) is a drastically simplified model for strongly correlated electron systems, very few exact statements can be made for lattice dimensions $d > 1$ (see above). Therefore, in general, additional approximations have to be made. One particularly successful approximation for the Hubbard model and its extensions is the Dynamical Mean-Field Theory (DMFT), also known as the *limit of infinite dimensions* (Metzner and Vollhardt, 1989). In the following, we will first characterize the DMFT in comparison to other approaches used in the literature and discuss its unique advantages in the context of the paramagnetic metal-insulator transition (MIT) studied in chapter 3. We will then, in turn, explore the implications of a large coordination number $Z$ for spin models, noninteracting fermions, and diagrammatic perturbation expansions for the Hubbard model and see that (for appropriate scaling) the Hubbard model indeed remains nontrivial. It can be mapped to a single impurity Anderson model (SIAM) plus a self-consistency condition in the limit $Z \to \infty$. Possible extensions of the theory towards lower dimensions are referred to App. A.1.

Approximate analytic methods are only reliable if they are controlled by a parameter which becomes small in some limit of the problem. Typical examples are weak-coupling perturbation theories in $U/t$ of first order, i.e., Hartree-Fock theory (HF), a *static* mean-field approximation (Penn, 1966), or of second order (Georges and Yedidia, 1991; van Dongen, 1994*a*). Further examples are strong-coupling expansions in $t/U$ (Harris and Lange, 1967; Takahashi, 1977; van Dongen, 1994*b*), and expansions for high temperature or for low density $n$, i.e., the $T$-matrix approximation for $d \geq 2$ (Abrikosov and Khalatnikov, 1958; Galitskii, 1958; Landau, 1959; Engelbrecht, Randeria, and Zhang, 1992).

The DMFT is a controlled approach in the above sense; its small parameter is the inverse coordination number $1/Z$. It is a conserving approximation which guarantees thermodynamic consistency. Due to its nonperturbative character, the DMFT is not restricted to small or large values of $U$. Since the theory is formulated in the thermodynamic limit, it yields continuous spectra and is per se free from finite-size problems. The local dynamics of the correlation problem, i.e., on-site correlations, are retained; only off-site short-range correlations are neglected. The mapping to a SIAM implies a considerable reduction in technical complexity. Still, the resulting problem is too complicated for a general analytic or computationally inexpensive numerical solution.

For a description of the MIT in $d = 3$ (where $1/Z$ has a value, e.g., of 1/6, 1/8, and 1/12 for the simple cubic, body centered cubic, and face centered cubic lattice, respectively) the limit of large coordination number seems a reasonable starting point. Only a nonperturbative method which keeps the local dynamical correlations can reliably characterize the MIT, since this transition is intrinsically an intermediate-coupling phenomenon (in contrast, e.g., to the transition to antiferromagnetic order which can be understood both at weak and strong coupling within perturbation theory). The thermodynamic limit is particularly important in the context of MITs as finite systems are always insulators. Provided that short-range fluctuations do not completely change the physics, we can hope to capture the essential correlation physics within the DMFT, at least when the impurity part is solved numerically exactly as in the QMC algorithm used throughout this work (see Sec. 1.3).

### 1.2.1 Limit $Z \to \infty$ for Spin Models

It has been known for a long time that the results of the Weiss mean-field approximation (Weiss, 1907) become exact for models with localized spins in the limit of infinite coordination number $Z \to \infty$ (Brout, 1960).[10] The Hamilton operator of the isotropic spin-1/2 nearest-neighbor Heisenberg model reads

$$\hat{H}_{\text{Heisenberg}} = -J \sum_{\langle ij \rangle} \hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j = -J \sum_{\langle ij \rangle} \left( \hat{S}_i^z \hat{S}_j^z + \frac{1}{2}(\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+) \right) , \tag{1.15}$$

where the components of the spin operators obey the commutation rules $[\hat{S}_i^z, \hat{S}_j^\pm]_- = \pm\delta_{ij}\hat{S}_i^\pm$ and $[\hat{S}_i^+, \hat{S}_j^-]_- = 2\delta_{ij}\hat{S}_i^z$. In order for the energy per lattice site to remain finite when $Z \to \infty$, one has to scale the exchange interaction as $J = J^*/Z$, where $J^*$ is independent of $Z$. Defining the average spin $\hat{\boldsymbol{h}}_i$ of the nearest neighbors (NN) of lattice site $i$,

$$\hat{\boldsymbol{h}}_i = \frac{1}{Z} \sum_{j \text{ NN of } i} \hat{\boldsymbol{S}}_j \,, \tag{1.16}$$

one finds that[11] $[\hat{\boldsymbol{h}}_i, \hat{\boldsymbol{h}}_j] = \mathcal{O}(1/Z)$(since only $Z$ terms are nonvanishing) and that $\hat{\boldsymbol{h}}_i$ commutes with the Hamiltonian (1.15) to leading order (see, e.g., Gebhard, 1997).

[10]Note that Brout does not explicitly make the connection from the limit of infinite coordination number to the limit of infinite dimensionality but rather to the limit of infinite range of the exchange interaction which he assumes to be approached in the physical high-density limit of spin systems.

[11]Here and in the following, $\mathcal{O}(x^n)$ means "of the order of $x^n$", i.e., that the functional dependence on $x$ (in the limit $x \to 0$) does not contain terms with lower powers than $n$.

Thus, the operators $\hat{\boldsymbol{h}}_i$ correspond to conserved quantities and may be replaced by their expectation values. Choosing the $z$ axis as the quantization axis, this procedure is equivalent to an application of the Hartree decoupling scheme $\hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j \longrightarrow \hat{S}_i^z \hat{S}_j^z \longrightarrow \langle \hat{S}_i^z \rangle \hat{S}_j^z + \hat{S}_i^z \langle \hat{S}_j^z \rangle - \langle \hat{S}_i^z \rangle \langle \hat{S}_j^z \rangle$ to (1.15) which leads to the mean-field Hamiltonian

$$\hat{H}_{\mathrm{MF}} = -2J^* \sum_i \hat{S}_i^z h_i + J^* \sum_i \langle \hat{S}_i^z \rangle h_i \ . \tag{1.17}$$

The same form is obtained in the $Z \to \infty$ limit of the Ising Hamiltonian $\hat{H}_{\mathrm{Ising}} = -J \sum_{\langle ij \rangle} \hat{S}_i^z \hat{S}_j^z$, the anisotropic limit of the Heisenberg model.

On the mean-field level, the thermodynamics of the ferromagnetic ($J > 0$) and antiferromagnetic ($J < 0$) models are the same. However, for the quantum Heisenberg model in $d = 3$, ferromagnetic and antiferromagnetic order differ substantially in character since the fully polarized antiferromagnet is not an eigenstate. The mean-field critical temperature of $T_{\mathrm{c}}^{\mathrm{MF}}/J^* = 0.5$ can be compared to the numerically exact Curie temperature of the ferromagnetic 3-$d$ Heisenberg model where $T_{\mathrm{c}}/J^* = 0.28$, 0.325, and 0.346 for simple cubic (sc), body centered cubic (bcc), and face centered cubic (fcc) lattices, respectively (Kittel, 1996). For the Ising model, a thermodynamically consistent approximation which also captures $\mathcal{O}(1/Z)$ corrections to the mean-field theory is given by the so-called spherical model (Berlin and Kac, 1952) which also applies to the Heisenberg model for $T > T_{\mathrm{c}}$ (Brout, 1960; Brout, 1961).[12] An application of this $\mathcal{O}(1/Z)$ theory to the 3-d ferromagnetic Heisenberg model yields $T_{\mathrm{c}}/J^* = 0.37$, 0.36, and 0.33 for sc, bcc, and fcc lattices, respectively (Brout, 1960). Here, $1/Z$ corrections already greatly improve the mean-field estimates for $Z \sim 10$.

### 1.2.2 Limit $Z \to \infty$ for Fermions

In order to define a nontrivial $Z \to \infty$ limit of the Hubbard model (1.14) one has to scale its parameters so that the competition between kinetic and potential energy is retained. Since $U$ parameterizes the local interaction, it is independent of the coordination number. The scaling (Metzner and Vollhardt, 1989)

$$t = t^*/\sqrt{Z} \tag{1.18}$$

of the nearest neighbor (NN) hopping amplitudes ensures constant, finite variance of its energy eigenvalue spectrum,

$$< \epsilon^2 >= t^{*2} \tag{1.19}$$

as derived for arbitrary (pseudo-) lattice type, dimension, and hopping range in subsection 2.2.2. A convenient class of lattices with arbitrarily high (even) coordination number is formed by the hypercubic (hc) lattices which for dimension $d$ (with $Z = 2d$)

[12] Among the pathologies encountered in earlier $1/Z$ expansions for spin models is a violation of the sum rule $\sum_{i=1}^{N} (S_i^z)^2 = N/4$. An enforcement of this equality using a Lagrange parameter leads to a reselection of diagrams of order $1/Z^2$. Only then one observes, e.g., a divergence of the susceptibility for $T \to T_{\mathrm{c}}^+$ as required for a 2[nd] order phase transition (Brout, 1961).

and unit lattice spacing have the dispersion

$$\epsilon_{\boldsymbol{k}} = -2t \sum_{l=1}^{d} \cos(k_l) \ . \tag{1.20}$$

In the limit $d \to \infty$ and for generic $\boldsymbol{k}$, this expression turns into an infinite sum over statistically independent contributions to which the central limit theorem applies. Consequently, the noninteracting density of states (DOS) acquires a Gaussian form,

$$\rho^{\mathrm{hc}}(\epsilon) = \frac{1}{\sqrt{2\pi}\, t^*} \exp\left(-\frac{\epsilon^2}{2t^{*2}}\right) \ . \tag{1.21}$$

One might be disturbed by the observation that this DOS is unbounded both at low and high energies which makes the low-density limit $n \to 0$ unphysical since the kinetic energy per electron diverges. This pathology is easily linked to the additive independent energy contributions from each dimension to the total dispersion for the hc lattice.[13] In fact, an infinite number (of measure 0) of so-called nongeneric momenta $\boldsymbol{k}$ with $\epsilon_{\boldsymbol{k}} = \pm\infty$ exists such as the center $\mathbf{0} := (0,0,0,\dots)$ of the Brillouin zone or the antiferromagnetic wave vector $\boldsymbol{Q} := (\pi,\pi,\pi,\dots)$ (for lattice spacing $a = 1$). For finite electron density $n$, however, all states with $\epsilon_{\boldsymbol{k}} = -\infty$ ($\epsilon_{\boldsymbol{k}} = \infty$) remain full (empty) at all finite temperatures and values of the interaction $U$ and thus essentially drop out of the problem. Also, the energy per lattice site remains finite for arbitrary filling. An observation related to the appearance of infinite energy contributions is that, for $d \to \infty$, the 2-particle density of states with momentum transfer $\boldsymbol{q}$

$$D_{\boldsymbol{q}}(\epsilon_1, \epsilon_2) := \frac{1}{L} \sum_{\boldsymbol{k}} \delta(\epsilon_1 - \epsilon_{\boldsymbol{k}+\boldsymbol{q}})\delta(\epsilon_2 - \epsilon_{\boldsymbol{k}}) \tag{1.22}$$

only depends on $\boldsymbol{q}$ via a scalar $\eta$ and reads (for hc DOS)

$$D_{\boldsymbol{q}}(\epsilon_1, \epsilon_2) = \frac{\exp\left(-\frac{(\epsilon_1+\epsilon_2)^2}{4t^{*2}(1+\eta_{\boldsymbol{q}})}\right)}{t^*\sqrt{2\pi(1+\eta_{\boldsymbol{q}})}} \; \frac{\exp\left(-\frac{(\epsilon_1-\epsilon_2)^2}{4t^{*2}(1-\eta_{\boldsymbol{q}})}\right)}{t^*\sqrt{2\pi(1-\eta_{\boldsymbol{q}})}} \ , \tag{1.23}$$

which implicitly defines $\eta_{\boldsymbol{q}}$. Specifically, $\eta_{\boldsymbol{q}} = \epsilon_{\boldsymbol{q}}/(t^*\sqrt{2d})$ for the hc lattice. Evidently, $\eta_{\boldsymbol{q}}$ vanishes generically (i.e., for all momenta with finite energy) and thus $D_{\boldsymbol{q}}$ factorizes for all generic $\boldsymbol{q}$ (Müller-Hartmann, 1989*a*; Müller-Hartmann, 1989*b*; van Dongen, Gebhard, and Vollhardt, 1989)

$$D_{\boldsymbol{q}}(\epsilon_1, \epsilon_2) = \rho(\epsilon_1)\, \rho(\epsilon_2) \quad \text{for generic } \boldsymbol{q}. \tag{1.24}$$

Physically, it is important that some correlations remain in $D_{\boldsymbol{q}}(\epsilon_1, \epsilon_2)$ even in the limit $Z = \infty$. There is necessarily always full correlation for $\boldsymbol{q} = \mathbf{0}$: $\eta(\mathbf{0}) = 1$. Furthermore, perfect nesting implies $\eta(\boldsymbol{Q}) = -1$ for any bipartite lattice.[14]

[13] We will explore alternative lattice types in chapter 2 and find that sharp band edges can be obtained in the limit $Z \to \infty$ not only for the Bethe pseudo-lattice, but also for regular lattices (with full momentum space). For the rest of this section, however, we focus on the more conventional hc lattice.

[14] Note that the presence of extended hopping along the axes can destroy perfect nesting (Müller-Hartmann, 1989*a*) without changing the DOS in $d = \infty$; see subsection 2.1.3.

Little is usually said in the literature about the distribution of nongeneric momenta and the topology of Fermi surfaces in high dimensions. Even Gebhard's (1997) motivation for the random dispersion approximation (RDA, see below) which is based on the generic factorization (1.24) of the 2-particle DOS is somewhat vague and partially incorrect in this respect. In App. A.2, we attempt a more careful characterization of generic and nongeneric momenta. While the modifications relative to Gebhard's view do not explain the discrepancy between RDA result and all recent DMFT results for the MIT (see chapter 3), they shed more light on foundations of the DMFT which will prove useful for our development of a generalized dispersion formalism in Sec. 2.3.

### 1.2.3 Simplifications for the Hubbard Model in $Z \to \infty$

Up to now we have concentrated on the noninteracting tight-binding part of the Hubbard model Hamiltonian, since the local Coulomb interaction remains unscaled for $Z \to \infty$. We have seen that the dependence of the noninteracting 2-particle DOS (1.22) on the momentum transfer $\boldsymbol{q}$ becomes essentially irrelevant in this limit (i.e., for generic $\boldsymbol{q}$). This observation can be generalized to higher order correlation functions (see Sec. 4.3). The connection to the interacting problem is made by the real-space diagrammatic perturbation expansion where properties of the full interacting system such as its single-particle Green function $G_{ij,\sigma}$ are expressed in terms of bare (noninteracting) Green functions $G^0_{ij,\sigma}$ and interactions. For the Hubbard model with only local interactions, interaction lines reduce to local 4-leg vertices.[15]

Let us show by power counting that for $Z = \infty$ all pairs of vertices $i$ and $j$ must coincide when they are connected by at least three independent lines: since the hopping is scaled as $t = t^*/\sqrt{Z}$, each (direct) path made of fermion propagators connecting $i$ to $j$ involves a factor of order $Z^{-\|i-j\|/2}$. For fixed site $i$, we have at most $Z^{\|i-j\|}$ choices for a "first" path connecting $i$ to $j$ or a topologically equivalent site. In all regular lattices, this first path selects $\|i-j\|$ directions. Since $j$ is now fixed, the freedom of choice for the second and additional paths connecting $i$ and $j$ is generically reduced to a permutation of these directions which introduces additional prefactors independent of $Z$. For a total of $P_{ij}$ independent paths the scaling factor[16] is then $\mathcal{O}(Z^{\|i-j\|(1-P_{ij}/2)})$ and vanishes in high dimensions for $P_{ij} \geq 3$.[17] Thus, the proper self-energy becomes local for $Z \to \infty$,

$$\Sigma_{ij,\sigma}(\omega) \stackrel{Z\to\infty}{\longrightarrow} \delta_{ij}\Sigma_{\sigma}(\omega), \tag{1.25}$$

which can be directly used to eliminate one summation over lattice sites in the fol-

[15]NN interactions of the density-density type [cf. (A.7)] reduce to their Hartree contribution and can, thus, be easily included in all computations (Müller-Hartmann, 1989*a*).

[16]Obviously, these considerations do not directly apply to the fully frustrated model with infinite range hopping which has been studied numerically in the disordered case (Georges, Kotliar, and Krauth, 1993) and been analytically solved in the clean case (van Dongen and Vollhardt, 1989).

[17]Note that this argument relies on the dimension being much larger than $P_{ij}$ and, ultimately, larger than the order of the diagram. Conversely, a generalization to finite dimensions would have to stop at finite-order perturbation theory (cf. footnote 5 on page 30).

lowing general real-space Dyson equation (for a time-invariant system),

$$G_{ij,\sigma}(\omega) = G^0_{ij,\sigma}(\omega) + \sum_{kl} G^0_{ik,\sigma}(\omega)\Sigma_{kl,\sigma}(\omega)G_{lj,\sigma}(\omega)\,. \tag{1.26}$$

This is also the defining equation for the proper self-energy.[18] For translationally invariant systems one can simplify this equation by Fourier transformation,

$$G^{-1}_\sigma(\boldsymbol{k},\omega) = (G^0_\sigma(\boldsymbol{k},\omega))^{-1} - \Sigma(\boldsymbol{k},\omega); \quad \Sigma(\boldsymbol{k},\omega) \equiv \Sigma(\omega)\,. \tag{1.27}$$

Consequently, the local Green function can be expressed in terms of the noninteracting DOS and the self-energy alone,

$$G_\sigma(\omega) \equiv G_{ii,\sigma}(\omega) = \frac{1}{V_B}\sum_{\boldsymbol{k}} \frac{1}{\omega+\mu-\epsilon_{\boldsymbol{k}}-\Sigma_\sigma(\omega)} \tag{1.28}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}\epsilon \frac{\rho(\epsilon)}{\omega+\mu-\epsilon-\Sigma_\sigma(\omega)}\,. \tag{1.29}$$

An alternative derivation of the $\boldsymbol{q}$-independence of the proper self-energy is based on the observation that the Laue function $\delta^*(\boldsymbol{q}) = \sum_{\boldsymbol{K}} \delta(\boldsymbol{q}+\boldsymbol{K})$ (with reciprocal lattice vectors $\boldsymbol{K}$) becomes effectively momentum-independent when applied to the internal vertices of a diagram (such as a vertex of the inserted proper self-energy). Thus, the self-energy cannot depend on the external momentum and is consequently momentum-independent (Müller-Hartmann, 1989*a*).

The absence of momentum dependence in the self-energy greatly simplifies the treatment of the Hubbard model. One may, in fact, single out one of the lattice sites and replace the influence of its neighbors by the interaction with a single, frequency-dependent bath, i.e., map the Hubbard model onto a single impurity Anderson model (SIAM) in the limit $Z \to \infty$. In order to restore the periodicity of the original lattice, this medium has to be determined self-consistently (Jarrell, 1992; Georges and Kotliar, 1992; Janiš and Vollhardt, 1992; Georges, Kotliar, Krauth, and Rozenberg, 1996). Written in terms of fermionic Matsubara frequencies[19] $\omega_n = (2n+1)\pi T$, self-energy $\Sigma_{\sigma n} \equiv \Sigma_\sigma(i\omega_n)$, and Green function $G_{\sigma n} \equiv G_\sigma(i\omega_n)$ the resulting coupled equations read

$$\boxed{\begin{aligned} G_{\sigma n} &= \int_{-\infty}^{\infty} \mathrm{d}\varepsilon\, \frac{\rho(\varepsilon)}{i\omega_n+\mu-\Sigma_{\sigma n}-\varepsilon} && (1.30)\\ G_{\sigma n} &= -\langle \psi_{\sigma n}\psi^*_{\sigma n}\rangle_{\mathcal{A}}. && (1.31)\end{aligned}}$$

[18] A self-energy is called proper if its external vertices cannot be be separated by cutting a single Green function line which implies that they are connected by at least three such lines.

[19] We here chose an imaginary-time formulation for brevity and in anticipation of its use in the context of the imaginary-time quantum Monte Carlo algorithm discussed in Sec. 1.3.

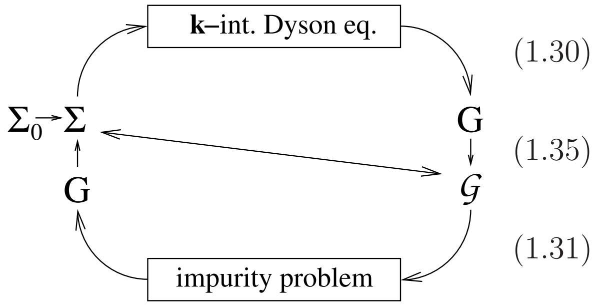



**Figure 1.1:** DMFT self-consistency cycle in conventional form: starting, e.g., with an initial guess for self-energy $\Sigma$, the $\boldsymbol{k}$ integrated Dyson equation (1.30) yields the lattice Green function $G$. Both $\Sigma$ and $G$ are used to compute the bath Green function $\mathcal{G}$ via (1.35) which defines the impurity problem (1.31). Its solution using QMC (or IPT, NCA, ED, NRG, etc.) provides a new estimate for $G$. The cycle is then closed by application of (1.35) to the new $G$ and the *old* $\Sigma$. For an alternative iteration scheme, see Sec. 3.2.

---

in the homogeneous phase. Here, properties of the lattice only enter via the DOS $\rho(\epsilon)$ of the noninteracting electrons. The thermal average $\langle \hat{C} \rangle_{\mathcal{A}}$ of some observable $\hat{C}$ is defined as a functional integral over Grassmann variables $\psi, \psi^*$,

$$\langle \hat{C} \rangle_{\mathcal{A}} = \frac{1}{\mathcal{Z}} \int \mathcal{D}[\psi]\, \mathcal{D}[\psi^*]\, C[\psi, \psi^*]\, \mathrm{e}^{\mathcal{A}[\psi,\psi^*,\mathcal{G}]}, \tag{1.32}$$

using the partition function

$$\mathcal{Z} = \int \mathcal{D}[\psi]\, \mathcal{D}[\psi^*]\, \mathrm{e}^{\mathcal{A}[\psi,\psi^*,\mathcal{G}]} \tag{1.33}$$

and the single-site action

$$\mathcal{A}[\psi, \psi^*, \mathcal{G}] = \sum_{\sigma,n} \psi^*_{\sigma n} \mathcal{G}^{-1}_{\sigma n} \psi_{\sigma n} - \frac{U}{2} \sum_{\sigma\sigma'} \int\limits_0^\beta \mathrm{d}\tau\, \psi^*_\sigma(\tau) \psi_\sigma(\tau) \psi^*_{\sigma'}(\tau) \psi_{\sigma'}(\tau)\,. \tag{1.34}$$

Here, $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature (in the following, we set $k_{\mathrm{B}} \equiv 1$) and

$$\mathcal{G}^{-1}_{\sigma n} = G^{-1}_{\sigma n} + \Sigma_{\sigma n} \tag{1.35}$$

is the effective local propagator.[20] Functions related by a Fourier transformation (here from imaginary time $\tau$ to fermionic Matsubara frequencies $\omega_n$ or vice versa) are denoted by the same symbol, but can be distinguished by their indices.

The solution of the DMFT problem by iteration is illustrated in Fig. 1.1. Here, the solution of the $\boldsymbol{k}$–integrated Dyson equation (1.30) is straightforward and can

[20] $\mathcal{G}_{\sigma n}$ may be regarded as a Weiss field in a (frequency-dependent) generalization of the usual static mean field. In contrast with spin models, for which the Weiss field replaces all (NN) interactions, the bath propagator replaces the hybridization of one site with the rest of the lattice.

be performed analytically for the semi-elliptic Bethe DOS (see Sec. 2.2) used in chapter 3. In contrast, the solution of the impurity problem (1.31) is highly nontrivial.[21] Most numerical methods developed for the treatment of SIAMs with fixed bath could be adapted to the DMFT problem, e.g., solutions based on exact diagonalization (ED) (Caffarel and Krauth, 1994; Georges et al., 1996), the noncrossing approximation (NCA) (Keiter and Kimball, 1970; Bickers, Cox, and Wilkins, 1987; Pruschke and Grewe, 1989; Pruschke, Cox, and Jarrell, 1993), the fluctuation-exchange approximation (FLEX) (Bickers, Scalapino, and White, 1989; Bickers and Scalapino, 1989; Bickers and White, 1991), the numerical renormalization group (NRG) (Wilson, 1975; Krishna-murthy, Wilkins, and Wilson, 1980; Costi, Hewson, and Zlatić, 1994; Bulla, 2000), and quantum Monte Carlo (QMC) algorithms. We will introduce the QMC method in Sec. 1.3 and use it throughout the numerical parts of this work; we will later also compare to results obtained from the other methods mentioned above as well as from iterated perturbation theory (IPT), which for half-filling ($n = 1$) is based on the following second-order approximation for the self-energy (Georges and Kotliar, 1992),

$$\Sigma^{\mathrm{IPT}}_{\sigma n} = \frac{U}{2} + U^2 \int_0^\beta \mathrm{d}\tau \, \mathrm{e}^{i\omega_n \tau} \big(\tilde{\mathcal{G}}_\sigma(\tau)\big)^3; \quad \tilde{\mathcal{G}}_\sigma(\tau) = \mathcal{G}_\sigma(\tau) - \frac{U}{2} \tag{1.36}$$

and becomes exact both at weak and strong coupling (Zhang, Rozenberg, and Kotliar, 1993).

While the DMFT is a very successful theory for correlated electron systems due to its nonperturbative character, it misses some important physics even in three dimensions and may fail completely for two dimensional systems. In particular, phenomena which involve strong momentum dependence such as $d$-wave superconductivity cannot be described within the DMFT. Attempts to build up a more widely applicable theory by systematically including $\mathcal{O}(1/Z)$ corrections have failed so far to produce breakthroughs for finite-dimensional systems (Gebhard, 1990; Vlaming and Vollhardt, 1992; van Dongen, 1994*c*; Schiller and Ingersent, 1995; Pruschke, Metzner, and Vollhardt, 2001).[22] However, significant progress has been made recently using schemes which interpolate between the DMFT single impurity problem and finite-dimensional clusters (with periodic boundary conditions). These fascinating extensions of the DMFT, namely the DCA and the CDMFT are reviewed in App. A.3, together with the RDA which was constructed as an alternative to the DMFT self-consistency approach.

[21] An exception is the application of (1.30) and (1.31) to a Lorentzian DOS $\rho(\epsilon) = t/(\pi(\epsilon^2 + t^2))$ which can be realized on lattices with long range hopping (Georges and Kotliar, 1992). For this DOS (which is clearly pathological due to its infinite variance), the Weiss function is independent of $U$; furthermore (1.31) is solvable by Bethe ansatz in this case so that many properties can be obtained analytically.

[22] Recently, an improved version of the Schiller-Ingersent scheme was applied for the hypercubic lattice using the FLEX approximation. Although no violation of causality was observed in this case, the causality of the method in general is still unclear (Zaránd, Cox, and Schiller, 2000).

## 1.3 Quantum Monte Carlo Algorithm

In this section, we will discuss the auxiliary-field quantum Monte Carlo (QMC) algorithm used in this work for solving the impurity model (1.31). It was originally formulated for treating a small number of magnetic impurities in metals (Hirsch and Fye, 1986) and later applied to arbitrary hybridization functions, i.e., in the form required for the solution of the DMFT problem (Jarrell, 1992; Rozenberg, Zhang, and Kotliar, 1992; Georges and Krauth, 1992; Ulmke, Janiš, and Vollhardt, 1995). Before we discuss technical details, let us point out some general important consequences of using this particular method for solving the impurity problem: First of all, the QMC method is formulated in imaginary time, i.e., $G(\tau)$ is evaluated (as a functional of $\mathcal{G}(\tau)$ and the model parameters) which implies that dynamical information can be obtained directly only for imaginary Matsubara frequencies and that analytic continuation is required for getting information, e.g., the spectrum, at real frequencies (see Sec. 1.4). A second implication is that Fourier transforms are necessary in order to connect the imaginary-time impurity part with the Dyson equations (1.30) and (1.35) which are diagonal in frequency space. Finally, the QMC method introduces a discretization $\Delta\tau$ of the imaginary time $\beta$ which not only necessitates an extrapolation of all DMFT(QMC) results to the physical limit $\Delta\tau \to 0$ (using converged DMFT solutions for different values of $\Delta\tau$) and restricts the method to relatively high temperatures, but also complicates the Fourier transformations. Here, we will concentrate on the solution of the impurity problem and discuss problems related to the Fourier transforms in Sec. 3.4. We also specialize to the single-band homogeneous case; for the generalization to the multi-band cases used in chapter 5, we refer to the literature (Rozenberg, 1997; Han, Jarrell, and Cox, 1998; Held and Vollhardt, 1998).

### 1.3.1 Wick's Theorem for the Discretized Impurity Problem

The difficulty in solving the functional integral equation (1.31) arises from the non-commutativity of the kinetic term and the interaction term in the single-site action (1.34). These terms can be separated by use of the Trotter-Suzuki formula (Trotter, 1959; Suzuki, 1976) for operators $\hat{A}$ and $\hat{B}$:

$$\mathrm{e}^{-\beta(\hat{A}+\hat{B})} = \left(\mathrm{e}^{-\Delta\tau\hat{A}}\,\mathrm{e}^{-\Delta\tau\hat{B}}\right)^{\Lambda} + \mathcal{O}(\Delta\tau)\,, \tag{1.37}$$

where $\Delta\tau = \beta/\Lambda$ and $\Lambda$ is the number of (imaginary) time slices.[23] Rewriting the action (1.34) in discretized form

$$\begin{aligned}\mathcal{A}_\Lambda[\psi,\psi^*,\mathcal{G},U] \;&=\; (\Delta\tau)^2\sum_\sigma\sum_{l,l'=0}^{\Lambda-1}\psi^*_{\sigma l}(\boldsymbol{\mathcal{G}}_\sigma^{-1})_{ll'}\,\psi_{\sigma l'}\\ &\quad-\Delta\tau U\sum_{l=0}^{\Lambda-1}\psi^*_{\uparrow l}\,\psi_{\uparrow l}\,\psi^*_{\downarrow l}\,\psi_{\downarrow l},\end{aligned} \tag{1.38}$$

[23]Since $\beta = 1/k_BT$ is the inverse temperature, small $\Delta\tau$ on each "time slice" corresponds to a higher temperature, for which the operators effectively decouple. Thus, we may view the Trotter approach as a numerical extension of a high-temperature expansion to lower temperatures.

where the matrix $\boldsymbol{\mathcal{G}}_\sigma$ consists of elements $\mathcal{G}_{\sigma l l'} \equiv \mathcal{G}_\sigma(l\Delta\tau - l'\Delta\tau)$, we apply (1.37) and obtain to lowest order

$$\begin{aligned}\exp\left(\mathcal{A}_\Lambda[\psi,\psi^*,\mathcal{G},U]\right) &= \prod_{l=0}^{\Lambda-1}\Bigg[\exp\Big((\Delta\tau)^2\sum_\sigma\sum_{l'=0}^{\Lambda-1}\psi^*_{\sigma l}(\boldsymbol{\mathcal{G}}_\sigma^{-1})_{ll'}\,\psi_{\sigma l'}\Big)\\ &\qquad\times\;\exp\left(-\Delta\tau\,U\,\psi^*_{\uparrow l}\,\psi_{\uparrow l}\,\psi^*_{\downarrow l}\,\psi_{\downarrow l}\right)\Bigg].\end{aligned} \tag{1.39}$$

Shifting the chemical potential by $U/2$, the four-fermion term can be rewritten as a square of two-fermion terms, which makes it suitable for the following discrete Hubbard-Stratonovich transformation (Hirsch, 1983):

$$\exp\left(\frac{\Delta\tau U}{2}(\psi^*_{\uparrow l}\,\psi_{\uparrow l} - \psi^*_{\downarrow l}\,\psi_{\downarrow l})^2\right) = \frac{1}{2}\sum_{s_l=\pm 1}\exp\left(\lambda s_l(\psi^*_{\uparrow l}\,\psi_{\uparrow l} - \psi^*_{\downarrow l}\,\psi_{\downarrow l})\right) \tag{1.40}$$

with $\cosh\lambda = \exp(\Delta\tau U/2)$. Here, the interaction between electrons is replaced by the interaction with an auxiliary binary field $\{s\}$ with components $s_l$ for $0 \le l \le \Lambda$. Acting like a local, but time-dependent magnetic field, $\{s\}$ can be regarded as an ensemble of Ising spins.

These transformations yield an expression for the functional integral

$$G_{\sigma l_1 l_2} = \frac{1}{\mathcal{Z}}\sum_{\{s\}}\int\mathcal{D}[\psi]\,\mathcal{D}[\psi^*]\;\psi^*_{\sigma l_1}\psi_{\sigma l_2}\exp\Big(\sum_{\sigma,l,l'}\psi^*_{\sigma l}M^{s_l}_{\sigma l l'}\psi_{\sigma l'}\Big), \tag{1.41}$$

with[24]

$$M^{s_l}_{\sigma l l'} = (\Delta\tau)^2(\boldsymbol{\mathcal{G}}_\sigma^{-1})_{ll'} - \lambda\sigma\delta_{ll'}s_l, \tag{1.42}$$

where in (1.41) the sum is taken over all configurations of the Ising spin field, and each term of the sum involves independent fermions only. Now Wick's theorem (see, e.g., Negele and Orland, 1987) can be applied to get the solution

$$G_{\sigma l l'} = \frac{1}{\mathcal{Z}}\sum_{\{s\}}\big(\boldsymbol{M}_\sigma^{\{s\}}\big)^{-1}_{ll'}\det\boldsymbol{M}_\uparrow^{\{s\}}\det\boldsymbol{M}_\downarrow^{\{s\}}, \tag{1.43}$$

where $\boldsymbol{M}_\sigma^{\{s\}}$ is the matrix with elements $M^{s_l}_{\sigma l l'}$, and the partition function has the value

$$\mathcal{Z} = \sum_{\{s\}}\det\boldsymbol{M}_\uparrow^{\{s\}}\det\boldsymbol{M}_\downarrow^{\{s\}}. \tag{1.44}$$

Computing one of the $2^\Lambda$ terms in (1.43) directly from definition (1.42) is an operation of order $\mathcal{O}(\Lambda^3)$. If the terms are ordered in a way so that successive configurations $\{s\}$ and $\{s\}'$ only differ by one flipped spin $s_l \to -s_l$ then all matrices and determinants can be updated at a cost of $\mathcal{O}(\Lambda^2)$ (Blankenbecler, Scalapino, and Sugar, 1981). Only for $\Lambda \lesssim 24$ can all terms be summed up exactly. Computations at larger $\Lambda$ are made possible by Monte Carlo importance sampling which reduces the number of terms that have to be calculated explicitly from $2^\Lambda$ to order $\mathcal{O}(\Lambda)$.

[24] A more precise form including subleading corrections is $M^{s_l}_{\sigma l l'} = (\Delta\tau)^2\,(\boldsymbol{\mathcal{G}}_\sigma^{-1})_{ll'}\,e^{\lambda\sigma s_{l'}} + \delta_{ll'}\left(1 - e^{\lambda\sigma s_l}\right)$ (Held, 1999; Georges et al., 1996).

### 1.3.2 Monte Carlo Importance Sampling

Monte Carlo (MC) procedures in general are stochastic methods for estimating large sums (or high-dimensional integrals) by picking out a comparatively small number of terms (or evaluating the integrand only for a relatively small number of points). Let us assume we want to compute the average $X := \frac{1}{M}\sum_{l=1}^{M} x_l$, where $l$ is an index (e.g., an Ising configuration $l \equiv \{s\}$) and $x$ some observable with the (true) variance $v_x = \frac{1}{M}\sum_{l=1}^{M}(x_l - X)^2$. In a simple MC approach (as used, e.g., for the computation of densities of states in chapter 2), one may select a subset of $N \ll M$ indices independently with a uniform random distribution $P(l_j) = \text{const.}$ (for $1 \leq j \leq N$),

$$\left(X_{\text{MC}}\right)^2 \quad = \quad \frac{1}{N}\sum_{j=1}^{N} x_{l_j} \tag{1.45}$$

$$\Delta X_{\text{MC}} \quad := \quad \langle (X_{\text{MC}} - X)^2 \rangle = \frac{v_x}{N} \approx \frac{1}{N(N-1)}\sum_{j=1}^{N}(x_{l_j} - X_{\text{MC}})^2 \,. \tag{1.46}$$

Here, the averages are taken over all realizations of the random experiment (each consisting of a selection of $N$ indices). In the limit of $N \to \infty$, the distribution of $X_{\text{MC}}$ becomes Gaussian according to the central limit theorem. Only in this limit is the estimate of $v_x$ from the QMC data reliable.

Smaller errors and faster convergence to a Gaussian distribution for the estimate may be obtained by importance sampling. Here, the function $x_l$ is split up,

$$x_l = p_l\, o_l; \quad p_l \geq 0; \quad \sum_{l=1}^{M} p_l = c \,, \tag{1.47}$$

where we may regard $p_l$ as a (unnormalized) probability distribution for the indices and $o_l$ as a remaining observable. If both the normalization $c$ is known (i.e, the sum over the weights $p_l$ can be performed exactly) and the corresponding probability distribution can be realized (by drawing indices $l$ with probability $P(l) = p_l/c$), we obtain

$$X_{\text{imp}}^{\text{MC}} = \frac{c}{N}\sum_{j=1}^{N} o_{l_j} \quad \text{and} \quad \Delta X_{\text{imp}}^{\text{MC}} = c\sqrt{\frac{v_o}{N}} \,. \tag{1.48}$$

Thus, the error can be reduced ($v_o < c^2 v_x$), when the problem is partially solvable, i.e., the sum over $p_l$ with $p_l \approx x_l$ can be computed.[25] Since this is not possible in general, one usually has to treat the normalization $c$ as an unknown and realize the probability distribution $P(l_j) = p_{l_j}/c$ in a stochastic Markov process: Starting with some initial configuration $l_1$, a chain of configurations is built up where in each step only a small subset of configurations $l'$ is accessible in a "transition" from configuration $l$. Provided that the transition rules satisfy the detailed balance principle,

$$p_l\, \mathcal{P}(l \to l') = p_{l'}\, \mathcal{P}(l' \to l) \,, \tag{1.49}$$

[25]The possible reduction of the variance is limited when $x_l$ is of varying sign. This "minus-sign problem" seriously restricts the applicability of Monte Carlo methods for finite-dimensional fermion problems.

and the process is ergodic (i.e., all configurations can be reached from some starting configuration), the distribution of configurations of the chain approaches the target distribution in the limit of infinite chain length. Since the normalization remains unknown, importance sampling by a Markov process can only yield ratios of different observables evaluated on the same chain of configurations. Another consequence of using a Markov process is that initial configurations have to be excluded from averages since the true associated probabilities might be vanishingly small. They would otherwise be overrepresented in any run of finite length. Consequently, we will later distinguish "warmup sweeps" from "measurement sweeps".

For the computation of errors, one has to take into account the finite autocorrelation induced by the Markov process, i.e., correlation between subsequent measurements. This correlaton may be characterized by the autocorrelation time[26] $\kappa_o \geq 1$ which effectively reduces the number of independent samples, so that $\Delta X = c\sqrt{v_o \kappa_o / N}$. The numerical effort necessary to reach some target statistical accuracy $\Delta X$ nevertheless increases only as $(1/\Delta X)^2$.

Returning to the evaluation of the Green function using (1.43) and (1.44), the obvious choice is to sample configurations $\{s\}$ according to the (unnormalized) probability

$$P(\{s\}) = \left| \det \boldsymbol{M}_\uparrow^{\{s\}} \det \boldsymbol{M}_\downarrow^{\{s\}} \right| . \tag{1.50}$$

The Green function can then be calculated as an average $\langle \ldots \rangle_s$ over these configurations:

$$G_{\sigma l l'} = \frac{1}{\tilde{\mathcal{Z}}} \left\langle \left(\boldsymbol{M}_\sigma^{\{s\}}\right)^{-1}_{ll'} \operatorname{sign}\left(\det \boldsymbol{M}_\uparrow^{\{s\}} \det \boldsymbol{M}_\downarrow^{\{s\}}\right) \right\rangle_s , \tag{1.51}$$

$$\tilde{\mathcal{Z}} = \left\langle \operatorname{sign}\left(\det \boldsymbol{M}_\uparrow^{\{s\}} \det \boldsymbol{M}_\downarrow^{\{s\}}\right) \right\rangle_s . \tag{1.52}$$

Here, $\tilde{\mathcal{Z}}$ deviates from the full partition function by an unknown prefactor which cancels in (1.51). The inability to compute the partition function is a consequence of the importance sampling and is thus a general characteristic of QMC methods. This will have severe consequences for the study of phase transitions (see chapter 3).

In this work, the transition probability is chosen according to the Metropolis transition rule (Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller, 1953)[27]

$$\mathcal{P}(x \to y) = \min\{1, P(y)/P(z)\} \tag{1.53}$$

for transitions from a spin configuration $x$ to a configuration $y$, which clearly fulfills detailed balance (1.49). Configuration updates are always performed by sweeping

[26] For a set $\{o_1, o_2, \ldots, o_N\}$ of measurements, the autocorrelation function (for the observable $o$) is $c_l^o = \langle (o_k - \langle o \rangle)(o_{k+l} - \langle o \rangle) \rangle_k$. An associated autocorrelation time may then be defined as $\kappa_o = c_0^o + 2\sum_{l=1}^{N_0} c_l^o$, where the cutoff $N_0$ is determined by $c_l^o > 0$ for $l \leq N_0$ and $c_{N_0+1} < 0$.

[27] Since in the present problem most of the computational cost is associated with the spin update and only has to be paid for accepted spin flips, the Metropolis rule, which has the highest acceptance ratio compatible with detailed balance, might not be optimal. Our tests with the symmetric heat-bath rule and generalizations of it which further suppress transitions between states with similar probability did not, however, lead to significant saving of computer time at constant accuracy.

through the full spin field, i.e., by attempting to flip spins $s_0$ through $s_{\Lambda-1}$ one after the other and then starting the next sweep. Roughly, the number of such sweeps has to be held constant (rather than the number of attempted single-spin flips) for constant statistical error when the number of spins $\Lambda$ is varied. Since the systematic error depends on $\Delta\tau$, the necessary number of time slices is proportional to the inverse temperature, $\Lambda \propto \beta$, so that the total numerical cost for given accuracy is proportional to $\beta^3$.

In this work, up to $\Lambda = 400$ time slices and up to $10^7$ measurement sweeps were used. A single iteration for $\Lambda = 400$ with 200000 sweeps took more than 7 CPU hours on a Cray T90. Since typically 10-20 iterations are needed for well-converged solutions (even when initializing the runs with data obtained for larger $\Delta\tau$), the computation of each such data point took several weeks of real time and used up about a monthly contingent of supercomputer time. No minus-sign problem was encountered, i.e., the sign in (1.51) and (1.52) was always positive.

## 1.4 Maximum Entropy Method

Within the DMFT, all single-particle properties can be expressed in terms of the local single-particle spectral function (also called "full" or "interacting" density of states) which is proportional to the imaginary part of the local retarded Green function,

$$A(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,G(\omega + i0^+)\,. \tag{1.54}$$

The spectral function $A(\omega)$ is accessible experimentally: measured (angular averaged) photoemission spectra (PES) can under certain simplifying assumptions (see subsection 5.4.1) be identified with $A(\omega)$ multiplied by the Fermi function $n_{\mathrm{f}}(\omega-\mu)$. Correspondingly, inverse photoemission spectra or X-ray absorption spectra (XAS) can be identified with $A(\omega)$ multiplied by $1-n_{\mathrm{f}}(\omega-\mu)$. Furthermore, nonlocal spectral functions and the optical conductivity $\sigma(\omega)$ can, within the DMFT, be calculated from $A(\omega)$ (see chapter 4).

In QMC calculations, however, the Green function $G$ (and thus the spectral function $A(\omega)$) cannot be directly computed on the real axis. Instead, real-time dynamical information has to be extracted from imaginary-time data $G(\tau)$ (or, equivalently, from the Fourier transformed Matsubara-frequency data $G(i\omega_n)$) via analytic continuation. This is in principle possible through inversion of the spectral representation for $G(\tau)$, i.e.,

$$G(\tau) \;=\; \int\limits_{-\infty}^{\infty} \mathrm{d}\omega\; K(\tau,\omega-\mu;\beta)\; A(\omega)\,, \tag{1.55}$$

$$K(\tau,\omega;\beta) \;:=\; \frac{\exp(-\tau\omega)}{1+\exp(-\beta\omega)}\,, \tag{1.56}$$

but poses an ill-conditioned problem since $G(\tau)$ is only measured on a grid $\tau_l = l\Delta\tau$ (where $\Delta\tau = \beta/\Lambda$) and since the kernel (1.56) becomes exponentially small for generic

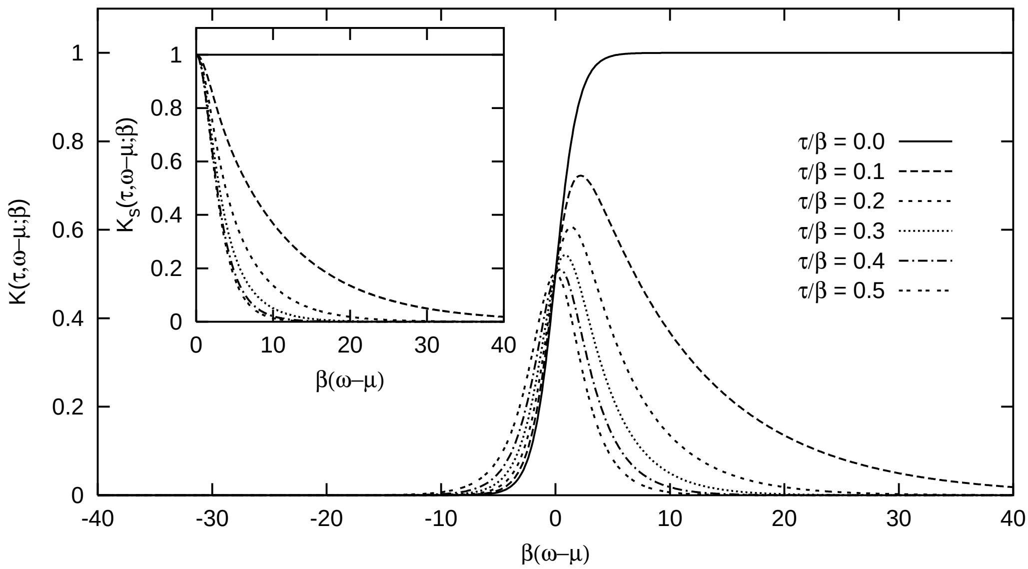



**Figure 1.2:** General fermion kernel (1.56) for an equidistant set of values of $\tau/\beta \leq 1/2$. Except for $\tau = 0$, large frequencies are suppressed exponentially. Inset: symmetrized fermion kernel (1.57).

values of $\tau$ at large absolute frequencies $|\omega|$ as illustrated in Fig. 1.2. For a symmetric problem, i.e., symmetric noninteracting DOS and $n = 1$, the integral in (1.55) can be restricted to positive frequencies when the symmetrized fermion kernel

$$K_s(\tau, \omega; \beta) = K(\tau, \omega; \beta) + K(\tau, -\omega; \beta) \tag{1.57}$$

shown in the inset of Fig. 1.2 is used. We stress that a very small number $\Lambda = \mathcal{O}(10)$ of time slices poses a more serious limitation for obtaining reliable spectra with good resolution than the exponential nature of the kernel since the number of degrees of freedom which can be reliably resolved in a spectrum is obviously much smaller than the number of data points $\{G(l\Delta\tau)\}$ and since there always exists an infinite number of spectra which correspond to the same data. Still, in general, the resolution is much better at smaller frequencies and, according to Nyquist's theorem, essentially no information can be obtained from QMC for $\omega > \Lambda\pi/\beta = \pi/(\Delta\tau)$. Before we address the full analytic continuation problem and introduce the maximum entropy method, we collect some useful relations (denoting the $m^{\text{th}}$ derivative as $G^{(m)}$),

$$G(\beta) = n, \qquad G(0^+) = 1 - n, \tag{1.58}$$

$$G^{(m)}(0) + G^{(m)}(\beta) = (-1)^m \langle(\omega - \mu)^m\rangle_{A(\omega)} \tag{1.59}$$

$$G(\beta/2) \approx \frac{\pi}{\beta}\,\overline{A(\omega)}\Big|_{|\omega-\mu|\lesssim\pi/\beta}\,. \tag{1.60}$$

Since the filling given by (1.58) is known in the symmetric case, the value of $G(0)$ then provides no useful information which is also seen in the inset of Fig. 1.2. Equation (1.59) also shows the loss of high-frequency information from the discretization of imaginary time: for a finite grid the error in estimating derivatives $G^{(m)}$ increases rapidly with order $m$; thus, the determination of high order moments $\langle(\omega-\mu)^m\rangle_{A(\omega)}$ of the spectrum is in general an ill-posed problem. At low temperatures, $G(\beta/2)$ gives a hint as to the weight of a quasiparticle peak or the existence of a gap via (1.60) since its value is proportional to the value of the spectral function near the Fermi energy, averaged over an inverse hyperbolic cosine with width $\pi/\beta$.

First attempts to address the analytic continuation problem for QMC data included least-squares fits, Padé approximants, and regularization (for references, see the pedagogical and concise review by Jarrell (1997)). Least-squares fits of spectra approximated as a set of box functions are inherently unstable. Padé approximations for $G(i\omega_n)$ only work well for very precise data (e.g., in the context of Eliashberg equations), but not for QMC. Regularization of the kernel (1.56) tends to produce overly smeared-out spectra. What is needed instead is a regularization of the solution $A(\omega)$ that only shows features which are supported by the data, but is as smooth as possible otherwise. This is essentially the idea of the maximum entropy method (MEM) of finding the most probable spectrum compatible with the data.

The MEM is a very general approach for reconstructing continuous, positive semidefinite functions (i.e., densities or spectra) from incomplete, noisy, and possibly oversampled[28] data. It has a long history in the context of image reconstruction in such diverse disciplines as radio aperture synthesis, optical deconvolution, $X$-ray imaging, structural molecular biology, and medical tomography (see Skilling and Bryan (1984) and references therein). A preliminary study of its usefulness in the context of the analytic continuation problem (Silver, Sivia, and Gubernatis, 1990) was soon followed by applications to spectra (Silver, Gubernatis, Sivia, and Jarrell, 1990; Jarrell, Sivia, and Patton, 1990) and correlation functions (Deisz, Jarrell, and Cox, 1990). An efficient and general formulation of a MEM algorithm for analytic continuation of (oversampled) QMC data was then given by Gubernatis, Jarrell, Silver, and Sivia (1991), closely following the general state-of-the-art approach by Bryan (1990).

For simplicity, we will first assume that the QMC simulation produces $N_d$ measurements of the observables $G_l \equiv G(l\Delta\tau)$, $0 \leq l < \Lambda$ which are independent both in computer time and in imaginary time, i.e., without significant autocorrelation between subsequent measurements or between adjacent imaginary time slices.[29] If we further disregard systematic errors and assume some "true" spectrum $A$ [which corresponds to a discretized Green function $G_l$ via (1.55)] as well as a Gaussian distribution of statistical errors, the probability distribution for the observed averaged

[28]Data is called *oversampled* when measurements of different data points (here: for different $\tau$) are not statistically independent, but correlated (Bryan, 1990). In this case, the number of "good" degrees of freedom (entering the error statistics) is reduced.

[29]Typically, one of the "measurements" of a set $\{G_l\}$ is generated by binning a macroscopic number of measurements within QMC. Autocorrelation in computer time, i.e., between different bins vanishes in the limit of infinite bin size.

values $\bar{G} \equiv \{\bar{G}_l\}$ (where $\bar{G}_l = \sum_{i=1}^{N_d} G_l^i$) is

$$P(\bar{G}|A) \propto \mathrm{e}^{-\frac{1}{2}\chi^2}; \quad \chi^2 = \sum_{l=0}^{\Lambda-1} \frac{\left(\bar{G}_l - G_l\right)^2}{\sigma_l^2} \,. \tag{1.61}$$

Here, $\sigma_l$ can be estimated from the data alone, $\sigma_l^2 \approx \sum_{i=1}^{N_d} \left(\bar{G}_l - \bar{G}_l^i\right)^2 / (N_d(N_d-1))$. Obviously, the likelihood function $P(\bar{G}|A)$ is not directly helpful; in order to find the most probable spectrum given the measured data, we need the converse probability $P(A|\bar{G})$. These probabilities are related by Bayes' theorem,

$$P(A|\bar{G})\, P(\bar{G}) = P(\bar{G}|A)\, P(A) \,. \tag{1.62}$$

Since the data $\bar{G}$ is constant in the search for an optimal $A$, the associated probability $P(\bar{G})$ drops out of the problem. For the probability $P(A)$ of a spectrum $A(\omega)$ in absence of data, an entropic ansatz is made where prior knowledge can also be incorporated by choosing an appropriate (positive semidefinite) default model $m(\omega)$,

$$P(A) = \mathrm{e}^{\alpha S[A(\omega), m(\omega)]} \,. \tag{1.63}$$

Here, $\alpha$ is a numerical parameter while $S$ is a generalized Shannon-Jaynes entropy,

$$S[A, m] = \int \mathrm{d}\omega \Big( A(\omega) - m(\omega) - A(\omega) \ln \big( A(\omega)/m(\omega) \big) \Big) \,. \tag{1.64}$$

For a constant default model (within some finite frequency range), the entropic form (1.64) clearly favors smooth spectra. This is also true for a general smooth default model. It also enforces positivity of $A$ and pushes the solution towards the (normalized) default model in absence of data. From (1.61), (1.62), and (1.63), the posterior probability can be read off as

$$P(A|\bar{G}, m, \alpha) = \mathrm{e}^{\alpha S[A,m] - \frac{1}{2}\chi^2[\bar{G}, A]} \,. \tag{1.65}$$

The balance between a tight match of data and a high entropy is calibrated by the Lagrange parameter $\alpha$ which may be chosen so that $\chi^2 = \Lambda$ (historic MEM). Alternatively, one may use the value of $\alpha$ with the highest probability $P(\alpha|\bar{G}, A, m)$ which can approximately be calculated within the method (classic MEM). Given the QMC data, a default model, a representation of the spectrum (i.e., a possibly inhomogeneous grid of $\omega_j$ of frequencies for which $A$ is going to be computed), and a starting guess for $\alpha$, a simple MEM program thus both searches for the spectrum $\{A(\omega_j)\}$ with maximum probability $P(A|\bar{G}, m, \alpha)$ for given $\alpha$ using, e.g., the Newton-Raphson method and, in an outer loop, searches for the best value of $\alpha$.

More advanced aspects of the MEM, in particular the breakdown of some assumptions made above and improved formulations are discussed in the context of the analytic continuation of Green functions for correlated systems in subsection 3.8.1.

# Chapter 2

# Lattice and Density of States

Within the DMFT, properties of the lattice enter the computation of single-particle properties via the "noninteracting" density of states (DOS) $\rho(\epsilon)$ only.[1] Thus, in addition to the Hubbard interaction $U$, the temperature $T$, and the density $n$, the DOS constitutes an infinite set of input parameters for the DMFT treatment of the Hubbard model. While some aspects of a given DOS are easily characterized (with, e.g., the variance setting the energy scale for $U$ and $T$), many of its degrees of freedom may influence physical properties of the interacting system in a complicated way. More generally, the choice of hopping matrix elements or, equivalently, the choice of a dispersion $\epsilon_{\boldsymbol{k}}$ may be crucial for properties of Hubbard type models in any dimension. In most studies of the Hubbard model in dimensions $d > 1$, the hopping range has been restricted to nearest neighbors (NN) and, in some cases, next-nearest neighbors (NNN). While using only these largest hopping matrix elements is certainly a reasonable starting point, the remaining part can still be qualitatively important, in particular in high dimensions.

Three related problems motivate the studies presented in this chapter. First, up to now there has been no fully satisfactory way of treating the Hubbard model rigorously in $d = \infty$. As shown in Sec. 1.2, NN hopping on the hypercubic (hc) lattice leads to a Gaussian DOS without band edges in $d \to \infty$, which is unphysical since in the low-density limit the energy per electron becomes infinite and because an unbounded band violates the single-band assumption of energetically well-separated bands. Inclusion of NNN hopping not only leads to magnetic frustration and an asymmetric DOS but (for an appropriate sign of the hopping amplitudes) cures the problem of infinite energy in the low-density limit by introducing a sharp lower band edge in $d \to \infty$.[2] Another conventional way of establishing band edges in $d = \infty$ is taking this limit on the Bethe lattice which yields the semi-elliptic DOS used throughout chapter 3. Here, problems arise from the fact that the Bethe lattice is not a regular lattice but a treelike structure which has no loops and no well-defined

[1] The quotation marks, which we will drop from now on, are a reminder that (at least) interactions with lattice ions and core electrons are already included in the "noninteracting" dispersion $\epsilon_{\boldsymbol{k}}$ and, consequently, determine the "noninteracting" DOS.

[2] This cure is, however, not stable upon further extension of the hopping range: hopping to third-nearest neighbors generically reintroduces the low-energy tail extending to $\epsilon = -\infty$.

momentum space.

The second motivation for studying the relation between lattices and corresponding densities of states is that the knowledge of the DOS is not sufficient for the computation of two-particle properties like susceptibilities and, in particular, transport properties. In fact, a large part of chapter 4 is devoted to the question of how the optical conductivity $\sigma(\omega)$ can be defined and computed for the Bethe lattice (or, more generally, for a lattice with a semi-elliptic DOS in the limit $d = \infty$). It would be desirable to find a regular isotropic[3] and translationally invariant lattice (with a well-defined $\boldsymbol{k}$-space) with the same single-particle properties (i.e., the same DOS) as the Bethe lattice. Transport properties could then be computed directly from first principles. As shown in Sec. 2.4 and in subsection 4.4.5, this is indeed possible. Another important application of the new method developed in this work is the computation of the optical conductivity for a given LDA DOS in Sec. 5.4. Although here, in principle, the full dispersion $\epsilon_{\boldsymbol{k}}$ could be calculated, an educated guess for the DMFT estimate of $\sigma(\omega)$ will be obtained without the numerically complicated evaluation of $\nabla\epsilon_{\boldsymbol{k}}$.

A third point relates to the question of how relevant results calculated in $d = \infty$ are to the finite-dimensional systems of the real world. Since the DOS represents the noninteracting limit of the model Hamiltonian in any dimension, it seems questionable to rely on features of the DOS which are not robust under change of dimensions, i.e., cannot be obtained under similar assumptions concerning the hopping matrix elements in at least any dimension $3 \leq d \leq \infty$. One example where a feature is not generic although it occurs in $d = 3$ and $d = \infty$ is the singularity of the fcc DOS at the lower band edge. This feature is not present in any finite dimension $d > 3$ unless NNN hopping is chosen appropriately; see, e.g., Ulmke (1998) and Sec. 2.1. Even more importantly, our study of the paramagnetic Mott metal-insulator transition performed in chapter 3 would not be complete without a discussion of microscopic mechanisms for frustration and their implications for the DOS and for transport properties in infinite and finite dimensions.

The numerically exact studies of tight-binding models in large but finite dimensions presented in this chapter were made possible by the development of a computer program which allows for the computation of the DOS in arbitrary (finite) dimensions given the dispersion $\epsilon_{\boldsymbol{k}}$. This computation is performed by Monte Carlo sampling of the $d$-dimensional integral (volume $V_B$ of the Brillouin zone)

$$\rho(\epsilon) = \frac{1}{V_B}\int \mathrm{d}\boldsymbol{k}\ \delta(\epsilon - \epsilon_{\boldsymbol{k}}) \tag{2.1}$$

on a discretized energy interval. Originally, the computation of the DOS was only a byproduct (at no additional cost) of computations of

$$\tilde{\rho}(\epsilon) = \frac{1}{V_B}\int \mathrm{d}\boldsymbol{k}\ \delta(\epsilon - \epsilon_{\boldsymbol{k}})|\boldsymbol{v}_{\boldsymbol{k}}|^2 \tag{2.2}$$

[3]Here and in the following, we use the term *isotropic* for a lattice with hypercubic symmetry since in that case the cartesic axes are equivalent and all transport properties are isotropic in the long-wavelength limit.

which is the essential lattice input to the optical conductivity within the bubble approximation (see chapter 4 where results for $\tilde{\rho}(\epsilon)$ are also presented). However, a fast and reliable way of computing the DOS is always useful since in spite of considerable analytic effort a numerical integration is unavoidable even for the hypercubic lattice with NN hopping in $d \geq 2$ [see, e.g., Schlipf (1998)].

The most important results of this chapter are the exact analytic expression and numerical estimates for extended hopping matrix elements defined on a hypercubic lattice which yield a semi-elliptic DOS in the limit $d \to \infty$. In fact, the new method developed in this chapter for determining dispersion, transport properties, and matrix elements in $d = \infty$ is very general; it can be applied for any arbitrary target DOS. The method is based on a generalization of the observation that the dispersion of the fcc lattice can be written as a function of the dispersion of the hc lattice. This expression, a (quadratic) polynomial in $\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}$, implies a simple transformation rule for the DOS. In the course of this thesis it became apparent that parts of the necessary formalism had already been worked out, but had not been published. Extending van Dongen's (2001) construction via expansion of the dispersion in Hermite polynomials we will show that both the hopping matrix elements and all information (e.g., the Fermi velocity) needed for the computation of transport properties can be obtained in a controlled scheme.

The structure of the remainder of this chapter is as follows: In Sec. 2.1 we will first discuss hypercubic lattices with NN hopping matrix elements only, then consider $t$-$t'$ hopping to NNs and NNNs and, finally, we will treat the fcc lattice which is equivalent to a hc lattice with NNN hopping only. Conventional ways of constructing a semi-elliptic DOS in $d = \infty$, i.e., via the fractal Bethe lattice and the disordered infinite-range hopping model are summarized in Sec. 2.2. The new general dispersion scheme is then developed in Sec. 2.3. As a first practical application we use this method for the construction of a regular lattice with Bethe semi-elliptic DOS and begin with the characterization of its transport properties in Sec. 2.4. The robustness of the new method is demonstrated by studying both the effects of truncation of the hopping range and of finite dimensionality. Our treatment of the hyperdiamond lattice, a regular non-Bravais lattice with a reduced number of loops, in App. B supplements the findings of this chapter and our discussion of loops in subsection 4.4.2.

## 2.1 Hypercubic Lattice and Extensions

### 2.1.1 Definitions and Analytical Considerations

The $d$-dimensional hypercubic (hc) lattice is spanned by $d$ orthonormal primitive vectors (setting the lattice spacing to $a = 1$) so that the cartesian components of each lattice vector are integers and the Euclidean length (or 2-norm) of a vector is given by

$$|\boldsymbol{v}| := \|\boldsymbol{v}\|_2 = \Big(\sum_{\alpha=1}^{d} v_\alpha^2\Big)^{1/2} \tag{2.3}$$

as a special case of the $l$-norm $\|\boldsymbol{v}\|_l := (\sum_{\alpha=1}^{d} |v_\alpha|^l)^{1/l}$. Clearly, the shortest lattice vectors are the unit vectors $\pm\boldsymbol{e}_\alpha$ with length 1 which define nearest neighbors (NN) while for $d > 1$ next-nearest neighbors (NNN) are connected by the sum of two different primitive lattice vectors $\pm\boldsymbol{e}_\alpha \pm \boldsymbol{e}_\beta$ with $\alpha \neq \beta$ and length $\sqrt{2}$. For $d \geq 3$, third-nearest neighbors have length $\sqrt{3}$ while among the fourth-nearest neighbors are those connected by extended hopping along the coordinate axis (length 2). While even for a full simple cubic lattice of Wannier orbitals some hopping matrix elements may exactly vanish for symmetry reasons, hopping to, e.g., 4$^{\text{th}}$ or 10$^{\text{th}}$ nearest neighbors (i.e., a 2-3 fold increase of the Euclidean hopping range compared with NN) is not necessarily negligible.[4]

### Scaling and Dispersion

Generically, the number of $D^{\text{th}}$-nearest neighbors in $d$ dimensions is of the order $(2d)^D$, provided[5] that $D \lesssim d$. In order for the variance of the DOS to remain finite the hopping matrix elements then have to be scaled[6] as $t_{ij} = t_{ij}^*(2d)^{-\|\boldsymbol{r}_i - \boldsymbol{r}_j\|/2}$. Here, we used a short notation for the 1-norm, also called “taxi cab metric” or “New York metric”,

$$\|\boldsymbol{v}\| := \|\boldsymbol{v}\|_1 = \sum_{\alpha=1}^{d} |v_\alpha| \,. \tag{2.4}$$

For the remainder of this section we restrict the matrix elements to NN and NNN hopping with amplitudes $t$ and $t'$, respectively, as illustrated in Fig. 2.1 for $d = 3$. This corresponds to a dispersion

$$\begin{aligned} \epsilon_{\boldsymbol{k}}^{tt'} &= -\frac{t}{L}\sum_{\langle i,j\rangle} \exp\left[i\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{R}_j)\right] - \frac{t'}{L}\sum_{\langle\langle i,j\rangle\rangle} \exp\left[i\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{R}_j)\right] && (2.5)\\ &= -2t\sum_{\alpha=1}^{d} \cos(k_\alpha) - 2t'\sum_{\substack{\alpha,\beta=1\\ \alpha\neq\beta}}^{d} \cos(k_\alpha)\cos(k_\beta), && (2.6)\end{aligned}$$

[4] Note that long-range hopping is compatible with a local Coulomb interaction provided the latter is understood as the non-Hartree part of the full interaction (whereas the Hartree part is absorbed into the lattice potential).

[5] This assumption arises in all DMFT calculations and is clearly generically violated in any finite dimension: In the DMFT perturbation expansion[] one assumes that at every vertex $\mathcal{O}(d)$ directions can be chosen independently which is only true if the dimension is larger than the number of vertices, i.e., the order of the diagram. In other words, for a given dimension, higher order diagrams are treated worse within a local approximation than lower order diagrams; the limits of infinite-order perturbation theory and infinite dimensionality do not commute. Physically, this means that there exists no critical dimension above which the DMFT is always qualitatively correct. For consequences of the violation of scaling relations in the context of our construction of general kinetic energies, see Sec. 2.3.

[6] A counter example to this general rule is extended hopping along the axes (Müller-Hartmann, 1989*a*) which only contributes to the energy if scaled as $t_{ij} = t_{ij}^*\,(2d)^{-1/2}$. It leads to a Gaussian DOS although, in general, the lattice is neither bipartite nor shows perfect nesting (cf. subsection 2.1.3).

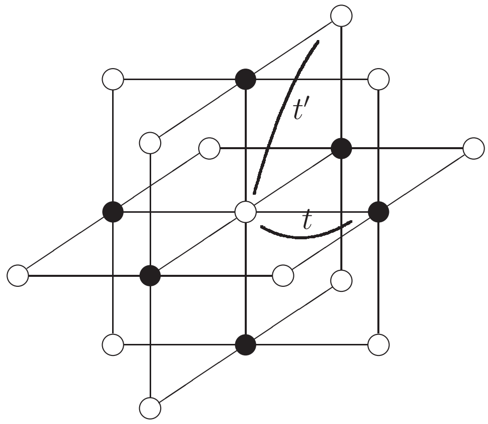



**Figure 2.1:** Cubic lattice with NN ($t$) and NNN hopping ($t'$); (from Schlipf (1998)).

where $L$ is the number of lattice sites $\boldsymbol{R}_i$ and $\langle i,j\rangle$ and $\langle\langle i,j\rangle\rangle$ denote NN pairs and NNN pairs, respectively. The variance of the resulting DOS can be easily computed for arbitrary dimension from the sum of squares of all hopping terms (cf. subsection 2.2.2):

$$\langle \epsilon^2\rangle_{\rho(\epsilon)} = \langle \epsilon_{\boldsymbol{k}}^2\rangle_{\boldsymbol{k}} = 2dt^2 + d(d-1)(t')^2. \tag{2.7}$$

Here, we have introduced the expectation values[7]

$$\langle f(\epsilon)\rangle_{\rho(\epsilon)} = \int_{-\infty}^{\infty} \mathrm{d}\epsilon f(\epsilon)\rho(\epsilon) \tag{2.8}$$

$$\langle f(\boldsymbol{k})\rangle_{\boldsymbol{k}} = \frac{1}{N}\sum_{\boldsymbol{k}} f(\boldsymbol{k})\,. \tag{2.9}$$

When $t$ and $t'$ are scaled with the inverse number of NN and NNN, respectively, and we introduce the ratio $a^*$ of these scaled amplitudes,

$$t = \frac{t^*}{\sqrt{2d}}, \quad t' = \frac{t'^*}{\sqrt{d(d-1)}}, \quad a^* = \frac{t'^*}{t^*} \tag{2.10}$$

we obtain for the variance

$$\langle \epsilon_{\boldsymbol{k}}^2\rangle_{\boldsymbol{k}} = (t^*)^2 + (t'^*)^2 = t^{*2}\left(1 + a^{*2}\right)\,. \tag{2.11}$$

For given ratio $a^* < \infty$ the variance can thus be fixed to 1 by setting $t^* = 1/\sqrt{1+a^{*2}}$.

[7] Note that the definitions are obviously consistent, i.e., $\langle f(\epsilon_{\boldsymbol{k}})\rangle_{\boldsymbol{k}} = \langle f(\epsilon)\rangle_{\rho(\epsilon)}$ for constant functional dependence $f$; consequently, the appropriate definition of the average usually follows from the argument and may be omitted.

**Bipartite Lattices and Perfect Nesting**

Independent of hopping range and dimension, a hc lattice may be split into two sublattices $A$ and $B$,

$$A = \{\boldsymbol{r}; (-1)^{\|\boldsymbol{r}\|} = 1\}, \quad B = \{\boldsymbol{r}; (-1)^{\|\boldsymbol{r}\|} = -1\}, \tag{2.12}$$

which is visualized in Fig. 2.1 by using empty circles for $A$ sites and filled circles for $B$ sites, respectively. Since the origin has no special meaning, both sublattices $A$ and $B$ are equivalent. Note that each sublattice constitutes a face centered cubic lattice in 3 dimensions so that we can use either of the two sets in (2.12) in order to define a generalized face centered cubic (fcc) lattice in arbitrary dimensions $d > 2$. The tendency towards breaking this $A - B$ symmetry by staggered charge or spin order is largest when hopping matrix elements are nonvanishing only between (as opposed to within) the sublattices. The simplest example for such a *bipartite lattice* is the hc with NN hopping only; the most general case (on a full hypercubic host lattice) can be expressed by the requirement $t_{ij} = 0$ for $(-1)^{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|} = 1$. It is easily seen that (for real $t_{ij}$) all such bipartite Bravais lattices have symmetric DOSs by performing the particle-hole transformation $\hat{c}_{i\sigma} \to \hat{c}^{\dagger}_{i\sigma}$ for $\boldsymbol{r}_i \in A$ and $\hat{c}_{j\sigma} \to -\hat{c}^{\dagger}_{j\sigma}$ for $\boldsymbol{r}_j \in B$. A prominent example for a non-bipartite lattice is the generalized fcc lattice mentioned above which corresponds to the limit $t \to 0$ (while $t' \neq 0$) of (2.5).

For the hypercubic lattice the particle hole symmetry can be expressed in $\boldsymbol{k}$-space ($\boldsymbol{Q} = (\pi, \pi, \dots)$)

$$\epsilon_{\boldsymbol{k}+\boldsymbol{Q}} = -\epsilon_{\boldsymbol{k}}\,, \tag{2.13}$$

which implies *perfect nesting* at half filling ($n = 1$): the Fermi surface (all $\boldsymbol{k}$ points with $\epsilon_{\boldsymbol{k}} = 0$) is mapped on itself for $\boldsymbol{k} \to \boldsymbol{k} + \boldsymbol{Q}$ which amounts to an effective doubling of the unit cell (or halving of the Brillouin zone) and again implies a potential instability against $A - B$ symmetry breaking. A trivial consequence of (2.13) is a symmetric DOS, i.e., $\rho(-\epsilon) = \rho(\epsilon)$.

### 2.1.2 Numerical Results

The DOS of the hc lattice with NN hopping is shown in Fig. 2.2. For $d \geq 3$, $\rho^{\rm hc}(\epsilon) < \infty$ at all energies. Apart from Van-Hove singularities with infinite slope for $d = 3$ and a jump in the slope for $d = 4$ the overall shape is quite regular with no visible qualitative changes for $d > 6$. For $d \to \infty$, more spectral weight is transferred from $|\epsilon| \approx 3/2$ to $\epsilon \approx 0$ while the exponentially small tails extend to $\pm\infty$. An application of $d = \infty$ results to $d = 3$ apparently misses the Van-Hove singularities (as it must for any lattice type) and (for $n = 1$) slightly overestimates the DOS at the Fermi level but does not introduce artificial features – except for the infinite energy tails.

Switching on additional NNN hopping destroys the symmetry of the DOS as illustrated in Fig. 2.3 for $a^* = (t')^*/t^* = -0.25$ which introduces an important aspect of realistic tight-binding calculations. Except for $d = 2$ and the Van-Hove singularities visible for $d = 3$ (at $\epsilon = \pm\sqrt{2/3}$) and $d = 4$ (at $\epsilon = 0$), Fig. 2.3 implies a rapid qualitative convergence with few differences between $d = 4$ and $d = 5$. This

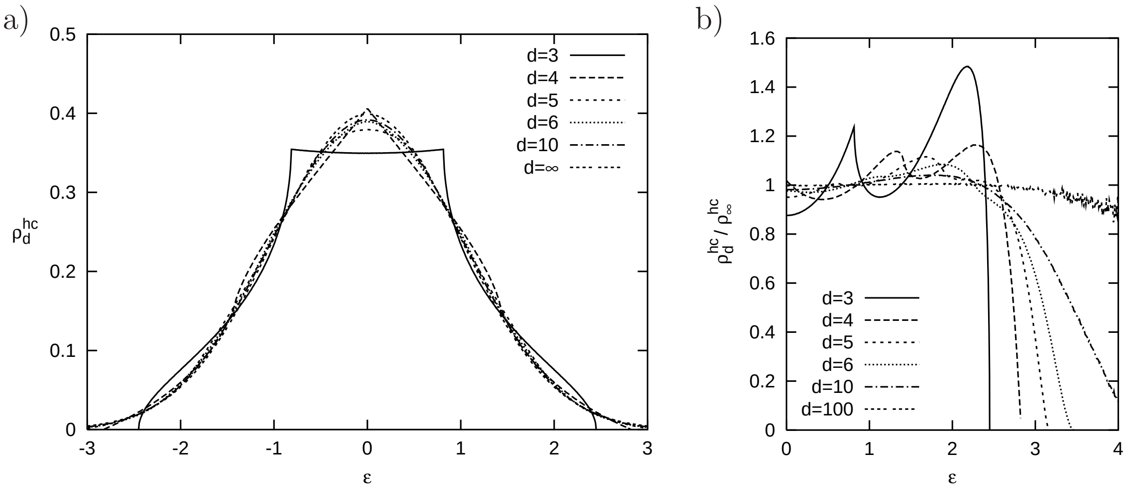



**Figure 2.2: a)** DOS of the hypercubic lattice with NN hopping in $d$ dimensions. **b)** Ratio of the DOS for finite $d$ to its $d = \infty$ limit (only shown for $\epsilon \geq 0$).

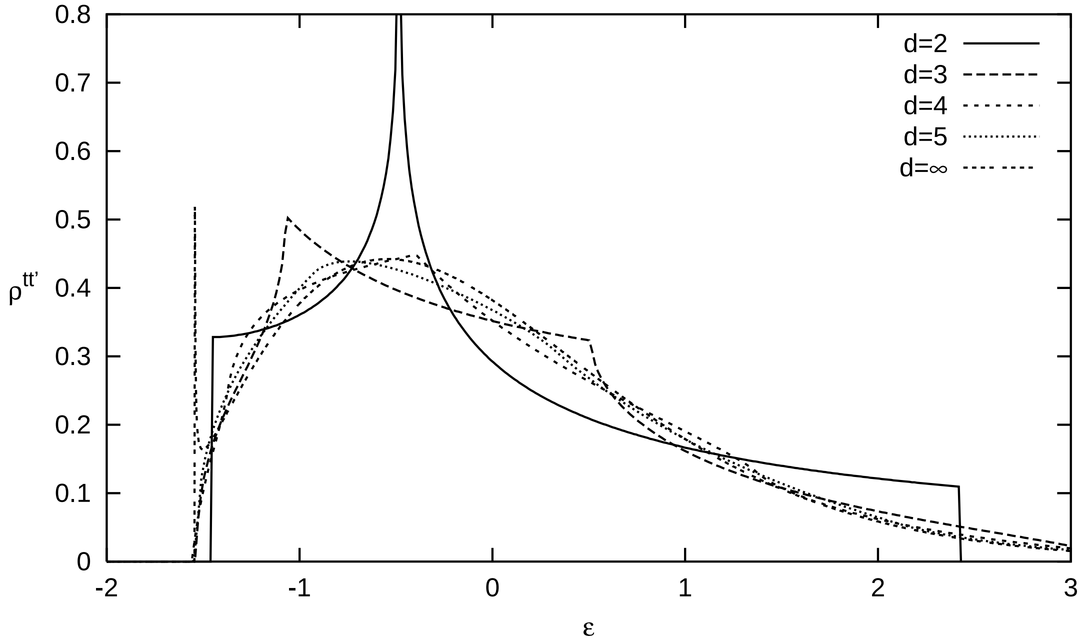



**Figure 2.3:** DOS of the hypercubic lattice with NN and NNN hopping, $a^* = -0.25$ for low dimensions. The limit $d = \infty$ is included for comparison.

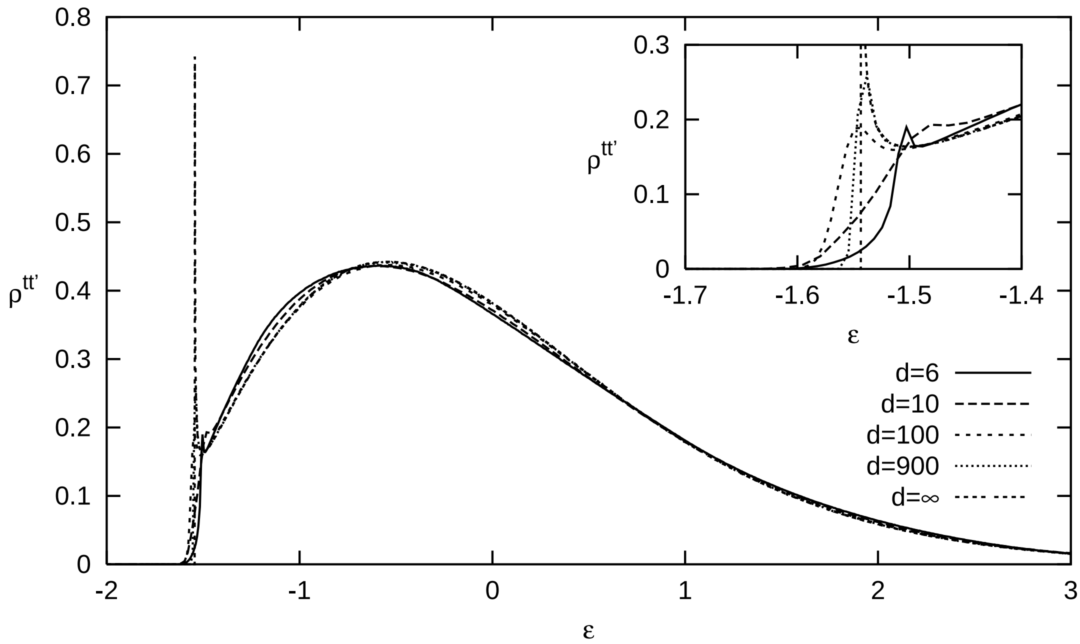



**Figure 2.4:** DOS of the hypercubic lattice with NN and NNN hopping, $a^* = -0.25$ for high dimensions. As seen in the inset, a sharp singular lower band edge develops for $d \to \infty$.

picture changes, however, when $d$ is further increased. As shown in Fig. 2.4, for $d \geq 6$ a maximum appears[8] near the lower band edge which develops into a sharp band edge with a square-root singularity (i.e., with the asymptotic form of an inverse square root) for $d \to \infty$. Increasing the absolute value of $a^*$ increases these tendencies as depicted for $a^* = -0.3333$ in Fig. 2.5 and Fig. 2.6. Here, a broad maximum at low energies develops already for $d = 2$ which becomes narrow and distinguishable from the main maximum for $d \approx 10$ and approaches a square-root singularity with a height substantially exceeding the main maximum for $d \gtrsim 100$. We observe that for finite dimensions there is a low energy tail of relative width $1/\sqrt{d}$ which only vanishes in $d = \infty$. An application of $d = \infty$ with NNN hopping to $d = 3$ captures the aspect of non-particle-hole symmetric Hamiltonians, but the removal of the tail extending to negative infinite energy comes at the price of an unphysical singularity which is characteristic of very high dimensions $d \gtrsim 100$ only.

Similar observations can be made for the fcc lattice (formally $a^* = -\infty$) where in $d = 3$ one of the Van-Hove singularities is at the lower band edge leading to a logarithmic divergence. Hence, in this case, sharp band edges with singularities exist in $d = 3$ and $d = \infty$. At intermediate dimensionality, e.g., for $d = 4$, $d = 5$, and $d = 6$, however, the overall shape changes little with dimension while for $d > 10$ the

[8]This feature is absent in $d = 6$ for the alternative scaling $t' = t'^*/d$ (no results shown) to which (2.10) converges in the limit $d \to \infty$ . Scaling (2.10) is, however, strongly favorable since it ensures the same ratio of NN and NNN contributions to the variance of the energy in all dimensions; see (2.11).

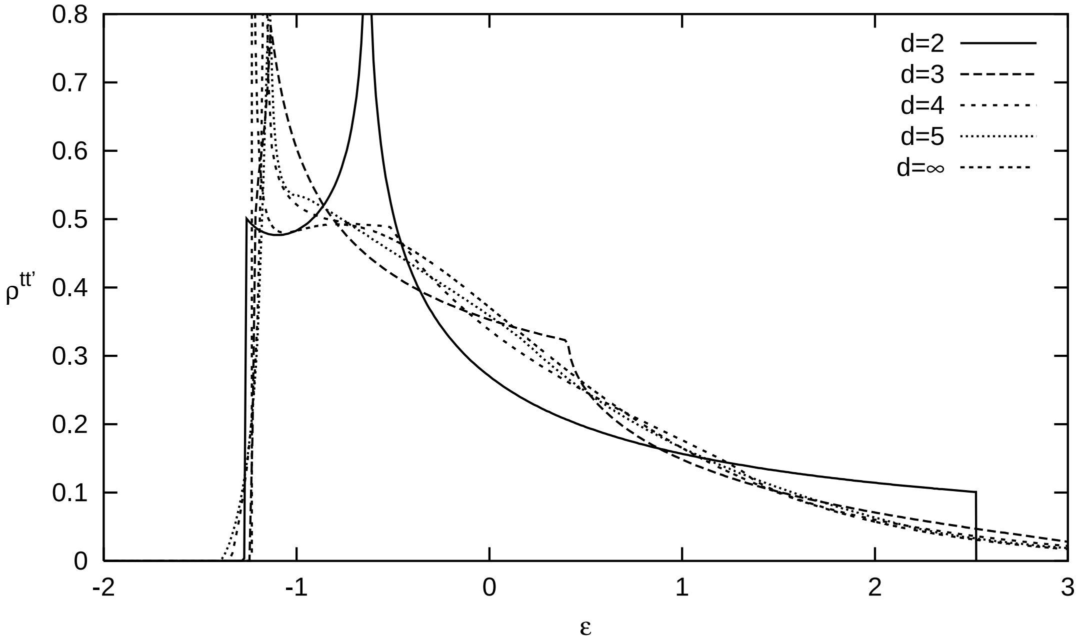



**Figure 2.5:** DOS of the hypercubic lattice with NN and NNN hopping, $a^* = -0.3333$ for low dimensions. The limit $d = \infty$ is included for comparison.

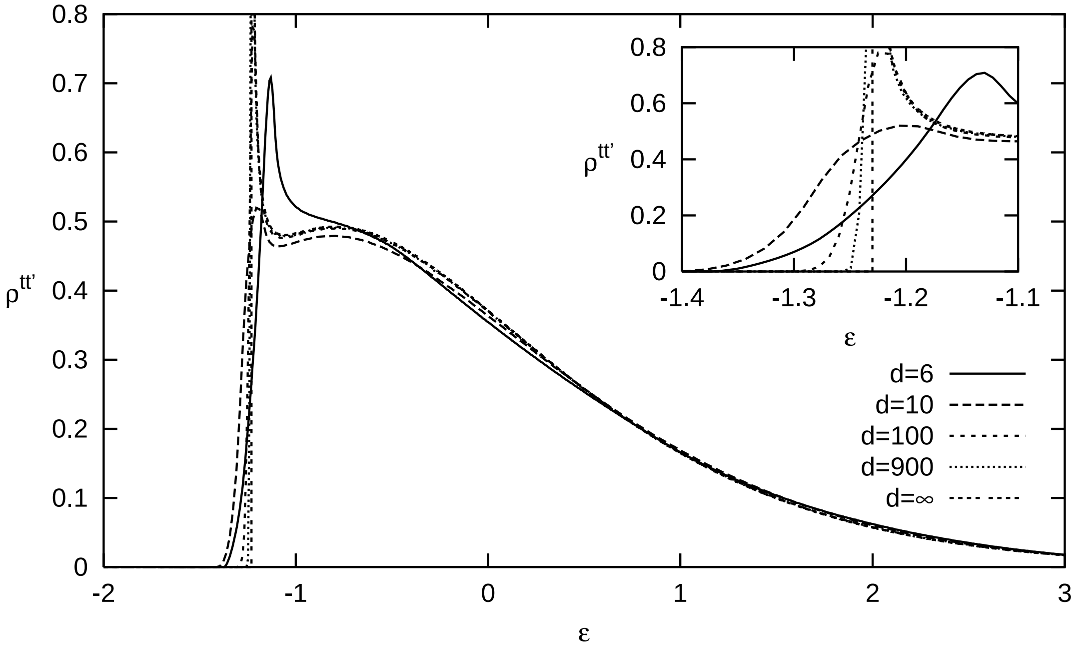



**Figure 2.6:** DOS of the hypercubic lattice with NN and NNN hopping, $a^* = -0.3333$ for high dimensions. The magnified inset shows that the form at the lower band edge is not singular for moderately high dimensions $d \approx 10$.

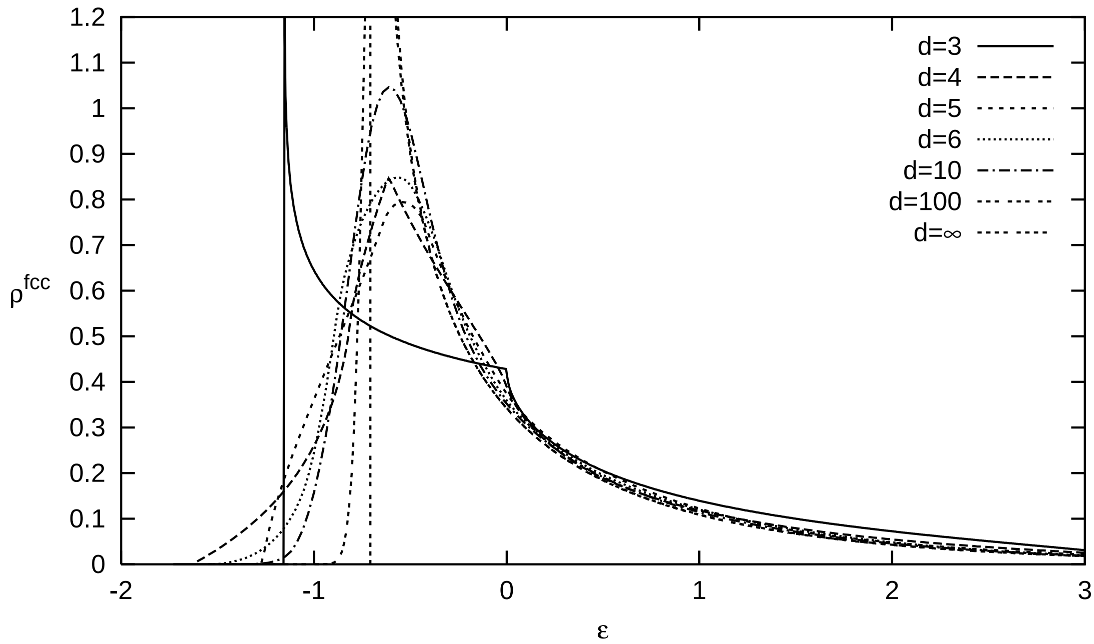



**Figure 2.7:** DOS of the fcc lattice.

maximum value of $\rho^{\text{fcc}}(\epsilon)$ increases significantly and develops into the square-root singularity observed at $d = \infty$ (see Fig. 2.7). Again, the singularity and the absence of the lower energy tail in this limit are not robust features when the dimensionality is decreased and thus results obtained in $d = \infty$ which depend on the singularity (as opposed to only a strong asymmetry) do not carry over to physically interesting dimensions.

The general behavior seen for $t$-$t'$ hopping on the hc lattice and for NN hopping on the fcc lattice is easily explained by looking at the expressions for the energies for the hypercube with NN hopping,

$$\epsilon_{\boldsymbol{k}}^{\text{hc}} = -2t \sum_{\alpha=1}^{d} \cos(k_\alpha) \tag{2.14}$$

and the fcc lattice,

$$\begin{aligned}
\epsilon_{\boldsymbol{k}}^{\text{fcc}} &= 2t' \sum_{\substack{\alpha,\beta=1 \\ \alpha\neq\beta}}^{d} \cos(k_\alpha)\cos(k_\beta) && (2.15) \\
&= 2t' \Big[ \Big( \sum_{\alpha=1}^{d} \cos(k_\alpha) \Big)^2 - \sum_{\alpha=1}^{d} \cos^2(k_\alpha) \Big] && (2.16) \\
&= dt' \Big[ \Big( \sqrt{\frac{2}{d}} \sum_{\alpha=1}^{d} \cos(k_\alpha) \Big)^2 - 1 - \frac{1}{d} \sum_{\alpha=1}^{d} \cos(2k_\alpha) \Big] \, . && (2.17)
\end{aligned}$$

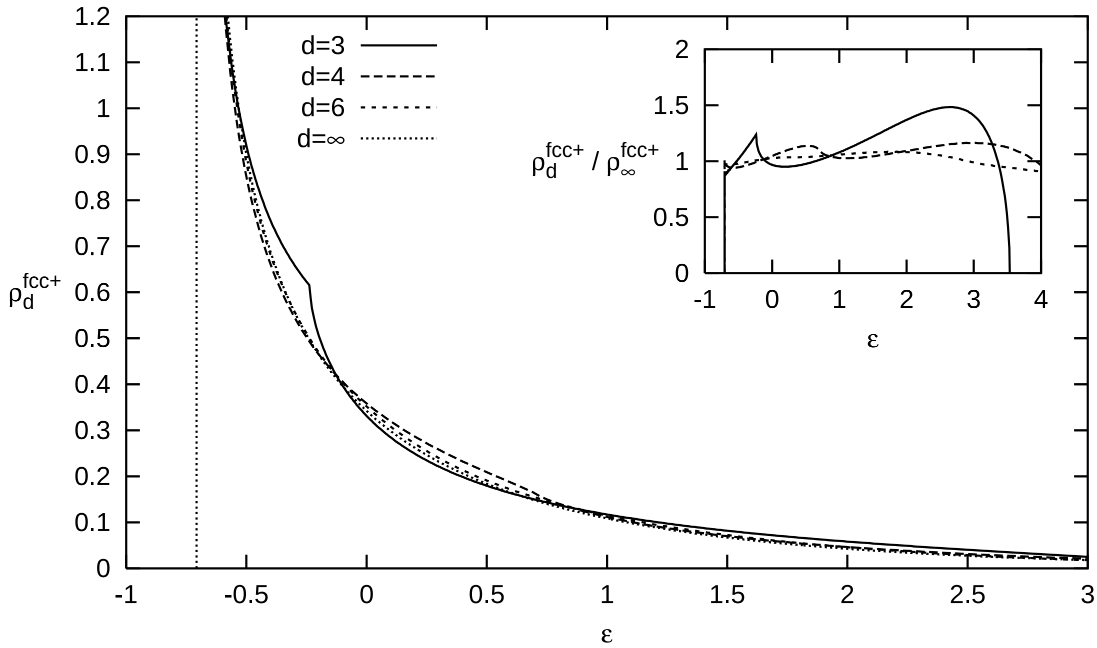



**Figure 2.8:** DOS of the fcc lattice with extended hopping along the axes.

Both the prefactor and the first two terms in the square bracket of (2.17) are $\mathcal{O}(1)$ while the third term is $\mathcal{O}(1/\sqrt{d})$ in high dimensions. Consequently, we can write

$$\epsilon_{\boldsymbol{k}}^{\text{fcc}} = \frac{(t')^*}{\sqrt{2}} \left[(\epsilon_{\boldsymbol{k}}^{\text{hc}}/t^*)^2 - 1\right] + \mathcal{O}(1/\sqrt{d}) \tag{2.18}$$

which for $d \to \infty$ implies[9] a square-root singularity in the DOS broadened by a Gaussian of width $1/\sqrt{d}$. Strictly in $d = \infty$, the DOS of the fcc lattice is given by

$$\rho^{\text{fcc}} = 2\frac{\mathrm{e}^{-(1+\sqrt{2}\epsilon)/2}}{\sqrt{1+\sqrt{2}\epsilon}\sqrt{2\pi}} \ . \tag{2.19}$$

Since the third term in (2.17) corresponds to extended hopping along the axes it can be cancelled by adding a corresponding term in the Hamiltonian in which case there is a square-root singularity in any dimension as shown in Fig. 2.8. Densities of states obtained via such a transformation $\epsilon_{\boldsymbol{k}} = f(\epsilon_{\boldsymbol{k}}^{\text{hc}})$ cannot show a stronger dimensional dependence than the hc DOS itself.[10] Indeed, the curves for $d > 3$ in Fig. 2.8 are hardly distinguishable. Even $d = 3$ only stands out only by the Van-Hove singularity at $\epsilon = -\sqrt{2}/6$, but has the same square-root singularity at the lower band edge.

While the dramatic changes of the DOS introduced by NNN hopping clearly show that extended hopping matrix elements are important we wish to point out that the

[9] As is apparent from (2.1), a DOS (which may be seen as probability distribution) transforms under change of variables like a delta-function, see (2.56).

[10] This fact will be utilized in Sec. 2.3.

inclusion of hopping elements $t_{ij}$ with odd range $D = \|\boldsymbol{r}_i - \boldsymbol{r}_j\| \geq 3$ will reintroduce infinite-ranged low-energy tails and thus destroy the sharp band edge, one of the prominent features of $t$-$t'$ DOS in $d = \infty$. This issue will be treated in a broader context in Sec. 2.3. For the DMFT computation of transport properties for the $t$-$t'$ lattice, see subsection 4.5.3.

For a discussion of the hyperdiamond lattice, see App. B.

### 2.1.3 Magnetic Frustration and Asymmetry of the DOS

One important question in the context of the DMFT is whether magnetic frustration, in particular the suppression of antiferromagnetism, is closely related to other properties like the DOS and the optical conductivity or whether it can be regarded as being essentially independent, at least in high dimensions. While it is immediately clear that unfrustrated particle-hole symmetric systems have a symmetric DOS, the converse statement is not generally true for $Z = \infty$. In fact, when extended hopping only along the cartesic axes is introduced for the hypercubic lattice, the DOS remains Gaussian while for suitably chosen hopping matrix elements antiferromagnetic order is suppressed (Müller-Hartmann, 1989*a*). In the following, we will outline an argument why this behavior (and not only the model) should be regarded as pathological. First, we observe that for fixed hopping matrix elements $t_D$ to $D^{\text{th}}$-nearest neighbors along the cartesic axes the DOS in $d$ dimensions can be obtained from folding the one-dimensional DOS $d-1$ times with itself. Since the DOS is symmetric in $d = 1$ if and only if the system is unfrustrated (i.e., $t_{2n} = 0$ for integer $n$) the same applies for all finite dimensions $d$. In Sec. 1.2 we have seen that perfect anticorrelation in the two-particle density of states $D_{\boldsymbol{q}}(\epsilon_1, \epsilon_2)$ at the antiferromagnet wave vector $\boldsymbol{Q}$, i.e., $\eta(\boldsymbol{Q}) = -1$ is associated with a strong antiferromagnetic instability due to particle-hole symmetry. More generally, an instability at vector $\boldsymbol{q}$ occurs for $\eta_{\boldsymbol{q}} \equiv \eta(\boldsymbol{q}) \approx -1$. It seems thus reasonable to associate magnetic frustration with the deviation of the minimum of $\eta$ from $-1$. For the lattice under consideration it is clear that all extrema are on the diagonal $\{\boldsymbol{q}; \boldsymbol{q} = \lambda\boldsymbol{Q}; 0 \leq \lambda \leq 1\}$. For this line, we have[11] (Müller-Hartmann, 1989*a*)

$$\eta(\lambda\boldsymbol{Q}) = \frac{1}{t^2}\sum_{D=1}^{\infty} t_D^2 \cos(D\lambda\pi)\,, \tag{2.20}$$

where $t^2 = \sum_{D=1}^{\infty} t_D^2$. First, we note that $\eta(\lambda\boldsymbol{Q}) \geq -1$ and that $\eta(\boldsymbol{Q}) = -1$ only in the unfrustrated case where $t_{2n} = 0$ for all $n$. For a fully quantitative analysis, we restrict hopping to NN and to NNN (along the cartesic axes). Then, the minimum of $\eta(\lambda\boldsymbol{Q})$ is for unit variance (${t_1^*}^2 + {t_2^*}^2 = 1$)

$$\eta_{\min} = \begin{cases} -1 + 2{t_2^*}^2 & \text{for } {t_2^*}^2 \leq 0.2 \\ -\frac{9{t_2^*}^4 - 2{t_2^*}^2 + 1}{8{t_2^*}^2} & \text{for } 0.2 < {t_2^*}^2 \leq 1 \end{cases}\,. \tag{2.21}$$

[11] Note that (in this special case) we can use either scaled or unscaled hopping matrix elements since the scaling factor is independent of the hopping range.

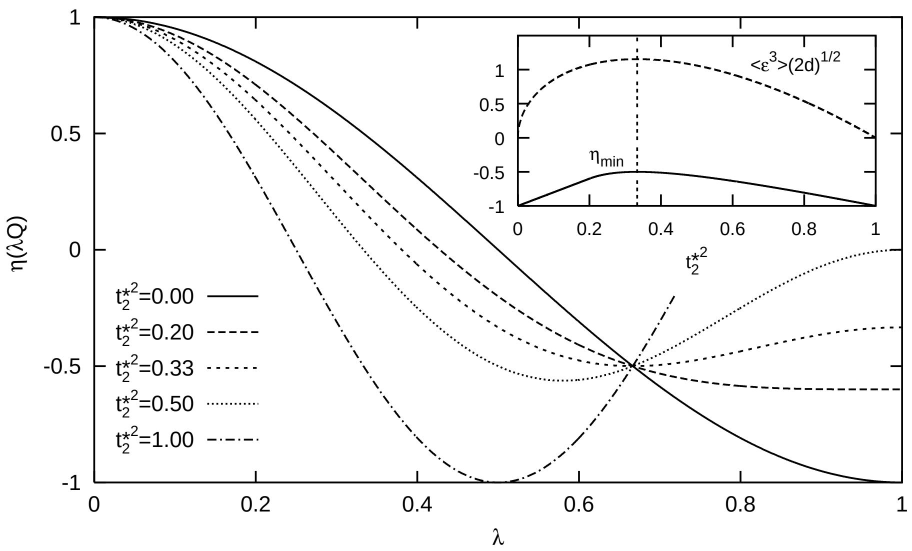



**Figure 2.9:** DMFT momentum parameter $\eta(\lambda \boldsymbol{Q})$ [(1.23) and (2.20)] for the hypercubic lattice with hopping to NN and to NNN along the cartesic axes using scaled amplitudes $1 - t_2^{*2}$ and $t_2^{*2}$, respectively. The symmetric graph for $t_2^{*2} = 1$ is not fully shown. Inset: minimal value of $\eta(\boldsymbol{q})$ and rescaled third moment of the DOS as a function of $t_2^{*2}$; both frustration and asymmetry of the DOS are maximal at $t_2^{*2} = 1/3$.

---

This value, i.e., the frustration is maximal for $t_2^{*2} = 1/3$ (when $\eta(\boldsymbol{q}) \geq -1/2$ for all $\boldsymbol{q}$). This is exactly the value for which also the asymmetry of the DOS, measured by its third moment, $\langle \epsilon^3 \rangle = 6d\, t_1^2 t_2 = 3(1 - t_2^{*2}) t_2^* / \sqrt{2d}$ acquires its maximum. Thus, for finite dimensions, maximal frustration coincides with maximal asymmetry of the DOS. This effect becomes only unobservable for $Z \to \infty$ since the associated weight becomes exponentially small.[12] The main part of Fig. 2.9 shows $\eta(\lambda \boldsymbol{Q})$ for different values of $t_2^{*2}$. The minimum is at $\lambda = 1$ for $t_2^{*2} \leq 0.2$ and approaches $\lambda = 0.5$ for $t_2^{*2} \to 1$ (which is again unfrustrated). The value of the minimum as a function of $t_2^{*2}$ is shown in the inset together with the scaled third moment which illustrates that maximal frustration coincides with maximal asymmetry for $t_2^{*2} = 1/3$. We will find more indications that magnetic frustration is generically associated with an asymmetry of the DOS by using the renormalized perturbation expansion in subsection 2.2.2.

## 2.2 Bethe Lattice, RPE, and Disorder

Within this work and the DMFT literature in general, the term "Bethe lattice" is most often used in the sense of a lattice having the "Bethe DOS", i.e., a semi-elliptic DOS, without paying too much attention to the particular properties of this pseudo-

[12] Note that the ratio of the band edges, i.e., of the maximum and minimum energies, is independent of $d$. With respect to this criterion, the asymmetry is maximal for $t_2^{*2} = 1/5$, where the ratio is 0.5 or 2 (depending on the signs of $t_1$ and $t_2$).

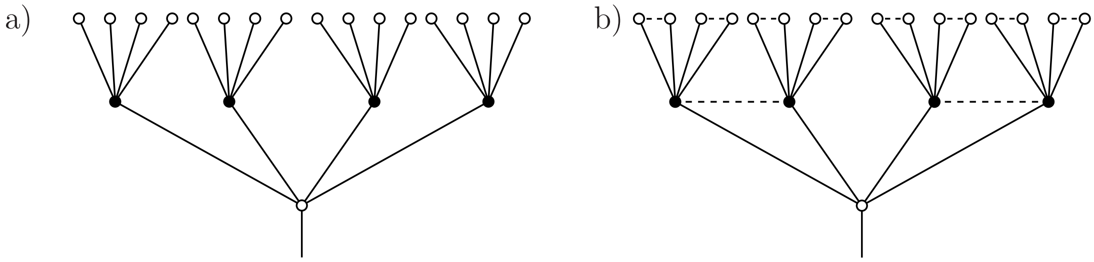



**Figure 2.10: a)** Scheme of Bethe tree for connectivity $K = 4$ (i.e., $Z = 5$). **b)** Scheme of triangular Husimi cactus for connectivity $K = 2$ (i.e., $Z = 6$) which differs locally from the Bethe tree by the presence of a single bond (dashed lines). Black and white dots denote A and B sites on the bipartite Bethe tree, respectively. Corresponding AF order is partially frustrated on the corresponding Husimi tree in finite dimensions.

lattice. In this section, we will take a closer look at the implicit definitions and come to the conclusion that the original construction need not and should not be taken too seriously in the context of applications of the DMFT method to correlated electron systems. We will also consider a Bethe lattice with NNN hopping and derive some more general results from an expansion in terms of self-avoiding loops, the RPE.

### 2.2.1 Bethe Tree, Cayley Tree, and Husimi Cactus

The Bethe lattice of connectivity $K$ is defined as a pseudo-lattice where each site is surrounded by $Z = K + 1$ nearest neighbors and where no closed loops are formed by NN bonds (Economou, 1979) which makes it a tree in the sense of graph theory.[13] While this construct coincides with the usual $d = 1$ lattice for $Z = 2$, it is not a regular lattice for $Z > 2$ as is visualized in Fig. 2.10 a for $K = 4$, i.e., $Z = 5$. Here it is shown as a directed graph where a level number $i \in \mathbb{Z}$ can be assigned to each site, so that a site at level $i$ is connected by one edge to a site in the lower level $i - 1$ and by $K$ edges to sites in the higher level $i + 1$. Its name stems from the fact that the Bethe-Peierls approximation (Bethe, 1935; Peierls, 1936) for crystalline alloys or Ising models becomes exact on this structure (Kurata, Kikuchi, and Watari, 1953).[14] The Bethe tree is often viewed as the infinite-size limit of a Cayley tree which is a finite mathematical tree where all nodes except for the leaves are connected by edges to $Z$ other nodes. The infinite-size limit of the Cayley tree is not unproblematic, since a "spherical" Cayley tree where all leaves have (taxi cab) distance $n$ from a selected central site consists of $N = [Z(Z - 1)^n - 2]/(Z - 2)$ sites of which $N_s =$

[13]A tree is best defined as a connected, planar graph without loops, where a graph is a topological structure which can be represented as a set of points (nodes) and a set of lines (edges) connecting a subset of the nodes. Final segments and the nodes at their ends are called leaves (see, e.g., Weisstein (1998)).

[14]It is easily seen that the Ising model without external field becomes trivial on a lattice without loops by replacing the site variables through (then independent) bond variables $\tau_{ij} = \sigma_i\sigma_j$. Also note that an extension of the Bethe-Peierls approximation to a model with diagonal disorder can be used in order to study Anderson localization (Kawarabayashi, 1993).

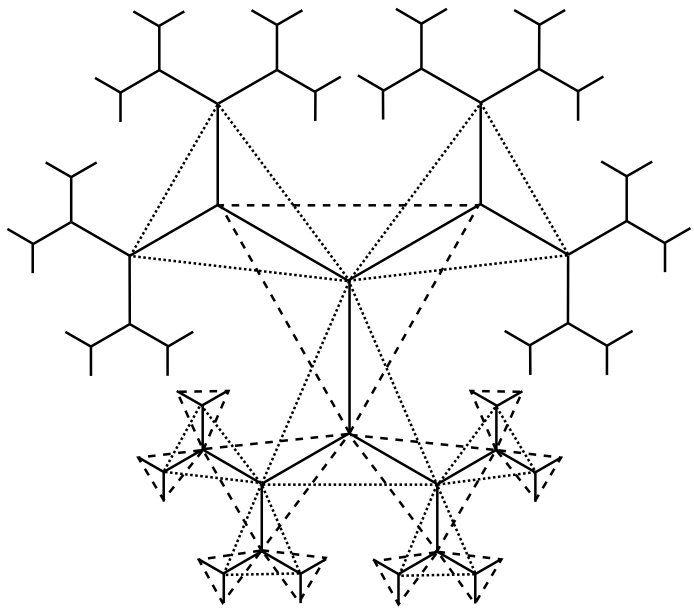


**Figure 2.11:** Scheme of spherical Cayley tree with connectivity 2 and radius 5 (black solid lines). A selected set of NNN bonds is marked by dashed and dotted lines. Each sublattice formed by the NNN bonds is fully frustrated with respect to AF order.

---

$Z(Z-1)^{n-1}$ constitute the surface (i.e., are leaves). This is illustrated in Fig. 2.11 schematically for radius $n = 5$ and $Z = 3$. In the limit of infinite radius $n$, a finite fraction $N_s/N \approx (Z-2)/(Z-1)$ of sites is on the surface (Thorpe, 1981). A tight-binding model defined on such a Cayley tree would be dominated by surface states, in particular in the limit $Z \to \infty$, where $N_s/N \to 1$. This problem can be avoided by using the bulk definition given at the beginning of the paragraph. The Bethe lattice is bipartite; therefore, the generic low-temperature phase of a Hubbard model defined on the Bethe lattice is antiferromagnetic at half filling.

The Husimi cactus is a generalization of the Bethe lattice where each bond is part of exactly one loop and is shown in its triangular form[15] in Fig. 2.10 b. For finite $Z$, these loops introduce some degree of frustration. Since a single loop becomes irrelevant for $Z \to \infty$, however, all properties of the Husimi cactus converge to that of the Bethe tree in this limit. In contrast, the inclusion of hopping along all NNN bonds (dashed lines in Fig. 2.11) leads to strong magnetic frustration in all dimensions and to an asymmetric DOS (see subsection 2.2.2).

The absence of loops greatly simplifies not only calculations for spin models, but also allows for an exact computation of the Green function for the noninteracting tight-binding model with nearest-neighbor hopping for arbitrary $Z$. Using a mean-field approach in the spirit of the Bethe-Peierls approximation (Thorpe, 1981) or the renormalized perturbation expansion around the atomic limit (Economou, 1979), one

[15]In the general definition of the Husimi cactus, both the type of polygons, i.e., the length of loops, and the connectivity for each site can be chosen arbitrarily.

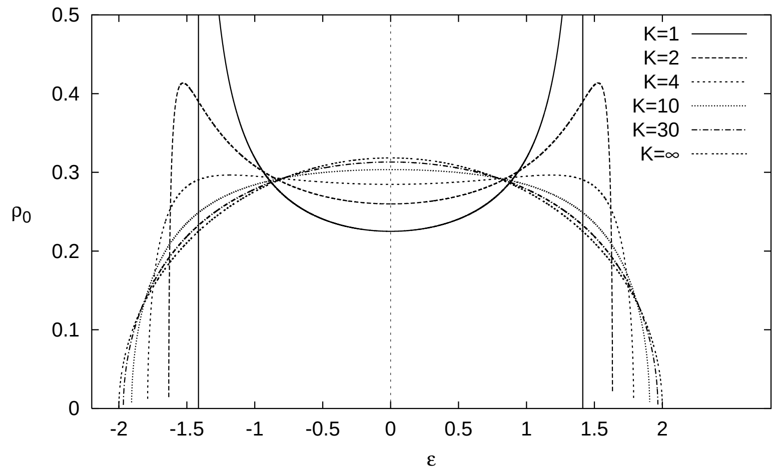



**Figure 2.12:** DOS (2.23) of the Bethe tree with finite connectivity $K$. It evolves from the singular form of a chain at $K = 1$ to a semi-elliptic form for $K \to \infty$.

obtains, e.g., for the diagonal element (hopping amplitude $t$, on-site energy $\epsilon_0 = 0$),

$$G_{00}(\omega) = \frac{2(Z-1)}{(Z-2)\omega + Z\sqrt{\omega^2 - 4(Z-1)t^2}}, \tag{2.22}$$

where the imaginary part of the square root is taken to be positive for a retarded Green function. The density of states,

$$\rho(\epsilon) = \frac{Z}{2\pi} \frac{\sqrt{4(Z-1)t^2 - \epsilon^2}}{(Zt)^2 - \epsilon^2}, \tag{2.23}$$

is shown in Fig. 2.12 for a hopping amplitude scaled as $t = t^*/\sqrt{Z}$ and $t^* = 1$. The striking feature of the Bethe DOS that the full bandwidth (including all states) is $\mathcal{O}(Z^{1/2})$ instead of $\mathcal{O}(Z)$ (which would be expected from Perron's theorem) can be attributed to the absence of long-wavelength excitations; see Thorpe (1981) and references therein. The spectral weight, concentrated at the band edges for low $Z$, shifts toward the center with increasing $Z$ until a semi-elliptic form,

$$\rho^{\text{Bethe}}(\epsilon) = \frac{1}{2\pi}\sqrt{4-\epsilon^2}\,. \tag{2.24}$$

is obtained for $Z = \infty$ which we will label as "Bethe DOS" in the remainder of this work.

While the reason for using this Bethe DOS within the DMFT is the presence of sharp band edges, the semi-elliptic form also has the advantage that the $\boldsymbol{k}$ integrated Dyson equation (1.30) can be performed and inverted analytically

$$G(\omega) \;=\; \frac{z(\omega)}{2}\left(1 - \sqrt{1 - \frac{4}{z^2(\omega)}}\right), \tag{2.25}$$

$$\Sigma(\omega) \;=\; \omega + \mu - G(\omega) - 1/G(\omega)\,, \tag{2.26}$$

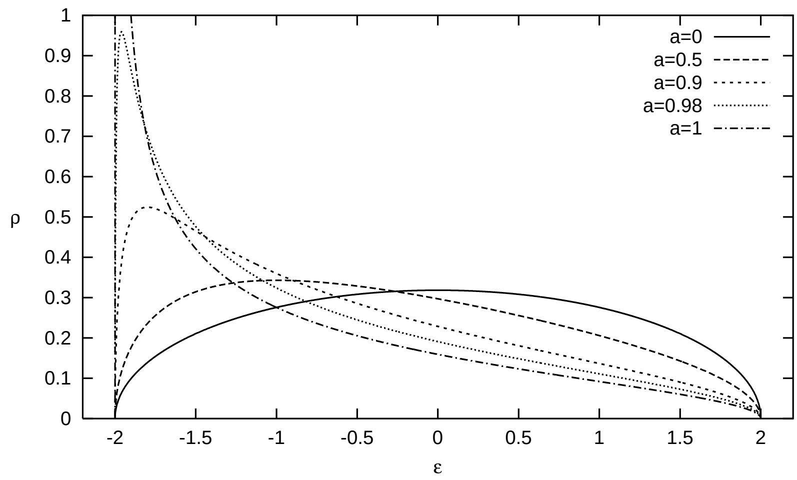



**Figure 2.13:** Asymmetric model DOS (2.27). For $a = 0$, the semi-elliptic DOS of the Bethe lattice (for $Z = \infty$) is recovered. As shown below, the extremely asymmetric case $a = 1$ corresponds to a Bethe lattice with only NNN hopping (also for $Z = \infty$).

---

where $z(\omega) = \omega + \mu - \Sigma(\omega)$.

Giving up the strict connection to a (somewhat unphysical) model, one may break the particle-hole symmetry and concentrate spectral weight near one of the band edges by introducing an asymmetry parameter $a$ (Wahle, Blümer, Schlipf, Held, and Vollhardt, 1998)

$$\rho^{\mathrm{a}}(\epsilon) = \frac{1 + \sqrt{1 - a^2}}{2\pi} \frac{\sqrt{1 - (\epsilon/2)^2}}{1 + a\epsilon/2} \,. \tag{2.27}$$

Here, $-1 \leq a \leq 1$; $a = 0$ corresponds to the semi-elliptic DOS. As apparent in Fig. 2.13, the asymmetry develops into a square-root divergence at one of the band edges for $|a| \to 1$. The particular advantage of (2.27) is that the analytic solution and inversion of the corresponding lattice Dyson equation is still possible.

### 2.2.2 Renormalized Perturbation Expansion

We will now derive very general relationships between weights of self-avoiding loops in lattices and moments of the corresponding DOS. Interesting general consequences for lattices with semi-elliptic DOS will follow in subsection 2.2.2.

We assume a finite total bandwidth, i.e., that all energy levels are in the range $[\epsilon_a, \epsilon_b]$. Even when the assumption is not always fulfilled for $Z = \infty$ it can be relaxed in the relevant cases by defining a suitable limiting process for $Z \to \infty$.[16] We use the spectral representation of the noninteracting Green function to write for

[16] One exception is the fully connected model, for which $\langle e \rangle \neq 0$ (van Dongen and Vollhardt, 1989); here no limiting form of the DOS exists for $Z \to \infty$ (and all moments $\langle \epsilon^i \rangle$ for $i > 1$ are infinite).

$\omega > \max\{|\epsilon_a|, |\epsilon_b|\}$

$$G(\omega) = \int_{-\epsilon_a}^{\epsilon_b} \mathrm{d}\epsilon \, \frac{\rho(\epsilon)}{\omega - \epsilon} = \int_{-\epsilon_a}^{\epsilon_b} \mathrm{d}\epsilon \, \rho(\epsilon) \frac{1}{\omega} \sum_{i=0}^{\infty} \left(\frac{\epsilon}{\omega}\right)^i = \sum_{i=0}^{\infty} \frac{\langle \epsilon^i \rangle}{\omega^{i+1}} \,. \tag{2.28}$$

Thus, the high-frequency behavior of $G$ is easily obtained from moments of the DOS and vice versa. Obviously, a necessary condition for (2.28) and the following conclusions to hold is that all moments $\langle \epsilon^i \rangle$ exist.

The Renormalized Perturbation Expansion (RPE) is an expansion in the hopping matrix elements around the atomic limit where all contributions to the local Green function for some site $i$ can be obtained by considering all decorated self-avoiding loops starting and ending at $i$. While in finite dimensions the decorations must be computed in terms of special self-avoiding loops which do not touch any site already contained in loops of lower hierarchy, this restriction becomes irrelevant in the limit $Z \to \infty$ where the decorations are nothing but full local Green functions. Thus, for the simplest case of full translational symmetry one arrives at a single closed self-consistency equation for some site $i$,

$$G^{-1}(\omega) = \omega - \left( G \sum_j t_{ij}^2 + G^2 \sum_{j \neq k} t_{ij} t_{jk} t_{ki} + G^3 \sum_{j \neq k \neq l; j \neq l} t_{ij} t_{jk} t_{kl} t_{li} + \dots \right) \tag{2.29}$$

$$=: \omega - \sum_{n=2}^{\infty} a_n G^{n-1} \,, \tag{2.30}$$

where the coefficients $a_n$ measure the weight of self-avoiding loops of length $n$. Note that when all nonzero hopping matrix elements are equal, the simpler expression $a_n = t^n N_n$ applies, where $N_n$ is the number of self-avoiding loops of length $n$ originating from site $i$ which implies a one-to-one correspondence between the DOS and the topology of the lattice.

Multiplying (2.30) by $G$ and inserting the expansion (2.28) we can compute moments of the DOS from the hopping weights order by order

$$\langle \epsilon^0 \rangle = 1 \,; \qquad \langle \epsilon^1 \rangle = 0 \,; \qquad \langle \epsilon^2 \rangle = a_2 \,. \tag{2.31}$$

Enforcing, for simplicity, unit variance by setting the weight $a_2$ of 2-loops to 1, the higher order moments read

$$\langle \epsilon^2 \rangle = 1 \,; \qquad \langle \epsilon^3 \rangle = a_3 \tag{2.32}$$

$$\langle \epsilon^4 \rangle = 2 + a_4 \,; \quad \langle \epsilon^5 \rangle = 2 + 3a_3 + a_5 \,. \tag{2.33}$$

It is easy to check that only (2.33) and corresponding higher order expressions are specific to $d = \infty$, while (2.32) also applies to finite dimensions: the third moment of any DOS is directly determined only by the weight of self-avoiding 3-loops (i.e., the scaled number of 3-loops if all nonvanishing matrix elements are equal for a translation invariant lattice, $\langle \epsilon^3 \rangle = t^3 N_3$). We can thus spot one of the most prominent mechanisms for frustration (which determines, e.g., the magnetic properties of fcc lattices) by just looking at the DOS and measuring its third moment.[17] More generally, a symmetric DOS implies irrelevance or cancellation of loops with odd length.

### Application to the Bethe lattice

An application of the above formalism to the Bethe lattice is particularly simple, since the only self-avoiding loops have length 2 and thus the series in (2.29) terminates at the first term,

$$G^{-1}(\omega) = \omega - G(\omega) \tag{2.34}$$

from which (2.25) follows directly. Introduction of NNN hopping on the Bethe lattice completely changes its topology and introduces loops of arbitrary order as is illustrated in Fig. 2.11. For $d = \infty$ and unit variance, we have in this case

$$a_3 = t'^*(t^{*2} + t'^{*2}); \quad a_4 = 8t^{*2}t'^{*2} + t'^{*4}; \quad a_5 = 20t^{*2}t'^{*3} + t'^{*5}\,. \tag{2.35}$$

Already at this point we can see that the third moment is nonvanishing for all finite $t'^*$ and consequently the DOS must be asymmetric. This result contradicts a claim of the Kotliar group (Rozenberg, Kotliar, Kajüter, Thomas, Rapkine, Honig, and Metcalf, 1995) that the DOS of the Bethe lattice with NN and NNN hopping should be always semi-elliptic.[18] While it might be possible to find a closed solution for $a_n$ for arbitrary $n$ and thus compute reliable numerical estimates via Padé approximants, it is clearly impossible to obtain a closed solution for $G$ and the DOS in the general case along this route (it would, in fact, be very surprising when the full exact result could be written in closed form at all). At least, we are able to present the full solution for the Bethe lattice with NNN, but no NN hopping (i.e., for $t^* = 0$). For this special case and unit variance we have $a_n = 1$ for all $n$ due to the very special topology in this limit where all sites visited by each self-avoiding loop of length $n \geq 3$ are NN of the same site (not on the loop). Therefore (2.29) reduces to a geometric series and we obtain for $Z = \infty$[19]

$$G(\omega) = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{1+\omega}} \tag{2.36}$$

$$\rho(\omega) = \frac{1}{2\pi}\frac{\sqrt{3-\omega}}{\sqrt{1+\omega}}, \tag{2.37}$$

which agrees with the model DOS (2.27) for the case of extreme asymmetry ($a = 1$) when the frequency is shifted by $\omega \to \omega+1$. Thus, we have shown that the model DOS

[17]This observation implies that 3-loops do not survive the limit $d \to \infty$ for the hc lattice with extended hopping along the axes (Müller-Hartmann, 1989*a*) which we verified in subsection 2.1.3 and makes the mechanism for magnetic frustration present in this lattice type appear highly pathological.

[18]In fact, our observation essentially invalidates all results concerning the microscopic treatment of frustration in the cited paper. We mention in passing that the sum rule for the optical conductivity stated in this paper is also incorrect, even for the Bethe lattice with only NN hopping (cf. Sec. 4.4).

[19]The generalization to finite $Z$ seems straightforward and might be considered in an upcoming publication.

interpolates between Bethe lattices for $Z = \infty$ with only NN and only NNN hopping. In order to exclude that there is an accidental connection to the microscopic model also for $0 < a < 1$, we determined the leading nontrivial corresponding moments of the DOS (2.27) for $a = 0.9$ (which are $\langle(\epsilon - \langle\epsilon\rangle)^3\rangle \approx 0.6268$, $\langle(\epsilon - \langle\epsilon\rangle)^4\rangle \approx 2.3929$, $\langle(\epsilon - \langle\epsilon\rangle)^5\rangle \approx 3.3802$) and fixed the ratio $t'^*/t^* \approx 0.2208$ to fulfill (2.32) and (2.35). This leads to $a_4 = 0.3568$, $a_5 = 0.1917$ and corresponds to $4^{\text{th}}$ and $5^{\text{th}}$ moments of 2.3568 and 4.0720, respectively. Clearly, the results are not compatible; $a_5$ would even have to have the opposite sign in order to reproduce the $5^{\text{th}}$ moment of the model DOS.

From the above we may further conclude that the weights $a_n$ for self-avoiding loops of length $n > 2$ must cancel exactly also for the tight-binding Hamiltonian having a semi-elliptic DOS which we will construct in the following section. While this consideration implies that the Bethe tree is the only (pseudo)lattice with semi-elliptic DOS and only one nonzero hopping parameter it could in principle also be used to verify or even construct Hamiltonians. In practice, however, there exist many contributions, e.g., to the sum over 4-loops on hypercubic lattices, even for a small set of longer range hopping terms. For hopping $t$ to NN and $t_3$ to $3^{\text{rd}}$-nearest neighbors we found on the hc lattice the leading contribution

$$a_4 \approx t^{*4} + 36t^{*2}t_3^{*2} + 4\sqrt{6}t^{*3}t_3^* + 48\sqrt{6}t^*t_3^{*3} + 108t_3^{*4} \qquad (2.38)$$

which adds up to -0.04 when the values from Table 2.1 (page 55) are used. Contributions from longer-range hopping are larger than might be expected from the small value of this partial sum. One contribution is, e.g., $12\sqrt{20}t^{*2}t_3^*t_5^* \approx -0.31$. We find it remarkable that the semi-elliptic DOS imposes such severe constraints (i.e., the cancellation of all loops of length greater than 2) on corresponding tight-binding models, even though this criterion appears not to be of general constructive value.

### Off-diagonal disorder

One further application where the RPE can provide much insight is off-diagonal disorder. For independent random variables $t_{ij}$, all contributions from self-avoiding loops of order $n > 2$ vanish exactly as long as the hopping elements vanish on average, $\langle t_{ij}\rangle = 0$. The remarkable conclusion is that noninteracting Green function and DOS agree with that of the Bethe lattice (with NN hopping) for arbitrary topology in this case. This statement, valid for $Z = \infty$, is more general than Wegner's treatment of "gauge invariant" tight-binding models which predicts local Bethe lattice properties only for systems with a fully symmetric distribution of hopping matrix elements, $\langle t_{ij}^{2n+1}\rangle = 0$ (Wegner, 1976). In our view, the tighter restrictions for finite $Z$ can be linked to the fact that in this case not only the weight of self-avoiding loops, but of all loops (of length $n > 2$) has to vanish in order to reduce the effective mean free path to one lattice spacing, which is the physics of all models with Bethe DOS. In the context of DMFT calculations, frustration caused by disorder of the above type was considered for a fully connected model (Georges and Krauth, 1993) and then also for the more general cases (Dobrosavljević and Kotliar, 1993; Dobrosavljević and Kotliar, 1994).

As a last conclusion from this section let us note that additional hopping (e.g., along NNN bonds) does *not* modify the semi-elliptic DOS of the Bethe lattice when these additional bonds are fully disordered (with zero mean hopping amplitudes). In contrast to the clean case it is, thus, possible to tune magnetic frustration in such a system by variation of $\langle t_2^{*2} \rangle$ at constant DOS.[20]

## 2.3 General Density of States in $d = \infty$

In this section, we will present a general algorithm for finding an isotropic, translationally invariant tight-binding Hamiltonian corresponding to an arbitrary DOS in $d = \infty$. While the availability of such a method to construct a microscopic Hamiltonian corresponding to some model DOS might be useful in any dimension, it is absolutely essential for considering spatially inhomogeneous phases (e.g, antiferromagnetism or charge density waves) or computing transport properties within the DMFT. The algorithm developed in this section will later enable us to perform the first fully consistent DMFT calculation of the optical conductivity for a periodic and isotropic lattice (cf. footnote 3 on page 28) with the Bethe semi-elliptic DOS (see Sec. 4.4).

**Scaling Relations**

We start by rewriting the usual noninteracting Hamiltonian with general (translationally invariant) hopping,[21]

$$\hat{H}_0 = \sum_{ij,\sigma} t_{ij}\, \hat{c}^\dagger_{\boldsymbol{R}_i\sigma} \hat{c}_{\boldsymbol{R}_j\sigma} = \sum_{i,\sigma} \sum_{\boldsymbol{\tau}} t_{\boldsymbol{\tau}}\, \hat{c}^\dagger_{\boldsymbol{R}_i\sigma} \hat{c}_{\boldsymbol{R}_i+\boldsymbol{\tau},\sigma} \tag{2.39}$$

$$= \sum_{\boldsymbol{k},\sigma} \epsilon(\boldsymbol{k}) \hat{n}_{\boldsymbol{k}\sigma}, \tag{2.40}$$

where

$$\epsilon(\boldsymbol{k}) = \sum_{\boldsymbol{\tau}} t_{\boldsymbol{\tau}} \mathrm{e}^{i\boldsymbol{\tau}\cdot\boldsymbol{k}} = \sum_{\boldsymbol{\tau}} \left( \frac{t_{\boldsymbol{\tau}} + t^*_{\boldsymbol{\tau}}}{2} \cos(\boldsymbol{\tau}\cdot\boldsymbol{k}) + i \frac{t_{\boldsymbol{\tau}} - t^*_{\boldsymbol{\tau}}}{2} \sin(\boldsymbol{\tau}\cdot\boldsymbol{k}) \right). \tag{2.41}$$

Classifying the hopping vectors $\boldsymbol{\tau}$ according to their length measured by the taxi-cab metric, we may split up the kinetic energy accordingly,

$$\epsilon(\boldsymbol{k}) = \sum_{D=1}^{\infty} \epsilon_D(\boldsymbol{k}), \qquad \epsilon_D(\boldsymbol{k}) = \sum_{||\boldsymbol{\tau}||=D} t_{\boldsymbol{\tau}} \mathrm{e}^{i\boldsymbol{\tau}\cdot\boldsymbol{k}}. \tag{2.42}$$

For $D > 1$, each set of vectors $\{\boldsymbol{\tau}; ||\boldsymbol{\tau}|| = D\}$ consists of different subsets which are pairwise inequivalent such as, for $D = 2$, the subsets $\{\boldsymbol{\tau}; \boldsymbol{\tau} = \pm\boldsymbol{e}_\alpha \pm \boldsymbol{e}_\beta; \alpha \neq$

[20] The same applies to the regular, translation invariant model with semi-elliptic DOS which we will construct in the next section: additional even-range disorder hopping terms could introduce frustration without changing the DOS.

[21] For sake of clarity, we write $\epsilon(\boldsymbol{k})$ instead of $\epsilon_{\boldsymbol{k}}$ for the full dispersion [i.e., the full sum in (2.42)] throughout this section.

$\beta\}$ and $\{\boldsymbol{\tau}; \boldsymbol{\tau} = \pm 2\boldsymbol{e}_\alpha\}$. Consequently, one may expect the corresponding hopping amplitudes to be different. We will choose to neglect all hopping vectors with $|\boldsymbol{\tau}\cdot\boldsymbol{e}_\alpha| > 1$ for any direction $\alpha$, i.e., only consider vectors of the form $\boldsymbol{\tau} = \sum_{i=1}^{D} \boldsymbol{e}_{\alpha_i}$ with pairwise different directions $\alpha_i$. This selection can be justified by the following observations:

- For fixed $|\boldsymbol{\tau}|$, the vectors considered are of minimal Euclidean length $|\boldsymbol{\tau}|$ hinting at maximal overlap, i.e., largest $|t_{\boldsymbol{\tau}}|$.
- The fraction of neglected vectors is $\mathcal{O}(1/d)$ in high dimensions (as long as $d \gg D$, cf. footnote 5 on page 30).

For real (and therefore symmetric) hopping amplitudes[22] and in the isotropic case, the amplitudes $t_{\boldsymbol{\tau}}$ are functions of $D = \|\boldsymbol{\tau}\|$ alone, $\epsilon_D(\boldsymbol{k}) = t_D \sum_{\|\boldsymbol{\tau}\|=D} \mathrm{e}^{i\boldsymbol{\tau}\cdot\boldsymbol{k}}$. As in the case of $t$-$t'$ hopping (which corresponds to $D = 1, 2$ in the present scheme), there are no cross terms in the variance

$$\langle \epsilon^2(\boldsymbol{k})\rangle = \sum_{\boldsymbol{\tau},\boldsymbol{\tau}'} t_{\boldsymbol{\tau}} t_{\boldsymbol{\tau}'} \langle \mathrm{e}^{i(\boldsymbol{\tau}-\boldsymbol{\tau}')\cdot\boldsymbol{k}}\rangle = \sum_{\boldsymbol{\tau}} |t_{\boldsymbol{\tau}}|^2 = \sum_{D} \langle \epsilon_D^2(\boldsymbol{k})\rangle \tag{2.43}$$

$$\langle \epsilon_D^2(\boldsymbol{k})\rangle = t_D^2 N_D; \quad N_D = 2^D \binom{d}{D} \approx \frac{(2d)^D}{D!}\,. \tag{2.44}$$

Here, $N_D$ is the number of hopping terms in shell $D$ and the approximation in (2.44) is valid for $d \gg D$. In order for the total variance to remain finite, the amplitudes $t_D$ must be scaled as $t_D = t_D^*/\sqrt{N_D}$, which is consistent with (2.10) and leads to

$$\langle \epsilon^2(\boldsymbol{k})\rangle = \sum_{D=1}^{\infty} \langle \epsilon_D^2(\boldsymbol{k})\rangle = \sum_{D=1}^{\infty} {t_D^*}^2\,. \tag{2.45}$$

### Hopping Matrix Elements and Dispersion; Transformation Function $\mathcal{F}(\tilde{\epsilon})$

Now we have

$$\begin{aligned} \epsilon_D(\boldsymbol{k}) &= t_D 2^D \sum_{\alpha_D > \alpha_{D-1} > \cdots > \alpha_1} \cos(k_{\alpha_D})\cos(k_{\alpha_{D-1}})\ldots\cos(k_{\alpha_1}) && (2.46)\\ &= t_D \frac{2^D}{D!}\left(\frac{d}{2}\right)^{D/2} B_D(\boldsymbol{k}), && (2.47)\end{aligned}$$

where the functions

$$B_D(\boldsymbol{k}) = \left(\frac{2}{d}\right)^{D/2} \sum_{\alpha_D \neq \alpha_{D-1} \neq \cdots \neq \alpha_1} \cos(k_{\alpha_D})\cos(k_{\alpha_{D-1}})\ldots\cos(k_{\alpha_1}) \tag{2.48}$$

[22] Note that while Eqn. (1.10) always yields real $t_{ij}$ for atomic orbitals, this is not generally true for Wannier orbitals where phase factors may arise. In a single-band description, however, a restriction to the $s$-band case of real and isotropic hopping amplitudes is natural since the full Hamiltonian should have the symmetry of the lattice.

fulfill a recursion relation in $d = \infty$,

$$\begin{aligned} B_{D+1}(\boldsymbol{k}) &= B_1(\boldsymbol{k})\, B_D(\boldsymbol{k}) - D\, B_{D-1}(\boldsymbol{k})\, \frac{2}{d} \sum_{\alpha=1}^{d} \cos^2(k_\alpha) + \mathcal{O}(1/d) \\ &= B_1(\boldsymbol{k})\, B_D(\boldsymbol{k}) - D\, B_{D-1}(\boldsymbol{k}) + \mathcal{O}(1/\sqrt{d})\,. \end{aligned} \tag{2.49}$$

To leading order, this relation is very similar to that of the Hermite polynomials[23]

$$\mathrm{He}_{n+1}(x) = x\, \mathrm{He}_n(x) - n\, \mathrm{He}_{n-1}(x)\,. \tag{2.50}$$

From (2.49), (2.50), and $B_0 = 1$ we conclude that $B_D(\boldsymbol{k}) = \mathrm{He}_D(B_1(\boldsymbol{k}))$. Furthermore, we can express the argument $x$ in (2.50) as the dispersion of the hypercubic lattice with a positive hopping term [i.e., $t^* = -1$ in (1.14)] and unit variance: $B_1(\boldsymbol{k}) = \epsilon_{\boldsymbol{k}}^{\mathrm{hc}}$. Consequently, we obtain

$$\epsilon_D(\boldsymbol{k}) = t_D \frac{2^D}{D!} \left(\frac{d}{2}\right)^{D/2} \mathrm{He}_D(B_1(\boldsymbol{k})) \tag{2.51}$$

and

$$\boxed{\epsilon(\boldsymbol{k}) = \sum_{D=1}^{\infty} \frac{t_D^*}{\sqrt{D!}}\, \mathrm{He}_D(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}) \quad =: \quad \mathcal{F}(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}})\,.} \tag{2.52}$$

Finding not only an effectively one-dimensional expression, but even a closed form (2.52) for the general dispersion of a hypercubic lattice with arbitrary-range hopping in terms of scaled hopping matrix elements is clearly a significant result. On the one hand, this expression applies to all such models where the hopping amplitudes depend on taxi-cab distance and therefore systematically classifies most models studied in the DMFT literature. On the other hand, (2.52) will enable us to explore information hidden in any given DOS, namely the dispersion and the scaled hopping matrix elements of a corresponding model (in $d = \infty$), as we will show in the following. Since any analytic function can be expanded in Hermite polynomials using the orthogonality relation

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{He}_n(x)\, \mathrm{He}_m(x)\, \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}} = n!\, \delta_{mn}, \tag{2.53}$$

we can find hopping amplitudes $t_D^*$ so that $\epsilon(\boldsymbol{k}) = \mathcal{F}(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}})$ for an arbitrary analytic function $\mathcal{F}$. As an illustration of this expansion, we rescale the Hermite polynomials in Fig. 2.14 a to

$$p_D(x) := \mathrm{He}_D(x)/\sqrt{D!}\,. \tag{2.54}$$

The full transformation function is then a linear combination of the basis functions $p_D$ where the Euclidean norm of the vector of expansion coefficients $t_D^*$ sets the energy scale for the resulting DOS.

[23] For the two different definitions of the Hermite polynomials with $H_n(x) = 2^{n/2}\mathrm{He}_n(\sqrt{2}x)$ see, e.g., Abramowitz and Stegun (1972).

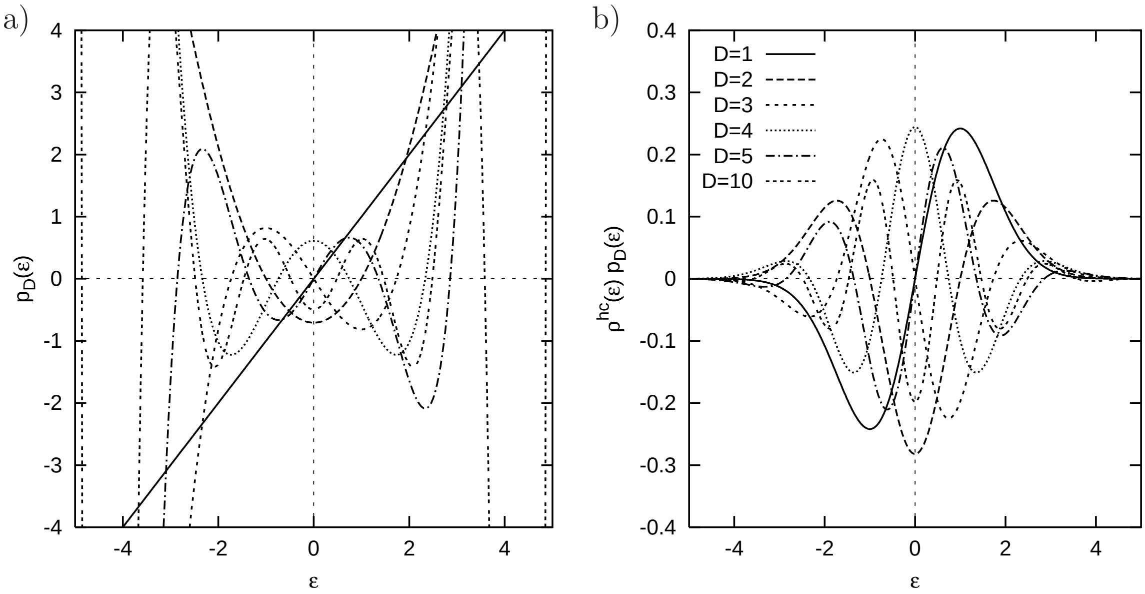



**Figure 2.14: a)** Rescaled Hermite polynomials $p_D(x) = \mathrm{He}_D(x)/\sqrt{D!}$ forming the basis set in expansion (2.52) of the generalized dispersion for small order $D$. **b)** Dual basis, i.e., same curves multiplied by $\rho^{\mathrm{hc}}$ as in (2.55). The legend in b) applies to both graphs.

The evaluation of the hopping matrix elements

$$t_D^* = \frac{1}{\sqrt{2\pi D!}} \int\limits_{-\infty}^{\infty} \mathrm{d}\epsilon\, \mathcal{F}(\epsilon)\, \mathrm{He}_D(\epsilon)\, \mathrm{e}^{-\epsilon^2/2} = \int\limits_{-\infty}^{\infty} \mathrm{d}\epsilon\, \mathcal{F}(\epsilon)\, p_D(\epsilon)\, \rho^{\mathrm{hc}}(\epsilon)\,, \tag{2.55}$$

where $\mathcal{F}(x)$ is yet unknown [to be obtained by numerical inversion of (2.61)], generically requires numerical integration. As shown in Fig. 2.14 b, the kernel $p_D(\epsilon)\rho^{\mathrm{hc}}(\epsilon)$ falls off exponentially with increasing $|\epsilon|$ for fixed hopping range $D$. Due to this fact and due to the increasingly oscillatory behavior, the matrix elements $t^*$ can be expected to fall off rapidly with $D \to \infty$ for any monotonic (or, weaker, a reasonably smooth) transformation function $\mathcal{F}$.

### Transformation Function $\mathcal{F}(\epsilon)$, DOS, and Fermi Velocity

For a given transformation function $\mathcal{F}(x)$, we can obtain the DOS by transforming the corresponding $\delta$-functions in (2.1),

$$\rho(\epsilon) = \sum_{\substack{i=1\\ \mathcal{F}(\tilde{\epsilon}_i)=\epsilon}}^{N(\epsilon)} \frac{1}{\mathcal{F}'(\tilde{\epsilon}_i)} \rho^{\mathrm{hc}}(\tilde{\epsilon}_i)\,, \tag{2.56}$$

($N(\epsilon)$ is the multiplicity of solutions of the equation $\mathcal{F}(\tilde{\epsilon}) = \epsilon$) which simplifies to

$$\rho(\epsilon) = \frac{1}{\mathcal{F}'(\mathcal{F}^{-1}(\epsilon))} \rho^{\mathrm{hc}}(\mathcal{F}^{-1}(\epsilon)) \tag{2.57}$$

for a monotonic (i.e., invertible) function[24] $\mathcal{F}(x)$. In order to solve for $\mathcal{F}(x)$ we form the integral

$$\int_{-\infty}^{\epsilon} \mathrm{d}\epsilon' \rho(\epsilon') = \int_{-\infty}^{\tilde{\epsilon}} \mathrm{d}\tilde{\epsilon}' \rho^{\mathrm{hc}}(\tilde{\epsilon}'); \quad \epsilon = \mathcal{F}(\tilde{\epsilon}) \,, \tag{2.58}$$

which leads to

$$R(\mathcal{F}(\tilde{\epsilon})) = \operatorname{erf}(\tilde{\epsilon}/\sqrt{2}), \quad R(\epsilon) := 2 \int_{-\infty}^{\epsilon} \mathrm{d}\epsilon' \rho(\epsilon') - 1 \,. \tag{2.59}$$

We obtain

$$\mathcal{F}(\tilde{\epsilon}) = R^{-1}(\operatorname{erf}(\tilde{\epsilon}/\sqrt{2})) \quad \Leftrightarrow \quad \mathcal{F}^{-1}(\epsilon) = \sqrt{2}\, \operatorname{erf}^{-1}(R(\epsilon)) \tag{2.60}$$

or, summarizing,

$$\boxed{\mathcal{F}^{-1}(\epsilon) = \sqrt{2}\, \operatorname{erf}^{-1}\left(2 \int_{-\infty}^{\epsilon} \mathrm{d}\epsilon' \rho(\epsilon') - 1\right) .} \tag{2.61}$$

This concludes our presentation of the general algorithm for constructing a tight-binding Hamiltonian, i.e., hopping matrix elements, from a given DOS in $d = \infty$ based on ideas developed by van Dongen (2001).

For later use we also evaluate the Fermi velocity, $\boldsymbol{v}_{\boldsymbol{k}}$. From (2.57) and (2.60) we have

$$\mathcal{F}'(\mathcal{F}^{-1}(\epsilon)) = \frac{\rho^{\mathrm{hc}}(\sqrt{2}\, \operatorname{erf}^{-1}(R(\epsilon)))}{\rho(\epsilon)} \tag{2.62}$$

which we can use for computing

$$\boldsymbol{v}_{\boldsymbol{k}} \;=\; \nabla \mathcal{F}(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}) = \mathcal{F}'(\mathcal{F}^{-1}(\epsilon)) \boldsymbol{v}_{\boldsymbol{k}}^{\mathrm{hc}} \tag{2.63}$$

by numerical integration and application of the inverse error function. As will be discussed in detail in subsection 2.4.1 and chapter 4, the behavior of the squared Fermi velocity $\langle |\boldsymbol{v}_{\boldsymbol{k}}|^2 \rangle(\epsilon) := \tilde{\rho}(\epsilon)/\rho(\epsilon)$ [cf. definitions (2.1) and (2.2)] dramatically changes from that observed in the hc lattice: for any nontrivial transformation it is no longer effectively constant but depends strongly on energy. In particular, the Fermi velocity vanishes at the band edges as physically required in any dimension.

A practical application of the general dispersion formalism proceeds as follows from a given target DOS $\rho(\epsilon)$:

1. compute $\mathcal{F}^{-1}(\epsilon)$ using (2.61)

2. invert numerically or analytically to obtain $\mathcal{F}(\epsilon)$

3. evaluate transport properties, e.g., $\langle |\boldsymbol{v}_{\boldsymbol{k}}|^2 \rangle(\epsilon)$ or $\tilde{\rho}(\epsilon)$ using (2.62) and (2.63)

[24] Here and in the following, $\mathcal{F}^{-1}$ denotes the functional inverse, not the fraction $1/\mathcal{F}$.

4. optionally: determine microscopic model parameters, i.e., the hopping matrix elements $t_D^*$ using (2.55)

The only choice inherent in this procedure beyond the usual assumptions for large dimensions is contained in step 1 which by construction produces a monotonic transformation function $\mathcal{F}$.

**Example: Flat-band System**

One interesting limiting case of a monotonic transformation function which we can treat analytically is the step function $\mathcal{F}(x) = 2\Theta(x) - 1$ corresponding to a flat band DOS, $\rho(\epsilon) = (\delta(\epsilon - 1) + \delta(\epsilon + 1))/2$. Due to the nonanalytic step at $\epsilon = 0$ we expect the decay of the hopping matrix elements $t_D^*$ to be much slower than in generic cases, at least for monotonic $\mathcal{F}$. Using the identity

$$\int_0^x \mathrm{d}y \; \mathrm{e}^{-y^2/2} \mathrm{He}_n(y) = \mathrm{He}_{n-1}(0) - \mathrm{e}^{-x^2/2} \mathrm{He}_{n-1}(x) \tag{2.64}$$

for $x = \infty$ and $\mathrm{He}_{2n}(0) = (-1)^n (2n-1)!!$, $\mathrm{He}_{2n+1}(0) = 0$ we can evaluate (2.55) for the case at hand (trivially, $t_{2n}^* = 0$):

$$\begin{aligned} t_{2n+1}^* &= \frac{2}{\sqrt{2\pi(2n+1)!}} \int_0^\infty \mathrm{d}\epsilon \, \mathrm{He}_{2n+1}(\epsilon) \, \mathrm{e}^{-\epsilon^2/2} & (2.65) \\ &= (-1)^n \sqrt{\frac{2}{\pi}} \frac{(2n-1)!!}{\sqrt{(2n+1)!}} & (2.66) \\ &\stackrel{n\to\infty}{\longrightarrow} (-1)^n (\pi n)^{-3/4} \, . & (2.67) \end{aligned}$$

The asymptotic form, obtained by applying the Stirling formula and retaining only leading powers of $n$, is accurate to within 1% for $n > 30$. As expected, $t_{2n+1}^*$ falls off slowly; the power $-3/4$ is only slightly smaller than the threshold value $-1/2$ required for the convergence of the sum $\sum_D {t_D^*}^2$. Correspondingly, the transformation function converges only slowly to its limiting form. Both the hopping matrix elements and the transformation functions, corresponding to truncated series expansions with $D \leq 2N+1$, are shown in Fig. 2.15. We do not attempt to calculate truncation effects for the DOS for this toy model which is complicated due to the many branches formed by the oscillatory truncated transformation functions. Still, the main features of the DOS for a model with truncated hopping range are clear: the two $\delta$-peaks (of the $D_{\max} = \infty$ model) are broadened with a potentially irregular shape, while a finite density of states remains at $\epsilon = 0$ for finite maximal hopping range. Furthermore, we can read off from (2.67) and (2.71) that the gap develops very slowly for $D_{\max} \to \infty$ with $\rho(0) \propto D_{\max}^{-1/2}$.

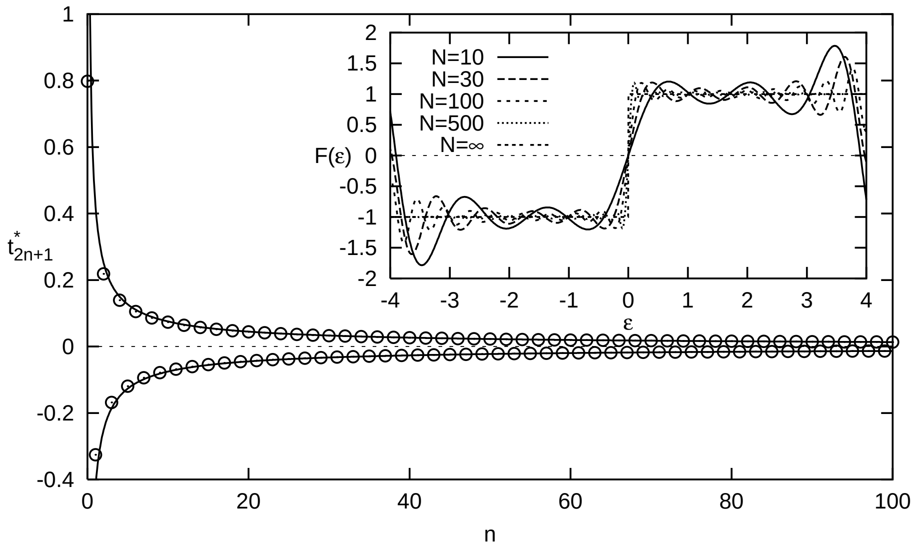



**Figure 2.15:** Hopping matrix elements corresponding to a flat DOS as given by (2.65). Solid lines: asymptotic form (2.67). Inset: transformation function as obtained from the truncated series with $D \leq D_{\max} = 2N + 1$.

## 2.4 Redefinition of the Bethe Lattice

In this section, we will apply the new formalism developed in Sec. 2.3 to the Bethe semi-elliptic DOS as given by (2.24) in order to determine a corresponding tight-binding Hamiltonian defined on the hypercubic lattice with the same local properties as the Bethe lattice (with NN hopping) in the limit $d \to \infty$. For the purpose of DMFT calculations, we can regard the new model as a redefinition of the Bethe lattice, since all properties which could rigorously be defined and calculated on the original Bethe tree are unaffected. In addition, properties which are not well-defined on the Bethe lattice such as transport properties are well-defined for the new model which has the full hypercubic symmetry.[25] Furthermore, these transport properties can be computed analogously to the hypercubic case which is the only case for which rigorous derivations have been previously published.

In the following, we will first derive and briefly characterize the new model Hamiltonian with semi-elliptic DOS and then discuss the behavior of the Fermi velocity in this model, a central result of this section. This will enable us to compute the corresponding optical conductivity in Sec. 4.4. Having demonstrated the usefulness of our approach, we will critically review possible shortcomings by studying the effects of truncation of the hopping range and of finite dimensionality in the context of its application to the Bethe DOS. The application of the method to the asymmetric model DOS (2.27) will illustrate the effects of asymmetry and allow us to make

[25]We note that in a simpler general scheme, transport properties can be defined and computed for arbitrary lattice models in $d = \infty$ when these are periodically stacked in one extra dimension (see Sec. 4.4). This procedure, however, leads to a highly anisotropic model.

contact with a related approach.

### 2.4.1 Model in $d = \infty$

A practical application of the theory outlined in Sec. 2.3 always proceeds in two steps: first, one has to find the energy transformation function $\mathcal{F}$ which relates the "target" DOS to the DOS of the hypercubic lattice [cf. (2.57)] and suffices for the computation of all transport related properties discussed in this work. In a second step one may fully determine the underlying microscopic tight-binding model by computing the corresponding scaled hopping matrix elements.

In order to determine the inverse of the transformation function via (2.60), we need to compute the function $R$ defined in (2.59) in terms of an integral over the DOS.[26] Using (2.24) for unit variance, $R$ can be obtained in closed form,

$$\begin{aligned} R(\epsilon) &= -1 + \frac{2}{\pi} \int_{-2}^{\epsilon} \mathrm{d}\epsilon' \sqrt{1 - \epsilon'^2/4} \\ &= \frac{1}{\pi} \left( \epsilon \sqrt{1 - \epsilon^2/4} + 2 \sin^{-1}(\epsilon/2) \right) . \end{aligned} \tag{2.68}$$

Since we defined the function $\mathcal{F}$ to be smooth and monotonic, it can easily be obtained from $\mathcal{F}^{-1}(\epsilon) = \sqrt{2}\, \mathrm{erf}^{-1}(R(\epsilon))$ using numerical root-finding algorithms. The transformation function shown as solid line in Fig. 2.16 a is monotonic, antisymmetric, and smoothly approaches an asymptotic value of $+2$ ($-2$) for very high (low) energies $\tilde{\epsilon}$. Keeping in mind that $\mathcal{F}$ was defined as a mapping function which when applied to the hc DOS (i.e., energy levels with a Gaussian distribution function) yields the target DOS, we see that the boundedness of $\mathcal{F}$ guarantees sharp band edges, while the linearity at small energies with a slope greater than one flattens out the central part of the DOS compared to the hc case [cf. (2.71)]. Slight variations of the transformation functions only lead to a slight redistribution of spectral weight. This can be illustrated by replacing the (analytically inaccessible) transformation function $\mathcal{F}$ by a hyperbolic tangent,

$$\mathcal{F}(\tilde{\epsilon}) \approx a \tanh\left( \sqrt{\frac{\pi}{2a}}\, \tilde{\epsilon} \right); \quad a \approx 2.176, \tag{2.69}$$

where the coefficients were chosen subject to the requirements that the corresponding DOS has unit variance and that it agrees with the Bethe DOS for $|\epsilon| \to 0$. Due to the boundedness we again obtain sharp band edges, now with finite slopes; otherwise the "tanh" DOS is hardly distinguishable from the semi-elliptic Bethe DOS, as seen in Fig. 2.16 b.

In order to determine the microscopic model, we have to apply (2.55) to the numerically evaluated transformation function $\mathcal{F}$. As shown in Table 2.1, the scaled hopping matrix elements fall off exponentially fast – only a fraction $10^{-3}$ of the total energy variance arises from hopping amplitudes beyond 3rd-nearest neighbors and

[26] Note that $R(\mu) = 2n(\mu) - 1$ when $n$ is the electron filling for a chemical potential $\mu$ at $T = 0$.

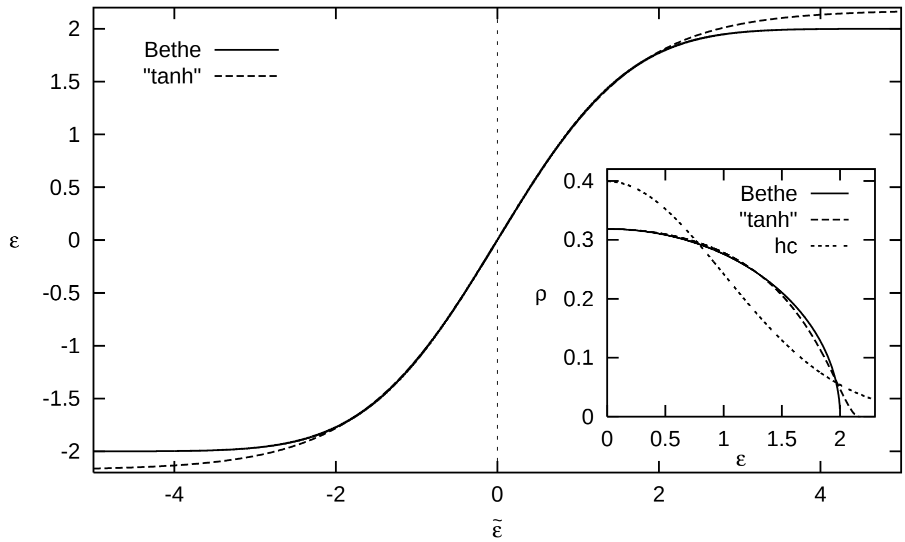



**Figure 2.16:** Transformation function $\epsilon = \mathcal{F}(\tilde{\epsilon})$ for the semi-elliptic Bethe DOS: numerically exact result and approximation by a hyperbolic tangent according to (2.69). The inset shows corresponding DOSs [cf. (2.57)] in comparison to the Gaussian DOS of the hc lattice.

| D | $t^*_D$ | $\sum_{n=1}^{D} t^{*\,2}_n$ | D | $t^*_D$ | $\sum_{n=1}^{D} t^{*\,2}_n$ |
|---|---|---|---|---|---|
| 1 | 0.98731 | 0.974773 | 9 | 0.00343 | 0.999999 |
| 3 | -0.15353 | 0.998345 | 11 | -0.00108 | 1.000000 |
| 5 | 0.03893 | 0.999861 | 13 | 0.00033 | 1.000000 |
| 7 | -0.01125 | 0.999987 | 15 | -0.00011 | 1.000000 |

**Table 2.1:** *Leading scaled hopping matrix elements corresponding to a semi-elliptic Bethe DOS. The columns consisting of partial sums indicate rapid convergence of the series: only a fraction* $1 - \sum_{n=1}^{D} t^{*\,2}_n$ *of the variance arises from hopping terms beyond* $D^{th}$ *neighbors. The mean squared hopping distance is* $\sum_{D=1}^{\infty} D\, t^{*\,2}_D \approx 1.05406$*.*

only a fraction $10^{-6}$ results from hopping beyond $9^{\text{th}}$-nearest neighbors. Since such a short average range is certainly within physical limits for any itinerant system, our model seems a no less natural description for systems involving narrow tight-binding orbitals than the Hubbard model (with only NN hopping) itself. The situation would be different if we had found slow algebraic decay of the hopping amplitudes like we have seen for the toy flat band example, (2.67), which made the latter appear rather unphysical.

Using (2.62) and (2.63), we can compute the averaged squared Fermi velocity $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$, where for each value of $\epsilon$ the average is taken over all momenta $\boldsymbol{k}$ with

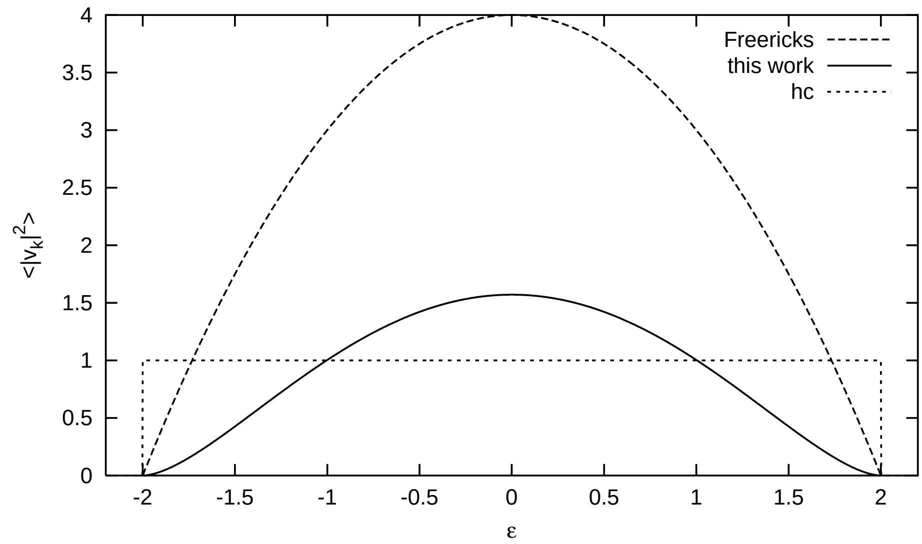



**Figure 2.17:** Averaged squared Fermi velocity of the isotropic lattice model with Bethe DOS (solid line). As required physically, it vanishes at the band edges. This has to be contrasted with the constant value used in calculations so far which is only exact for the hc lattice. Also shown is the parabolic form proposed by Chung and Freericks (1998*a*).

---

$\epsilon_{\boldsymbol{k}} = \epsilon$ in closed form for unit lattice spacing $a = 1$:[27]

$$\langle|\boldsymbol{v}_{\boldsymbol{k}}|^2\rangle(\epsilon) = \frac{\pi}{2(1-\epsilon^2/4)} \exp\left[-2\left(\mathrm{erf}^{-1}\Big(\frac{\epsilon\sqrt{1-\epsilon^2/4} + 2\sin^{-1}(\epsilon/2)}{\pi}\Big)\right)^2\right] \tag{2.70}$$

Here, we have used the fact that $\langle|\boldsymbol{v}_{\boldsymbol{k}}|^2\rangle(\epsilon)$ is effectively constant (and equals 1 for unit variance and lattice spacing) in the hypercubic case. The result, shown in Fig. 2.17, has all the qualitative features expected for this observable in any dimension: $\langle|\boldsymbol{v}_{\boldsymbol{k}}|^2\rangle(\epsilon)$ is maximal near the band center, strongly reduced for large (absolute) energies and vanishes at the band edges: states at the (noninteracting) band edge do not contribute to transport. For the hc lattice, the band edges are at $\pm\infty$, so that a constant $\langle|\boldsymbol{v}_{\boldsymbol{k}}|^2\rangle(\epsilon)$ is somewhat pathological, but at least consistent. This consistency is lost when a formalism derived for the hc lattice is applied to the Bethe DOS. While a more complete discussion of this point has to be postponed to chapter 4, we note that this problem (which had otherwise been ignored in the literature) led Freericks to propose the parabolic form[28] also shown in Fig. 2.17. The resulting computation of the optical conductivity was the first one to go beyond the hc lattice

[27]More formally, we may use equations (2.1) and (2.2) to define $\langle|\boldsymbol{v}_{\boldsymbol{k}}|^2\rangle(\epsilon) := \tilde{\rho}(\epsilon)/\rho(\epsilon)$.

[28]One may ask if one can find any model for which $|\boldsymbol{v}_{\boldsymbol{k}}|^2$ shows this behavior. Using our formalism, namely (2.63), we obtain a differential equation for the energy transformation function, $|\mathcal{F}'(\tilde{\epsilon})| = \sqrt{4-\mathcal{F}^2(\tilde{\epsilon})}$, which is solved by $\mathcal{F}(\tilde{\epsilon}) = 2\sin(\tilde{\epsilon} + \tilde{\epsilon}_0)$. This is by no means a monotonic function, however, which violates the assumption under which it was derived. Therefore, we conclude that the behavior assumed by Freericks cannot occur for the class of tight-binding Hamiltonians considered within our method.

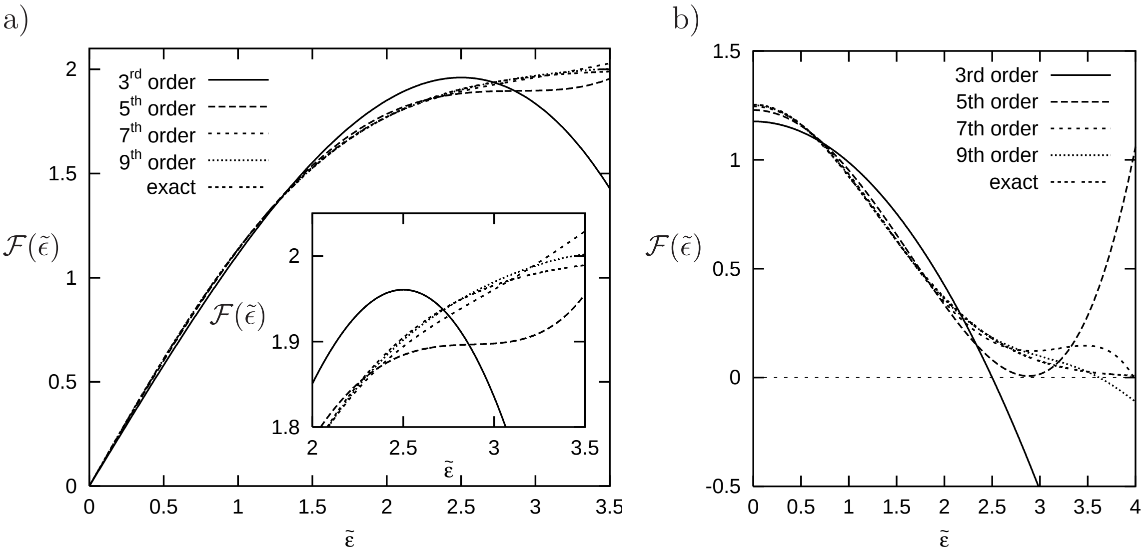



**Figure 2.18: a)** Transformation function $\epsilon = \mathcal{F}(\tilde{\epsilon})$ for the semi-elliptic Bethe DOS: approximations obtained by truncation of the hopping range in (2.52) to $3^{\text{rd}}$, $5^{\text{th}}$, $7^{\text{th}}$, and $9^{\text{th}}$ neighbors (rescaled to unit variance) in comparison with the exact function. **b)** Derivative $\mathcal{F}'(\tilde{\epsilon})$; its zeros lead to singularities in the corresponding DOSs.

(Chung and Freericks, 1998*a*). Evidently, our result deviates even more strongly from a constant (having zero slope at the edges) than Freericks' form. The most important difference, however, is that our approach identifies (for the first time) a model with semi-elliptic DOS for which all transport properties are well-defined. Therefore, we can use DMFT calculations performed for a semi-elliptic DOS for the computation of an isotropic optical conductivity in chapter 4 without ignoring the energy dependence of $\langle |\boldsymbol{v_k}|^2\rangle(\epsilon)$ or resorting to heuristic approaches.

### 2.4.2 Truncating the Hopping Range

By construction, the DOS of our tight-binding model is exactly semi-elliptic when all hopping matrix elements are taken into account, i.e., when the full series (2.52) contributes to the dispersion. Since the limit of an infinite hopping range implies infinite dimensionality (cf. footnote 5 on page 30) we cannot directly study the effect of finite dimensionality on the full model. Instead, we will first discuss the effects of a truncated hopping range in $d = \infty$ before we present the DOS of a truncated model, evaluated in finite dimensions.

Figure 2.18 a shows approximations to the transformation function $\epsilon = \mathcal{F}(\tilde{\epsilon})$ in comparison with the exact form. The finite-order approximations represent truncations of (2.52) with $D \leq D_{\max}$ in which the hopping matrix elements, obtained from application of (2.55) to the numerically exact transformation function, have been rescaled to $\sum_{D=1}^{D_{\max}} |t_D^*|^2 = 1$. Since the effect of the transformation function $\mathcal{F}$ on the level of the DOS is to just shift the energy levels away from their position in the NN-hopping hypercubic case and since there is only exponentially small weight in

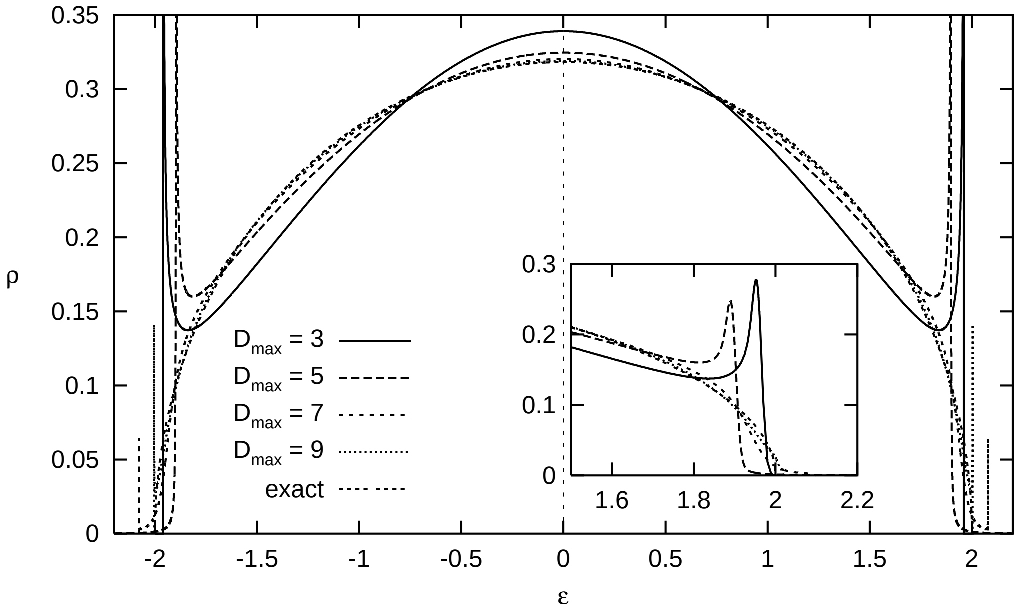



**Figure 2.19:** DOS for the tight-binding model truncated at different maximum hopping range $D_{\max}$ compared with the semi-elliptic DOS of the full model. Inset: same curves, broadened by a Gaussian of width 0.01.

---

the hc DOS at high energies (about 4.6% for $|\epsilon| > 2$ and 0.3% for $|\epsilon| > 3$, respectively), the high-energy part of the transformation function becomes quantitatively irrelevant. In Fig. 2.18 a, the truncation error is seen to be small already in the relevant range of low and intermediate energies for $5^{\text{th}}$ order and almost invisible for $9^{\text{th}}$ order. Therefore, we may also expect good overall agreement on the level of the DOS for the truncated model. Note, however, that truncated transformation functions need not, in general, be monotonic; they have extrema for $D_{\max} = 3, 7, 9$ while the transformation function for $D_{\max} = 5$ is very close to a saddle point (see derivatives $\mathcal{F}'(\tilde{\epsilon})$ shown in Fig. 2.18 b). Thus, on the level of the DOS, the truncation error will be large pointwise near the resulting singularities even though the associated weight is small.

This is exactly what is seen in in Fig. 2.19: the DOS of the truncated model becomes singular at the band edges with rapidly decreasing weight for increasing $D_{\max}$. The good overall agreement with a semi-elliptic form can be further improved for $D_{\max} = 7$ and $D_{\max} = 9$ by broadening the curves with a Gaussian of width 0.01 in order to mimic the effect of a finite imaginary part of the self-energy caused by interactions (see inset). Thus, practical DMFT calculations should not be able to distinguish between a Bethe DOS and our model, truncated at $D_{\max} \geq 9$. As we will see below, the effects of truncation are even less severe when the model is applied in finite dimensions. Note that in the symmetric case and for a monotonic transformation function (i.e., when $\mathcal{F}(0) = 0$), the DOS at the symmetry point is fully determined by the first derivative of the transformation function at the origin,

$$\frac{1}{\rho(\epsilon=0)} = \sqrt{2\pi}\,\mathcal{F}'(0) = \sqrt{2\pi}\sum_{n=0}^{\infty} t^*_{2n+1}(-1)^n \frac{(2n+1)!!}{\sqrt{(2n+1)!}}\,. \tag{2.71}$$

This series is quickly converging for the tight-binding model corresponding to a Bethe DOS as seen in Fig. 2.18 and Fig. 2.19.[29]

### 2.4.3 Finite Dimensionality

After having verified that the properties of the tight-binding model with semi-elliptic DOS derived in this section are robust with respect to truncation, we can now study the robustness with respect to application in finite dimensions and thereby distinguish generic from more artificial features. For this purpose, we explicitly calculate the dispersion for the redefined Bethe model [i.e., the model defined by (2.68)], truncated at 3$^{\text{rd}}$ or 5$^{\text{th}}$ order hopping. Here, we evaluate $B_D(\boldsymbol{k})$ as defined in (2.48) without resorting to any large-$d$ approximations for the involved sums or the number of equivalent hopping terms; in particular, the approximation in (2.44) is not used.[30] In the necessary summations, the exclusion of terms with equal indices poses a combinatorial problem which can be solved iteratively or as a linear system of equations. In order to illustrate the basic considerations, we show how the summation restrictions are replaced for $D = 2,3$: $\sum_{i\neq j} = \sum_{ij}(1-\delta_{ij})$; $\sum_{i\neq j\neq k} = \sum_{ijk}(1-\delta_{ij}-\delta_{ik}-\delta_{jk}+2\delta_{ijk})$, where the generalized $\delta$-function equals 1 if all indices are equal and is 0 otherwise. For the sake of brevity, we will in the following state the results for $A_D(\boldsymbol{k}) := (2/d)^{-d/2}B_D(\boldsymbol{k})$, from which the exact $d$-dimensional dispersion can be obtained in the form

$$\epsilon_D(\boldsymbol{k}) = t^*_D \frac{2^{D/2}}{D!\sqrt{\binom{d}{D}}} A_D(\boldsymbol{k}) \tag{2.72}$$

and which agrees with (2.47) in the limit $d \gg D$. The coefficient functions read

$$A_1(\boldsymbol{k}) = \sum_{\alpha=1}^{d}\cos(k_\alpha); \qquad A_2(\boldsymbol{k}) = \Big(\sum_\alpha \cos(k_\alpha)\Big)^2 - \sum_\alpha \cos^2(k_\alpha)$$

[29]Note that a gapped DOS as in the case of the flat band DOS considered above can only arise when the sum in (2.71) diverges which means that the hopping matrix elements $t^*_n$ decay more slowly than $n^{-5/4}$.

[30]We keep, however, the selection of hopping terms defined at the beginning of this section, i.e., do not consider extended hopping along the cartesic axes.

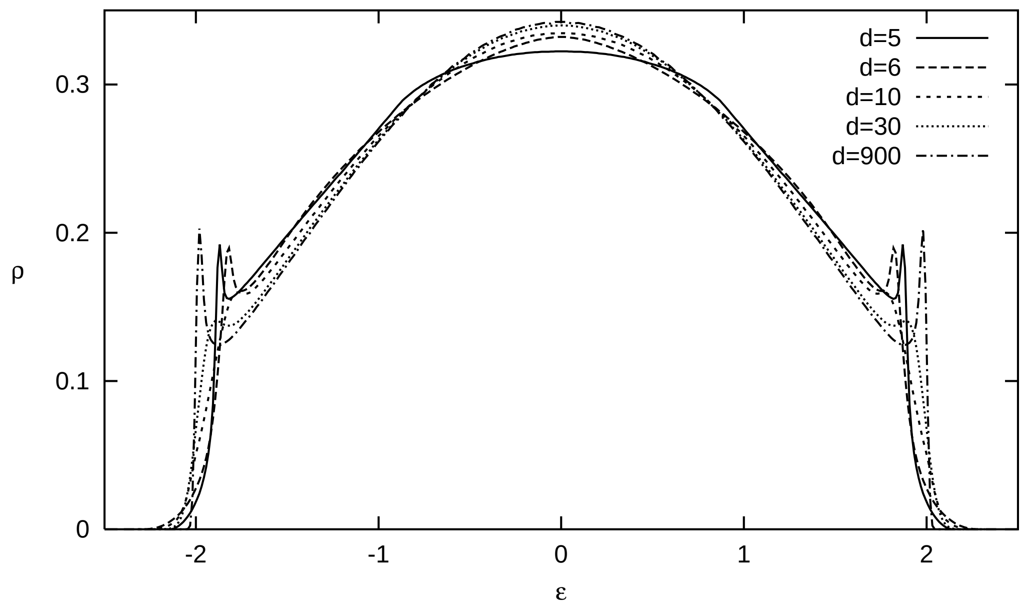



**Figure 2.20:** DOS of the isotropic tight-binding model with semi-elliptic DOS, truncated at 3$^{\text{rd}}$ order and applied in finite dimensions.

$$
\begin{aligned}
A_3(\boldsymbol{k}) &= \Big(\sum_\alpha \cos(k_\alpha)\Big)^3 - 3\sum_\alpha \cos^2(k_\alpha)\sum_\alpha \cos(k_\alpha) + \sum_\alpha \cos^3(k_\alpha)\\
A_4(\boldsymbol{k}) &= \Big(\sum_\alpha \cos(k_\alpha)\Big)^4 - 6\Big(\sum_\alpha \cos(k_\alpha)\Big)^2\sum_\alpha \cos^2(k_\alpha) + 3\Big(\sum_\alpha \cos^2(k_\alpha)\Big)^2\\
&\quad +8\sum_\alpha \cos^3(k_\alpha)\sum_\alpha \cos(k_\alpha) - 6\sum_\alpha \cos^4(k_\alpha)\\
A_5(\boldsymbol{k}) &= \Big(\sum_\alpha \cos(k_\alpha)\Big)^5 - 20\sum_\alpha \cos^2(k_\alpha)\sum_\alpha \cos^3(k_\alpha) - 30\sum_\alpha \cos(k_\alpha)\sum_\alpha \cos^4(k_\alpha)\\
&\quad +20\Big(\sum_\alpha \cos(k_\alpha)\Big)^2\sum_\alpha \cos^3(k_\alpha) + 15\sum_\alpha \cos(k_\alpha)\Big(\sum_\alpha \cos^2(k_\alpha)\Big)^2\\
&\quad -10\Big(\sum_\alpha \cos(k_\alpha)\Big)^3\sum_\alpha \cos^2(k_\alpha) + 24\sum_\alpha \cos^5(k_\alpha)\,.
\end{aligned}
$$

The resulting DOSs for various dimensions and a hopping range truncated at $D_{\max} = 3$ have been evaluated by Monte Carlo sampling of the momentum integrals and are shown in Fig. 2.20. As expected, the singularities in the DOS observed for $d = \infty$ are smoothed out in finite dimensions, so that the agreement with a Bethe DOS is even improved. The value at $\epsilon = 0$ is slightly too high (which is accidentally offset by a finite-$d$ error for $d = 5$). For even lower dimensions $d = 3$ and $d = 4$ (not shown, cf. subsection 4.4.5), the shape of densities of states is dominated by Van-Hove singularities rather than the influence of 3$^{\text{rd}}$-nearest–neighbor hopping.

Both the overall deviations from the Bethe DOS and the dimensional dependence

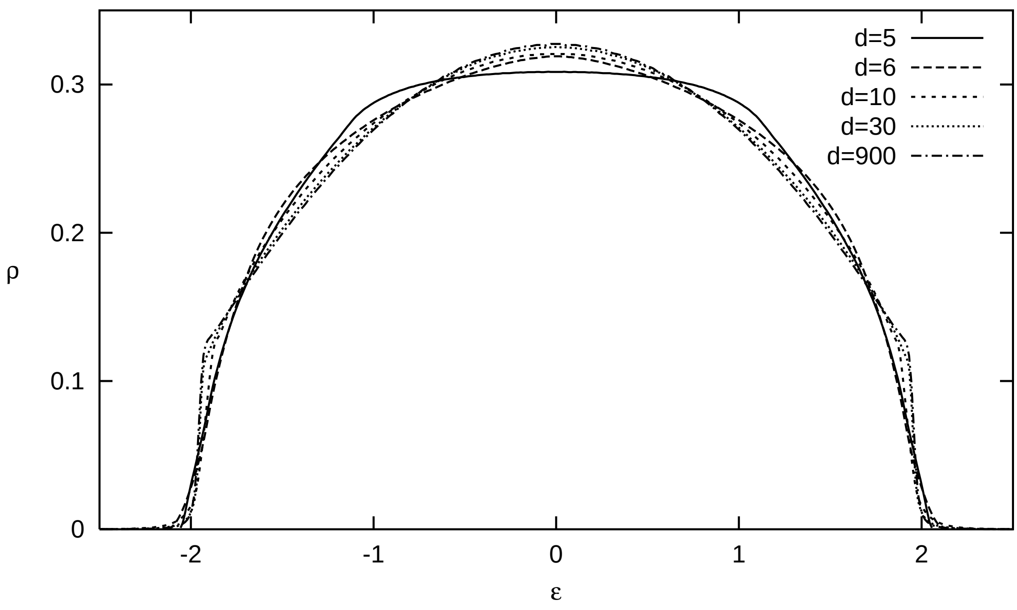



**Figure 2.21:** DOS of the isotropic tight-binding model with semi-elliptic DOS, truncated at $5^{\text{th}}$ order and applied in finite dimensions. Accidentally, the curve for $d = 10$ best approximates the semi-elliptic Bethe DOS.

---

are much smaller when the hopping range is extended to $5^{\text{th}}$-nearest neighbors as shown in Fig. 2.21. It is remarkable that a reasonable approximation to the Bethe DOS is obtained even for $d = 5$, although the underlying tight-binding model was derived for $d \gg D_{\text{max}}$. While the dimensional dependence is expected to further decrease significantly when the hopping range is extended to $D_{\text{max}} \geq 9$, we did not find it worth the effort to extend these finite-dimensional studies beyond $D_{\text{max}} = 5$ after having successfully demonstrated the overall robustness of the method. The weak dimensional dependence seen in Fig. 2.21 should be contrasted with the very strong dependence of the DOS on dimensionality observed for other lattice types, e.g., for the fcc lattice shown in Fig. 2.7.

One useful feature of our method is that (for a monotonic transformation function) the energy minimum remains associated with $\boldsymbol{k} \approx \boldsymbol{0}$ and the energy maximum associated with momenta $\boldsymbol{k} \approx \boldsymbol{Q}$. This can be illustrated in finite dimensions, where these vectors (which are nongeneric[31] in the limit $Z \to \infty$) still have finite weight. In Fig. 2.22 a and Fig. 2.22 b we show the full DOS of the hc lattice and of our model truncated at $5^{\text{th}}$ order for dimension $d = 10$, respectively (dotted lines). The solid and dashed lines in the same figures indicate partial DOSs which were obtained by taking the sum in (2.1) either only for momenta with $|\boldsymbol{k}| < 2d$, $|\boldsymbol{k} - \boldsymbol{Q}| < 2d$, or with $|\boldsymbol{k}|/\langle|\boldsymbol{k}|\rangle < 0.15$. Here, the average length of a momentum vector is approximated by its asymptotic for $Z \to \infty$, $\langle|\boldsymbol{k}|\rangle \approx \pi\sqrt{d/3}$ (cf. App. A.2). Evidently, generic states of average length are least affected by the transformation. Therefore, our model may

[31] For a discussion of generic versus nongeneric momenta, see App. A.2.

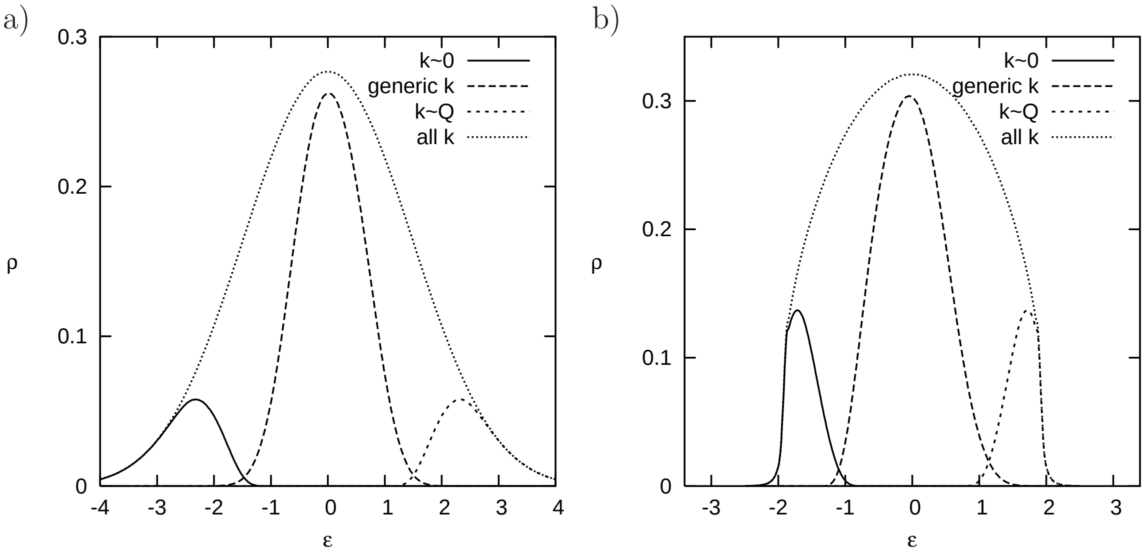



**Figure 2.22:** Partial DOSs **a)** of the hc lattice and **b)** of the $5^{\text{th}}$ order truncation of the redefined Bethe lattice in $d = 10$ (solid and dashed lines) in comparison with the full DOSs (dotted lines). The restriction to momenta with $|\boldsymbol{k}| \approx 0$, $|\boldsymbol{k}| \approx \langle|\boldsymbol{k}|\rangle$, and $|\boldsymbol{k} - \boldsymbol{Q}| \approx 0$ is explained in the text.

be regarded as a minimal modification of the NN-hopping hc lattice for which the DOS becomes semi-elliptic.

### 2.4.4 Application to Asymmetric Model DOS

It should be evident that our method is by no means restricted to the Bethe DOS or to a symmetric DOS in general. In fact, it gives a natural description for all nonsingular single-band DOSs. In order to illustrate the application to an asymmetric case, to draw a connection to earlier work (Wahle et al., 1998), and to compare with estimates for hopping amplitudes obtained within an alternative scheme (Kollar, 1998), we also applied our formalism to the asymmetric model DOS (2.27) which contains the asymmetry parameter $-1 \leq a \leq 1$ and coincides with the semi-elliptic Bethe DOS for $a = 0$.

As seen in Table 2.2, slight asymmetry ($a = 0.5$) leaves the scaled hopping matrix elements corresponding to odd hopping range $D = 2n + 1$ almost invariant while the even range hopping becomes relevant. The energy shift (i.e., finite first moment) implicit in (2.27) for $|a| > 0$ gives rise to a finite on-site amplitude $t_0^2$ which can be evaluated using (2.55) for $D = 0$ just like for finite hopping range $D > 0$. With increasing asymmetry ($a = 0.9$ and $a = 0.98$, cf. Fig. 2.13), the even range hopping elements become more dominant and also the decay is slower than for the symmetric Bethe DOS. Even an asymmetry ($a = 0.98$) strong enough to allow for ferromagnetism in the pure Hubbard model (Wahle et al., 1998) thus naturally fits into the present scheme. However, an application of the method (with a monotonic transformation function) would be questionable for the singular case $a = 1$, which

| | $a = 0.5$ | | $a = 0.9$ | | $a = 0.98$ | |
|---|---|---|---|---|---|---|
| D | $t^*_D$ | $\sum_{n=1}^{D} t^{*\,2}_n$ | $t^*_D$ | $\sum_{n=1}^{D} t^{*\,2}_n$ | $t^*_D$ | $\sum_{n=1}^{D} t^{*\,2}_n$ |
| 0 | -0.26795 | 0.000000 | -0.62679 | 0.000000 | -0.81735 | 0.000000 |
| 1 | 0.98364 | 0.967540 | 0.96671 | 0.934524 | 0.95139 | 0.905141 |
| 2 | 0.08255 | 0.974354 | 0.19413 | 0.972212 | 0.25457 | 0.969945 |
| 3 | -0.15137 | 0.997267 | -0.14143 | 0.992213 | -0.13252 | 0.987505 |
| 4 | -0.03143 | 0.998255 | -0.07444 | 0.997755 | -0.09835 | 0.997177 |
| 5 | 0.03775 | 0.999680 | 0.03236 | 0.998802 | 0.02759 | 0.997938 |
| 6 | 0.01273 | 0.999842 | 0.03046 | 0.999730 | 0.04063 | 0.999589 |
| 7 | -0.01062 | 0.999955 | -0.00776 | 0.999790 | -0.00528 | 0.999617 |
| 8 | -0.00534 | 0.999983 | -0.01294 | 0.999957 | -0.01748 | 0.999923 |
| 9 | 0.00310 | 0.999993 | 0.00160 | 0.999960 | 0.00032 | 0.999923 |
| 10 | 0.00229 | 0.999998 | 0.00566 | 0.999992 | 0.00777 | 0.999983 |
| 11 | -0.00091 | 0.999999 | -0.00012 | 0.999992 | 0.00053 | 0.999983 |
| 12 | -0.00101 | 1.000000 | -0.00253 | 0.999998 | -0.00355 | 0.999996 |
| 13 | 0.00024 | 1.000000 | -0.00018 | 0.999998 | -0.00050 | 0.999996 |
| 14 | 0.00045 | 1.000000 | 0.00116 | 1.000000 | 0.00167 | 0.999999 |
| 15 | -0.00006 | 1.000000 | 0.00016 | 1.000000 | 0.00032 | 0.999999 |

**Table 2.2:** *Leading hopping matrix elements corresponding to the generalized asymmetric model DOS (2.27) with parameter $a$ (cf. Table 2.1). The columns which consist of partial sums indicate rapid convergence of the corresponding dispersion and DOS: only a fraction $1 - \sum_{n=1}^{D} t^{*\,2}_n$ of the variance arises from hopping terms beyond $D^{th}$ neighbors. The "hopping" term with range $D = 0$ denotes an on-site energy, i.e., the first moment of the energy distribution.*

corresponds to a Bethe lattice with only NNN hopping, as we have shown in Sec. 2.2. We would have to overcome numerical stability issues just to reproduce a square-root singularity at the lower band edge which we regard as an artifact of $d = \infty$ (in complete analogy to the situation for the fcc lattice).

Earlier attempts to link the asymmetric model DOS (2.27) to some microscopic tight-binding model (Wahle et al., 1998; Kollar, 1998) have led to a related, but quite different construction scheme, which can be applied in low and infinite dimensions (Kollar, 2002). The straightforward application in $d = 1$ to the DOS (2.27) with $a = 0.98$ yields scaled hopping matrix elements which are well approximated by $t^*_{2n+1} = 0.67\,(2n)^{-1.6}$, $t^*_{2n} = -0.54\,(2n)^{-1.6}$ for small hopping range and decay as $t_D \sim D^{-5/3}$ for large range (Kollar, 1998). When generalized to higher dimensions, Kollar's construction is based on a dispersion $\epsilon(\boldsymbol{k}) = g(|k|)$ (for some function $g$) which is spherical except at the boundaries of the Brillouin zone. We regard it more as a proof of existence of a dispersion for a given DOS in arbitrary dimension than as a workable scheme for the following reasons: Most importantly, the above spherical form for the dispersion implies discontinuities in the Fermi velocity at the boundaries of the Brillouin zone for $d > 1$ which, if taken seriously, makes the computation of transport properties pointless. Due to the violation of symmetry between $\boldsymbol{k} = \mathbf{0}$ and $\boldsymbol{k} = \boldsymbol{Q}$, concepts like perfect nesting become meaningless. Additionally, the hopping is dominantly along the axes and is very slowly decaying (as $t_D \sim D^{-2}$, only slightly faster than in the one-dimensional case). Thus, our approach appears much more physical in higher dimensions which is not surprising since it was derived for the large-$d$ limit.[32]

## 2.5 Conclusion

In this chapter, we have carefully studied the relationships between lattice types, densities of states, and magnetic frustration in high dimensions. Evaluating Brillouin-zone integrals by Monte Carlo sampling, we have computed the (noninteracting) DOS of the hypercubic lattice with hopping to nearest and/or next-nearest neighbors as well as the DOS of the hyperdiamond lattice (see App. B) in low, intermediate, and high dimensions. Thereby and by comparing with the corresponding analytical $d = \infty$ results we have distinguished generic features from features that are artifacts of the $d = \infty$ limit. In particular, we have established that for regular lattices, magnetic frustration is generically associated with an asymmetric DOS or with off-diagonal disorder.

We have then characterized the Bethe "lattice" which is a tree in the sense of graph theory rather than a regular lattice. We have pointed out that a semi-elliptic DOS implies the absence or cancellation of all self-avoiding loops in a renormalized perturbation expansion for any lattice. By computing moments of the DOS of the Bethe lattice with hopping to nearest and next-nearest neighbors (NNN), we have contradicted an earlier claim that NNN hopping should leave the DOS invariant. Furthermore, we have presented the first (analytic) calculation of the (strongly asymmetric) DOS of the Bethe lattice with only NNN hopping.

We have developed a general method of constructing a tight-binding Hamiltonian which yields a given, arbitrary DOS in the limit of infinite dimensions. This method allows for the computation of all physical quantities that can be defined for the corresponding finite-dimensional model. Such a method is particularly useful in the context of the DMFT, where nonlocal properties can be derived by combining an essentially local property of the interacting system with a nonlocal property of the noninteracting system. Examples for such properties are the optical conductivity (which we will consider in chapter 4) and susceptibilities.

By applying the new general-dispersion formalism to the semi-elliptic "Bethe" DOS, we have found the first regular lattice with a bounded DOS in high dimensions, thereby removing one of the major artifacts previously regarded as inherent to the limit $d \to \infty$. We have verified that versions of this model with truncated hopping ranges still give reasonable descriptions. While the DOS is then no longer strictly bounded, the weight associated with long-frequency tails is suppressed exponentially with the number of hopping terms and becomes negligible for numerical purposes

[32] In contrast, only Kollar's construction can be applied in very low dimensions, in particular in one dimension where it becomes exact.

when hopping to $7^{\text{th}}$ or $9^{\text{th}}$ nearest neighbors is included. For the full model, the average squared hopping distance is 1.054, i.e., surprisingly close to the value 1 valid for the usual hypercubic lattice. The convergence of the method (with the number of nonzero hopping matrix elements and with dimensionality) is similarly good in terms of the distribution of spectral weight for other sufficiently smooth target DOSs. The fact that exactly vanishing even-range hopping matrix elements are predicted for any symmetric DOS follows necessarily for any consistent approach and is not specific to our method.

With the general dispersion method developed in Sec. 2.3, we have singled out one route to high dimensions with a series of very favorable properties (specializing on the case of a monotonic transformation function): (i) The approach does not require large long-range hopping amplitudes. (ii) Generically, the DOSs corresponding to finite-$d$ versions of the computed tight-binding model depend only weakly on dimensionality. (iii) The topology of the Fermi surface remains hypercubic-like at all fillings in arbitrary dimensions with the energy minimum of $\epsilon_{\boldsymbol{k}}$ at $\boldsymbol{k} = \mathbf{0}$ (and the maximum at $\boldsymbol{k} = \boldsymbol{Q}$). (iv) The full hypercubic symmetry is retained. (v) For the first time, transport properties for the Bethe DOS can be calculated in a fully consistent way (including an $f$-sum rule) without resorting to anisotropic models. (vi) The derived equations can be used as a heuristic approach for transport calculations in finite dimensions when only the DOS is known (e.g., from LDA calculations).

# Chapter 3

# Mott Metal-Insulator Transition in the $d \to \infty$ Hubbard Model

One of the most fascinating phenomena of strongly correlated electron systems is the interaction-induced metal-insulator transition. Such a transition is fundamentally different from filling-induced transitions between metals and band insulators since it cannot be understood in terms of (effectively) noninteracting electrons. Particularly challenging is the Mott transition, the interaction-induced transition between a paramagnetic metal (PM) and a paramagnetic insulator (PI).[1] As this transition does not break the translational symmetry of the lattice, its description needs to take strong electronic correlations into account. For this reason, proximity to a Mott transition may be even regarded as a defining property of strongly correlated electron systems. The explanation and modeling of the Mott transition is a long-standing important and difficult problem of condensed matter physics (Mott, 1968; Mott, 1990; Gebhard, 1997; Imada, Fujimori, and Tokura, 1998).

In this chapter, we present studies of the Mott metal-insulator transition (MIT) occurring in the infinite-dimensional half-filled Hubbard model. Specifically, results will be obtained for the Bethe lattice, i.e., for a semi-elliptic noninteracting DOS. In order to observe the transition of interest in the paramagnetic phase, antiferromagnetism is suppressed in the calculation, which corresponds to treating the lattice as fully frustrated. For this model, the Mott transition occurs for an interaction $U$ slightly larger than the bandwidth $W$ (see below). Consequently, perturbative methods are not adequate for its description and may only a posteriori be useful for comparison. We will argue later that even numerically exact methods such as QMC or ED may become problematic in the transition region if not used with great care. In fact, one of the achievements of this thesis is the settlement of a controversy about the qualitative shape of the MIT phase diagram of the specified model. Therefore, our treatment will have to at least partially follow the historical development.

In Sec. 3.1 we give an experimental motivation and discuss the applicability of the model. In Sec. 3.2, we discuss strategies for pinpointing and characterizing phase

[1] Interactions can alternatively drive the transition of a metal into an insulating ordered state, namely an antiferromagnetic Néel low-temperature phase (AFI); the AFI, however, may be understood as an effective band insulator in the halved Brillouin zone (see subsection 1.1.4).

transitions within the DMFT and specify potentially useful observables. A first look at the history of the phase diagram is given in Sec. 3.3 where it will be seen that early QMC studies collaboratively performed in this work correctly captured some aspects of the phase diagram while failing in describing others. In Sec. 3.4, we explain the reasons for this failure and correct a fundamental problem of the QMC code previously used within our group. In addition, we quantify the remaining numerical errors and present further improvements of the QMC algorithm and code developed as part of this thesis.

Accurate phase boundaries are then determined in Sec. 3.5. Here, we show that the total internal energy (per site) is best suited for distinguishing metallic from insulating solutions, in particular in the extrapolation $\Delta\tau \to 0$. The resulting phase diagram is constructed and the controversy resolved in subsection 3.5.4. The observation of a coexistence region of metallic and insulating solutions generates the need to determine the true thermodynamic phase transition line using free energy information. Prior to this work, this task had seemed too complicated to be accomplished on the basis of numerically exact techniques; even in the context of iterative perturbation theory, results for this line had been contradictory. Using newly developed schemes, we demonstrate in Sec. 3.6 how the first order phase transition line can be located based on QMC simulations; using low-temperature thermodynamic information and well-established numerical zero-temperature results, the thermodynamic phase transition line is even extended down to zero temperature. The presentation of the full phase diagram in subsection 3.6.3 is supplemented by a discussion of possible scenarios for partially frustrated models in subsection 3.6.4.

Implications of the Landau theory of phase transitions for the convergence of the DMFT iteration process and the suitability of the fitting functions introduced in Sec. 3.5 are discussed in Sec. 3.7. In addition, we propose a new concept for computing free energy differences by solving impurity models which are (in general) not solutions of the DMFT equations. In the final section 3.8, local spectral functions are computed using the maximum entropy method. These results provide additional insight in the nature of the phase transition and will later be needed for the computation of the optical conductivity in chapter 4.

## 3.1 Motivation

### 3.1.1 Experiment

Experimentally, metal-insulator transitions may be observed upon variation of temperature and pressure for a given sample or as a function of doping within a class of materials. Since the induced variations of energy scales are relatively small, only few classes of materials, such as the oxides of vanadium, titanium, and chrome as well as $NiSe_{2-x}S_x$ display a correlation-induced MIT (Heine and Mattheiss, 1971; Mattheiss, 1972; Mott, 1990). The compound $La_{1\text{-}x}Sr_xTiO_3$, which will be treated within LDA+DMFT in chapter 5, falls into this group; among the most prominent and best characterized examples, however, is $V_2O_3$ (Rice and McWhan, 1970; McWhan, Remeika, Rice, Brinkman, Maita, and Menth, 1971; McWhan, Remeika,

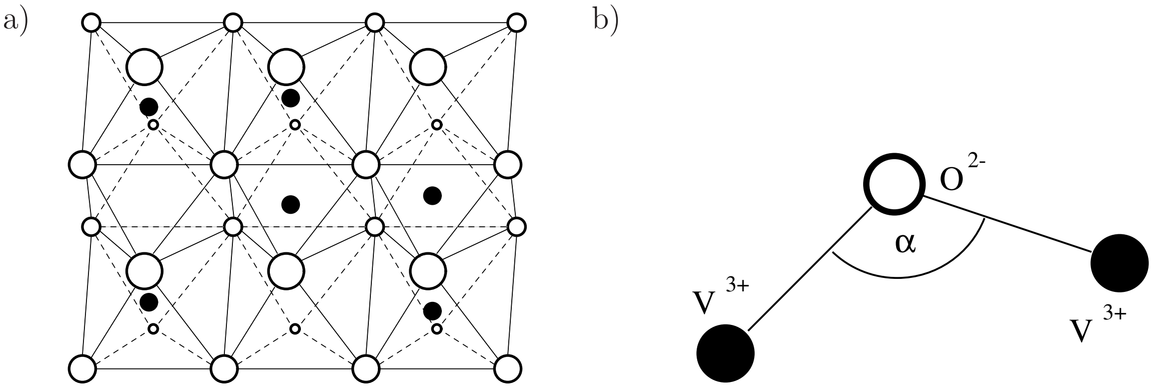



**Figure 3.1: a)** High-temperature corundum lattice structure of $V_2O_3$: oxygen ions (empty spheres) form an hcp lattice, vanadium ions or dopants (solid spheres) occupy 2/3 of the octahedral vacancies. **b)** In leading order, the binding angle $\alpha$ determines the effective hybridization between the vanadium sites.

Maita, Okinaka, Kosuge, and Kachi, 1973; Kuwamoto, Honig, and Appel, 1980; Bao, Broholm, Carter, Rosenbaum, Aeppli, Trevino, Metcalf, Honig, and Spalek, 1993; Bao, Broholm, Aeppli, Dai, Honig, and Metcalf, 1997; Bao, Broholm, Aeppli, Carter, Dai, Rosenbaum, Honig, Metcalf, and Trevino, 1998).

$V_2O_3$ crystallizes in the corundum structure as schematically shown in Fig. 3.1 a. Here, the oxygen ions (formally $O^{2-}$) form an hcp lattice while the vanadium ions, formally $V^{3+}$: $[\mathrm{Ar}](3d)^2$, fill two thirds of the octahedral vacancies. The asymmetry in the occupation of the octahedral sites induces a distortion. Upon doping, Ti or Cr ions replace a fraction of the V ions. While the dopants (formally isovalent to V) only slightly influence the charge distribution, the lattice distortion of the doped system changes significantly due to the different effective atomic radii. Consequently, doping influences the angle of V-O-V paths (cf. Fig. 3.1 b) just like the application of external pressure.[2] The potentially mobile electrons predominantly occupy orbitals on the V sites (associated with the $3d$ shell) while $O^{2-}$ (in noble gas configuration) may be modeled as chemically inert. The hybridization between the V orbitals via the O sites and, thus, the effective bandwidth increases with increasing $\alpha$.

Therefore, the experimental control over doping level and pressure may indeed be characterized as a bandwidth control. Under the assumption that interactions do not depend on substitution and pressure, the systems also qualify as interaction controlled when the bandwidth is taken as unit of energies. One should be aware of the fact that doping necessarily reduces the translational symmetry of the host lattice: either the positions of dopants form a regular superlattice or they introduce (predominantly offdiagonal) disorder. Both the modified interaction and the introduced disorder could in principle provide for quite different mechanisms for the metal-insulator transition.

One of the earliest accounts of the $V_2O_3$ phase diagram is reproduced in Fig. 3.2 a

[2] Note that uniform pressure also increases the overlap by decreasing the distances between sites.

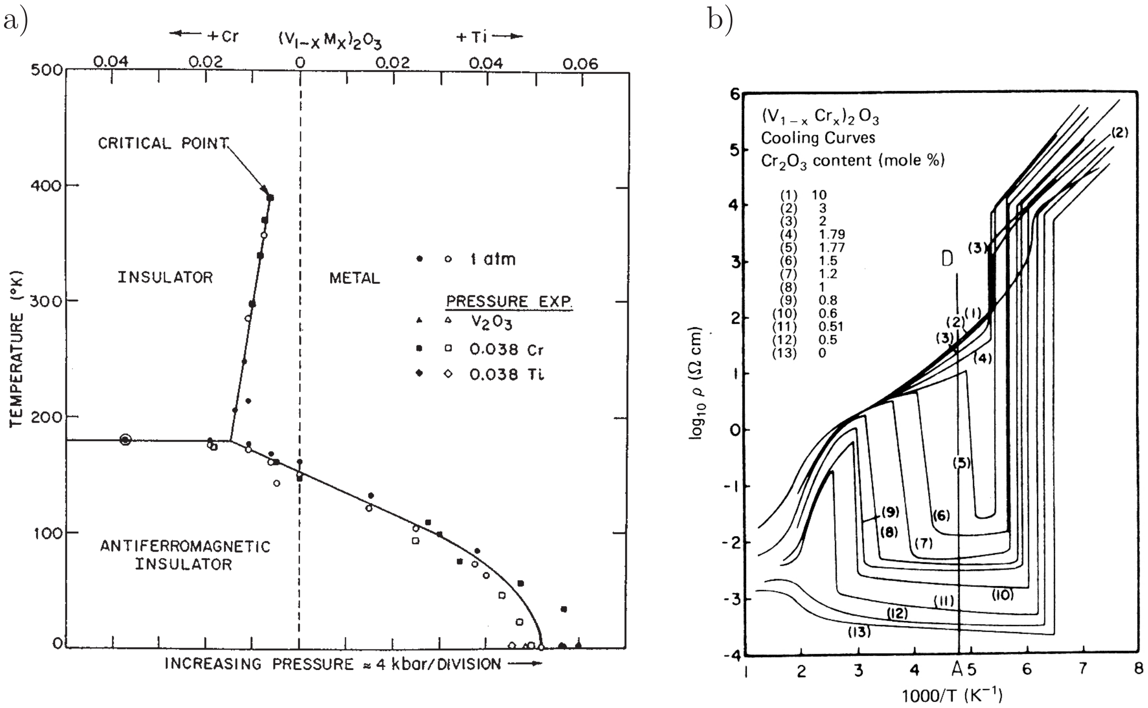



**Figure 3.2:** Experimental realization of Mott metal-insulator transitions in $V_2O_3$ doped with Cr or Ti. **a)** Phase diagram: transition temperatures as function of pressure and/or doping level (McWhan et al., 1971). **b)** Resistivity (log scale) as a function of inverse temperature (Kuwamoto et al., 1980).

where transition temperatures (vertical axis) are plotted as function of doping (upper horizontal axis) or pressure (lower horizontal axis). In this scaled version, both horizontal axes may be regarded as indicating the effective pressure or bandwidth. The low-temperature phase is seen to be antiferromagnetic and insulating for not too high effective pressure. Upon increasing the temperature, the system becomes paramagnetic and goes either into the PI or the PM phase. The latter two phases are separated by a roughly straight first-order phase transition line ending in a critical end point. Note that for ambient pressure the pure system is a metal which becomes insulating upon doping with Cr, quite contrary to the behavior for band insulators. In addition, the scaling law implicit in the figure makes disorder as the driving mechanism for localization of electrons less likely[3] since the influence of Cr doping can be offset by increased pressure which further supports the characterization of the MIT as being interaction driven. Of fundamental interest is the question whether the change in energy at the first-order transition line is of electronic origin. This problem is obscured experimentally by a simultaneous isostructural transition, i.e., a slight change of the lattice parameter at the MIT.

[3]While Anderson localization associated with diagonal disorder is, indeed, ruled out by the scaling law, offdiagonal disorder might be quantitatively important. On the other hand, disorder is clearly not an essential ingredient for a model description.

Figure 3.2 b shows the resistivity as a function of inverse temperature for various levels of Cr doping. The lowest curve corresponds to the pure case, i.e., the vertical dashed line in the phase diagram Fig. 3.2 a. This curve displays only one kink, a sharp increase of the resistivity by 7 orders of magnitude at the Néel temperature. For sufficiently large doping, a second jump by some 2 or 3 orders of magnitude appears at higher temperature (i.e., going left in the plot) which indicates the Mott metal-insulator transition which is the subject of interest within this chapter.

### 3.1.2 Theory

A theoretical description of the $V_2O_3$ phase diagram should at least reproduce its topology including the first-order Mott transition line. If such a description could be found in terms of a purely electronic model, this would be a strong indication for electronic correlations being the driving force behind the Mott transition. Lattice degrees of freedom would, then, appear less important in this respect. While orbital degrees of freedom certainly play an important role for $V_2O_3$, at least quantitatively [see Held, Keller, Eyert, Vollhardt, and Anisimov (2001) and references therein], the very essence of the Mott transition may already be captured qualitatively by the half-filled single-band Hubbard model. This model is a likely candidate since its paramagnetic phase is a (perfect) metal for $U = 0$ and an insulator for $U = \infty$. A transition would be expected for $U \approx W$ if $W$ is an appropriate measure of the free bandwidth. For the usually considered case of NN hopping on a bipartite lattice, however, the antiferromagnetic phase generically extends to high enough temperatures that the evolution from metal to insulator as a function of $U$ should occur as a smooth crossover. This expectation is indeed an established fact for the Bethe lattice as well as for the hc lattice with NN hopping in infinite dimensions (see below) and is supported by recent QMC results for the simple cubic lattice in $d = 3$ (Staudt, Dzierzawa, and Muramatsu, 2000). Therefore, the Mott transition can only be studied on the Hubbard model when the antiferromagnetic phase is suppressed.

The natural microscopic method of modeling magnetic frustration is the inclusion of hopping terms to (at least) next-nearest neighbors. As discussed in chapter 2, frustration then generically implies an asymmetric DOS. DMFT results for the hypercubic lattice with NN and NNN hopping are shown in Fig. 3.3 a. Here, symbols and interpolating lines indicate the Néel temperature as computed using QMC for an increasing degree of frustration from top to bottom (Schlipf, 1998). Alternatively, antiferromagnetism may be suppressed in analytical and in mean-field methods[4] by restricting solutions to the subspace with appropriate symmetries. For the semi-elliptic Bethe DOS, this choice corresponds, e.g., to solving a fully disordered model where the hopping terms obey $\langle t_{ij} \rangle = 0$ and $\langle t_{ij}^2 \rangle = t^{*2}/Z$ (cf. subsection 2.2.2). The phase diagram of this "fully frustrated Bethe lattice" is shown in Fig. 3.3 b as obtained by Georges et al. (1996) within iterated perturbation theory (IPT). According to this result, metallic and insulating solutions coexist at low temperatures over a

[4]Such a suppression is not possible in the context of QMC applied directly in finite dimensions. For this method, the inclusion of NNN is not a practical alternative since it leads to the minus-sign problem. Therefore, finite-dimensional QMC can hardly treat frustration at all.

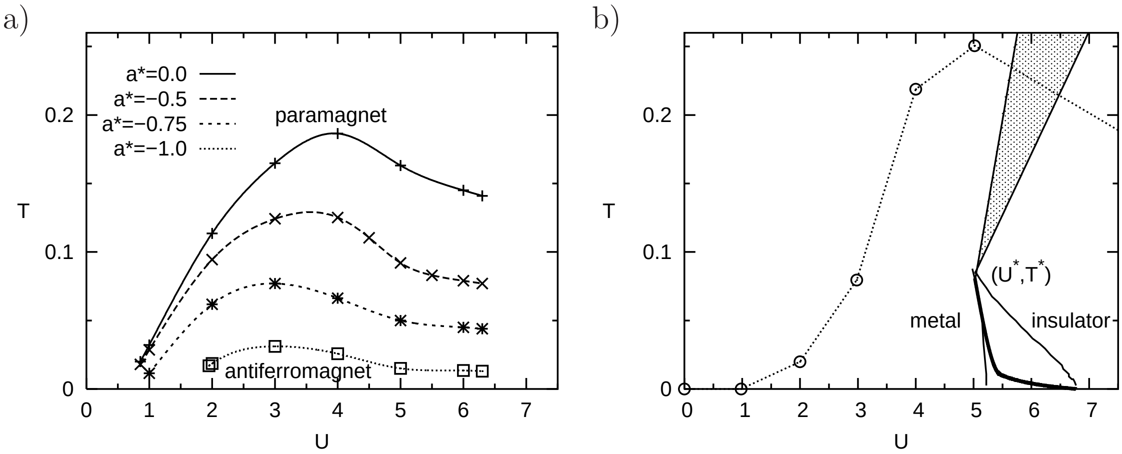



**Figure 3.3:** Magnetic and MIT phase diagrams for the Hubbard model. **a)** Néel temperature $T_\mathrm{N}$ for the hc lattice with NN and NNN hopping (Schlipf, 1998). **b)** MIT scenario for the fully frustrated Bethe lattice as computed within IPT (Georges et al., 1996). Here, a first-order phase transition is observed for $T < T^*$ at $U_\mathrm{c}(T)$ (thick line) within the coexistence region $U_{\mathrm{c}1}(T) \leq U \leq U_{\mathrm{c}2}(T)$ (thin lines); a crossover region (shaded) extends above $T^*$. Crosses (connected by a dotted line) indicate $T_\mathrm{N}$ for the unfrustrated Bethe lattice (Rozenberg et al., 1994); the AF then hides the MIT within IPT.

wide range of interaction strengths $U_{\mathrm{c}1}(T) \leq U \leq U_{\mathrm{c}2}(T)$; the coexistence region (narrow lines) extends up to $T^*_{\mathrm{IPT}} \approx 0.085$.[5] The shaded region at higher temperatures indicates a crossover region.

Schlipf's (1998) QMC calculations of the phase diagram for the hc lattice with NNN hopping showed no traces of a first-order MIT. Even with only NN hopping, no such transition appeared (for a calculation in the homogeneous phase). The contradiction with Rozenberg et al.'s (1995) results was perfect when a first-order MIT line could not even be found for the fully frustrated Bethe lattice at temperatures significantly below the IPT critical temperature (see Sec. 3.3). Thus, new studies became necessary which evolved into a project for this thesis. In this chapter, we will almost exclusively consider the Bethe lattice in the homogeneous phase which is enforced down to $T = 0$. Apart from being relevant on the physical grounds discussed in this section, this model has evolved into a benchmark for the methods of solving the DMFT. Since at least some of the same methods are used for multi-band models and in combination with LDA, the verification of their reliability has consequences far beyond the half-filled single band Hubbard model. The qualitative implications of finite frustration will be discussed in subsection 3.6.4.

[5] Here, energies are scaled to unit variance of the DOS for both models.

## 3.2 Characterization of Phase Transitions within the DMFT

In this section, we will discuss how phase transitions can be found and characterized within the DMFT. In subsection 3.2.1, we focus on the general implications of the mean-field character of the DMFT in connection with first-order transitions. A particularly important aspect is the question how the numerical stability of solutions in the iterative process is related to their thermodynamic (meta-) stability. This problem had been neglected in the literature so far and is now solved for a particular iteration scheme in subsection 3.2.2. Finally, we collect useful observables and discuss their general behavior in subsection 3.2.3.

### 3.2.1 Transitions of First or Higher Order

In a thermal statistical ensemble, properties such as the (free) energy per particle, average filling, or average double occupancy are defined in terms of traces over all possible states. Therefore, the observables are well-defined functions of a complete set of external parameters; for the Hubbard model on a given lattice (with unit variance of the DOS) these parameters are the on-site interaction $U$, the temperature $T$, and the chemical potential $\mu$. More precisely, the free energy $F$ is a continuous function while its derivatives may be discontinuous at some hypersurfaces, i.e., lines in a 2-dimensional subspace of parameters such as $U$ and $T$. The discontinuities indicate a phase transition; according to Ehrenfest, the order of the transition is defined as the order of the first discontinuous derivative of $F$. In principle, the search for a phase transition is, therefore, straightforward: sweeping through the parameter space, one searches for jumps or kinks in the derivatives of $F$ with respect to $U$ or $T$, i.e., the double occupancy $D$ or (minus) the entropy $S$ or related observables. This is the usual practice, e.g., for QMC in finite dimensions.[6]

In contrast, in mean-field theories such as the DMFT, the ensemble average is performed with the mean-field [here $\mathcal{G}(i\omega_n)$, cf. subsection 1.2.3] as an additional parameter which has to be determined self-consistently. The resulting fixed point problem may have one or more solutions for each point in the original parameter space. In regions of the phase space where only one solution is known to exist or where the derivatives of $F$ are found to be continuous, phase transitions can be pinpointed just like in direct methods by testing for higher order derivative discontinuities. If several solutions exist, however, the result of an iterative solution of the fixed point equations will depend on the initialization. For sequential solutions on a grid of phase points where the solution at one point initializes the iteration at the neighboring point, the solution within the coexistence region then depends on the sweeping direction: hysteresis occurs. For conserving approximations such as the DMFT, the free energy can be generalized as a Ginzburg-Landau functional $F_{\mathrm{GL}}$ of the physical parameters plus the mean field which agrees with the true free energy

[6] A serious complication for such methods based on the solution of finite size clusters is, however, the rounding of all phase transitions.

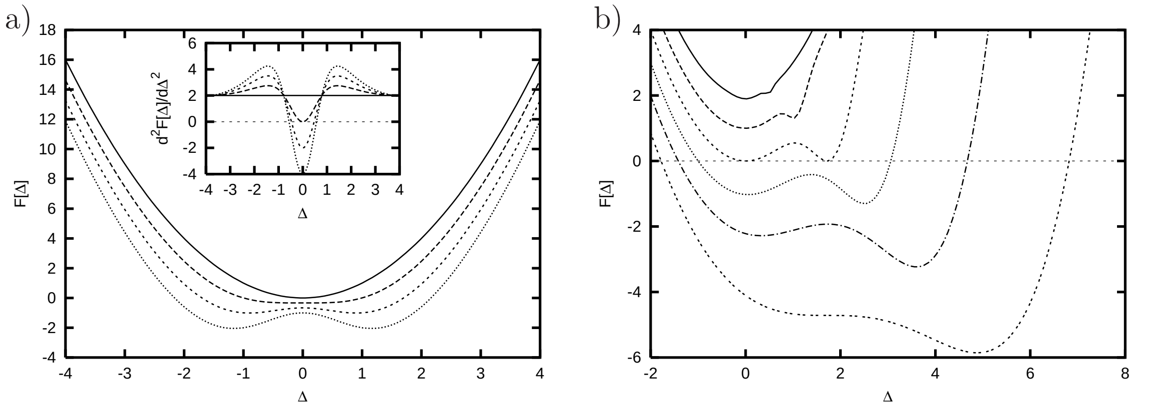



**Figure 3.4:** Schematic dependence of the free energy functional on an order parameter $\Delta$ near a first-order phase transition. **a)** For phase points above the critical point and along the transition line, one minimum splits into two minima of equal depth. The inset shows the second derivative of $F$. **b)** Evolution perpendicular to the transition line. The scenario in b) also serves as a cartoon of our concept of the Mott MIT; then, the right minimum corresponds to the metallic solution which exists almost until it equals the insulating solution while the insulating solution keeps its character until it disappears.

for the physical values of the mean field. Then, the physical solution is the one which minimizes $F_{\mathrm{GL}}$. Supposed we can identify an one-dimensional order parameter $\Delta$, the dependence of $F[U,T,\Delta]$ on $\Delta$ is similar to the curves shown in Fig. 3.4 for a first-order transition. The curves in Fig. 3.4 a illustrate the development along the first-order line (coming from above the critical point): a single minimum becomes flat (i.e., with zero curvature) at the second order end-point. It splits for lower $T$; the minima separate with increasing order parameter. Note that the region where the second derivative (inset of Fig. 3.4 a) is reduced may be quite small (cf. subsection 3.7.3). Therefore, asymptotically singular behavior may be visible only very close to the second-order end point. Figure 3.4 b depicts the evolution of the free energy functional perpendicular to the transition line: in the top curve, the left minimum corresponds to a stable solution while the right one is metastable. The third curve (from top) marks the MIT. In the bottom curve, only one minimum remains. Note the asymmetry in the coexistence region which will reappear in our results for the MIT.

Locating and characterizing phase transitions within the DMFT is potentially a two-step process. In a first step, solutions are obtained and observables are computed on a fine enough grid to identify nonanalytic behavior. If numerical discontinuities are detected, it is necessary to search for hysteresis effects with coexisting solutions. The result of the first step is then a diagram which indicates continuous transitions and/or coexistence regions. For the MIT on the Bethe lattice, such incomplete phase diagrams will be presented in Sec. 3.5. A second step is necessary for locating the first-order transition line within a coexistence region. This task is particularly difficult for QMC since this method is by construction unable to directly calculate the free

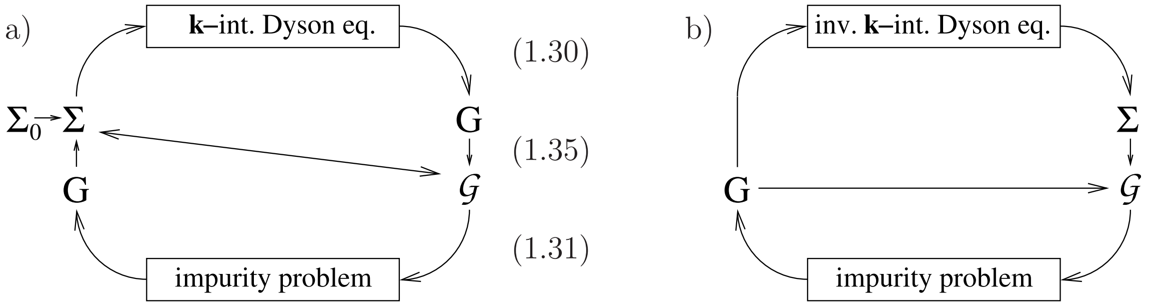



**Figure 3.5: a)** Conventional DMFT self-consistency cycle, replicated from Fig. 1.1. **b)** More direct iteration scheme, using the inverted lattice Dyson equation.

energy (which would be obtained by a full summation over the Hubbard-Stratonovich field). Our solution of this problem and its results will be discussed in Sec. 3.6.

### 3.2.2 Convergence of Fixed Point Methods

As discussed above and in Sec. 1.3, solutions are obtained within the DMFT in an iterative process. Evidently, there is a considerable degree of freedom for choosing a particular scheme. The conventional choice Fig. 3.5 a (here replicated from Fig. 1.1) is numerically convenient since an explicit inversion of the lattice Dyson equation is avoided. Its general usage in the literature implies the assumption that it converges to all solutions of the self-consistency equations if initialized appropriately and in the limit of high precision; i.e., that all solutions of the self-consistency equations are stable fixed points of the iteration scheme. This is, however, by no means guaranteed.[7] The only property that every iteration scheme fulfills by construction is that solutions of the self-consistency equations are (possibly unstable) fixed points. Hence, the iteration procedure for determining $U_{c1}$ and $U_{c2}$ can break down before the solutions themselves become thermodynamically unstable; as a consequence, the width of the coexistence region may be underestimated. At first sight, the iteration scheme Fig. 3.5 a might even imply the danger of introducing additional solutions since it incorporates a two-stage relaxation scheme: both the solution of the impurity problem and the solution of the Dyson equation are combined with old data for the next step in each cycle. Still, a full iteration may be written exclusively in terms of the "old" bath Green function which is updated as

$$\mathcal{G}^{-1} \longrightarrow \mathcal{G}^{-1} - \frac{1}{G[\mathcal{G}^{-1}]} + \frac{1}{\tilde{D}\big(\mathcal{G}^{-1} - \frac{1}{G[\mathcal{G}^{-1}]}\big)}\,. \tag{3.1}$$

Here, $G[\mathcal{G}^{-1}]$ denotes the solution of the impurity problem for given $\mathcal{G}^{-1}$, $\tilde{D}(\Sigma)$ denotes the solution of the lattice Dyson equation, i.e., the Hilbert transform for

[7]As a simple example, consider the reversal of any (locally) convergent scheme which clearly cannot converge to the same fixed point.

a given self-energy. At a fixed point, we have

$$G[\mathcal{G}^{-1}] = \tilde{D}\big(\mathcal{G}^{-1} - \frac{1}{G[\mathcal{G}^{-1}]}\big) \tag{3.2}$$

which is equivalent to

$$\mathcal{G}^{-1} = \tilde{D}^{-1}(G[\mathcal{G}^{-1}]) + \frac{1}{G[\mathcal{G}^{-1}]}\,. \tag{3.3}$$

Eqn. (3.3) involves the inversion of the Dyson equation and (using $\mathcal{G}^{-1} = \Sigma + 1/G$) is equivalent to the self-consistency equations. Therefore, no artificial solutions are introduced in scheme (3.1). Note that (3.3) can also be read as the prescription for the "direct" iteration scheme

$$\mathcal{G}^{-1} \longrightarrow \tilde{D}^{-1}(G[\mathcal{G}^{-1}]) + \frac{1}{G[\mathcal{G}^{-1}]}\,, \tag{3.4}$$

which is illustrated in Fig. 3.5 b.

The remaining question is whether the fixed points of both iteration schemes are also (locally) stable. Only then can the iteration schemes be expected to converge at least for a sufficiently high accuracy of the impurity solver. As we will see in subsection 3.7.1, the direct scheme (3.3) can be rigorously shown to be a downhill method. It is, thus, locally convergent at least in combination with underrelaxation. While the same has not yet been shown for the scheme (3.1) used throughout this work and in most of the literature, its stability is established empirically. Numerical tests performed in this work did not indicate an enhanced stability of the metallic solution for the direct scheme. Therefore, the ability to follow metastable solutions up to their stability edge seems primarily limited by the numerical accuracy of the impurity solver.

### 3.2.3 Observables

Evidently, the observables most directly available within the DMFT framework are the solutions $\Sigma(i\omega_n)$ and $G(i\omega_n)$ of the self-consistency equations themselves. One closely related quantity is the quasiparticle weight also known as mass renormalization factor $Z$, which is defined in terms of the real-frequency self-energy,

$$Z = \frac{m}{m^*} = \frac{1}{1 - \frac{\partial}{\partial\omega}\mathrm{Re}\,\Sigma(\omega)\big|_{\omega=0}}\,. \tag{3.5}$$

In the context of QMC simulations, one usually approximates this quantity by its discrete Eliashberg estimate

$$Z \approx \frac{1}{1 - \frac{\mathrm{Im}\,\Sigma(i\omega_1)}{\pi T}}\,. \tag{3.6}$$

Due to the rules for complex derivatives, both definitions agree in the limit $T \to 0$ as long as the Luttinger theorem is fulfilled, i.e., for $\mathrm{Im}\,\Sigma(\omega = 0) = 0$. For $U \to 0$, no mass renormalization takes place so that $Z = 1$. With increasing $U$, $Z$ decreases until the quasiparticle peak (in the spectrum) vanishes. Extending the Fermi liquid picture,

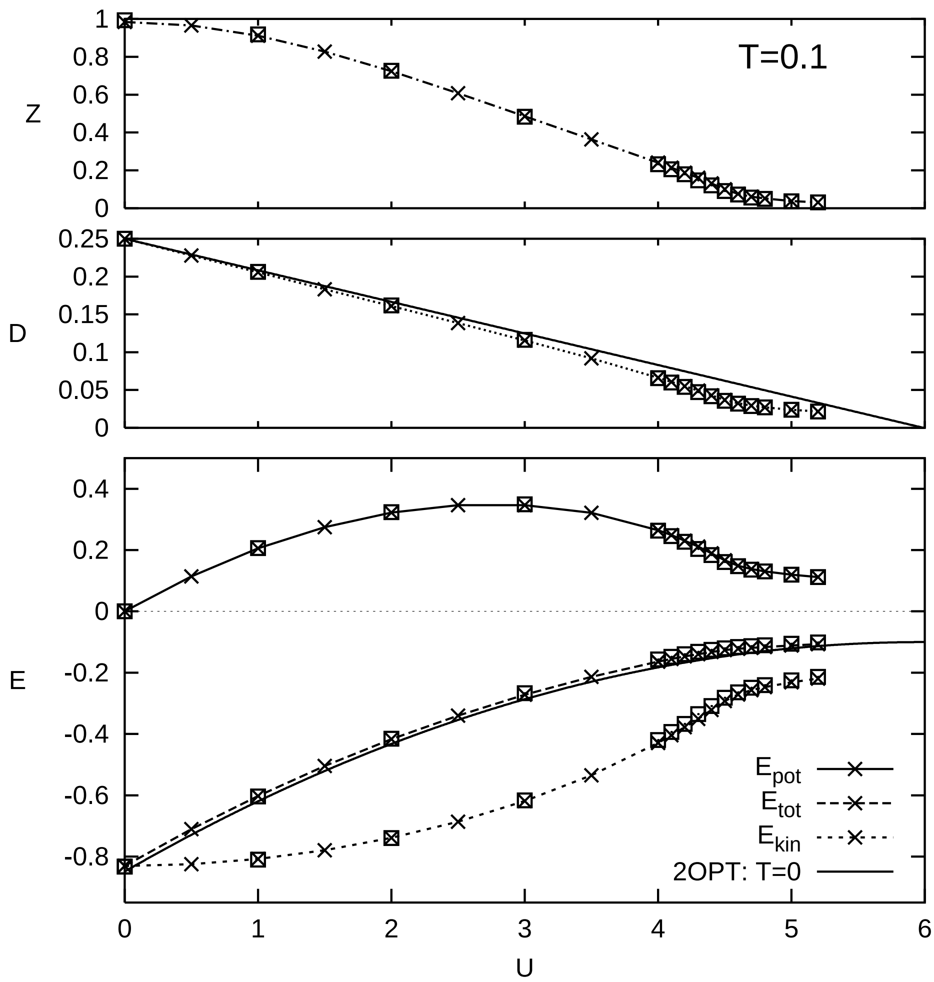



**Figure 3.6:** Quasiparticle weight $Z$ (Eliashberg estimate), double occupancy $D$, and energy contributions for $T = 0.1$. Crosses (connected with lines) denote QMC results for $\Delta\tau = 0.2$, squares are for $\Delta\tau = 0.125$. For comparison, results of second order perturbation theory (2OPT) are shown for the total energy and the double occupancy for $T = 0$ (solid black lines).

---

one usually associates the disappearance of the quasiparticle peak with $Z \approx 0$. $Z$ in the definition (3.5), however, is not positive near a metal-insulator transition so that its interpretation as a quasiparticle weight breaks down. In contrast, the discrete version (3.6) always leads to positive $\mathcal{Z}$ and may therefore appear more physical. In any case, $Z$ loses its theoretical foundation outside the Fermi liquid phase where it remains only a heuristic indicator of a metal-insulator transition. In the uppermost part of Fig. 3.6, $Z$ is shown for the relatively high temperature $T = 0.1$. A rapid change of slopes indicates a transition or crossover near $U \approx 4.7$.

Within the DMFT, the energy per lattice site is given as (Fetter and Walecka, 1971; Georges et al., 1996)

$$E = \lim_{\eta \to 0^+} T \sum_{n,\sigma} \int_{-\infty}^{\infty} \mathrm{d}\epsilon \frac{\mathrm{e}^{i\omega_n \eta}\, \epsilon\rho(\epsilon)}{i\omega_n - \epsilon - \Sigma(i\omega_n)} + \frac{1}{2} T \sum_{n,\sigma} \Sigma_\sigma(i\omega_n) G_\sigma(i\omega_n)\,. \tag{3.7}$$

Note the convergence factor $\mathrm{e}^{i\omega_n \eta}$ which is essential in order to get the correct result. Obviously, such a term is difficult to handle numerically; in practice, $\eta$ may be replaced, e.g., by the time discretization parameter $\Delta\tau$, which also determines the cutoff frequency in the infinite sum. This approximation can be avoided by evaluating the noninteracting part separately. For the kinetic energy, this implies

$$E_{\text{kin}} = \lim_{\eta \to 0^+} 2T \sum_{n=-\infty}^{\infty} \mathrm{e}^{i\omega_n \eta} \int_{-\infty}^{\infty} d\epsilon\, \epsilon\rho(\epsilon)\, \frac{1}{i\omega_n - \epsilon - \mu - \Sigma(i\omega_n)} \tag{3.8}$$

$$= 2 \int_{-\infty}^{\infty} d\epsilon \frac{\epsilon\rho(\epsilon)}{\mathrm{e}^{\beta\epsilon} + 1} + 2T \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} d\epsilon\, \epsilon\rho(\epsilon) \left( G_\epsilon(i\omega_n) - G^0_\epsilon(i\omega_n) \right) \tag{3.9}$$

$$\approx 2 \int_{-\infty}^{\infty} d\epsilon \frac{\epsilon\rho(\epsilon)}{\mathrm{e}^{\beta\epsilon} + 1} + 2T \sum_{n=-L/2+1}^{L/2} \int_{-\infty}^{\infty} d\epsilon\, \epsilon\rho(\epsilon) \left( G_\epsilon(i\omega_n) - G^0_\epsilon(i\omega_n) \right)\,, \tag{3.10}$$

where we have assumed the paramagnetic case and $\mu = 0$. Here, the interacting and noninteracting "momentum-dependent" Green functions read

$$G_\epsilon(i\omega_n) = \frac{1}{i\omega_n - \epsilon - \mu - \Sigma(i\omega_n)}\,; \qquad G^0_\epsilon(i\omega_n) = \frac{1}{i\omega_n - \epsilon - \mu}\,. \tag{3.11}$$

In (3.10), the terms in the Matsubara sum fall off at least as $1/\omega^2$, which makes it well-defined also without convergence factor. At the same time, the truncation error is reduced significantly. This is particularly important for our code, where the number of Matsubara frequencies is comparatively small since it equals the number of time slices $L$ (see Sec. 3.4). The complementary ingredient to the energy is the double occupancy $D$ with $E = E_{\text{kin}} + UD$. In the context of QMC calculations, this observable is best calculated directly from Wick's theorem when sampling over the auxiliary field. The overall behavior of $D$ and $E$ in the Hubbard model can be read off (for $T = 0.1$) from the middle and lower parts of Fig. 3.6, respectively. For small $U$, the kinetic energy increases quadratically while $D$ and, consequently, $E_{\text{pot}}$ and $E$ increase linearly. The potential energy reaches a maximum below $U = 3$. A region of strong curvature of $D$, $E_{\text{pot}}$, and $E_{\text{kin}}$ near $U = 4.6$ gives a rough indication of the metal-insulator crossover. The total energy $E$, however, hardly shows any anomalies at this scale. Note also that the solutions for $D$ and $E$ are quite close to the results of plain zero-temperature second-order perturbation theory. We will see later that the agreement actually becomes better for low-temperature QMC data, extrapolated to $T \to 0$. The offset of the curves for $E$ gives (for not too large $U$) an indication of the specific heat $c_V = \mathrm{d}E/\mathrm{d}T$ which is linear within the Fermi-liquid phase and is, in general, best evaluated by fitting the temperature dependence of $E$.

The direct evaluation of the compressibility (as of most other susceptibilities) is numerically costly since it requires the QMC computation of 2-particle vertex functions. The formalism is omitted here since we will not show corresponding results; it can be found, e.g., in (Georges et al., 1996; Blümer, 1996). Spectra and optical conductivity data also shed light on systems near an MIT. Since due to numerical uncertainties these observables cannot be used for a precise determination of the phase diagram, their discussion is postponed to Sec. 3.8 and chapter 4, respectively.

## 3.3 Phase Diagram: Development until 1999

In this section, we discuss early results for the fully-frustrated Bethe lattice in high dimensions, i.e., for the homogeneous phase of the Hubbard model with semi-elliptic DOS. The question of how the contrary scenarios of Brinkman and Rice (1970) and Hubbard (1964*b*) for the metal-insulator transition could be be realized in this model was among the first to be attacked numerically after the inception of the DMFT. Rozenberg et al. (1992) pinpointed the Mott-Hubbard transition at $U_\mathrm{c} = 6.2$ for $\beta = 32$ where they observed the disappearance of the quasiparticle peak using QMC.[8] Georges and Krauth (1992) reported $U_\mathrm{c} = 6$ for $\beta = 16$ with an analogous method. A coexistence region of metallic and insulating solutions extending from $U_{\mathrm{c}1} = 5.1$ to $U_{\mathrm{c}2} = 6.6$ at $T = 0$ was found by Georges and Krauth (1993) within iterative perturbation theory (IPT). The more detailed phase diagram by Rozenberg et al. (1994) is shown in two versions: Fig. 3.7 a is taken from the original publication while Fig. 3.7 b is reproduced from the review (Georges et al., 1996). The coexistence region is in good agreement with Georges and Krauth's (1993) earlier estimate. The position of the first-order phase transition line was apparently adjusted in the review on the basis of analytical considerations. The failure of the original IPT scheme to predict the first-order transition line $U_\mathrm{c}(T)$ qualitatively correctly[9] can be linked to the thermodynamic inconsistency of the method which is not a conserving approximation in contrast, e.g., to the DMFT framework itself. As shown in Fig. 3.8 a, the IPT overestimates the double occupancy near the Mott transition (in comparison with ED) at $T = 0$. Thus, the IPT scenario shown in Fig. 3.7 b cannot be quantitatively correct, but should overestimate $U_{\mathrm{c}1}$ and $U_{\mathrm{c}2}$. Otherwise, the IPT phase diagram Fig. 3.7 b, supported by QMC data as shown in Fig. 3.8 b, was considered authoritative.[10]

In fact, it stood until 1998 when our group (Schlipf, 1998) surprisingly could not stabilize a metallic phase in QMC calculations for the presumed coexistence region

[8] In his pioneering DMFT-QMC study, Jarrell (1992) demonstrated Mott-Hubbard behavior, i.e., a vanishing quasiparticle peak and the development of a pseudogap for the hypercubic lattice. The focus of that work, however, was on antiferromagnetism and on the exclusion of superconductivity.

[9] In this context, Georges et al. (1996) note that *"the free energy within the iterated perturbation theory must be evaluated from a functional that gives the mean-field equations by differentiation, rather than from the one-particle Green's function. This is crucial in order to obey the exact property of a second-order transition"* (at $T = 0$).

[10] In the review, Georges et al. (1996) write: *"Detailed comparisons between the iterated perturbation theory approximation and the QMC method at finite temperature have shown excellent agreement for $T/D$ of the order or $1/50$ or higher. In this region, the iterated perturbation theory phase diagram is quite reliable."*

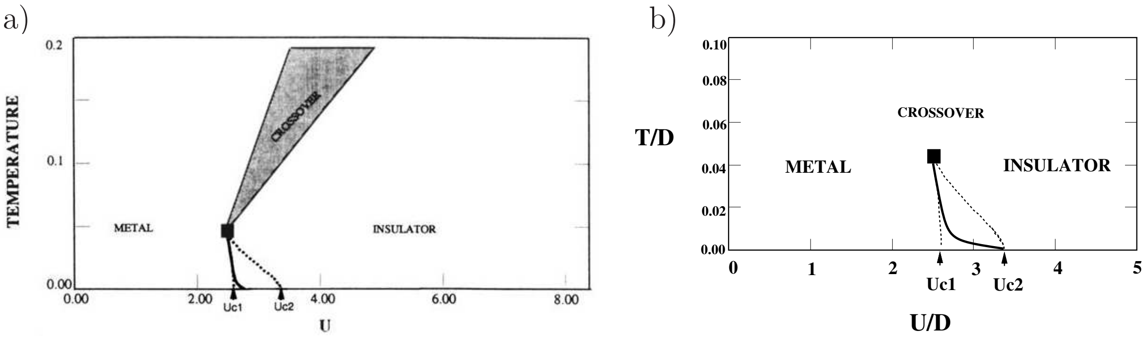



**Figure 3.7:** IPT phase diagram of the fully frustrated Bethe lattice. A coexistence region extends from $U_{c1}(0)/D = 2.6$ to $U_{c2}(0)/D = 3.3$ at $T = 0$. The second order end point of the first-order MIT (thick line) is at $T^*_{\rm IPT}/D = 0.044$, $U^*_{\rm IPT}/D = 2.51$. **a)** Original version with $U_c(T) \approx U_{c1}(T)$ for all temperatures (Rozenberg et al., 1994). **b)** Version in review with $U_c(0) = U_{c2}(0)$ (Georges et al., 1996). Scales in a) correspond to $t^{*2} = 1/2$. In both plots, the numerical values have to be doubled for the scaling $t^{*2} = 1$ used in this work (half bandwidth $D = 2$).

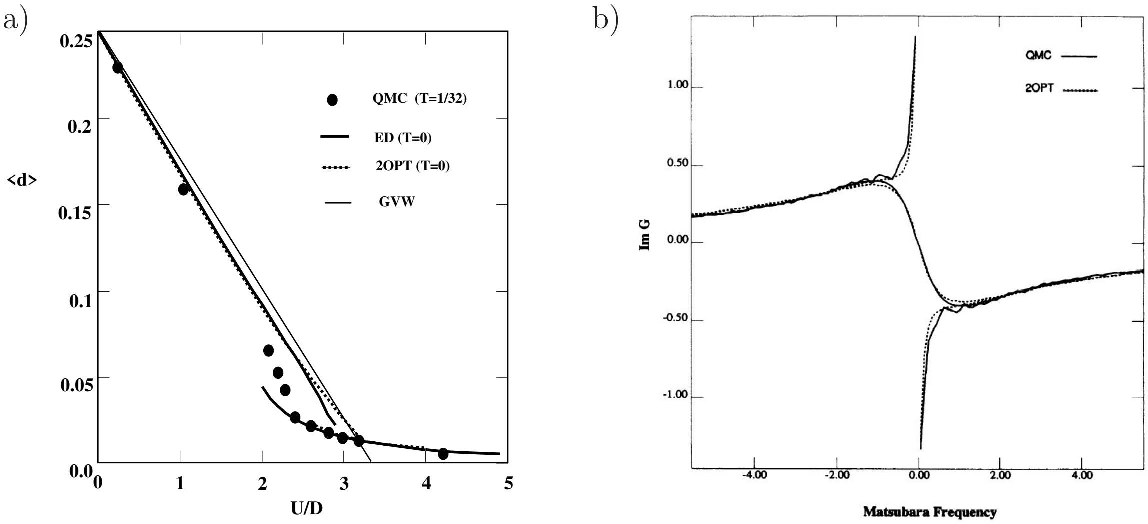



**Figure 3.8:** **a)** Double occupancy, here denoted as $\langle d \rangle$. Zero temperature results from IPT (here denoted as "2OPT") and ED of an 8-site cluster are compared to an estimate using the Gutzwiller wave function (GVW) and a finite-temperature QMC calculation at $T/D = 1/32$ (Georges et al., 1996). Note the agreement of IPT, ED, and even (finite-temperature) QMC in the insulating phase. Significant discrepancies between ED and IPT arise in the interesting region $U/D \approx 3$. **b)** Coexisting metallic and insulating solutions $\operatorname{Im} G(i\omega_n)$ for $U/D = 2.8$ and $T/D = 1/64$ in good agreement between IPT (dotted smooth lines) and QMC (solid noisy lines) (Rozenberg et al., 1994).

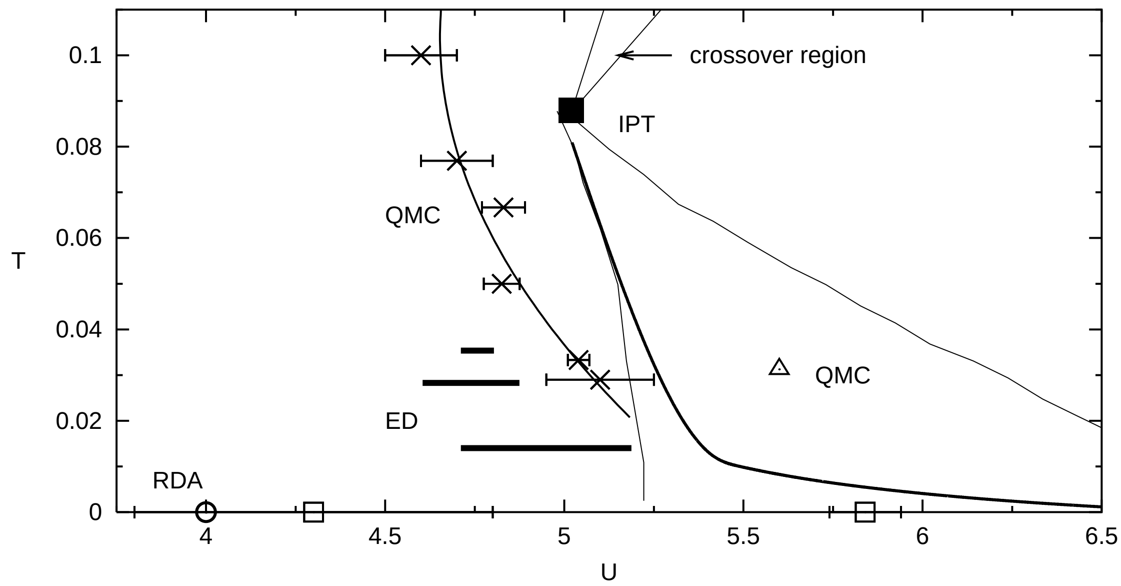



**Figure 3.9:** Status of the phase diagram of the fully frustrated Bethe lattice in early 1999: new QMC results [crosses; Schlipf et al. (1999)] indicate that the metallic phase breaks down before the IPT coexistence region (cf. Fig. 3.7) with the supporting QMC coexistence point (triangle, cf. Fig. 3.8 b) is reached. As an estimate for $U_{c2}$, the new QMC results are roughly consistent with ED for finite $T$ (Hofstetter, 1998) and $T = 0$ (Georges et al., 1996); the transition is, however, observed as continuous in QMC in contrast to ED and IPT. The RDA result (circle) also supports a continuous transition, but for much smaller interaction $U$.

(down to $T = 0.05$). Initial hysteresis near the computed transition line was found to disappear upon sufficiently long iteration. Schlipf also noted that no ED results for the MIT at finite temperature had been published. Using Krauth's ED code as made available in connection with the review, Hofstetter (1998) computed a rough estimate of the coexistence region (thick horizontal lines in Fig. 3.9). The observed coexistence region is much smaller than in IPT and shifted to smaller values of $U$. Due to this disagreement and since also the ED could in principle fail near the MIT when the width of the quasiparticle peak becomes smaller than its energy resolution, it appeared conceivable that the IPT scenario was even *qualitatively* wrong. The QMC calculations were refined and extended to lower temperatures (as part of this thesis). In order to exclude the possibility of errors in the implementation as the reason for this discrepancy, the collaboration was extended to include Jarrell who performed independent QMC runs. As a final conclusion, the metal-insulator transition was found to be continuous at least down to $T = 1/35$, about a factor of 3 below the IPT estimate for $T^*$ (Schlipf et al., 1999).[11] These QMC results (crosses) are also shown in Fig. 3.9. Here, the error bars for $U_c(T)$ extend in horizontal direction since

[11]Note that the estimate for $T^*$ is even higher in more recent IPT calculations (Kotliar, Lange, and Rozenberg, 2000); see Table 3.2.

the QMC hysteresis runs were performed at constant temperature. The interpolating line is a guide to the eye only; it is certainly compatible with the $T = 0$ estimates for $U_{c2}$ by ED and PSCT (Moeller, Si, Kotliar, Rozenberg, and Fisher, 1995). Still, the scenario of a continuous transition is in fundamental qualitative disagreement with both ED (squares in Fig. 3.9 indicate the zero-temperature results, note the large error bars) and IPT.

Our controversial result stirred new interest in the problem. Qualitatively, the continuous scenario was supported by results from the random-dispersion approximation (RDA), evaluated using ED on a small cluster. Quantitatively, however, the agreement was hardly better than, e.g., with IPT, since the (zero-temperature) RDA result of $U_c = 4$ (shown as a circle in Fig. 3.9) is clearly too small to be compatible with the QMC data.

The result of a new QMC study by Rozenberg, Chitra, and Kotliar (1999), in contrast, was quantitatively almost compatible with our data (within double error bars[12]). It also confirmed that the QMC coexistence point computed earlier by Rozenberg et al. (1994) and shown as triangle in Fig. 3.9 was indeed incorrect. However, the new study demonstrated the coexistence of metallic and insulating solutions (triangles in Fig. 3.10) in a temperature range where Schlipf et al. (1999) had ruled out a first-order transition. Further support for the scenario of a first order transition came from zero-temperature NRG studies (Bulla, 1999). The resolution of this contradiction, obtained for the same model (and now even using the same numerical methods) became an important project for this thesis and will be the subject of the remainder of this chapter. The first part, the verification of coexistence, was accomplished in the fall of 1999, soon after Rozenberg made available a set of QMC estimates for $G(\tau)$ and $\mathcal{G}(\tau)$ in both the metallic and insulating phase for the single parameter set $T = 1/32$, $U = 4.92$ (cf. Fig. 3.12).[13] At this point, all involved groups accepted that a finite, but small coexistence region was established in the lowest temperature range studied so far using QMC. Still, the shapes of the coexistence region were unclear. The thick solid lines in Fig. 3.10 represent Kotliar's perception that the region should broaden significantly for lower $T$, reaching the RDA point for $U_{c1}$ and matching the (practically identical) estimates for $U_{c2}$ obtained using ED, PSCT, and NRG (see Table 3.1 on page 112).

## 3.4 Discussion of QMC Algorithms

In the main part of this section, we will discuss two aspects of the quantum Monte Carlo (QMC) algorithm for the DMFT self-consistency problem: the discrete Fourier transformations between imaginary time and Matsubara frequencies and the search strategy for solutions of the DMFT equations. Both aspects are directly relevant for the detection of the Mott transition and the associated coexistence region. The different algorithms for handling the discrete Fourier transform used in the codes

[12]The point at $T = 1/30$ with the small error bar was not computed as part of this thesis.

[13]Note that this phase point is not among the previously published data, but more in the center of the claimed coexistence region.

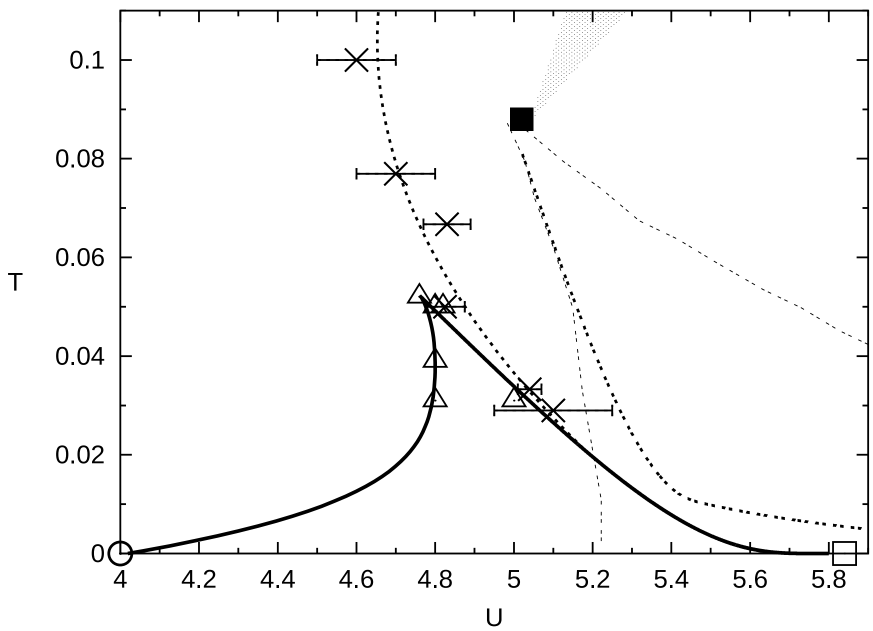



**Figure 3.10:** Status of the phase diagram of the fully frustrated Bethe lattice in summer 1999: QMC simulations (triangles) by Rozenberg et al. (1999) establish coexisting solutions for $T < 0.05$ on the metallic side of Schlipf et al.'s (1999) transition line. The solid lines indicate Kotliar's new suggestion for the coexistence region which should reach both the RDA point and the NRG/ED/PSCT point at $T = 0$ (cf. Fig. 3.9) and is considerably smaller and shifted in comparison with the IPT prediction (thin dashed lines).

---

by Ulmke (our group), Krauth and Rozenberg (as published in the review Georges et al. (1996)), and Jarrell are compared in subsection 3.4.1. There, we will see that the "Ulmke" code contains an avoidable error proportional to the discretization $\Delta\tau$ which may prevent the detection of a second stable phase. This fact had not been caught before, since the algorithm is still correct for $\Delta\tau \to 0$ in single-phase regions (with increased numerical effort for the same accuracy). This deficiency will be cured in a modified scheme which brings the resulting code approximately on par with Krauth's code. Jarrell's code, however, will be seen to be still superior, in particular with respect to treating the Fourier transform. Therefore, the fact that Jarrell apparently missed a (small) coexistence region at $T = 1/30$ cannot be explained from the code alone. In subsection 3.4.2, we will identify the search strategy for converged solutions as the problem. Insight gained in this subsection will also put the numerical significance of Rozenberg et al.'s (1999) results in perspective. The final two subsections are more general: subsection 3.4.3 contains a discussion of errors and their estimation while subsection 3.4.4 shortly explains the MPI parallelization of the QMC code accomplished in this thesis.

### 3.4.1 Fourier Transformation and Smoothing

A QMC simulation within the DMFT framework consists of a simultaneous solution of two principal equations: the lattice Dyson equation (1.30) and the defining expression for the impurity Green function (1.31). The IPT and QMC solutions of the impurity problem are formulated in imaginary time, i.e., the bath Green function $\mathcal{G}$ is needed as function of $\tau$ and the result is expressed as $G(\tau)$. In contrast, the Dyson equation is formulated (and is local) in the frequency domain, here for Matsubara frequencies $i\omega_n = i(2n+1)\pi T$. Therefore, two Fourier transformations (from frequency to imaginary time and vice versa) per self-consistency cycle are necessary, which for $G$ read

$$G(i\omega_n) = \int_0^\beta \mathrm{d}\tau\, \mathrm{e}^{i\omega_n\tau}\, G(\tau) \tag{3.12}$$

$$G(\tau) = \frac{1}{\beta}\sum_{n=-\infty}^{\infty} \mathrm{e}^{-i\omega_n\tau}\, G(i\omega_n)\,. \tag{3.13}$$

Note that (3.13) implies antiperiodicity of $G(\tau)$ for translations $\beta$ (since $\mathrm{e}^{i\omega_n\beta} = -1$) and allows for a discontinuity of $G(\tau)$ (at $\tau = 0$) since the number of terms is infinite. The spectral representation of $G$ implies a decay of $G(i\omega_n)$ as $1/i\omega_n$ for $|n| \to \infty$ like for any normalized analytic function. Furthermore, $G(i\omega_n)$ is purely imaginary when $G(\tau) = G(\beta - \tau)$ as in the case of interest.

#### Discretization Problem

Numerically, however, the integral in (3.12) needs to be discretized and the Matsubara sum in (3.13) has to be truncated. Since the numerical effort in QMC scales with the number $\Lambda = \beta/\Delta\tau$ of discretized time slices at least[14] as $\Lambda^3$, this method is presently restricted to $\Lambda \lesssim 400$. Typically, 200 time slices and less are used. A naive discrete version of the Fourier transform,

$$\tilde{G}(i\omega_n) = \Delta\tau\Big(\frac{G(0) - G(\beta)}{2} + \sum_{l=1}^{\Lambda-1} \mathrm{e}^{i\omega_n\tau_l}\, G(\tau_l)\Big); \quad \tau_l := l\Delta\tau \tag{3.14}$$

$$\tilde{G}(\tau_l) = \frac{1}{\beta}\sum_{n=-\Lambda/2}^{\Lambda/2-1} \mathrm{e}^{-i\omega_n\tau_l}\, G(i\omega_n)\,, \tag{3.15}$$

fails for such a coarse grid. The problems are that the Green function $\tilde{G}(\tau)$ estimated from a finite Matsubara sum can not be discontinuous at $\tau = 0$ (as required analytically for $G(\tau)$) while the discrete estimate $\tilde{G}(i\omega_n)$ oscillates with periodicity $2\pi i\Lambda/\beta$ instead of decaying for large frequencies. This implies a large error of $G(i\omega_n)$ when $|\omega_n|$ approaches or exceeds the Nyquist frequency $\pi\Lambda/\beta$. Both (related) effects would

[14] In practice, the scaling is even worse on systems with a hierarchy of memory systems of increasing capacity and decreasing speed.

make the evaluation of the corresponding self-consistency equations pointless. In particular, in the naive scheme, the self-energy diverges near the Nyquist frequency. Finally, the sum in (3.15) is numerically somewhat unstable since $\widetilde{G}(\tau)$ also oscillates between the grid points $\tau_l$.

### Splining Method

Fortunately, there is physical information left that has not been used in the naive scheme: $G(\tau)$ is known to be a smooth function. In fact, it follows from (1.55) and (1.56) that $G(\tau)$ and all even derivatives are positive definite and, consequently, reach their maxima at the edges $\tau = 0$ and $\tau = \beta$. This knowledge of "smoothness" can be exploited in a spline interpolation of the QMC result $\{G(\tau_l)\}_{l=0}^{\Lambda}$ by piecewise cubic polynomials. The resulting functions may then either be used for oversampling, i.e., for generating $G(\tau)$ on a finite grid or for piecewise direct analytic Fourier transforms. In both cases, $G(i\omega_n)$ can be calculated for a much larger frequency range than before. Jarrell, Akhlaghpour, and Pruschke (1993) used oversampling with typically 800 Matsubara frequencies. The second choice mentioned above is implemented in Krauth's program: for any choice of $\Lambda$, $G(i\omega_n)$ is evaluated for a very large number (by default $N = 2^{13} = 8192$) of Matsubara frequencies. While the spline interpolation incorporates some physical information, it is clearly not exact; therefore, the method used on its own has its limits. Jarrell noted that the usual cubic splines with continuous second derivative were dangerous in this respect since they tend to generate artificial high-frequency features. This effect was found to be reduced using Akima splines. Both implementations use splines with natural boundary conditions, i.e., with vanishing second derivative at the end points.

As we noted recently, this is in general not a good choice; after all, the true second derivatives are maximal at the end points. The bad fit at the boundaries then leads to ringing (which decays faster for Akima splines). An implementation with free boundary conditions for the second derivative which are determined in an iterative minimalization scheme was found to be vastly superior over both schemes at least for a restricted set of test cases. The minimization procedure can even be avoided by computing the second derivative of $G(\tau)$ at $\tau = 0$ exactly. A related method for improving on the natural spline scheme is to interpolate not $\{G(\tau_l)\}_{l=0}^{\Lambda}$, but only the difference with respect to some reference Green function obtained from, e.g., plain second order perturbation theory or IPT. This approach, discussed in a general context by Deisz, Hess, and Serene (1995), is implemented in Jarrell's code. In addition, the high-frequency part may be directly stabilized by supplementing the QMC estimates with IPT (using low-pass/high-pass filters). Since each of the approaches discussed so far generates a number of Matsubara frequencies which is much greater than the number of time slices, the inverse Fourier transform is unproblematic: the rounding-off near $\tau = 0$ is already small for the first grid point $\tau = \Delta\tau$ while at the end points $\tau = 0$ and $\tau = \beta$, the numerical Fourier transforms can be shifted by the value $1/2$ to exactly cancel the rounding effect. For more information on recent developments, see App. C and Knecht's (2002) diploma thesis.

### Ulmke's Smoothing Trick

A completely different approach was taken by Ulmke (1995). Here, the number of Matsubara frequencies is chosen equal to the number of time slices. In order to enforce the correct analytical behavior of the Fourier transforms, a "smoothing" trick is employed: the naive Fourier transform is combined with an approximate correcting transformation which approaches the identity for $\Delta\tau \to 0$. The algorithm consists of the following two steps:

1. Asymmetric naive Fourier transform $\tilde{G}(i\omega_n) = \Delta\tau \sum_{l=0}^{\Lambda-1} \mathrm{e}^{i\omega_n \tau_l}\, G(\tau_l)$. The difference to the symmetric transform (3.14) is a constant finite real term $\Delta\tau/2$.

2. Smoothing transformation
$$G(i\omega_n) = \frac{\Delta\tau}{\ln(1 + \Delta\tau/\tilde{G}(i\omega_n))}\,. \tag{3.16}$$

This corrected Green function fulfills $\mathrm{Re}\, G(\omega) = 0$ (for half filling and symmetric DOS) and (approximately) $G(i\omega_n) \propto 1/(i\omega_n)$ for $\omega \to \pm\omega_{\mathrm{max}}$. By using the inverse correction transform $\tilde{G}(\omega) = \Delta\tau/(\exp(\Delta\tau/G(\omega) - 1))$ before the inverse Fourier transform, also the artifacts in $G(\tau)$ (or $\mathcal{G}(\tau)$) are removed. Ulmke's (1995) claim that the transformation scheme would be exact in the noninteracting and in the atomic limit is, however, incorrect. The assumed asymptotic form $G(i\omega_n) = 1/(i\omega_n)$ for hopping $t = 0$ applies only to the degenerate $U = 0$ case. Otherwise, $G(i\omega_n)$ is linear for small frequencies in the insulating regime. This misconception already points towards the main deficiency of the scheme: in order to cure a problem at high frequencies, a transformation is applied which changes $G$ irrespective of the frequency. In fact, all values of $G$ are modified significantly which are not large compared to $\Delta\tau$. Among this group, however, are not only points near the Nyquist frequency, but also those with small $i\omega_n$ (in the insulator). Consequently, Ulmke's approach treats insulating solutions much worse than metallic solutions.

In order to illustrate this effect and for a general quantitative comparison, we used the insulating solution $G(\tau)$ for $\beta = 32$, $U = 4.92$, and $\Delta\tau = 0.25$ as provided by Rozenberg as input for the various Fourier transformation schemes. Each output $G(i\omega_n)$ was then transformed to the corresponding self-energy $\Sigma(i\omega_n)$ via the inverted Bethe lattice Dyson equation. As shown in Fig. 3.11, only Jarrell's algorithm yields a self-energy with the correct analytic form, i.e., with negative imaginary part for all positive $\omega_n$.[15] While Ulmke's algorithm just stops at the Nyquist frequency $\pi/\Delta\tau$, Krauth's code overshoots substantially (see also right inset).[16] At small frequencies, the results of Krauth's method and of Jarrell's method agree very well (left inset),

[15] Slight deviations from the quantitatively correct asymptotic form for very large frequencies (not shown) could have been avoided by using the exact asymptotic expression $\Sigma(i\omega_n) \to U^2/(4i\omega_n)$ in this regime; cf. App. C.

[16] This overshooting can be attributed to the use of natural splines; it might have been overlooked since the original code does not output frequency information. The origin of slight deviations for intermediate frequencies $\omega_n \approx 1$ (in comparison with the curves for Jarrell's method and the modified Ulmke method) is unclear.

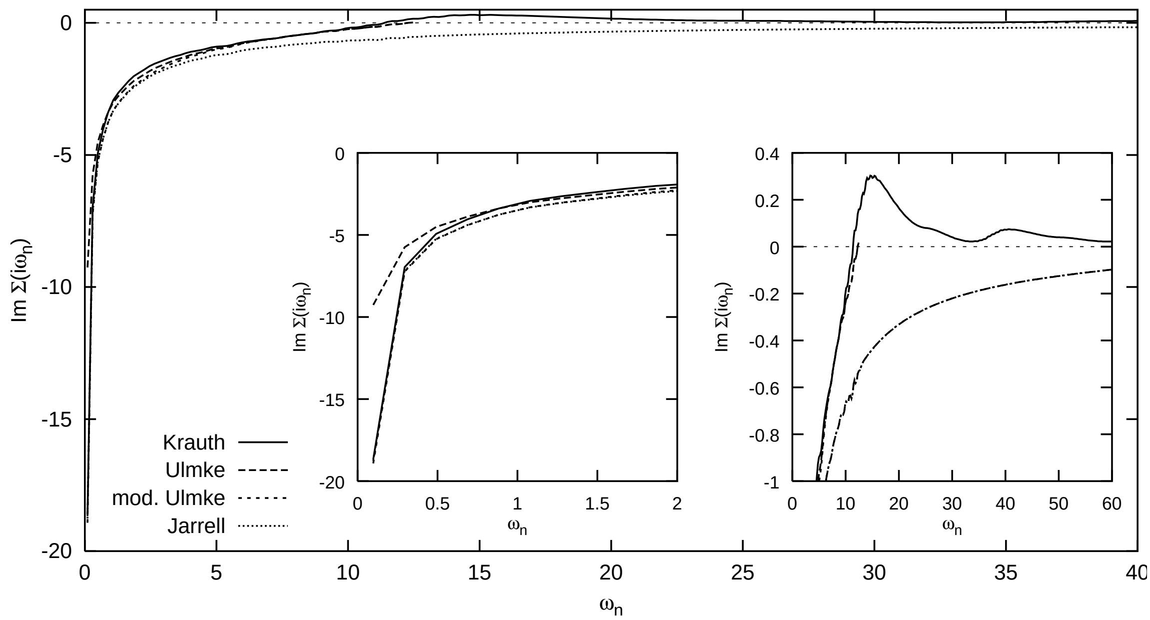



**Figure 3.11:** Comparison of Fourier-transform schemes for the step $G(\tau) \longrightarrow G(i\omega_n)$; here represented as $\Sigma(\omega) = \omega - G^{-1}(\omega) - (W/4)^2 G(\omega)$ for better resolution of the differences. For all methods, the imaginary-time solution $G(\tau)$ for the insulating phase at $\beta = 32$, $U = 4.92$, and $\Delta\tau = 0.25$ provided by Rozenberg (see text) was used as input. The left inset concentrates on low frequencies where the smoothing artifact of Ulmke's original method (long-dashed line) is seen clearly; the modified version does not show this artifact. At high frequencies (right inset), Krauth's curves overshoot substantially while Jarrell's method yields $1/\omega$ decay; Ulmke's method (both versions) makes no predictions beyond the Nyquist frequency.

---

while Ulmke's code is off by a factor of about 2. As we will check below, this striking disagreement completely destabilizes any insulating solution (for $\Delta\tau = 0.25$).

### Improved Smoothing Method

In order to further isolate the inaccurate treatment of $G(i\omega_n)$ at small frequencies as the cause of this problem, we searched for a specific cure within the general framework of the smoothing trick. Obviously, the smoothing transformation (3.16) has to be replaced by an explicitly frequency-dependent formulation. While interpolation (using the result of the "naive" transform at small frequencies) is in principle possible, our choice is better controlled: we replaced $\Delta\tau$ in (3.16) by $\big(1 - (\omega_n\Delta\tau/\pi - 1)^8\big)\Delta\tau$. The additional factor fulfills two requirements: it is very close to the value 1 when $\omega_n$ is of the order of the Nyquist frequency $\pi/\Delta\tau$ and of the order $\Delta\tau$ for $\omega_n \to 0$ (while always being smaller than 1). Therefore, the smoothing trick can still work for large frequencies while the error at low frequencies is now $\mathcal{O}(\Delta\tau^2)$ instead of $\mathcal{O}(\Delta\tau)$. The dramatic improvement is apparent in Fig. 3.11: the short-dashed line ("mod. Ulmke") is indistinguishable from the result of Jarrell's method for small frequencies

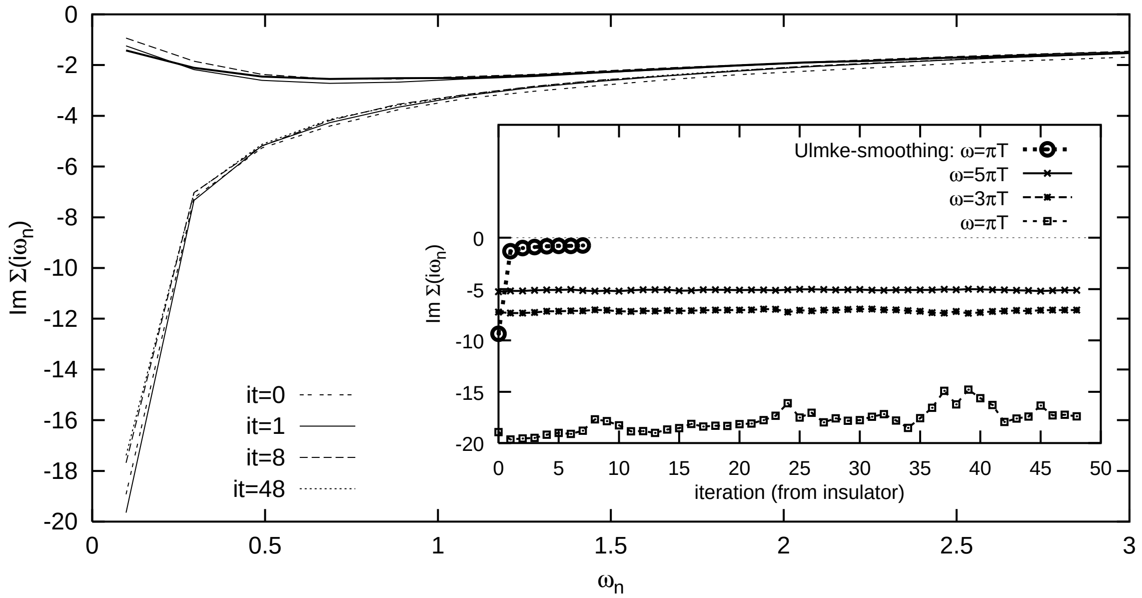



**Figure 3.12:** Behavior of the smoothing-corrected Ulmke code upon iteration, initialized with Rozenberg's solutions for $\beta = 32$, $U = 4.92$, and $\Delta\tau = 0.25$. The main part shows the input (it=0) plus snapshots after 1, 8, and 48 iterations. In the inset, $\mathrm{Im}\,\Sigma$ of the insulating solution for the first 3 Matsubara frequencies is presented as function of iteration number. For $\omega_n = \pi T$, results of the uncorrected ("Ulmke") code are also included.

while it shares the high-frequency behavior with the result of the original version ("Ulmke"). For additional comparisons, see App. C.

It can be seen in Fig. 3.12 that this improvement indeed stabilizes the insulating solution. Using Rozenberg's solutions (both for metal and insulator) as input, 8 iterations using the improved code were performed for the metal and 48 for the insulator. Both solutions are indeed stable in spite of relatively large fluctuations in the insulating phase.[17] In contrast, the insulating solution is lost and approaches the metallic solution within a few iterations when the code with uncorrected smoothing ("Ulmke") is used.

The insulating solution was also found to be stable upon iteration with Krauth's and Jarrell's code (not shown). Consequently, the coexistence of metallic and insulating phases for $\beta = 32$, $U = 4.92$, and $\Delta\tau = 0.25$ is established as a result of this subsection. Furthermore, we have identified and corrected the problem in Ulmke's code. Since our improved version of the code works so well even in the numerically difficult coexistence region and for a comparatively large discretization $\Delta\tau$, we continue to use it for the rest of our numerical studies.

[17]For the insulating phase, the accuracy was increased in later iterations. In iterations 1-16, 10 starting configurations of the Ising field with 5000 warmup and 45000 measurement sweeps each were used. Iterations 17-32 used 10 times 5000+150000 sweeps, iterations 33-48 used 10 times 15000+150000 sweeps.

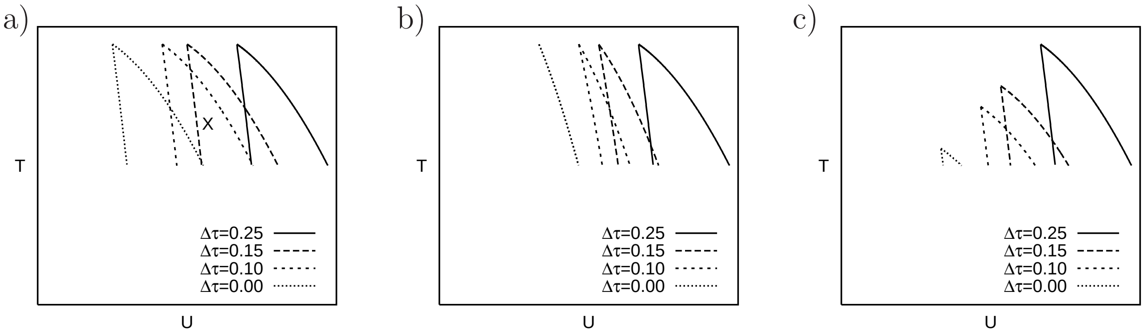



**Figure 3.13:** Scenarios for the effect of a $\Delta\tau \to 0$ extrapolation on the coexistence region: **a)** coexistence region shifts to smaller $U$ at constant shape, **b)** coexistence region shrinks with constant critical temperature $T^*$, **c)** coexistence region shifts to smaller $U$ and $T$, i.e., shrinks within the attainable temperature range.

## 3.4.2 Overrelaxation and Sweeping Strategies

One point that we have not discussed so far is the dependence of the shape and extent of the coexistence region on the discretization variable $\Delta\tau$. It is important to remember that the numerical results are only physically relevant after an extrapolation to $\Delta\tau = 0$. While Rozenberg et al.'s (1999) results for $\Delta\tau \geq 0.25$ make it likely that the observed coexistence persists at $\Delta\tau = 0$, both the remaining errors in the Fourier transformation and the inevitable Trotter error could have influenced the phase diagram considerably. Using the knowledge that $U_{c2}$ decreases with decreasing $\Delta\tau$ [this is already apparent from Schlipf's (1998) results], the scenarios depicted schematically in Fig. 3.13 appear possible: If the coexistence region is merely shifted to lower $U$ at roughly constant shape (Fig. 3.13 a), the conclusions drawn from $\Delta\tau = 0.25$ may be regarded as essentially correct. If, in contrast, the coexistence region shrunk with decreasing $\Delta\tau$ and vanished for $\Delta\tau = 0$ (Fig. 3.13 b), the picture at finite $\Delta\tau$ would be qualitatively incorrect. The existence of a first-order transition would then appear as an artifact for finite $\Delta\tau$. As an intermediate case, the coexistence region could also shrink with the critical temperature $T^*$ being considerably reduced for $\Delta\tau \to 0$. Then, the conclusions from $\Delta\tau = 0.25$ would still be qualitatively correct for the full phase diagram, but not for the temperature region under consideration. In subsection 3.5.4, will show that the $\Delta\tau \to 0$ extrapolation corresponds in leading order to a shift to smaller $U$, but also to slightly smaller $T$ which is somewhere in between the scenarios Fig. 3.13 a and Fig. 3.13 c.

Still, the question remains why Jarrell did not see the small coexistence region for $T = 1/30$ in his contribution to our paper (Schlipf et al., 1999). The answer includes two aspects. Both are related to methods that usually improve the convergence process of DMFT simulations both in speed and in quality, but may fail to detect multiple solutions. The first method essentially incorporates an extrapolation to small $\Delta\tau$ within the iteration process: starting with relatively large $\Delta\tau$, some iterations are performed; then, $\Delta\tau$ is decreased before continuing the iteration process. Thus, costly iterations at small $\Delta\tau$ are usually only performed when the input is already close to the converged solution. This strategy may fail, however, in the MIT coexistence region. For illustration, let us assume that the target discretization is $\Delta\tau = 0.1$ and that we are studying a phase point which corresponds to the black cross in Fig. 3.13 a, where both metal and insulator are stable (for $\Delta\tau = 0.1$). If the convergence process is started using $\Delta\tau = 0.25$, the intermediate result is a metallic solution, since the insulator is, then, unstable. Obviously, subsequent iterations for smaller $\Delta\tau$ can only increase the precision of the metallic solution, but are unlikely to cause "jumps" towards the insulating solution. A second, comparatively minor aspect of the explanation is that Jarrell initially accelerates the convergence of the iteration scheme using overrelaxation. At least in regions of the phase space where the insulator is thermodynamically metastable, this might also render insulating solutions numerically unstable and, eventually, unobservable.

### 3.4.3 Estimation of Errors

Generally, QMC estimates for any observable have both statistical and systematic errors. Systematic errors which arise from using finite $\Delta\tau$ such as the remaining error in the Fourier transformation or the inevitable Trotter error will be explicitly considered in Sec. 3.5. More precisely, most results will first be presented for a range of finite $\Delta\tau$. Error bars will indicate the statistical errors as determined empirically from a set of measurements (taking autocorrelation into account). Only in a second step, conclusions are drawn for the physical limit $\Delta\tau = 0$.

The statistical error bars computed for finite $\Delta\tau$ are only a good approximation for the true error when other sources of errors are removed. One possible source is the thermalization of the auxiliary Ising field used for the MC solution of the impurity problem. In our code, this field is initialized randomly for each iteration. Generically, these initial configurations have a vanishingly small true probability of occurring in an infinitely long Monte Carlo run.[18] Consequently, they would be overweighted in any run of finite length. As a resolution, a certain number of Monte Carlo "warm-up" sweeps after initialization is discarded before starting measurements. Schlipf (1998) also introduced the possibility of using multiple walkers in each iteration, i.e., of averaging over completely independent solutions of the impurity problem (with identical $\mathcal{G}(\tau)$). Thereby, the result of an iteration is less influenced by a particularly "bad" initial Ising configuration. We typically used 2000-4000 warm-up sweeps and 6 or 8 walkers, in parallel runs up to 32 walkers (i.e., starting points). A study on the influence of the number of warm-up sweeps is shown in Fig. 3.14. Here, the double occupancy is computed for a relatively high temperature ($\beta = 10$) using between 0 and 10000 warm-up sweeps. Since the curves are practically on top of each other in the main panel, the differences with respect to the unweighted average of the results (for 2000, 3000, 5000, and 10000 warm-up sweeps) are plotted in the inset for a magnified view. Runs without any warm-up sweeps show a significant systematic error within the metallic phase; all other runs agree within error bars. Here, the

[18] Note that this effect is already strongly reduced by random initialization compared with using an ordered initial Ising field. After all, the usage of only the initial Ising fields would correspond to a simple Monte Carlo procedure without importance sampling.

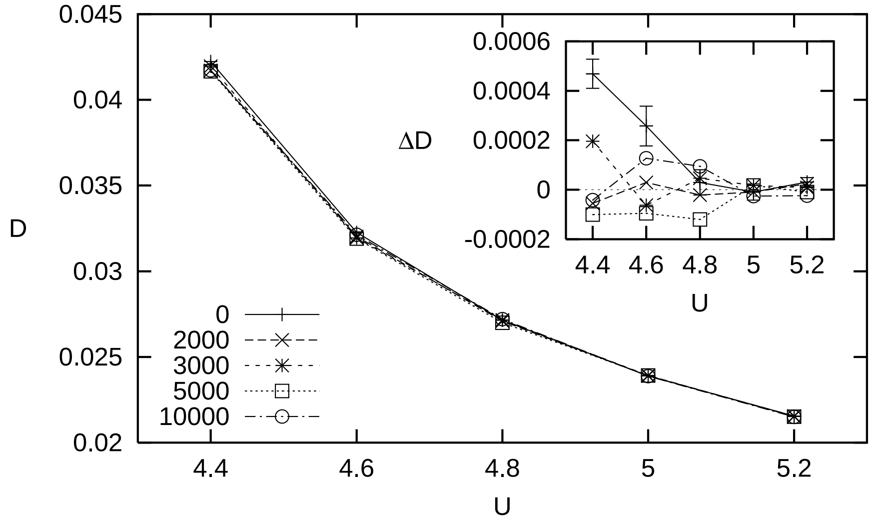



**Figure 3.14:** Double occupancy for $\beta = 10$ and $\Delta\tau = 0.125$: comparison of QMC runs with different number of warm-up sweeps (8 walkers with $n$ warm-ups plus 20000 measurements each; about 20 iterations). In the inset, differences to the average (of all results with $n > 0$) are shown. No systematic dependence on $n$, except for $n = 0$ in the metallic phase.

deviations arise both from statistical and convergence errors. Only the production runs (with 2000 warm-up sweeps; crosses) were long enough to reduce the total error far below $10^{-4}$.

The use of insufficiently converged solutions is potentially a very significant source of errors. It is important to realize that in principle measurements have to be performed exactly at the solution of the self-consistency equations, i.e., for the exact bath Green function. Averages over measurements performed for different impurity models corresponding to approximate solutions do not necessarily converge to the exact answer in the limit of an infinite number of models (i.e., iterations) and measurements.[19] Still, the most important practical point when computing observables is that only runs are included in averages which are close to the solution in comparison to the asymptotic statistical error. For higher precision, the number of sweeps (and not only the number of measurements) must be increased. In this work, we usually used between $10^5$ and $5 \cdot 10^5$ sweeps, in particular cases more than $10^6$ sweeps. If the convergence of solutions was unclear, we started new iteration chains; partially with the same initial self-energy (but different random numbers), partially using specially prepared self-energies.[20] The good convergence of our solutions will also become

[19]Trivially, a measurement of the free energy $F$ itself (using a suitable impurity solver) is a good example. Since $F$ is minimal for the true solution, all measurements taken for approximate solutions will be too large. The correct answer can, therefore, not be approached by averaging over many measurements, but only by reducing the deviations from the exact solution.

[20]We will prove in subsection 3.7.1 that the iteration process is convergent at least for the direct iteration scheme since (for an exact solution of the impurity problem) each step is downhill in the space of hybridization functions with the step length being proportional to the slope of the free

apparent from the high quality of the fits in the following sections.

### 3.4.4 Parallelization

A more technical part of this work was the parallelization of the single-band version of the QMC code. This effort was necessary in order to use all available computational resources efficiently and to reach the high accuracy necessary for computing the first-order metal-insulator phase transition line. We chose to implement only coarse-grained parallelism where the same impurity model is (approximately) solved independently in parallel processes. The self-consistency conditions (including the lattice Dyson equation) are evaluated by a master process after the Green functions obtained in each solution have been averaged. While this approach practically limits the number of processors that can be efficiently used to about 16 or 32 (otherwise the relative cost for warm-ups is too high), it has the advantage that the communication overhead is negligible and that its implementation using the Message Passing Interface (MPI) is straightforward. Note that it is a nontrivial task to ensure independent random-number streams for the parallel runs. We used the Scalable Parallel Random Number Generator (SPRNG) library, version 1.0 (Ceperley, Mascagni, Mitas, Saied, and Srinivasan, 1998) for this purpose.

In order for the new code to be executable in serial environments without retaining separate serial versions, we instrumented the program with compiler switches so that both MPI extensions and the random number generator (as well as the maximum value of $\Lambda$) can be chosen at compile time. While the code originally already ran on a range of architectures (Cray T90, IBM AIX, Fujitsu VPP), manual setup and knowledge of the suitable compiler options was required for each new compilation. In order not to further complicate matters with the inclusion of new architectures (Intel Linux, Cray T3E) in serial or parallel environments, we created makefiles and an autodetection script. This setup toolbox automatically executes the necessary compilers with appropriate options and includes libraries depending on the detected architecture. Furthermore, the whole code (including tools and required precompiled libraries) can be packed automatically into a single self-extracting file; this feature simplifies code distribution significantly. All of these improvements were later integrated into the multi-band versions of the QMC code by Held and Keller.

Ironically, the most demanding QMC calculations (for $\Lambda \gtrsim 300$) were run in serial mode. The reason is that problems of this size can only be treated efficiently on vector supercomputers with very fast memory access. In principle, the VPP as a vector parallel computer with clock speed memory access would have been ideal for the purpose. Despite an investment of some hundred hours of development time, however, we could not get the MPI version of our program to run on this machine. As the same code worked on all other architectures that we had tested it on, we suspect a bug in the MPI implementation for the VPP.

---

energy functional. Therefore, the iteration process cannot stick in unphysical regions; it always drifts towards a valid solution. Numerical problems in this respect can only arise near points in the extended $T, U, \Delta$ phase space where new minima of the free energy functional appear (as a functional of the hybridization function $\Delta$).

## 3.5 Results: Coexistence Region

In this section, we will begin the systematic construction of the phase diagram of the fully frustrated Hubbard model with semi-elliptic DOS in $d \to \infty$ at half filling. For this purpose, only those observables need to be considered which can be directly computed from standard QMC solutions of the DMFT self-consistency equations. As shown in subsection 3.5.1, the internal energy $E$ is very well suited for determining the metal-insulator transition in the physical limit $\Delta\tau \to 0$, much better than other observables considered so far. In subsection 3.5.2, we establish and parametrize universal, i.e., temperature-independent, properties of the insulating phase. Measurements of the energy $E$ presented in subsection 3.5.3 are used for the construction of a preliminary phase diagram including a crossover region at high temperatures, a critical end point, and a region of coexisting metallic and insulating solutions at low temperatures in subsection 3.5.4.

Measurements of the double occupancy $D$ and its (difficult) extrapolation for $\Delta\tau \to 0$ in the metallic phase are discussed in subsection 3.5.5. The very precise extrapolated QMC estimates for $E$ and $D$ obtained in this section will allow for an implicit comparison of the free energies of the metal and the insulator (within the coexistence region) and will lead to the full phase diagram in Sec. 3.6.

### 3.5.1 Choice of Observables and Extrapolation

In this subsection, we determine which of the observables introduced in subsection 3.2.3 are best suited for indicating and characterizing the metal-insulator transition. Using QMC results for the relatively high temperature $T = 1/15 \approx 0.067$ as a test case, we discuss in particular the difficulties of extrapolations to the physical limit $\Delta\tau = 0$.

The double occupancy for $T = 0.067$ is shown in Fig. 3.15 for a range of values of the discretization parameter $\Delta\tau$. Here, the symbols, which reflect QMC measurements, have been connected by straight line segments in order to guide the eye. As expected on physical grounds, the double occupancy decreases significantly with increasing $U$, i.e., with decreasing metallicity. Evidently, the slope $|\mathrm{d}D/\,\mathrm{d}U|$ reaches a maximum at $4.55 \lesssim U \lesssim 4.65$ (depending on $\Delta\tau$). One would be hard pressed, however, to decide if there is a continuous crossover or a sharp transition, even for finite $\Delta\tau$. Furthermore, the dependence of $D$ on $\Delta\tau$ is seen to be very irregular: coming from the metallic side the $\Delta\tau$ error increases up to $U \approx 4.6$ while it is invisible on the insulating side.[21] One might suspect that the primary effect of a finite $\Delta\tau$ error is a horizontal shift of the curves towards larger values of $U$. Still, results of an extrapolation to $\Delta\tau = 0$ would strongly depend on the assumed (nontrivial) functional forms and, thus, be unreliable without additional input.

A similar picture emerges for the quasiparticle weight $Z$ as defined in (3.6) and shown in Fig. 3.16. This observable is in principle a very good candidate for characterizing the metal-insulator transition since, within the Fermi liquid picture, the

[21] The absence of significant $\Delta\tau$ dependence for the QMC estimate of $D$ in the insulator may be specific to the particular algorithms used in our code.

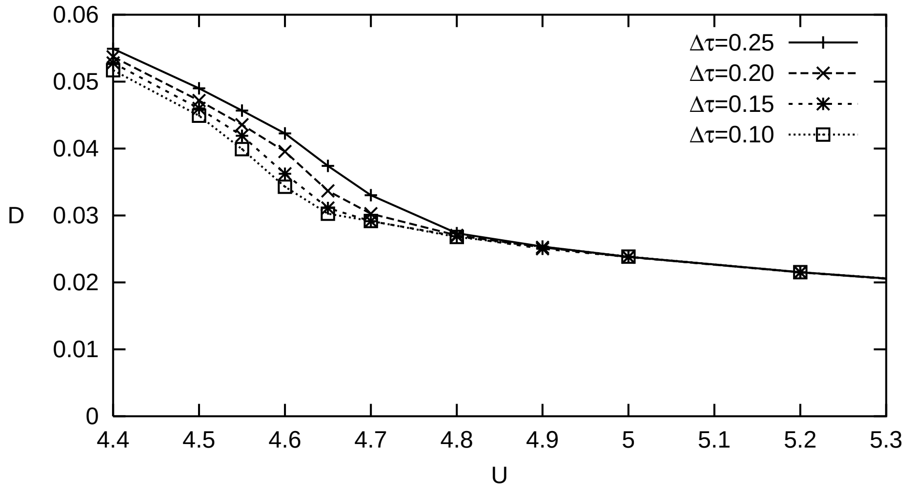



**Figure 3.15:** Double occupancy for $T = 1/15$ as a function of interaction $U$ for various values of the imaginary time discretization parameter $\Delta\tau$. Lines are guides to the eye only.

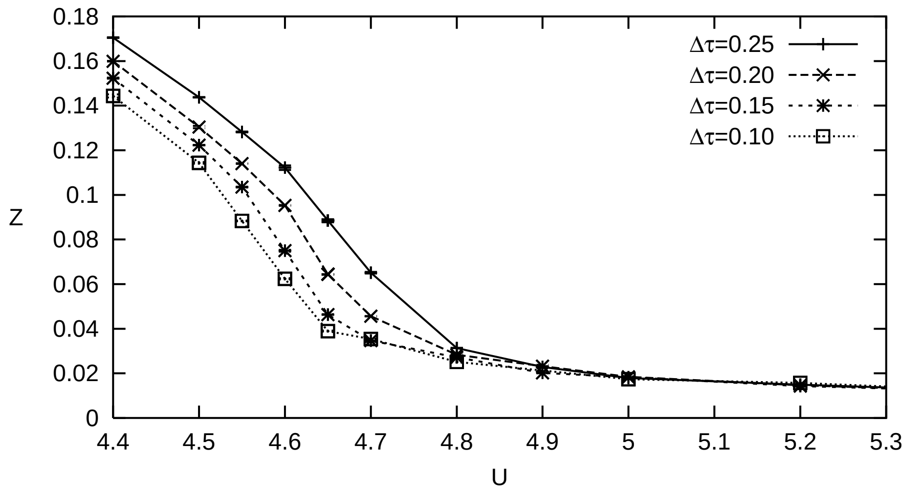



**Figure 3.16:** Discrete estimate for the quasiparticle weight $Z$ for $T = 1/15$. Lines are guides to the eye only.

effective mass in the metal diverges as $m^*/m = 1/Z$ for $Z \to 0$ so that $Z = 0$ defines the MIT. As discussed in subsection 3.2.3, however, $Z$ cannot be associated with a physical quasiparticle weight at finite temperatures very near to or beyond the MIT. In fact, it is seen in Fig. 3.16 that the (discrete) QMC estimates for $Z$ remain finite even for $U \gtrsim 5$. Again, the $\Delta\tau$ error is quite irregular: it is significant only on the metallic side and maximal in the transition region $U \approx 4.6$.

The potential energy $E_{\mathrm{pot}} = UD$ and the kinetic energy $E_{\mathrm{kin}}$ are presented in Fig. 3.17 a and Fig. 3.17 b, respectively. Obviously, our above remarks on the general shape and on the $\Delta\tau$ error of $D$ directly carry over to $E_{\mathrm{pot}}$. In comparison, the

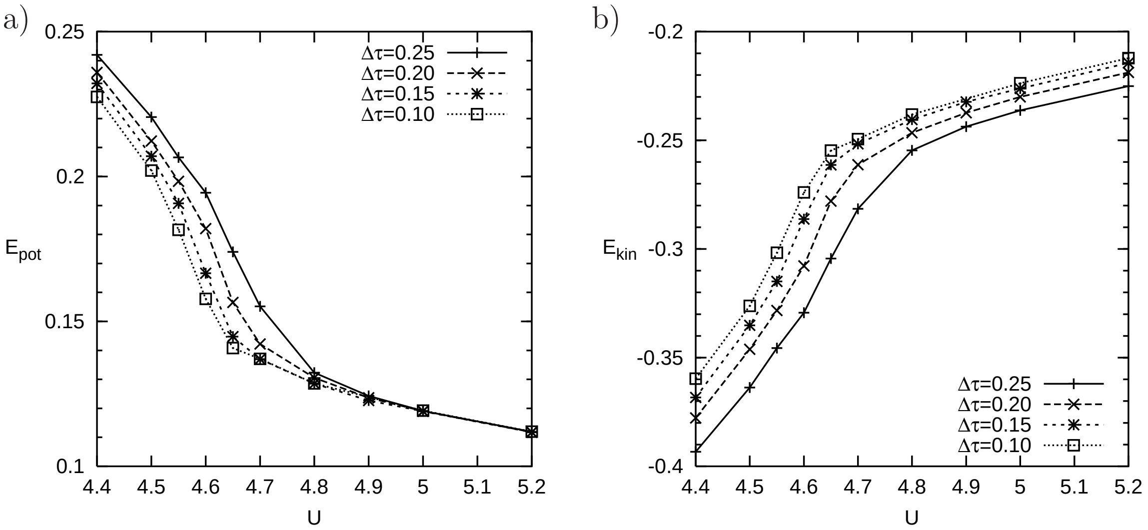



**Figure 3.17: a)** Potential energy, **b)** kinetic energy for $T = 1/15$. Lines are guides to the eye.

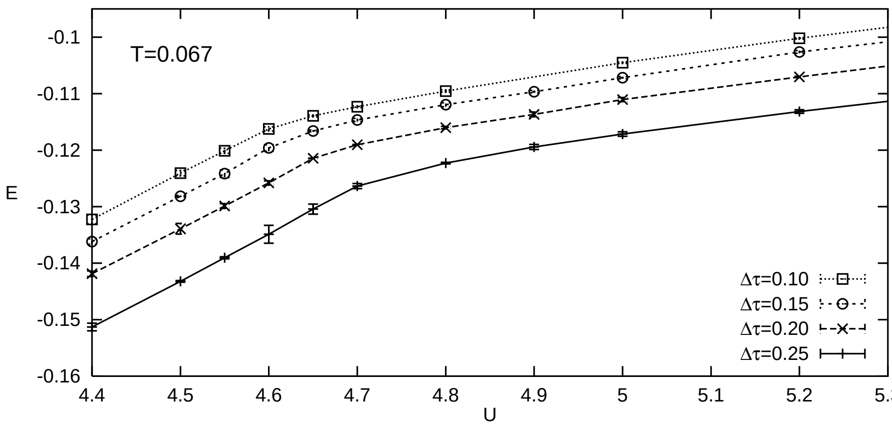



**Figure 3.18:** Total internal energy for $T = 1/15$. Lines are guides to the eye only. The kink extrapolating to $U \approx 4.6$ for $\Delta\tau \to 0$ indicates a transition or narrow crossover region.

---

kinetic energy shows a more pronounced kink near $U \approx 4.6$, in particular for small $\Delta\tau$. The $\Delta\tau$ error is still maximal in this region, but significant for all values of $U$. Since the $\Delta\tau$ dependence has opposite signs for $E_{\text{pot}}$ and $E_{\text{kin}}$, the total internal energy $E = E_{\text{pot}} + E_{\text{kin}}$ will have a *smaller* $\Delta\tau$ error than its constituents. Note also that the curvature has opposite sign for $E_{\text{pot}}$ and $E_{\text{kin}}$, both in the metallic and in the insulating region. Thus, some cancellation will occur so that $E(U)$ is also less curved than its constituents which may help in pinpointing a transition point.

This is indeed seen in Fig. 3.18. For each value of $\Delta\tau$, $E(U)$ is linear within (statistical) error bars on the metallic side up to some transition point which decreases from roughly $U \approx 4.7$ for $\Delta\tau = 0.25$ to $U \approx 4.6$ for $\Delta\tau = 0.1$. On the insulating side, the curves are again almost linear with significantly reduced slope; here, however, a small curvature remains visible. The comparatively sharp kinks unambiguously indicate the positions of phase transitions (or narrow crossover regions) for each value of the time discretization $\Delta\tau$. A detailed analysis using suitable fit functions will be presented in the next subsection. Furthermore, the dependence of the curves on $\Delta\tau$ is very regular both on the metallic and on the insulating side. While we will look at the precise extrapolation laws below, it is apparent from the figure that the $\Delta\tau$ error is about a factor of 3/2 larger in the metallic phase. This explains the irregularity of the $\Delta\tau$ dependence within the transition region where the phase transition (or crossover) line is passed upon varying $\Delta\tau$ for constant $U$. As long as this problematic region is excluded, very precise extrapolations $\Delta\tau \to 0$ can be expected. Finally, the regular shape of each curve also significantly reduces the convergence error since measurements obtained from unconverged solutions can be detected as deviations from the regular pattern and then corrected by performing additional simulations.

If high-frequency tail corrections in the energy estimators are properly taken into account (cf. subsection 3.2.3), low-frequency errors associated with the discrete Fourier transforms (cf. subsection 3.4.1) are avoided, and no phase transition lines are crossed, the error in the estimate for $E$ is dominated by the Trotter error (cf. Sec. 1.3) which is $\mathcal{O}(\Delta\tau^2)$. Therefore, a reasonable starting point for a quantitative analysis is a purely quadratic extrapolation $\Delta\tau \to 0$ (i.e., a linear least-squares fit in $\Delta\tau^2$), both on the metallic and on the insulating side. The intersection of linear fits for $E_{\mathrm{m}}(U)$ and $E_{\mathrm{i}}(U)$ of the metallic and insulating solution then determines a first estimate of the transition point.

We found that the accuracy both in the estimates for energies and for transition points could be improved by an order of magnitude by extending the least squares fits to fourth order in $\Delta\tau$ (i.e., by using fits of the form $E_U(\Delta\tau) = E_U(0) + c_2\Delta\tau^2 + c_4\Delta\tau^4$) and by extracting a more complicated, more accurate, and thermodynamically consistent expression for the energy in the insulator.

### 3.5.2 Properties of the Insulating Phase

Leaving properties of the metallic phase for the next subsections, we will now take a detailed look at the insulating phase. The symbols in Fig. 3.19 represent all of our QMC measurements of $E$ inside the insulating phase for various temperatures $T$ and interactions $U$ as a function of $\Delta\tau^2$. The most important observation is the absence of any significant dependence of $E$ on $T$. Such universal behavior is expected deep inside the insulating low-temperature phase where a well-developed gap in the spectrum exists at the Fermi energy (cf. Sec. 3.8) and where thermal excitations are suppressed exponentially. Evidently, the specific heat is already reduced by several orders of magnitude compared to the metallic case in the low-temperature region of interest very close above the MIT as soon as a pseudo-gap appears. One practical consequence is that for fixed interaction $U$ all QMC data points may be included in a least squares fit for the extrapolation $\Delta\tau \to 0$, irrespective of the associated temperature $T$. The solid lines in Fig. 3.19 reflect such extrapolations; here all three

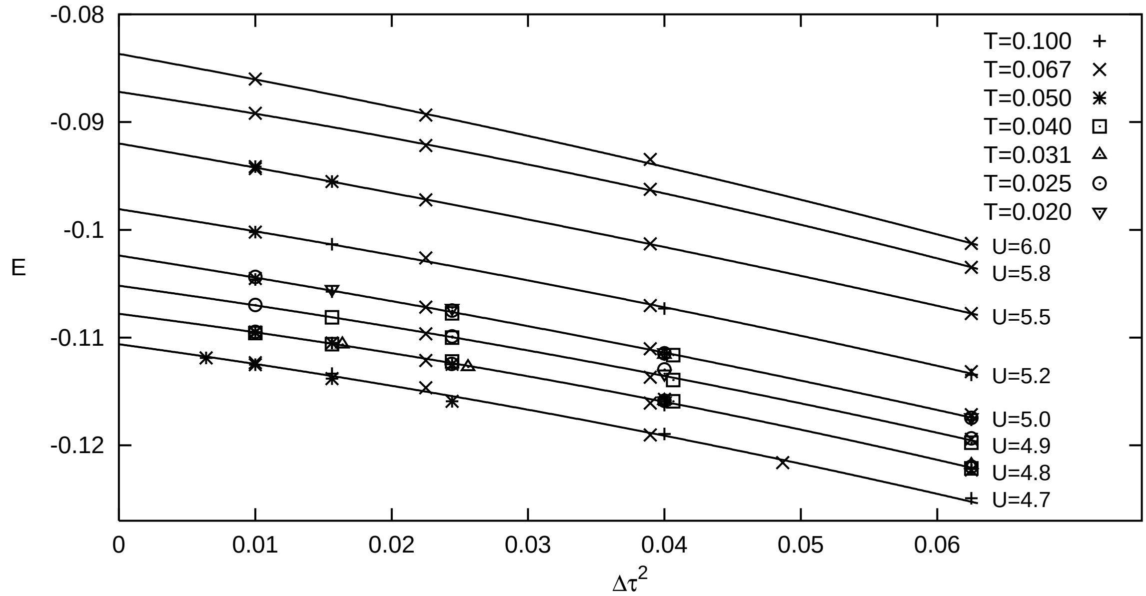



**Figure 3.19:** Total internal energy in insulating phase vs. $(\Delta\tau)^2$. Since the QMC estimates (symbols) do not significantly depend on the temperature $T$ for fixed interaction $U$, they can be collectively extrapolated to $\Delta\tau \to 0$ in quadratic least squares fits in $(\Delta\tau)^2$ (solid lines) for each value of $U$.

---

coefficients in the quadratic fit (in $\Delta\tau^2$) were taken as free parameters. Due to the small curvature, the extrapolations are well conditioned.

The extrapolated values for $E$ are shown as large circles in the lower part of Fig. 3.20 as a function of $U$. The small colored symbols in the same plot represent a scaled version of the measurements shown in Fig. 3.19 according to the “average” extrapolation law

$$E_U(\Delta\tau) = E_U(0) - 0.036\, U\Delta\tau^2 - 0.9\,\Delta\tau^4 \quad \text{(insulator)}\,, \tag{3.17}$$

which was extracted from the data shown in Fig. 3.19. The solid line in the lower part of Fig. 3.20 indicates our best fit for the extrapolated data,

$$E_{\mathrm{i}}(U) = -\frac{0.515}{U} - \frac{0.0027}{U-3.95} - 0.0071 + 0.00325\,U - 0.00025\,U^2\,, \tag{3.18}$$

where the index i stands for “insulator”. It is evident that this particular form was not chosen for elegance. Instead, we sought a curve that is smooth both within and near the parameter range $U = 4.7 \ldots 6.0$, is easily differentiable analytically, and accurately represents our extrapolated QMC data. As a slight bias we also tried to produce a reasonable approximation to data for the crossover region ($T = 0.1$, $U \lesssim 4.5$). The success is evident from the upper part of Fig. 3.20 where only the difference of the extrapolated or rescaled data to the fit function is plotted. Except for very small $U$ [where the scaling function (3.17) for the dependence on $\Delta\tau$ is no longer expected to hold], the vast majority of scaled data points deviates by less

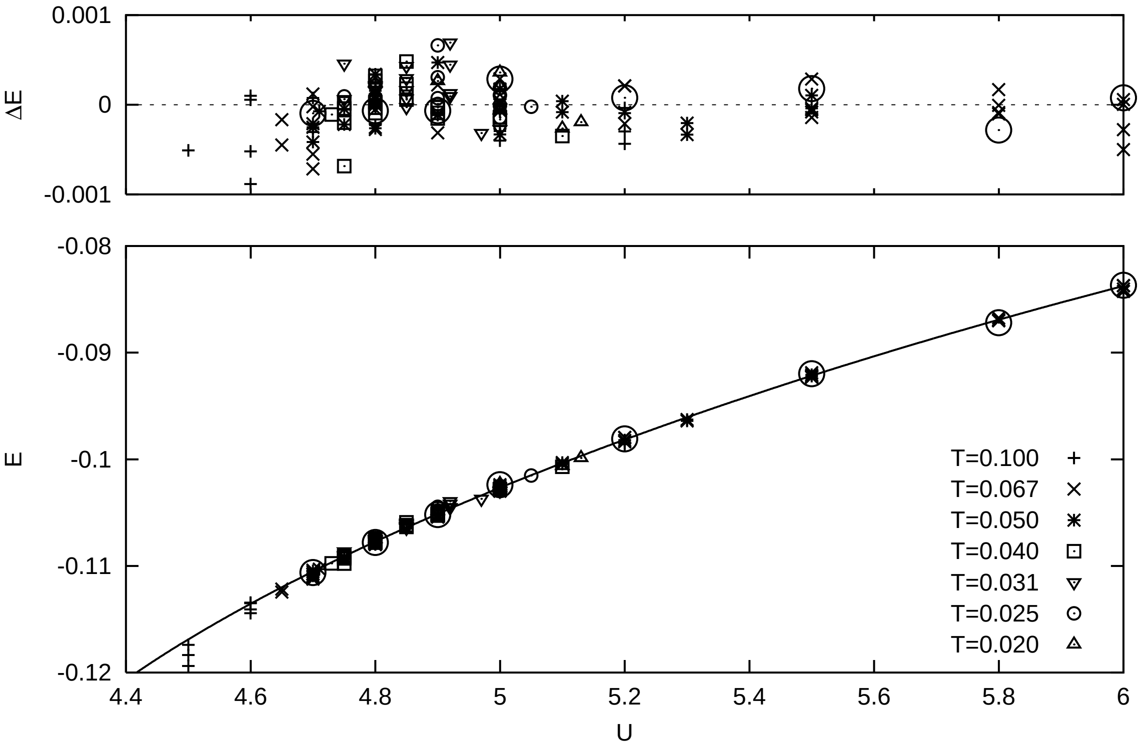



**Figure 3.20:** Total energy in insulating phase vs. $U$. In the lower plot, the large circles denote the results of the extrapolations shown in Fig. 3.19. The small symbols correspond to the rescaled QMC data, according to the extrapolation law (3.17) for $\Delta\tau \to 0$. The solid line indicates the ($T$-independent) fit function (3.18) which is consistent with the solid line for $D$ shown in Fig. 3.21. The upper part gives a magnified view on the (small) discrepancies between data and fit curve.

than about $4 \cdot 10^{-4}$. The deviations of the extrapolated data (large circles) are even smaller. In fact, we can trace the outliers at $U = 5.0$ and $U = 5.8$ to statistical deviations in the extrapolation process. This explanation is supported by the fact that the scaled data points [using the extrapolation law (3.17)] are much closer to the fit than the result of the free extrapolation. In summary, we expect the error of our fit (3.18) to be close to $10^{-4}$ at least for $4.75 \le U \le 6.0$.[22]

Such a high accuracy seems particularly impressive since the $T$-independent energy in the low-temperature insulating phase is a ground state property. This fact also allows for a numerical check since, at $T = 0$, the internal energy and the double occupancy are simply related: $D(T, U = 0) = \mathrm{d}E(T, U = 0)/\mathrm{d}U$. Therefore, the derivative of our fit (3.18) for $E(U)$ should represent the QMC data for $D(U)$. This

[22]The negative deviations for $T = 0.05$, $U = 4.7$ and for $U < 4.7$ are related to the proximity of the phase boundaries. It is unclear if the effects are fully physical; most probably, at least some part arises from a bias in the statistical error which may be significant for any almost unstable solution. On the other hand, the finite heat capacity implied by negative deviations with lower temperatures is also plausible on physical grounds. In any case, the effects are too small to affect our numerical results.

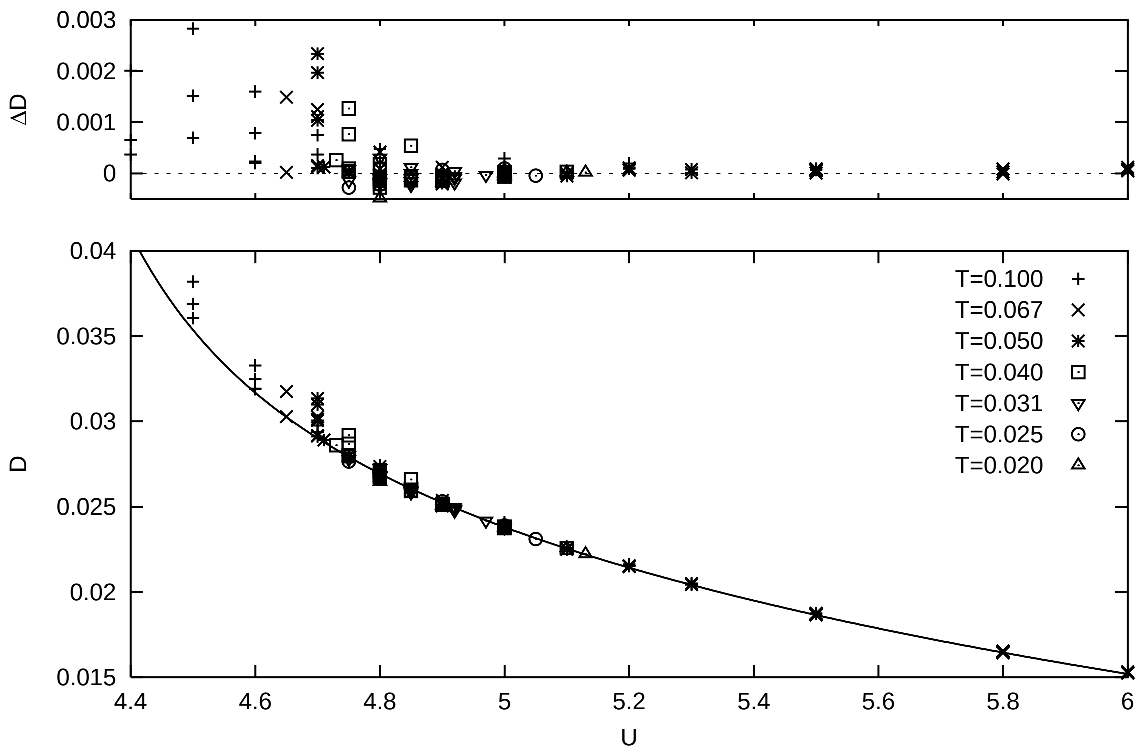



**Figure 3.21:** Double occupancy in insulating phase vs. $U$. The symbols in the lower part represent the QMC data (without extrapolation $\Delta\tau \to 0$). The fit curve (solid line), given by (3.19), is thermodynamically consistent with the fit for $E$ shown in Fig. 3.20. The upper part shows the deviations of the QMC data from the fit curve which is significant only for small $U$ [i.e., for $U \to U_{c1}(T, \Delta\tau)$] and dependent on $\Delta\tau$ (not encoded in the figure).

can be easily verified since the QMC measurements for $D$ are practically free of $\Delta\tau$ errors in the insulating phase so that no extrapolation $\Delta\tau \to 0$ or scaling is necessary. The "raw" QMC estimates are shown in the lower part of Fig. 3.21; here, the solid line is given by the function

$$D_{\rm i}(U) = \frac{0.515}{U^2} + \frac{0.0027}{(U-3.95)^2} + 0.00325 - 0.0005\, U\,. \tag{3.19}$$

Again, the deviations of the measurements from the fitting function are magnified in the upper part of the figure. The overall agreement is very good; the deviations at small $U$ indicate a strongly increased influence of the $\Delta\tau$ error near the phase boundary at not too low temperatures (see subsection 3.5.5).

As a result of this subsection, we have identified the internal energy $E$ as an appropriate observable for a QMC study of the metal-insulator transition. In addition, we have established functional expressions for $E$ and $D$ within the insulating phase. These will not only be used for the determination of the "coexistence" phase diagram in the next subsection, but also help determine the low-temperature behavior of the thermodynamic phase transition line in Sec. 3.6.

### 3.5.3 Internal Energy

In the following, we present QMC data for the internal energy $E$, fits for finite $\Delta\tau$, and extrapolations to the physical limit $\Delta\tau = 0$, ordered by decreasing temperature. The extrapolated data is summarized at the end of this subsection; the resulting phase diagram will be shown in the next subsection 3.5.4.

Since the insulating side has already been analyzed in the previous subsection, the remaining difficulty is the treatment of the metallic phase and the determination of the phase boundaries. Within the metallic phase, we use quadratic least-squares fits in $\Delta\tau^2$ for fixed $U$. Due to the small number of data points, the error in the quadratic coefficient (i.e., the $\Delta\tau^4$ term) is, however, of the same order of magnitude as the estimate itself. Furthermore, the estimate may shift depending on whether or not data points with small $\Delta\tau$ are included. In order to minimize the resulting errors, we have assumed slow variation of this coefficient as a function of $U$ and $T$ and have performed a global fit. The result is that the $\Delta\tau^4$ coefficient for the metal is always close to the value 0.95 computed for the insulating phase. The observation of a slight variation as a function of $T$ lead to a linear interpolation with a value of 1.2 for $T = 0.1$ and 0.8 for $T = 0.02$.[23] This coefficient was then held fixed for all extrapolations within the metallic phase.

The results of such extrapolations to $\Delta\tau = 0$ at $T = 0.1$, the highest temperature considered here, are denoted as solid circles in Fig. 3.22. The $\Delta\tau = 0$ values for $U = 4.1$, $U = 4.2$, and $U = 4.3$ fall on a straight line which extends to $U = 4.4$ and is partially drawn as a thick and partially as a thin line. The parallel lines connecting data points for finite $\Delta\tau$ are offset according to a scaling law which not only includes the globally determined $\Delta\tau^4$ coefficient, but also a $\Delta\tau^2$ coefficient which has been averaged from the extrapolations for $U = 4.1$, $U = 4.2$, and $U = 4.3$. Each of these "metallic" lines crosses the corresponding "insulating" line (beginning as thick curve at large $U$ and extending into a thin line towards the metallic phase) at some point. In this region, however, the QMC data deviates systematically from the fit curves, indicating a smooth crossover rather than a phase transition.[24]

An analogous plot for $T = 0.067$ (more accurately: $\beta = 15$, i.e., $T = 1/15$) is shown in Fig. 3.23.[25] Here, the transition is (almost) sharp within the numerical accuracy, i.e., the data follows linear fits in the metallic phase and the universal form (3.18) in connection with (3.17) in the insulating phase, respectively, for all values of the interaction $U$. Consequently, the transition points can be determined and extrapolated very accurately: $U_c = 4.59$. Only extremely close to the transition points, i.e., for $U = 4.65$ at $\Delta\tau = 0.15$ and $U = 4.7$ at $\Delta\tau = 0.2$, do slight deviations of the QMC data from the fits indicate some rounding effect. This observation, in combination with conclusions from lower temperatures, will later lead to the conclusion that $T = 0.067$ is close to, but slightly above, the critical temperature.

[23] Note that while the relative error for the coefficient is relatively large, its influence is small. The intrinsic uncertainty of about 0.2 translates, e.g., into a possible additional error in the energy estimates of $2 \cdot 10^{-5}$ for $\Delta\tau = 0.1$ and $3 \cdot 10^{-4}$ for $\Delta\tau = 0.2$.

[24] As the precise behavior of $E$ in the crossover region will be irrelevant in the following, we only sketched the assumed form by manually adapting cubic splines (thick lines).

[25] Compare also the data for this temperature shown in the previous subsection.

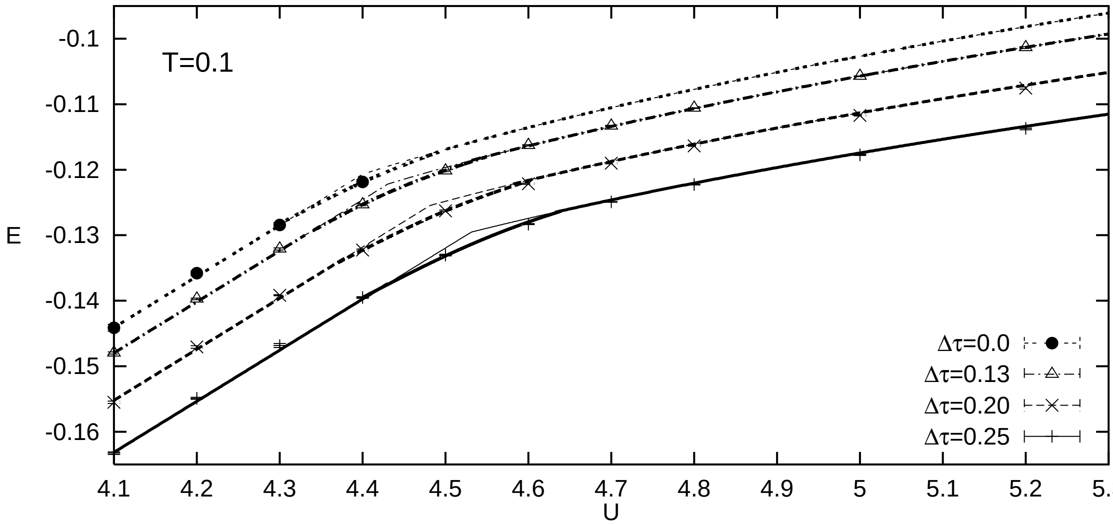



**Figure 3.22:** Total energy for $T = 1/10$: the QMC data (open symbols) and its extrapolation to $\Delta\tau = 0.0$ (solid circles) is well represented by linear fits (thin lines, partially coinciding with thick lines) for $U \lesssim 4.3$ (and beyond, depending on $\Delta\tau$). For $U \gtrsim 4.6$, the data agrees with fit curves for the insulator given by (3.18) and (3.17) (thin/thick lines). Deviations at intermediate $U$ (thick line segments below kinks of thin lines) indicate crossover regions.

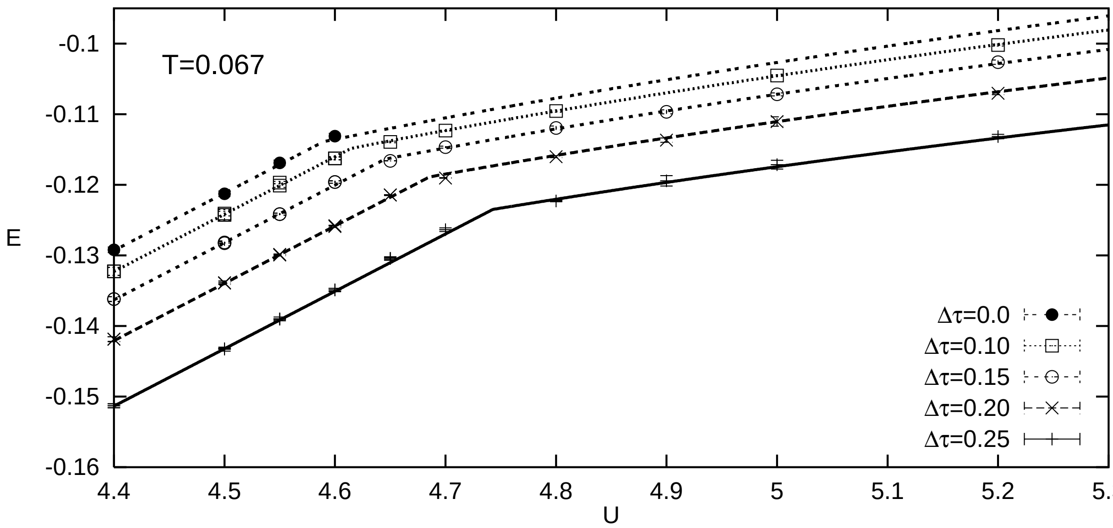



**Figure 3.23:** Total energy for $T = 1/15$: the QMC data (open symbols) and its extrapolation to $\Delta\tau = 0.0$ (solid circles) is well represented by linear fits on the metallic side and by (3.18) and (3.17) on the insulating side except for data right at the kinks. This indicates a very narrow crossover close to a sharp transition.

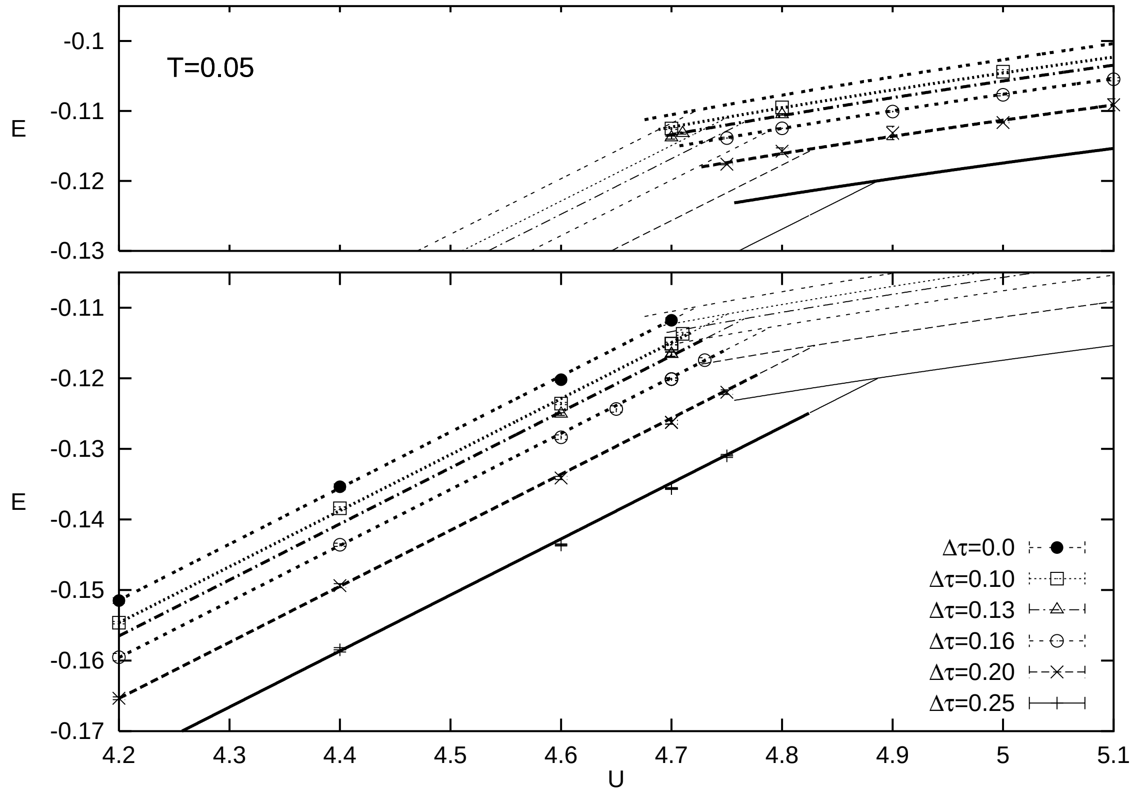



**Figure 3.24:** Total energy for $T = 1/20$: in order to avoid confusion within the coexistence region, QMC data for the metallic phase is presented in the lower plot and the data for the insulator in the upper plot (with identical scales). As in the previous plots, the fit functions for the insulating phase (thick lines in upper plot) have the universal form given by (3.18) and (3.17); their extent reflects our estimate of its stability region. In the lower plot, the linear fits for the metal are drawn as thick lines below $U_{c2}$, where the metallic phase is stable, but extended as thin lines up to $U_{c2b}$, where they cross the corresponding fit for the insulator.

In contrast, a small coexistence region of metallic and insulating solution appears for $T = 0.05$ as shown in Fig. 3.24. Here, we have split the presentation of the QMC data for the metallic phase (lower part) and insulating phase (upper part). The fit curves are repeated in both parts with emphasis (thick lines) on the phase under consideration. While the extent of the fit curves for the insulator reflects our estimate of the extent of its stability region (which depends on $\Delta\tau$), the metallic curves are always drawn until they cross the corresponding insulating solutions. These crossing points, which can be determined with great precision, will later be used as parameters for fitting the double occupancy. The metallic curves are drawn as thick lines within (our estimate of) their stability range. Note that the QMC data generally agrees well with the fit curves; only right at the stability edges, systematic deviations towards the more stable phase are seen. While it is plausible that this effect might be partially genuine, we observed that it can be reduced by improving the accuracy of

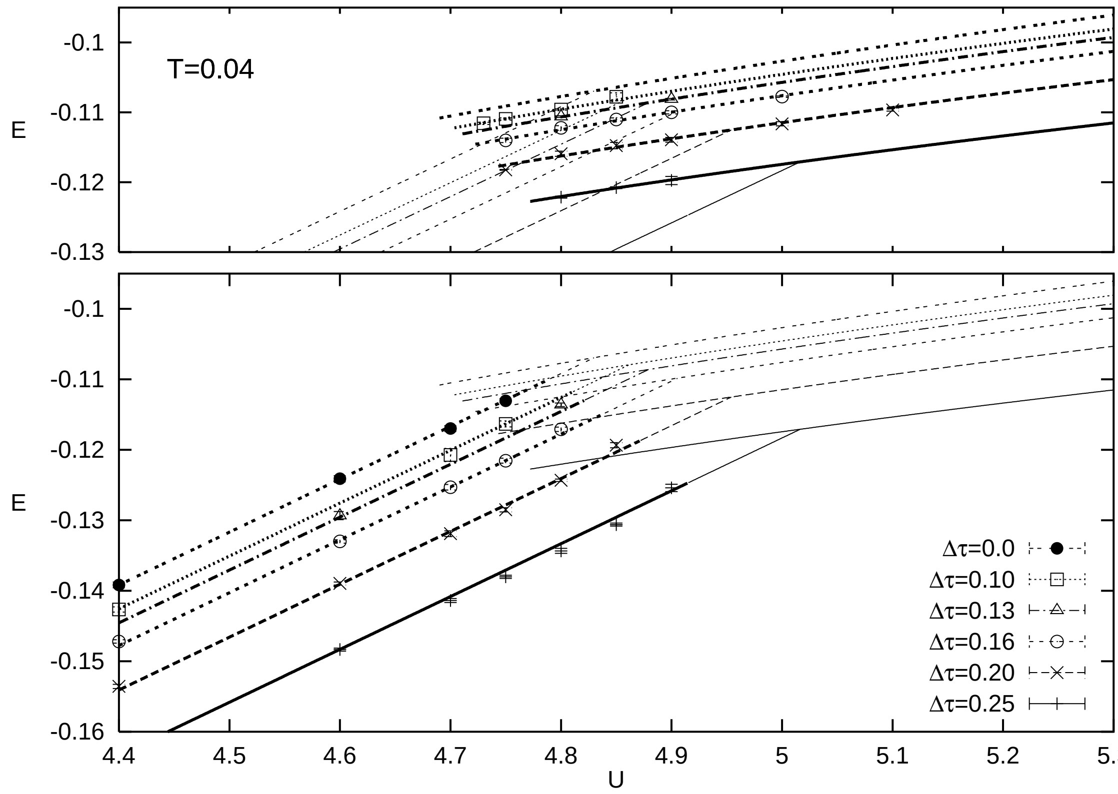



**Figure 3.25:** Total energy for $T = 1/25$, analogous to Fig. 3.24.

the QMC solution of the impurity model, i.e., by using more sweeps. Therefore, we attribute most of the visible deviations between the QMC data and the fit functions to remaining numerical errors.

For $T = 0.04$, the coexistence region is already much wider (in comparison with $T = 0.05$) as is apparent from Fig. 3.25. The linear dependence of $E$ on $U$ in the metallic phase clearly extends far into the coexistence region. The data points shown at the edges of the stability regions were only computed with high enough accuracy to verify their stability. Since our top priority was the accurate determination of the first-order phase transition line (which is not directly influenced by this data), additional available numerical resources were only spent for a selected set of temperatures.

The temperature $T = 0.031$ (i.e., $\beta = 32$) is already low enough that the extent of the coexistence region is large even on our usual grid with steps of $\Delta U = 0.1$ or $\Delta U = 0.05$. On the other hand, the temperature is still high enough that (on vector machines and supercomputers), a larger number of simulations could be run even for the smallest discretization regularly used in our study, $\Delta\tau = 0.1$. Therefore, we have chosen this temperature for a careful analysis within and near the edges of the stability regions of both metal and insulator. The impressive agreement with our fit curves (which are fixed for the insulator and adapted only for the metal) is clearly seen in Fig. 3.26. The data shown here in connection with data for parameter points

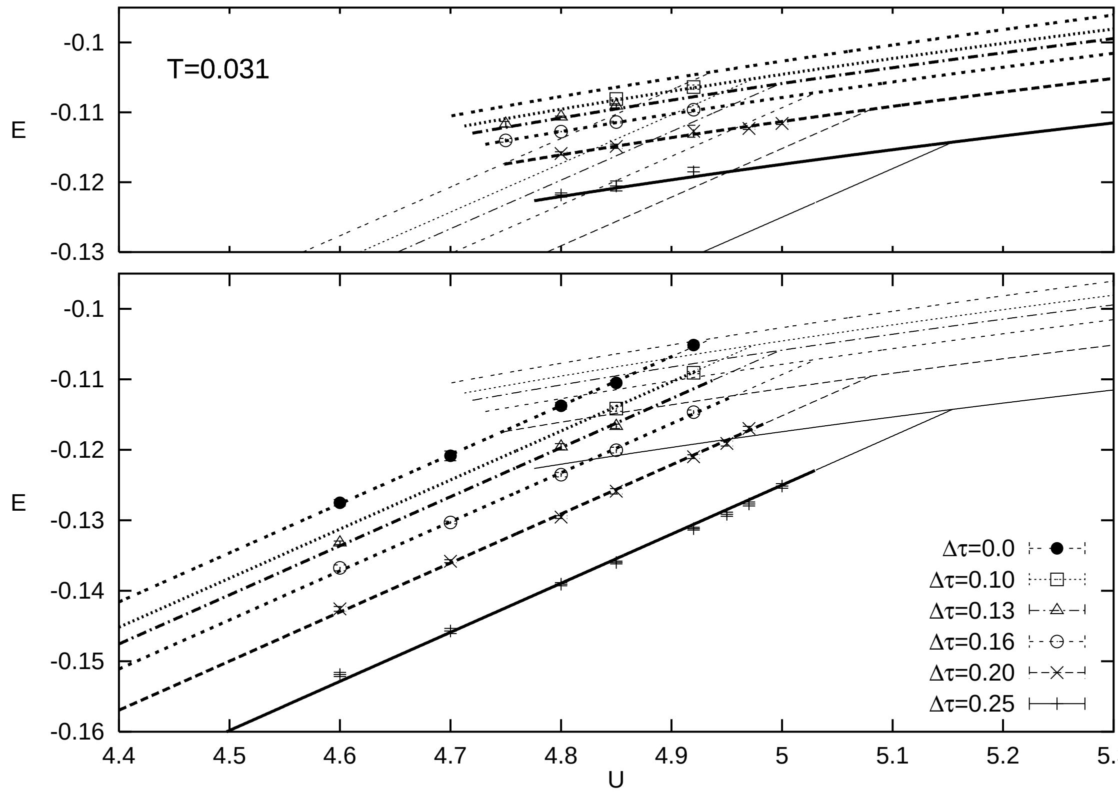



**Figure 3.26:** Total energy for $T = 1/32$, analogous to Fig. 3.24. The QMC data, here with increased precision near the edges of the coexistence region, is in very good agreement with the fit curves.

where metallic or insulating solutions were found to be unstable (not shown) was not only used for determining the stability edges for this temperature, but also for establishing the extrapolation laws (for $\Delta\tau \to 0$) for the other temperatures.

Note that the metal remains stable and closely follows the assumed linear dependence of $U$ until its energy almost reaches the energy of the insulating solution at $U_{c2}$. The solutions are very different energetically, however, when the insulator becomes unstable at $U_{c1}$. Qualitatively, the same behavior is also demonstrated for $T = 0.025$ in Fig. 3.27 and for $T = 0.02$ in Fig. 3.28. As we will see in subsection 3.5.5, the double occupancy of the metal also approaches that of the insulator near $U_{c2}$. Such behavior is not generically expected for first-order transitions: since the order parameter is necessarily different between coexistent solutions even at the stability edge of one solution, most properties (not directly related to the free energy) will usually also be quite different. On the other hand, the observed behavior is clearly compatible with a Landau theory as has been illustrated in Fig. 3.4 b. More light can be shed on the observed behavior by considering the qualitative physics of the $d = \infty$ Hubbard model near the MIT. Since the formation of quasiparticles lowers the total energy and since the space of solutions in the homogeneous phase may be regarded

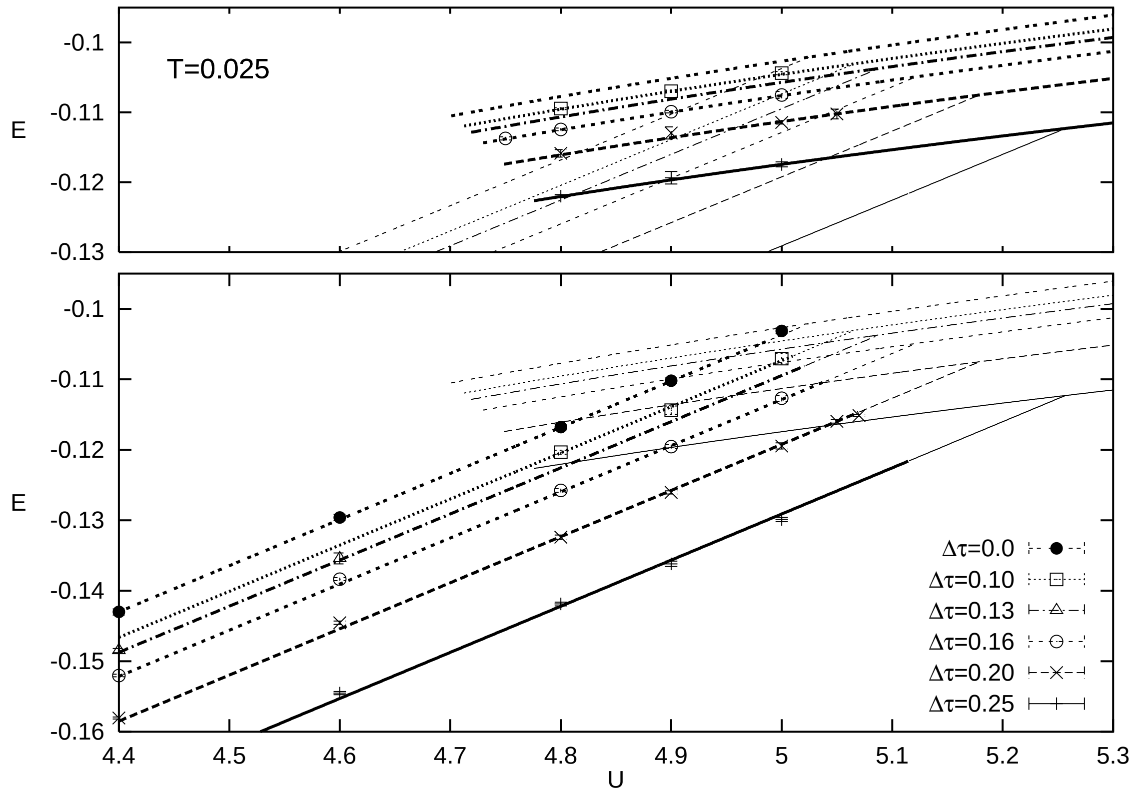



**Figure 3.27:** Total energy for $T = 1/40$, analogous to Fig. 3.24.

as predominantly one-dimensional,[26] the energy of the metallic phase is expected to be always lower than that of the insulator and that no crossing occurs in any relevant property. We will later make use of the fact established by Moeller et al. (1995) that, at $T = 0$, the metal approaches the insulator continuously. In this special ground state situation, both the energy (as always) and its derivative $D$ vary smoothly at $U_{c2}$. Approaching the MIT from the metallic side, the Fermi-liquid quasiparticle peak decreases in width (and weight) until it vanishes smoothly at the MIT. Right at the MIT, the difference between metallic and insulating spectrum is only a peak of zero width exactly at the Fermi energy. Since this difference is of measure zero, all observables that can be expressed via integrals over the spectral function must agree. Still, the presence or absence of the quasiparticle peak is a qualitative difference and associated with the fact that the entropy vanishes in the metal, but jumps to a finite value of $\ln 2$ at the MIT (Georges et al., 1996; Gebhard, 1997; Kalinowski and Gebhard, 2002). Thus, our observation that the metallic phase closely approaches the insulating phase in terms of energy and other observables is consistent with general properties of the model, in particular in the low-temperature limit $T \to 0$.

[26] If this remark could be made fully quantitative, we would have found the order parameter. In a qualitative sense, it is based on the smoothness and the observed behavior of $G(\tau)$ and $\operatorname{Im}\Sigma(i\omega_n)$: Solutions (for fixed $U$) are essentially characterized by $G(\beta/2)$ or $\operatorname{Im}\Sigma(i\pi T)$ (due to the smoothness and since the asymptotic behavior for $\tau \to 0$ and $|\omega_n| \to \infty$ is determined by $U$ alone).

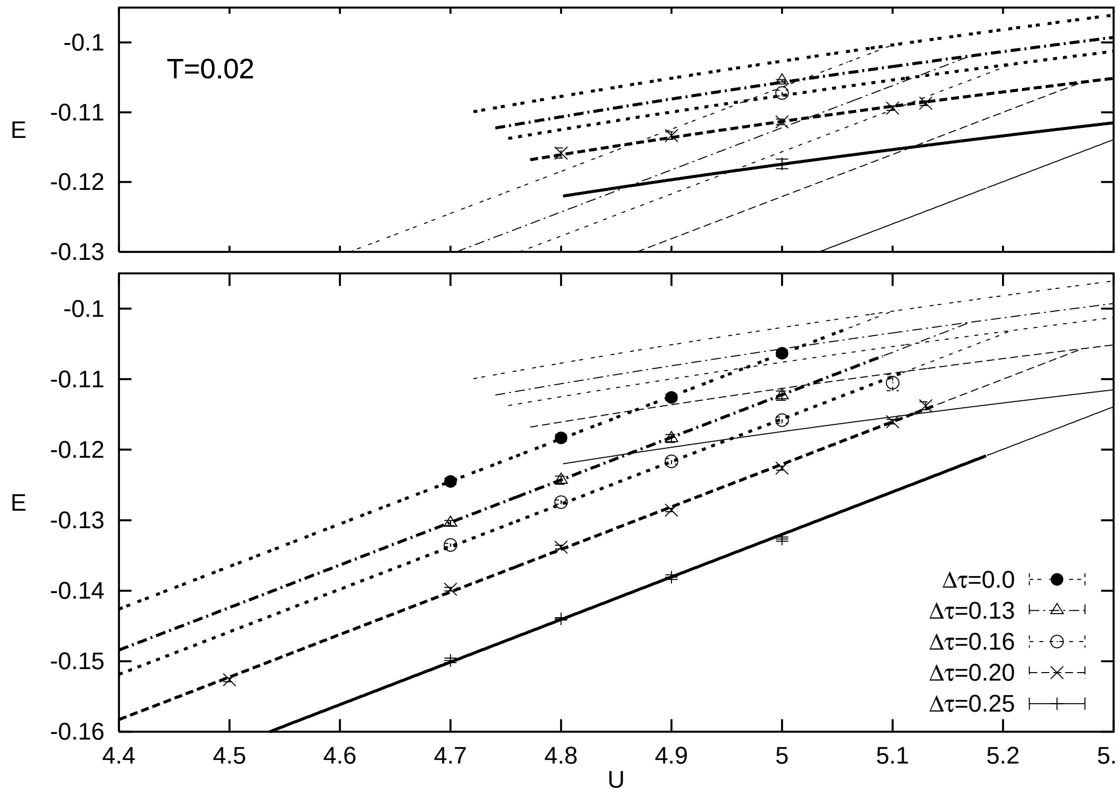



**Figure 3.28:** Total energy for $T = 1/50$, analogous to Fig. 3.24.

Figure 3.27 shows results for $T = 0.025$ and includes data (for $\Delta\tau = 0.1$) with $L = 400$ time slices which was the maximum used in this study and is significantly larger than values previously reported in the literature. Again, the agreement with the fit curves is seen to be very good. Note that $U_{c1}$, the stability edge for the insulator, agrees with the estimate obtained for $T = 1/32$. In contrast, lowering $T$ significantly reduces the energy of the metallic phase as the specific heat is then very significant. Therefore, $U_{c2}$, the stability edge of the metal, and $U_{c2b}$, the point where the extrapolated energies of metallic and insulating phases cross, further increase upon lowering $T$.

These trends continue for $T = 1/50$ as seen in Fig. 3.28. At this temperature, no simulations can be performed at $\Delta\tau = 0.1$; instead the maximal number of time slices is limited[27] to $\Lambda = 400$ leading to $\Delta\tau \geq 0.125 \approx 0.13$. Since the temperature-independence of the energy in the insulating phase has already been established in subsection 3.5.2, the number of data points in this phase is kept to a minimum and the numerical effort is concentrated instead on the metallic phase within the coexistence

[27]Note that the scaling of the computational effort as $\Lambda^3$ established in Sec. 1.3 is only valid for supercomputers where the memory system works at cpu speed. For regular workstations, the scaling at large $\Lambda$ is closer to $\Lambda^4$ so that the computational cost for $\Lambda = 400$ is already more than two orders of magnitude higher than for $\Lambda = 100$, e.g., for $\Delta\tau = 0.1$ at $T = 0.1$. See also subsection 1.3.2.

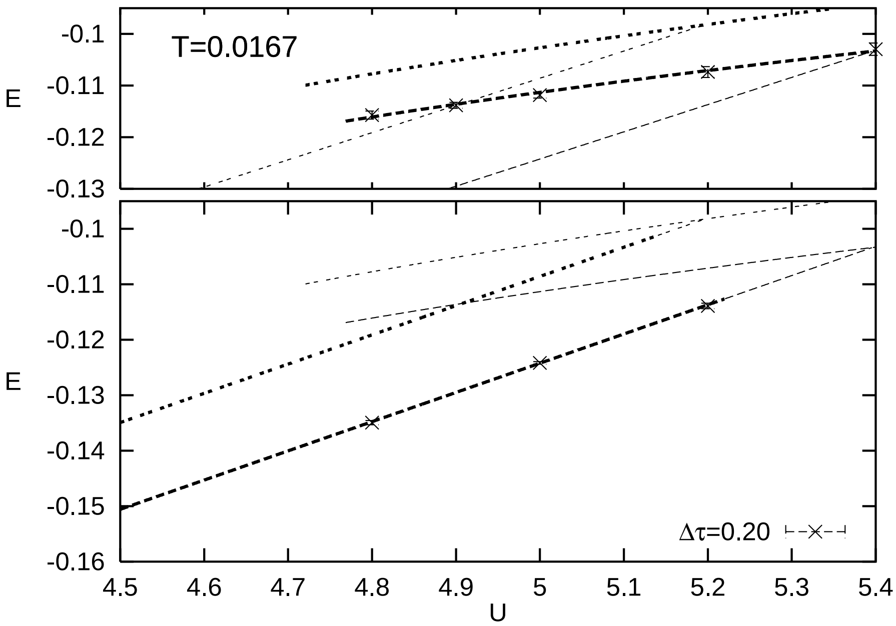



**Figure 3.29:** Total energy for $T = 1/60$. Due to lack of sufficient data, the coefficients for the $\Delta\tau \to 0$ extrapolation are estimated from the results at higher temperature. Otherwise, the plot is analogous to Fig. 3.24.

---

region. Again, the excellent agreement between the QMC data and the fit curves implies high confidence in the measured energies as well as the predictions for $U_{c2}$ and $U_{c2b}$.

The QMC data for $T = 0.0167$ (i.e., $\beta = 60$) presented in Fig. 3.29 should be regarded more as the basis of an extrapolation towards lower temperatures than as a complete, independent set of measurements. In fact, simulations were only performed at $\Delta\tau = 0.2$ on a sparse grid. The fit to the metallic solution was then scaled to $\Delta\tau = 0$ with a function whose coefficients in $\Delta\tau^2$ and $\Delta\tau^4$ were both extrapolated from the higher-temperature data. Due to the larger intrinsic error, these results are only used as an indication of the phase boundaries and not for the determination of the first-order phase transition line.

Figure 3.30 summarizes the QMC measurements of the full internal energy, extrapolated to the physical limit $\Delta\tau = 0$. Here, the symbols correspond to the data points labeled as "$\Delta\tau = 0.0$" in Fig. 3.23 – Fig. 3.28. In contrast to these previous plots, however, only those extrapolated data points are included which fall into our estimate of the stability region of the metallic phase. These data points are essentially independent from each other (up to a globally interpolated coefficient for the $\Delta\tau^4$ term in the extrapolation $\Delta\tau \to 0$). The straight narrow lines are the corresponding fits; they are cut off where the metal becomes unstable (towards larger $U$), or where the linearity is expected to break down (towards smaller $U$). The thick black line corresponds to the insulating phase. Its range of stability is not encoded in the plot; the slight increase of its lower edge $U_{c1}$ with increasing $T$ will be seen in Fig. 3.31.

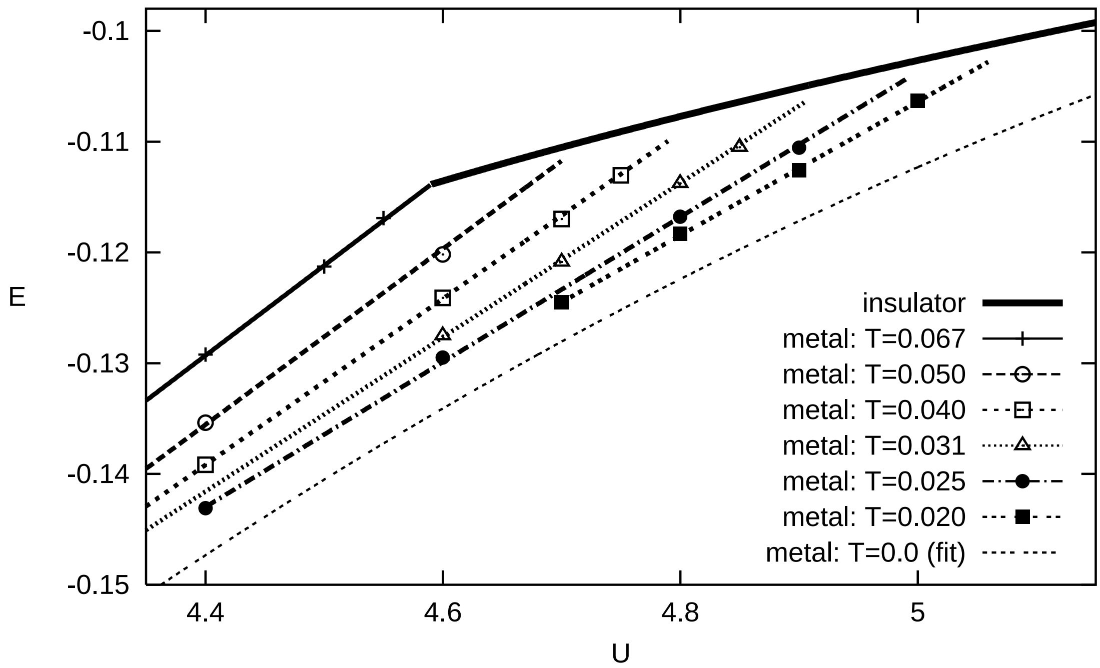



**Figure 3.30:** Internal energy $E$ vs. interaction $U$. The symbols and thin colored lines indicate the results of extrapolations $\Delta\tau \to 0$ for QMC data within the metallic phase and the linear fit curves (cf. previous plots). The thick solid line denotes the energy of the insulator within its $T$-dependent stability region. The thin bottom line marks the ground state energy in the metallic phase as determined in Sec. 3.6.

---

Also included in the graph is an extrapolation of the energy of the metal to $T = 0$ (see Sec. 3.6).

The most important result obtained so far in this section is certainly the observation of coexisting metallic and insulating solutions for $T \lesssim 1/15$. The corresponding phase diagram, to be shown and discussed in the following subsection 3.5.4, qualitatively agrees with earlier suggestions (cf. Sec. 3.3). A second important point that has not been previously discussed in the literature is the behavior of the energy near the MIT: the dependence of $E$ on $U$ is almost linear both in the metallic phase and in the insulating phase with the slope being significantly lower in the latter. Furthermore, the difference in $E$ (and other properties) is small when the metallic phase becomes unstable at $U_{c2}$ (see the discussion starting on page 104). A third point related to the linearity is the very regular dependence of $E$ on $\Delta\tau$ within each phase. It allows us to determine the phase boundaries with higher precision than has been achieved by other groups using different criteria.

### 3.5.4 Coexistence Phase Diagram

The implications of the data presented so far for the phase diagram are shown in Fig. 3.31. Here, large crosses mark the boundaries $U_{c1}$ and $U_{c2}$ of the coexistence

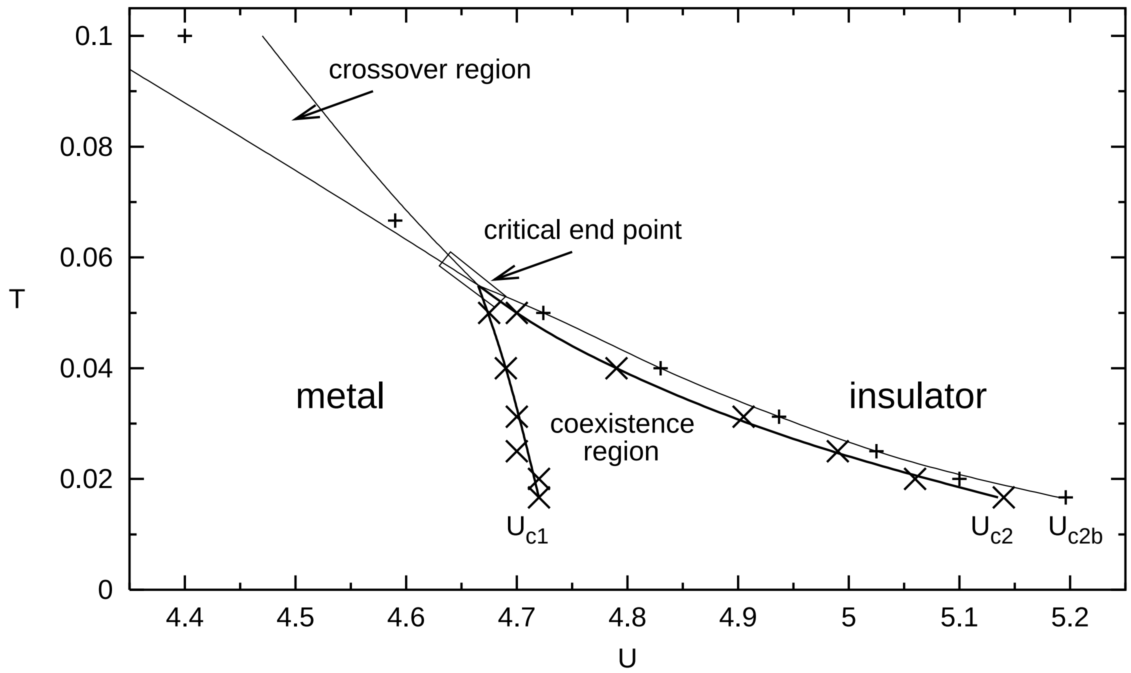



**Figure 3.31:** MIT phase diagram: critical end point at $U^* = 4.665$, $T^* = 0.055$ (approximate error bounds are indicated as green rectangle), crossover region at higher $T$, and coexistence region at lower $T$. Large tilted crosses indicate $U_{c1}$ and $U_{c2}$, the lower and upper edge of the coexistence region, respectively. Smaller upright crosses denote $U_{c2b}$, where the fits for the energies of metallic and insulating phases coincide (cf. subsection 3.5.3). Symbol widths roughly indicate the associated errors (in $U$).

region as estimated from the QMC simulations, extrapolated to $\Delta\tau = 0.0$. These data points correspond to the end points of the fit curves for the energy estimates presented in Fig. 3.24 – Fig. 3.29 (thick part of the fit curves for the metallic phase). Small crosses in Fig. 3.31 indicate the crossing points $U_{c2b}$ of the (extrapolated) fit curves for metal and insulator. While these crossing points lie just outside the stability region of the metallic phase at low temperatures, they are within the crossover region at higher temperatures. The lines shown in Fig. 3.31 do not only guide the eye, but also represent our best estimates of the phase boundaries. In particular, their common crossing point defines the critical point, i.e., the upper tip of the coexistence region at $U^* = 4.665$ and $T^* = 0.055$. At the same time, it also forms the lower tip of the crossover region. The thin blue lines represents our definition of the crossover region: between these lines, the QMC energy estimates deviate both from the linear fit curves on the metallic side and the universal form (3.18) on the insulating side (cf. Fig. 3.22 and Fig. 3.23). The estimate of the coexistence region, denoted by thick blue lines, was obtained from interpolation of a subset of the QMC data points, using a spline slightly biased to be consistent with the extrapolation of $U_{c2b}$ (thin grey line) and with the boundaries of the crossover region. The errors in the critical temperature and in the critical interaction are highly correlated as indicated by the

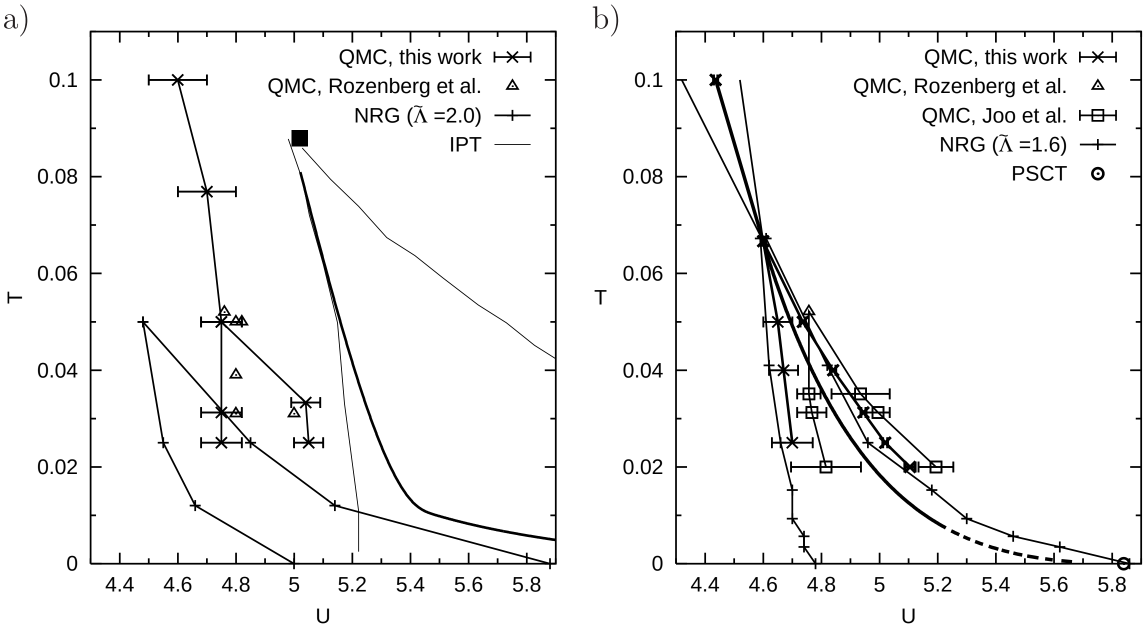



**Figure 3.32:** MIT phase diagram: **a)** Status as of Spring 2000 as established by this work, Rozenberg et al. (1999), and by preliminary NRG results for discretization $\tilde{\Lambda} = 2.0$ (Bulla, unpublished) in comparison with the IPT scenario [Georges et al. (1996); cf. Fig. 3.7]. **b)** Status as of Fall 2000: improved QMC estimates (this work), preliminary NRG results for discretization $\tilde{\Lambda} = 1.6$ (Bulla, unpublished), and new QMC boundaries by Joo and Oudovenko (2001) which include the critical point determined by Rozenberg. The NRG estimate for $U_{c2}(T = 0)$ is in excellent agreement with PSCT (Moeller et al., 1995).

green tilted rectangle. A reasonable numeric representation is $\Delta T = 0.004$, $\Delta U \lesssim 0.03$. The coexistence region shown in Fig. 3.31 is one of the main results of this thesis: It resolves the dispute described in subsection 3.1.2 and Sec. 3.3 since it essentially confirms Rozenberg et al.'s (1999) earlier and Joo and Oudovenko's (2001) independent QMC results and shows excellent agreement with independent NRG results (Bulla, Costi, and Vollhardt, 2001). In addition, our result has the smallest errors.

Before we further substantiate these claims, let us take a short look at the historical development. The tilted crosses in Fig. 3.32 a represent early QMC results obtained using the corrected code (cf. subsection 3.4.1) as presented at the DPG spring meeting on March 30, 2000. Here, lines are guides to the eye only. Within their (large) error bars, these preliminary results for the boundaries of the coexistence region were fully consistent with the coexistence points (triangles) previously found by Rozenberg et al. (1999). At the time of the conference, finite-temperature NRG estimates (Bulla, unpublished[28]) were not yet reliable: the horizontal crosses in Fig. 3.32 a correspond to an energy discretization parameter $\tilde{\Lambda} = 2.0$, where the spectrum is modeled by a discrete set of $\delta$-peaks at energies $\pm E_0 \, \tilde{\Lambda}^n$; the physical

[28]The NRG points at $T = 0$ coincide with previously published data (Bulla, 1999).

continuum limit is recovered for $\tilde{\Lambda} \to 1$. Attributing the differences between NRG and QMC to the known systematic underestimation of critical interactions by NRG at large $\tilde{\Lambda}$ (and finite temperature), the agreement appeared reasonable, in particular when compared with IPT (thin lines).

Figure 3.32 b represents the status of the phase diagram of late October 2000.[29] Here, our QMC estimates (crosses) are already based on measurements of the full internal energy (cf. subsection 3.5.1). This fact and a large number of simulations on a finer grid result in an accuracy which is much higher than that of the preliminary results shown in Fig. 3.32 a and also significantly higher than that of Joo and Oudovenko's (2001) results shown as squares in Fig. 3.32 b (which were first available as preprints in late September 2000).[30] The surprisingly good agreement between our QMC results and the improved NRG results [which are close to, but not identical with the data shown in the subsequent publication (Bulla et al., 2001)] is partially accidental as the final results for QMC and NRG show slightly larger differences (cf. subsection 3.6.3). Still, the good overall agreement between these results and (within larger error bars) with the independent QMC results obtained by Joo and Oudovenko (2001) indicated a resolution of the controversy towards a reliable phase diagram. In particular, the improved NRG estimate $U_{c2}(T=0) = 5.86$ is practically identical with the PSCT/ED estimate $U_{c2}(T=0) = 5.84$. These and all other published numerical estimates for the boundaries of the MIT coexistence region at $T = 0$ that we are aware of are reproduced in Table 3.1. A corresponding summary for the coordinates of the second-order critical end point is given in Table 3.2; the latter also includes our new QMC estimate.

The thick line in Fig. 3.32 b which evolves into a dashed line at low temperatures reflects our estimate (as of October 2000) of the true first-order phase transition line; the final estimate (which is hardly distinguishable from the curve shown here) will be presented after detailing the methodology in Sec. 3.6. The low-temperature part of this curve is partially based on the assumption that $U_{c2}(T = 0) \approx 5.85$ which is certainly justified based on the excellent agreement between PSCT, ED, and NRG.

Leaving the historical review, let us discuss the possible origin of the remaining discrepancies between the various QMC estimates of the boundaries of the coexistence region. Tilted and upright crosses in Fig. 3.33 represent our extrapolated QMC data (here repeated from Fig. 3.31) and QMC data corresponding to $\Delta\tau = 0.25$, respectively. Evidently, the extrapolation $\Delta\tau \to 0$ is important; otherwise (i.e., at $\Delta\tau = 0.25$), the coexistence region is shifted by about $\Delta U \approx 0.1$ towards larger interaction and by about $\Delta T \approx 0.005$ towards higher temperatures.[31] Interestingly,

[29] The comparison shown here was prepared for a conference in Grenoble (starting on October 30, 2000); on that occasion, a paper copy was also handed out to Georges. The same data (including our estimate for the first-order phase transition line) was presented in talks in condensed matter theory seminars in Augsburg and Mainz on 11/29/00 and 12/05/00, respectively and appeared in a report sent to the van Neumann Institute for Computing in November 2000.

[30] Note that Joo and Oudovenko (2001) did not calculate the tip of the coexistence region. Instead, they used Rozenberg et al.'s (1999) earlier estimate.

[31] Note that in our computations, the range of discretization parameters includes $0.1 \leq \Delta\tau \leq 0.25$ even for relatively low temperatures $T \geq 0.025$ ($\Delta\tau_{\min} = 0.125$ for $T = 0.02$). Therefore, the deviation of our best "raw" results from the extrapolated results is much smaller than shown in this

| $U_{c1}$ | $U_{c2}$ | method | factor | citation |
|---|---|---|---|---|
| 5.1 | 6.6 | IPT | $\sqrt{2}$ | Georges and Krauth (1993), p. 7172 |
| 5.2 | 6.6 | IPT | 2 | Rozenberg et al. (1994), Fig. 10 |
| 4.7 | | QMC | 2 | Rozenberg et al. (1994), p. 10186 |
| | 5.8 | PSCT | 2 | Moeller et al. (1995), p. 2084 |
| $4.3 \pm 0.5$ | $5.84 \pm 0.1$ | ED,PSCT | $2D$ | Georges et al. (1996), p. 66 |
| $4.0 \pm 0.4$ | $4.0 \pm 0.4$ | RDA+ED | $4W$ | Noack and Gebhard (1999), p. 1917 |
| 5 | 5.88 | NRG | $4W$ | Bulla (1999), p. 138/137 |
| 4.78 | 5.86 | NRG | $4W$ | Bulla et al. (2001), p. 8, Fig. 9 |

**Table 3.1:** *MIT at $T = 0$ in the half-filled fully frustrated Hubbard model with semi-elliptic DOS for $d \to \infty$. All values for $U_{c1}$ and $U_{c2}$ have been rescaled to unit variance of the DOS by multiplying the values stated in the publications by the appropriate factor (fourth column).*

| $U^*$ | $T^*$ | method | citation |
|---|---|---|---|
| 5.02 | 0.088 | IPT | Rozenberg et al. (1994), Fig. 10 |
| $4.76 \pm 0.04$ | $0.052 \pm 0.006$ | QMC | Rozenberg et al. (1999), p. 3500 |
| 4.9263 | 0.09379 | IPT | Kotliar et al. (2000), Fig. 2 |
| 4.68 | 0.050 | ED | Tong, Shen, and Pu (2001), p. 3 |
| | 0.08 | NRG | Bulla et al. (2001), p. 6 |
| $4.665 \pm 0.03$ | $0.055 \pm 0.004$ | QMC | this work |

**Table 3.2:** *Second order critical end point of the MIT in the half-filled fully frustrated Hubbard model with semi-elliptic DOS for $d \to \infty$. Rescaling to unit variance required multiplication by a factor of $4W$ for NRG and a factor of 2 for the other cited results.*

our estimates for $\Delta\tau = 0.25$ practically match the final results published by Joo and Oudovenko (2001) and are fully consistent with Rozenberg et al.'s (1999) data. In contrast, the agreement between the published data and our extrapolated estimates is much worse. In fact, no extrapolation is mentioned in Rozenberg et al.'s (1999) paper and an extrapolation $\Delta\tau \to 0$ is explicitly stated only for two selected phase points by Joo and Oudovenko (2001). We stress that (at $\Delta\tau = 0.0$) the metallic solution is not stable at four out of six points where coexistence has been claimed by Rozenberg et al. (1999). Thus, this paper contains quantitative errors; even its qualitative correctness, i.e., the evidence for the existence of a coexistence region (at $\Delta\tau = 0.0$) cannot be derived from data shown in the publication. Only our work, Joo and Oudovenko's (2001) selective $\Delta\tau$ studies, and Bulla's NRG results demonstrate that among the various scenarios shown in Fig. 3.13 for the $\Delta\tau$ dependence of

comparison. However, Rozenberg et al.'s (1999) best calculations at $T = 0.031$ were performed with $\Delta\tau = 0.25$ and the same can be assumed for the bulk of Joo and Oudovenko's (2001) calculations.

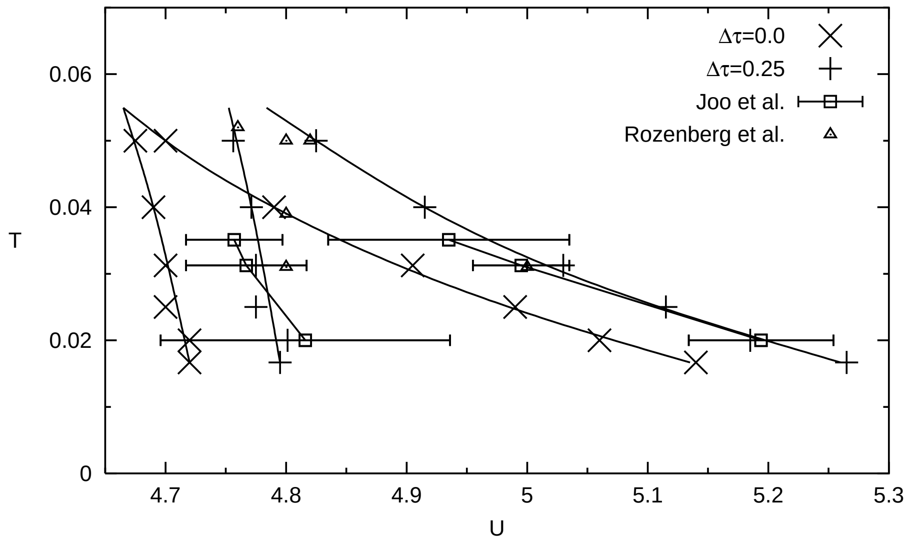



**Figure 3.33:** MIT phase diagram: dependency of the boundaries $U_{c1}(T)$ and $U_{c2}(T)$ of the coexistence region on $\Delta\tau$. Final results of this work for $\Delta\tau = 0.0$ (already shown in Fig. 3.31) and $\Delta\tau = 0.25$ in comparison with Joo and Oudovenko's (2001) results and the coexistence points established by Rozenberg et al. (1999).

---

the coexistence region which appear *a priori* possible, the scenario of Fig. 3.13 a is realized: finite $\Delta\tau$ primarily shifts the coexistence region towards larger values of $U$. Consequently, the coexistence found by Rozenberg can now *a posteriori* be assessed as genuine (and not as an artifact of using finite $\Delta\tau$).

Another fact established by these recent studies is that $U_{c1}$ is almost independent of $T$, only slightly increasing when the temperature is lowered from $T = T^*$ to $T = 0$. This finding is clearly incompatible with the MIT scenario Fig. 3.10 suggested by Kotliar which partially reconciled QMC results (at relatively high temperature) with (zero-temperature) RDA results by assuming that $U_{c1}(T = 0) = 4.0$. Ironically, our correction of this perception of the shape of the coexistence region validates *a posteriori* the qualitative features of the IPT scenario (cf. Fig. 3.7 and Fig. 3.32), originally supported by Kotliar (and others).

### 3.5.5 Double Occupancy

In this subsection, we construct an extrapolation scheme for the double occupancy $D$ in the metallic phase. In connection with the expression (3.19) established for the double occupancy of the insulating phase in subsection 3.5.1, the resulting estimates will allow for a determination of the difference $\Delta D(T, U)$ in double occupancy between metal and insulator within a few percent. This data will be used in combination with corresponding estimates energetic differences $\Delta E(T, U)$ (obtained in subsection 3.5.3) for the construction of a differential equation governing the thermodynamic first-order phase transition line $U_c(T)$ in Sec. 3.6. In addition to our best

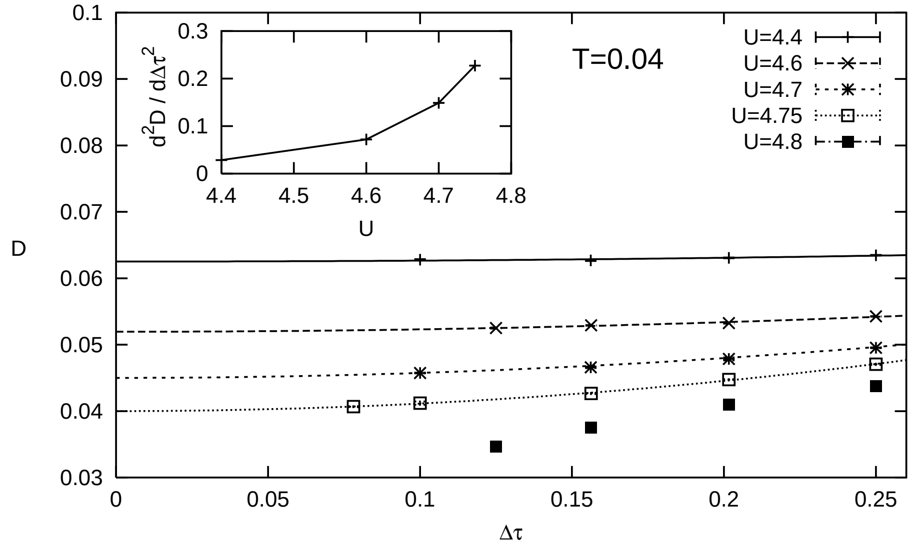



**Figure 3.34:** Extrapolation $\Delta\tau \to 0$ of double occupancy $D$ in metallic phase for $T = 0.04$: At small enough $U$, the QMC data points agree with purely quadratic least-squares fits (lines). Since the quadratic (in $\Delta\tau$) extrapolation law is seen to break down for $U = 4.8$, it is also unreliable at slightly smaller $U$. The inset shows the (second order) fit coefficient which increases significantly for $U \to U_{c2}$.

---

estimate for the double occupancy in the metal, we will also indicate approximate functional forms for $D_{\mathrm{m}}(T, U)$ within the coexistence region which correspond exactly to a linearized, interpolated version of this differential equation.

The difficulty in extrapolating $D$ to the physical limit $\Delta\tau \to 0$ is illustrated for $T = 0.04$ in Fig. 3.34. Well within the metallic phase, the dependence of $D$ on $\Delta\tau$ follows quadratic fit curves (i.e., without linear terms). Within the range of $\Delta\tau$ shown in the plot, this observation even remains true (with strongly increased $\Delta\tau$ dependence; see inset of Fig. 3.34) up to $U = 4.75$. This value of the interaction is just below $U_{c2} = 4.8$, where the metallic phase becomes unstable, and $U_{c2b} \approx 4.83$, where the extrapolated energy (see Fig. 3.25) meets that of the insulator. At $U = 4.8$, however, the purely quadratic extrapolation law for the double occupancy clearly breaks down for $\Delta\tau \lesssim 0.16$. The observed irregularity not only makes it impossible to perform the extrapolation $\Delta\tau \to 0$ for fixed interaction $U$ very close to the MIT, but also makes the results of extrapolations for slightly smaller $U$ suspicious. After all, deviations from the quadratic extrapolation law might occur for $\Delta\tau \ll 0.1$ so that the true value of $D$ could be lower than the extrapolated value. Therefore, reliable extrapolations can only be performed on the basis of appropriate fit functions, each of which approximate the QMC estimates for $D(U)$ for some value of $\Delta\tau$ in the metallic phase. Extrapolations to the limit $\Delta\tau = 0$ are then no longer local in $U$ and are thus much less affected by the difficulties discussed above.

As in subsection 3.5.3, we present the results for the double occupancy sorted from high to low temperatures, starting with $T = 0.1$. Due to the broad crossover at this

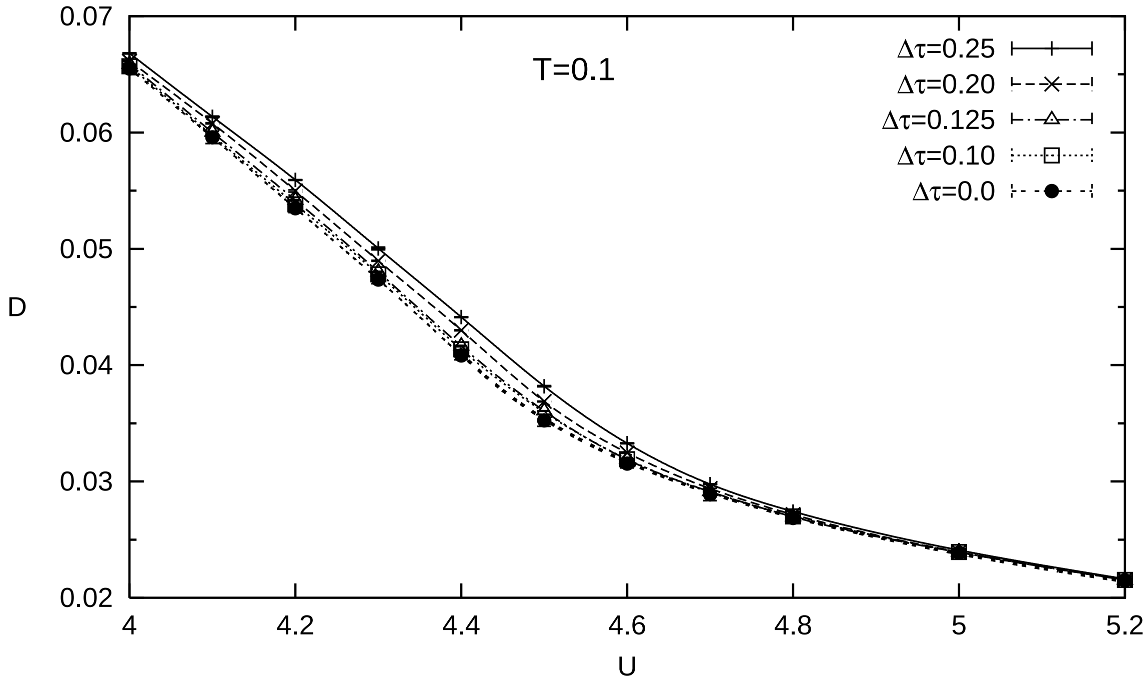



**Figure 3.35:** Double occupancy for $T = 1/10$. Filled circles indicate the results of purely quadratic $\Delta\tau \to 0$ extrapolations of the QMC data (open symbols). Thin lines are guides to the eye only. The thick portion of the double-dashed line for $\Delta\tau = 0$ (and $U \geq 4.4$) follows the universal expression (3.19) for the insulator.

temperature, no fits for the metallic-like region were attempted in Fig. 3.35. In this figure, all thin lines are guides to the eye only. The data labeled as $\Delta\tau = 0.0$ (solid circles) was obtained by quadratic extrapolation. The thick black line visible for $U > 4.4$ represents our universal fit (3.19) for the double occupancy in the insulator. The maximum slope of $D(U)$ for $U \approx 4.4$ supports our earlier estimate $U_{c2b} = 4.4$ for the center of the crossover region at this temperature.

In Fig. 3.36, the almost abrupt change of slopes for $D(U)$ indicates the proximity to a sharp transition for $T = 0.067$. Here and in all following plots, the lower curve for $\Delta\tau = 0.0$ is given by the temperature-independent fit (3.19) for the insulator. The curves for the metallic region are parameterized in terms of their difference with respect to the expression (3.19) for the double occupancy in the insulating phase:

$$D_{\mathrm{m}}(T, U, \Delta\tau) = D_i(U) + A_{T,\Delta\tau} \frac{U_{c2b}(T, \Delta\tau) - U}{1 + B_{T,\Delta\tau}(U_{c2b}(T, \Delta\tau) - U)} \,. \tag{3.20}$$

In particular, the differences vanish for $U = U_{c2b}(T, \Delta\tau)$ so that at these points not only the (extrapolated) fits for the energy of metal and insulator, but also the fits for the double occupancy of both phases coincide. Consequently, the ratio of the differences of both pairs of observables $\Delta E/\Delta D$ remains finite within (and beyond) the coexistence region which will prove useful later. The initial slope and curvature of $\Delta D$ at $U = U_{c2b}$ is controlled by the two free parameters (per curve), $A_{T,\Delta\tau}$ and

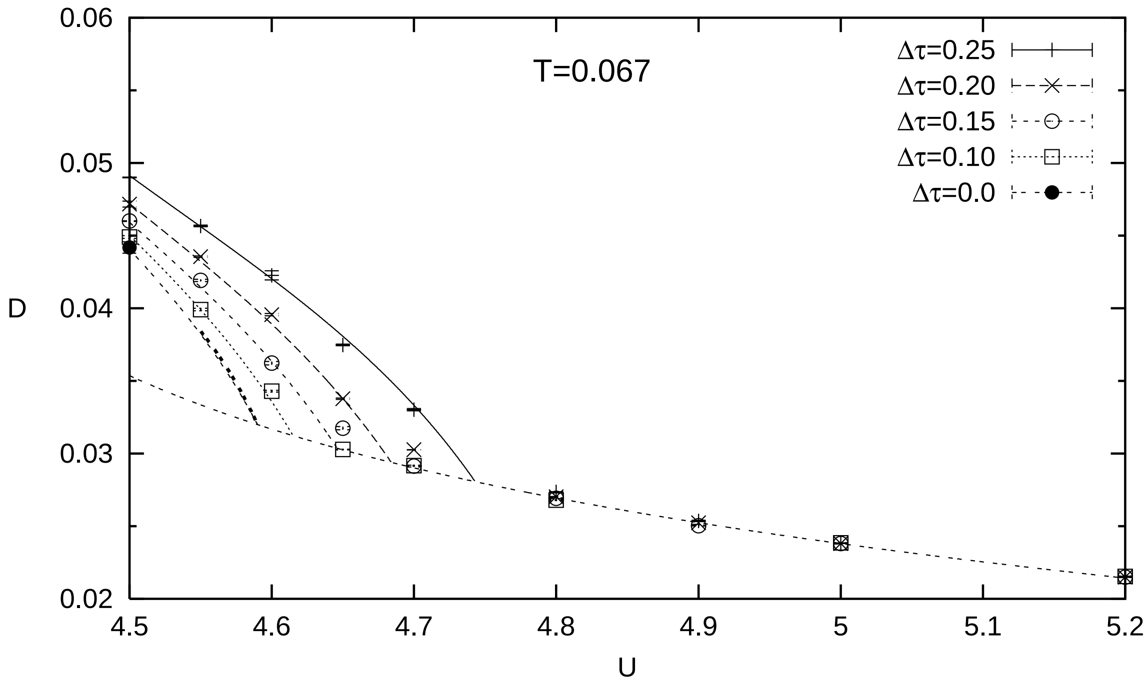



**Figure 3.36:** Double occupancy for $T = 1/15$. The lowest line follows the temperature-independent law (3.19); all other thin lines are fits according to (3.20) to the QMC data for the metallic region (open symbols) or its extrapolation to $\Delta\tau = 0$ (filled circle). An additional fit (thick line below $U = 4.59$) is related to the linear approximation of a differential equation for $U_c(T)$ (see text and Sec. 3.6). The deviations between QMC data and fit curves at the crossing points $U_{c2b}$ indicate a narrow crossover from metal to insulator.

$B_{T,\Delta\tau}$. Apart from generally fitting the data extremely well, the precise form chosen here stands out by the fact that it would lead to an exactly linear differential equation for the first-order transition line $U_c(T)$ in the limit of vanishing curvature of $E_i(U)$. Except for the close neighborhood of approximately $U_{c2b} - 0.02 \lesssim U \lesssim U_{c2b} + 0.02$ around the crossing points, the QMC data is seen to be reproduced well by the fits in Fig. 3.36. The deviations of the points for $\Delta\tau = 0.15$ at $U = 4.65$ and $\Delta\tau = 0.2$ at $U = 4.7$, however, are significant. This adds further support to our previous conclusion that this temperature is just above the critical temperature. The thick line extending below $U = 4.59$ indicates an additional fit to be used later in a linear differential equation for the first-order line $U_c(T)$.[32]

The analogous data for $T = 0.05$ is presented in Fig. 3.37. The only significant discrepancy between data and fits is seen for the insulating solution at $U = 4.7$. While the deviation might be slightly exaggerated due to using a finite number of sweeps, it is above error bars at all finite $\Delta\tau$. Note, however, that the distance between data and

[32]Since the coexistence region does not extend up to the present temperature, the differential equation will find no application here. Still, some continuity is expected and found with respect to the behavior of $D$ and $E$, so that the associated data point can still be used for interpolation.

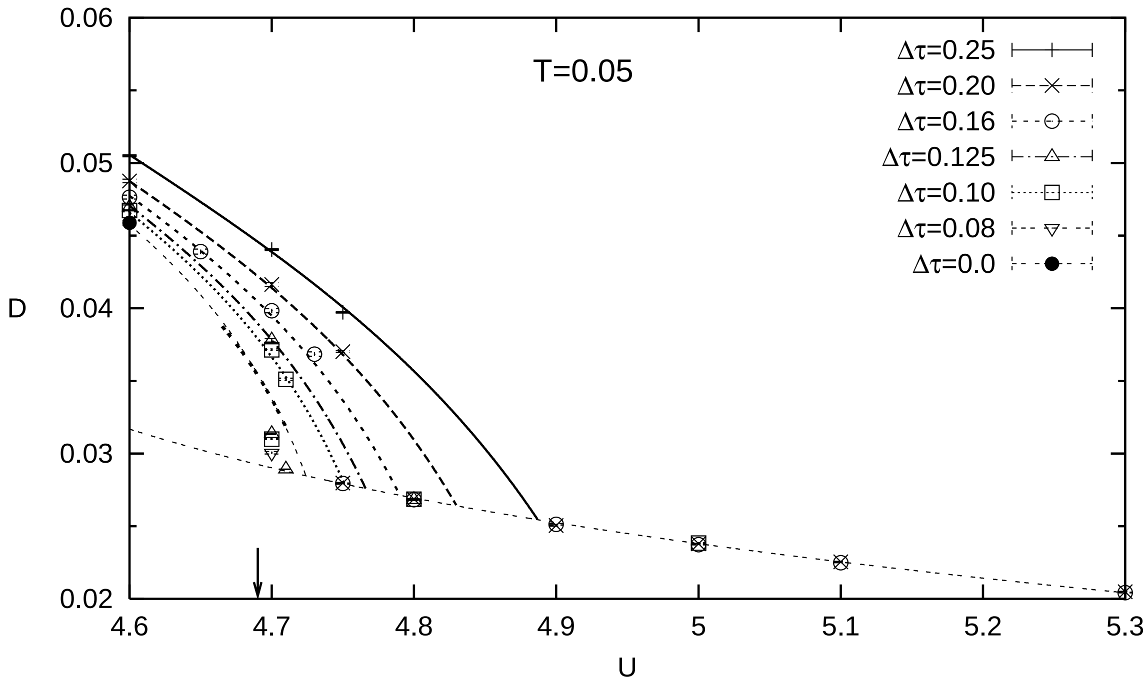



**Figure 3.37:** Double occupancy for $T = 1/20$. QMC data (symbols) and fit curves in analogy to Fig. 3.36. The extent of the "linearized" fit (thick curve) represents the width of the coexistence region (at $\Delta\tau = 0$). The arrow shows the estimate for $U_c$ obtained in Sec. 3.6. The $\Delta\tau$ dependence of the QMC data for the insulator is significant only in the coexistence region at $U = 4.7$.

the fit curve for the insulating solution decreases significantly with decreasing $\Delta\tau$.[33] Therefore and in view of the behavior of $D$ observed at lower $T$, it is still reasonable to assume that the $\Delta\tau = 0$ curve is correct also for the insulator. Further support for this scenario will arise from the shape of the coexistence region. As before, the thick double-dashed line corresponds to the linearized differential equation for $U_c(T)$ to be discussed in Sec. 3.6.[34] Here and in the following plots, the extent of the thick fit curve denotes the extent of the coexistence region. The arrow pointing towards the abscissa represents the estimate for $U_c$ (for $T = 0.05$) to be derived in Sec. 3.6. It is included in this and the following plots in order to indicate the most important parameter region for each temperature; the accuracy of the final phase diagram will predominantly depend on small errors in $\Delta D$ and $\Delta E$ in these regions.

For $T = 0.04$, the good agreement between QMC data and fit functions is evident

[33] Obviously, the correct extrapolation could be found by performing additional, careful simulations at very small $\Delta\tau$. The computational cost, however, would be enormous even for going to $\Delta\tau = 0.05$ since the proximity of the stability edge requires using a much larger number of sweeps than needed at lower $T$ (where up to 400 time slices were indeed used). Furthermore, the impact of the data at this high temperature on $U_c(T)$ is essentially local, i.e., will only affect it visibly within the tip of the coexistence region.

[34] The fact that this thick line closely matches the original fit for $\Delta\tau = 0.0$ (thin line) implies that no accuracy will be lost by the linearization.

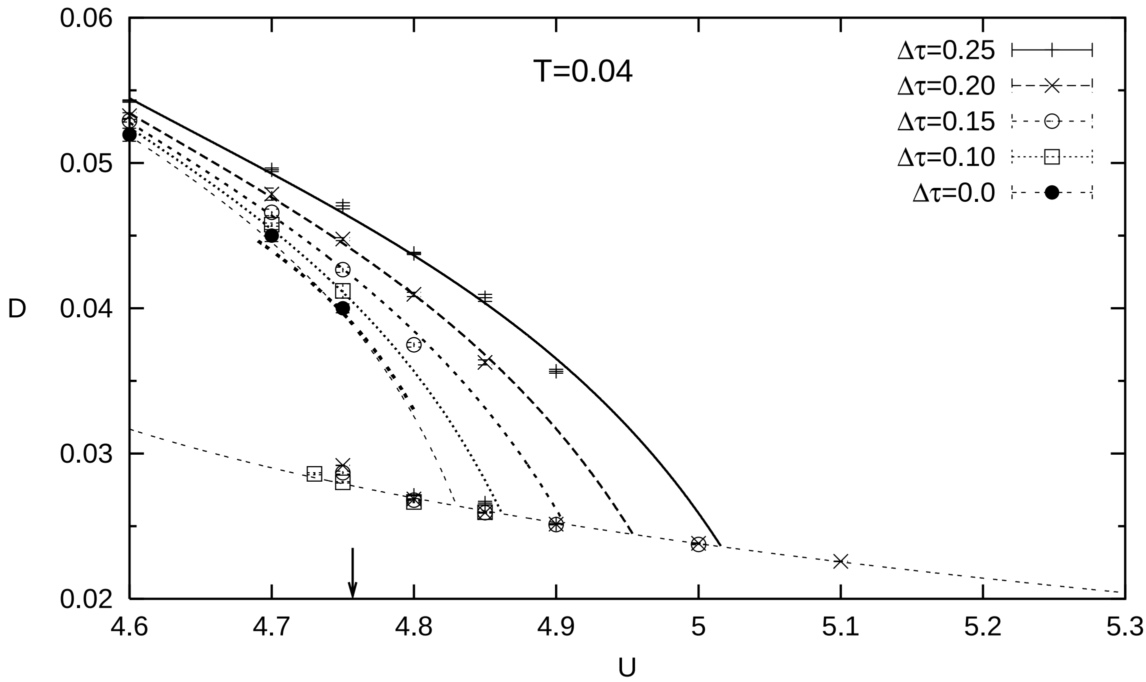



**Figure 3.38:** Double occupancy for $T = 1/25$, analogous to Fig. 3.37.

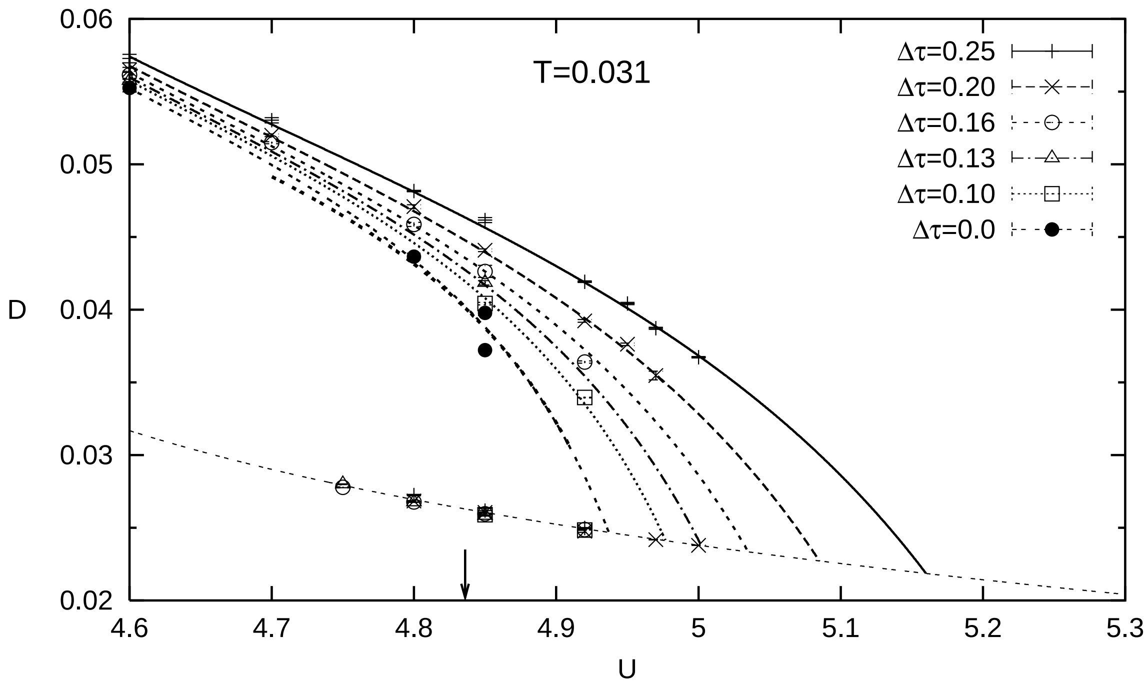



**Figure 3.39:** Double occupancy for $T = 1/32$, analogous to Fig. 3.37. For $U = 4.85$, where the quality of a purely quadratic least-squares fit for $D$ in the metallic phase (upper solid circle) is already bad, a fit including a linear term has been added (lower solid circle). Both estimates bracket the fit functions for $\Delta\tau = 0.0$ (double-dashed lines).

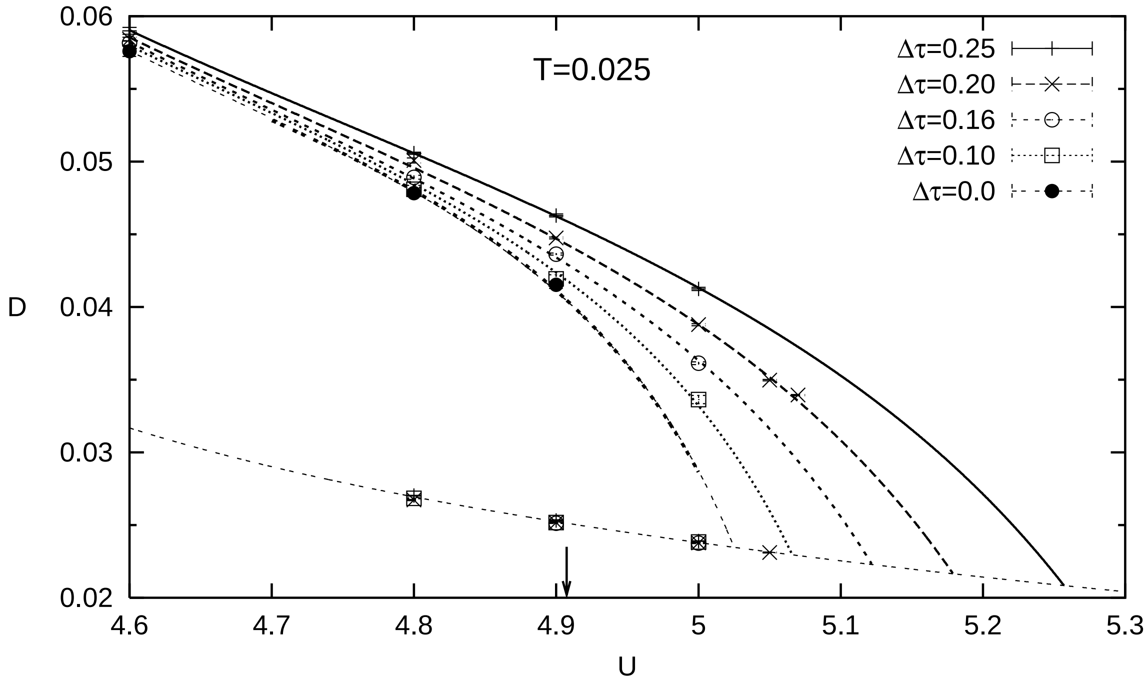



**Figure 3.40:** Double occupancy for $T = 1/40$, analogous to Fig. 3.37.

also within the coexistence region from Fig. 3.38. Note that for $U = 4.75$ (which is also our estimate for $U_c$, see arrow), the deviation from the insulating curve is significantly smaller for $\Delta\tau = 0.15$ than for $\Delta\tau = 0.2$ and essentially zero already for $\Delta\tau = 0.1$. The quadratic extrapolation for the metallic solution at $U = 4.75$ (solid circle) should already be close enough to $U_{c2b}$ to overestimate the true value of $D$; our globally systematic fit curve and its "linearized" approximation (thick portion) are more reliable.

The data for $T = 0.031$ presented in Fig. 3.39 follows the expected pattern. Note that all QMC points in the insulator are on the fit curve and that among the metallic solutions, only the point for $U = 4.92$ and $\Delta\tau = 0.16$ shows negative deviations. Since for the same interaction even $\Delta\tau = 0.1$ is on (or slightly above) its fit curve, a remaining convergence error or the influence of a too low number of sweeps is much more likely than an unsuitability of the fit functions. At $U = 4.85$, we have not only included the result of a purely quadratic fit of the metallic solution to $\Delta\tau = 0$ (upper solid circle), but also a least squares fit including a linear term (lower circle) in order to indicate the range suggested by a local analysis. Again, the global fit is seen to be superior since it takes the regular shape observed at all finite $\Delta\tau$ into account and keeps this shape also for $\Delta\tau = 0.0$.

For $T = 0.025$, the most important region for the determination of $U_c$ near $U = 4.9$ (arrow) is already far enough below $U_{c2}$ and $U_{c2b}$ so that an accuracy within the low percent range could also be obtained from regular quadratic extrapolation. The deviation at $U = 4.9$ and $\Delta\tau = 0.1$ apparent in Fig. 3.40 is slightly outside the statistical error bars but is within the range of longer-scale fluctuations (associated

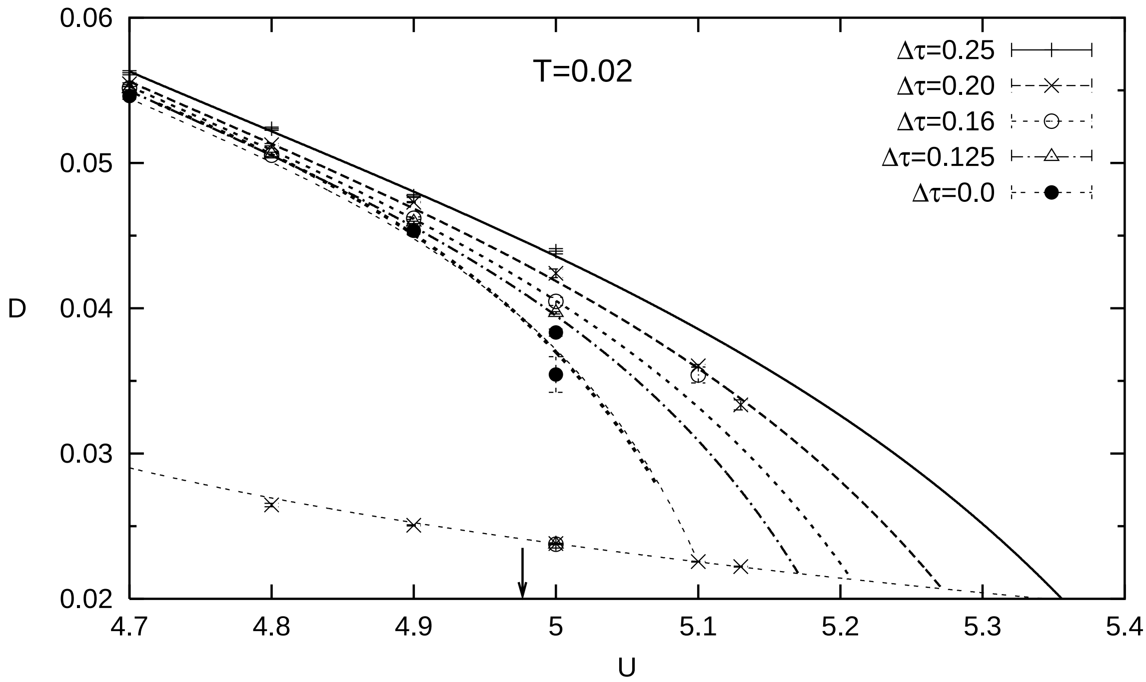



**Figure 3.41:** Double occupancy for $T = 1/50$, analogous to Fig. 3.37. For $U = 5.0$, where the quality of a purely quadratic least-squares fit for $D$ in the metallic phase (upper solid circle) is already bad, a fit including a linear term has been added (lower solid circle). Both estimates bracket the fit functions for $\Delta\tau = 0.0$ (double-dashed lines).

with using only 120000 sweeps) and is too small to impact our final results visibly. Note the good agreement of the "linearized fit" (thick line) with the original fit across the range $U_c = 4.7$ to $U_{c2} = 4.99$.

In Fig. 3.41, we have again included the result both of a purely quadratic and linear+quadratic extrapolation, here for $U = 5.0$. Note that the quality of least-squares fits is generally worse in this plot since simulations at $\Delta\tau = 0.1$ could not be afforded. Longer and even more precise simulations would be required for understanding or correcting the deviation of the metallic solution for $\Delta\tau = 0.16$ at $U = 5.1$. Still, the accuracy of the $\Delta\tau$ extrapolated fits is very satisfactory, considering that we are looking at the strongly correlated region at very low temperatures.

The summary of our results for the double occupancy is given in Fig. 3.42, where the fit functions for the metal at $\Delta\tau = 0$ are drawn (thin colored lines) together with QMC data (symbols) for interactions deep enough in the metallic phase so that quadratic least-squares fits are reasonable. In analogy to our summarizing result Fig. 3.30 for the energy, the thick line shows the double occupancy for the insulating phase. The nearly straight thin line at the top indicates the result of an extrapolation to $T = 0$ for the metallic phase to be detailed in Sec. 3.6.

The relative error of the estimates for $\Delta D$ implied by the results shown in Fig. 3.42 is about $10^{-2}$ in the proximity of the first-order phase transition (see arrows in Fig. 3.36 – Fig. 3.41) for $T \le 0.04$ and is only slightly larger for $T = 0.05$. Such a high

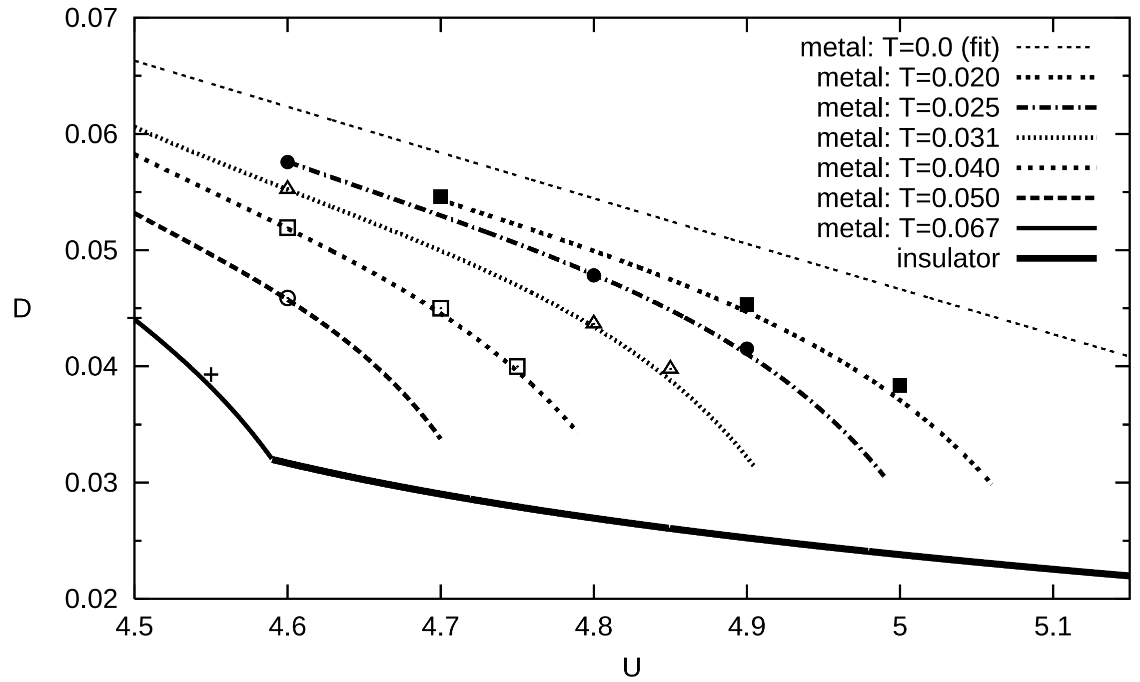



**Figure 3.42:** Double occupancy $D$ vs. interaction $U$ at $\Delta\tau = 0$. The symbols indicate extrapolated QMC data for the metallic phase; colored lines indicate "linearized" fit curves (thick lines in Fig. 3.36 – Fig. 3.41). The thick solid line denotes the double occupancy of the insulator within its $T$-dependent stability region. The thin top line marks the ground state double occupancy in the metal (to be determined in Sec. 3.6).

accuracy may seem surprising in view of our discussion initiated in subsection 3.5.1 and continued in the beginning of this subsection. However, the choice of suitable fit functions has made it possible to perform a nonlocal $\Delta\tau$ extrapolation (i.e., for all values of $U$ at the same time for fixed $T$), including both physical intuition and prior knowledge that had been obtained in subsection 3.5.3. Thus, reliable estimates of $D$ could be obtained even close to the MIT where naive, local extrapolations would have failed (as illustrated in Fig. 3.34).

# 3.6 Results: Thermodynamic Phase Transition Line

In this section, a differential equation will be derived for the transition line $U_c(T)$ between metal and insulator based on the fact that both phases have the same free energy along the transition line.

## 3.6.1 Differential Equation for $\mathrm{d}U_c/\mathrm{d}T$ and Linearization

For the Hubbard model, the free energy $F$ and its differential $\mathrm{d}F$ can be expressed in terms of energy $E$, temperature $T$, entropy $S$, double occupancy $D$, and interaction strength $U$ as

$$F = E - TS; \qquad \mathrm{d}F = -S\,\mathrm{d}T + D\,\mathrm{d}U. \tag{3.21}$$

These expressions are not directly useful in the context of QMC simulations. As pointed out in Sec. 1.3, the free energy cannot be estimated using QMC since the importance sampling technique by construction samples the partition sum only up to an unknown prefactor. Since $E$ can be measured within QMC (and $T$ is an external parameter), this argument and (3.21) also imply the inability of QMC to measure $S$ directly. However, using

$$\left.\frac{\partial \beta F}{\partial \beta}\right|_U = F + \beta \left.\frac{\partial F}{\partial \beta}\right|_U = F - T \left.\frac{\partial F}{\partial T}\right|_U = F + TS = E \tag{3.22}$$

with $\beta = 1/T$, the differential may be rewritten as

$$\mathrm{d}(\beta F(\beta, U)) = E(\beta, U)\,\mathrm{d}\beta + \beta D(\beta, U)\,\mathrm{d}U\,, \tag{3.23}$$

i.e., expressed in terms of observables which we have determined in Sec. 3.5. Provided that both the metallic and the insulating solution can be continuously followed (in general on different paths[35] in the $(U,\beta)$ plane) up to a point $(U_1,\beta_1)$ in the crossover region where they coincide, then by integrating equation (3.23) one can calculate $\beta F_{\mathrm{m/i}}(\beta, U) - \beta_1 F(\beta_1, U_1)$ within the coexistence region and, in principle, determine the transition line.[36] In practice, however, it is likely that along these different paths different systematic errors will contribute to the free energy estimates for metal and insulator, respectively. Consequently, the resulting estimate for the transition line may be systematically biased.

Systematic errors can be kept to a minimum in a scheme where differences in observables (between metal and insulator) are only computed locally, i.e., for fixed $U$ and $\beta$. This may be achieved by rewriting equation (3.23) for metal/insulator

$$\mathrm{d}\big(\beta F_{\mathrm{m/i}}(\beta, U)\big) = E_{\mathrm{m/i}}(\beta, U)\,\mathrm{d}\beta + \beta D_{\mathrm{m/i}}(\beta, U)\,\mathrm{d}U \tag{3.24}$$

and taking the difference

$$\mathrm{d}\big(\beta \Delta F(\beta, U)\big) = \Delta E(\beta, U)\,\mathrm{d}\beta + \beta \Delta D(\beta, U)\,\mathrm{d}U\,. \tag{3.25}$$

Here, $\Delta F(\beta, U) = F_{\mathrm{m}}(\beta, U) - F_{\mathrm{i}}(\beta, U)$; corresponding differences for the energy and double occupancy are denoted by $\Delta E$ and $\Delta D$, respectively. Since the transition line $U_{\mathrm{c}}(\beta)$ is smooth and determined by $\Delta F(\beta, U_{\mathrm{c}}(\beta)) = 0$, all derivatives of $\Delta F$ also vanish along the line $U_{\mathrm{c}}(\beta)$. This is true, in particular, for the differential:

$$\mathrm{d}(\beta \Delta F(\beta, U))|_{U=U_{\mathrm{c}}(\beta)} = 0. \tag{3.26}$$

From (3.25) and (3.26) we get

$$0 = \Delta E(\beta, U_{\mathrm{c}}(\beta))\,\mathrm{d}\beta + \beta \Delta D(\beta, U_{\mathrm{c}}(\beta))\,\mathrm{d}U_{\mathrm{c}} \tag{3.27}$$

[35] The necessity of using different paths for metallic and insulating solutions arises from the shift of the coexistence region with $\Delta\tau$ and from the larger numerical errors of each solution near its stability edge.

[36] Here and in the following, the indices m and i indicate metal and insulator, respectively.

which is equivalent to

$$\frac{\mathrm{d}U_\mathrm{c}(\beta)}{\mathrm{d}\beta} = -\frac{\Delta E(\beta, U_\mathrm{c}(\beta))}{\beta \Delta D(\beta, U_\mathrm{c}(\beta))} \tag{3.28}$$

$$= \frac{1}{\beta}\left(-\frac{\Delta E_\mathrm{kin}(\beta, U_\mathrm{c}(\beta))}{\Delta D(\beta, U_\mathrm{c}(\beta))} - U_\mathrm{c}\right). \tag{3.29}$$

While the second form (3.29) is somewhat more elementary, it conveys the wrong impression that large errors should be expected due to the subtraction of numbers of the same order of magnitude. This problem does not occur here since the numerical error of $E$ is significantly smaller (by about a factor of two or three) than the error of $E_\mathrm{kin}$. Therefore, we will in the following proceed with the total energy, i.e., use (3.28). Changing variables from $\beta$ to $T$ we arrive at the final form of the differential equation:[37]

$$\boxed{\frac{\mathrm{d}U_\mathrm{c}(T)}{\mathrm{d}T} = f(T, U_\mathrm{c}(T)); \qquad f(T,U) := \frac{\Delta E(T,U)}{T\,\Delta D(T,U)}\,.} \tag{3.30}$$

Since $U_\mathrm{c}(T)$ has to go through the tip of the coexistence region ($T = T^*$) where $U_{\mathrm{c}1} = U_{\mathrm{c}2} = U^*$, the true phase transition line $U_\mathrm{c}(T)$ is uniquely defined as the solution of (3.30) satisfying the initial condition $U_\mathrm{c}(T^*) = U^*$:

$$U_\mathrm{c}(T) = U^* + \int_{T^*}^{T} \mathrm{d}T' f(T', U_\mathrm{c}(T')); \quad T < T^*. \tag{3.31}$$

The slope of this curve is negative everywhere as we can see from the following argument: In going from an insulator to a metal, the entropy decreases. In order for the free energies to coincide at $U = U_\mathrm{c}(T)$, the change in total energy $\Delta E(T,U)$ has to be negative. Since the double occupancy increases at the same time, (3.30) yields $\mathrm{d}U_\mathrm{c}(T)/\mathrm{d}T < 0$.

In principle, $f(T,U)$ can be evaluated pointwise at finite $\Delta\tau$ by calculating the differences of $D$ and $U$ for fixed $U$ and $T$. This was, in fact, our initial approach.[38] It is, however, not applicable near the edges and in the tip of the coexistence region, since the phase boundaries shift significantly with changing $\Delta\tau$. Consequently, in general, not enough (if any) data points are available for an extrapolation $\Delta\tau \to 0$ at fixed $U$.[39] Furthermore, an interpolation between the grid points (in $U$ and $T$) at some stage is unavoidable in order to integrate up the formal solution (3.31).

[37]Note that this differential equation is in direct analogy to the Clausius-Clapeyron equation $\mathrm{d}p/\mathrm{d}T = \Delta Q/(T\Delta V)$ for the vapor pressure curve $p(T)$, where $\Delta Q$ is the latent heat and $\Delta V$ the volume difference between the two phases.

[38]This strategy yields accurate results deep within the coexistence region, where the resulting estimates for $f(T,U)$ are easily extrapolated to $\Delta\tau = 0$. The nearly linear dependence of $f(T,U)$ on $U$ observed using this method motivated the linearization of the differential equation which we will introduce below.

[39]The difficulties associated with an extrapolation of $D_\mathrm{m}$ to $\Delta\tau = 0$ at fixed $U$ clearly add to the problem (cf. subsection 3.5.5)

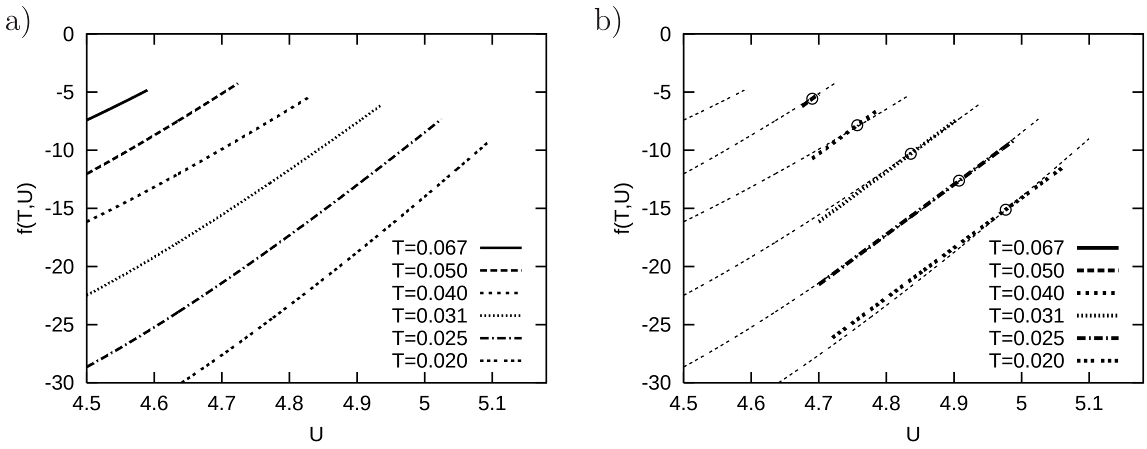



**Figure 3.43:** Function $f(T,U)$ which coincides with the derivative $\mathrm{d}U_\mathrm{c}/\mathrm{d}T$ on the phase transition line $U_\mathrm{c}(T)$. **a)** "Raw data" extracted from fits of QMC data for $E$ and $D$. **b)** Linearized fits with the functional form $f(T,U) = A(T)+43\,U$. Circles indicate $\mathrm{d}U_\mathrm{c}/\mathrm{d}T$ of the final solution.

Therefore, very accurate results in the whole coexistence region can only be achieved using the fit functions for $E_\mathrm{i}(U)$, $D_\mathrm{i}(U)$, $E_\mathrm{m}(T,U)$, and $D_\mathrm{m}(T,U)$ established in subsections 3.5.2, 3.5.3, and 3.5.5. Among these data sets, by far the largest error is associated with the double occupancy $D_\mathrm{m}(T,U)$ in the metallic phase. Figure 3.43 a shows "raw" results for $f(T,U)$ which follow from the QMC data when the best estimate for $D_\mathrm{m}$ according to the 2-parameter fits (3.20) (shown as thin lines in Fig. 3.36 – Fig. 3.41) are used. Here, the fit functions are not only shown within (our estimate of) the coexistence region, but extend up to $U = U_{\mathrm{c}2b}(T)$ (where the extrapolated energy of the metal crosses the energy of the insulator). Evidently, the dependence on $U$ is almost linear for fixed $T$. Furthermore, the slope of the curves shown in Fig. 3.43 a is only very weakly dependent on $T$, at least within the coexistence region. Consequently, only a small error is made when these raw results are replaced by straight lines with identical slopes throughout the coexistence region. This is seen in Fig. 3.43 b, where the thick lines (shown only for $U$ within the coexistence region) follow the form

$$f(T,U) = A(T) + 43\,U\,. \tag{3.32}$$

While the discrepancy between this linearized form and the original raw data is in the percent range throughout the coexistence region, it is even smaller near the most important points touched by the solution (circles) to be determined below. This fact alone ensures that the linearized form of $f(T,U)$ is a good interpolation for the purpose of finding the solution $U_\mathrm{c}(T)$. The particular practical advantage of the linear form (3.32) is that the solution of the corresponding differential equation $\mathrm{d}U_\mathrm{c}/\mathrm{d}T = A(T)+43U_\mathrm{c}(T)$ is given by a regular integral which can even be performed analytically using the interpolation for $A(T)$ that will be established later in this section.

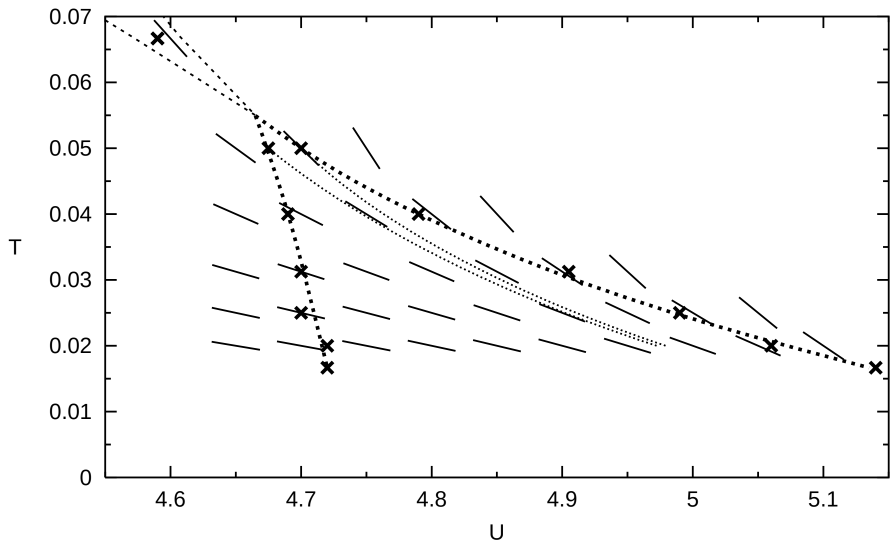



**Figure 3.44:** Gradient field for $U_c(T)$ as implied by the differential equation (3.30), symbolized by short solid line segments. The symbols and short-dashed lines represent the boundaries of coexistence region and crossover region (cf. Fig. 3.31). The dotted lines show solutions of the (linearized) differential equation with the initial condition $U_c = U_{c1}$ and $U_c = U_{c2}$ for $T = 0.05$, respectively, which bracket the final solution.

Figure 3.44 illustrates that the procedure of integrating the solution towards lower temperatures, starting at the tip of the coexistence region is inherently stable. Here, the crosses and short-dashed curves indicate the coexistence and crossover regions determined in Fig. 3.31. The slopes of each short solid line with center $U_i$, $T_i$ is given by the value $f(T_i, U_i)$ which can be read off from Fig. 3.43 a. Since the slope of any solution of the differential equation (3.30) has to agree with $f(T, U)$ at each point, solutions practically cannot cross the slopes shown in Fig. 3.44 (within their error bars and in the limit of vanishing lengths). The fast increase of the absolute slope from left to right makes them act like a funnel for the integration of solutions towards lower $T$. Therefore, the impact of errors in the position of the tip of the coexistence region (and also of $f(T, U)$ at high $T$) on the solutions decays rapidly with decreasing $T$. This is illustrated by the two test solutions (dotted lines) shown in Fig. 3.44 which fulfill the initial conditions $U(0.05) = U_{c1}(0.05)$ and $U(0.05) = U_{c2}(0.05)$, respectively. Although these lines, which bracket the physical solution $U_c(T)$, are far apart initially, their distance is already very small for $T = 0.02$.

### 3.6.2 Low-temperature Asymptotics of $U_c(T)$

While the data presented so far suffices for an accurate determination of the thermodynamic phase boundary $U_c(T)$ within the range of temperatures $T \geq 0.02$ covered by QMC simulations, the quality of an extrapolation to even lower temperatures based on the QMC data for $f(T, U)$ alone would be questionable. By including additional low-temperature information, however, we will be able to reliably integrate the solution down to $T = 0$. More precisely, we will first derive the asymptotic form of $U_c(T)$ for $T \to 0$ analytically [which implies corresponding asymptotics for $f(T, U)$] and fix its numerical parameter in a second step using various extrapolated QMC data. This procedure will not only allow for a controlled construction of the full phase diagram, but validate the internal consistency of our results and some underlying assumptions at the same time.

### Consequences of Fermi Liquid Theory

As we have established in subsection 3.5.3, the energy in the (low-temperature) insulating phase is independent of the temperature.[40] By virtue of (3.21), the same applies to the entropy in this phase. In addition, $S_i$ is independent of $U$; its value per lattice site equals $S_0 = \ln 2$ (Georges et al., 1996; Gebhard, 1997; Kalinowski and Gebhard, 2002). Thus, the free energy of the insulator at low $T$ is determined by

$$E_i(T, U) = E_i^0(U); \quad S_i(T, U) = S_0 \,. \tag{3.33}$$

Fermi liquid properties imply for the metallic phase that

$$E_m(T, U) = E_m^0(U) + \frac{1}{2}\gamma(U)T^2; \quad S_m(T, U) = \gamma(U)T \,, \tag{3.34}$$

where $\gamma(U)$ is the linear coefficient of the zero-temperature specific heat. Using the projective self-consistent technique (PSCT), Moeller established that $U_c(0) = U_{c2}(0)$, i.e., that the metallic solution is always stable within the coexistence region at $T = 0$ (Moeller et al., 1995; Georges et al., 1996). According to the PSCT, the energy difference between metal and insulator vanishes quadratically for $U \to U_c^0$ while the specific heat diverges,

$$E_i^0(U) - E_m^0(U) = \frac{a}{2}(U - U_c^0)^2; \quad \gamma(U) = \frac{\gamma_0}{U_c^0 - U} \,. \tag{3.35}$$

Since the PSCT is reliable only for exponents, not for absolute prefactors, $a$ and $\gamma_0$ should for the moment be regarded as unknown parameters. Equating the free energies, we conclude

$$0 = \Delta F(T, U_c) = \frac{a}{2}(U - U_c^0)^2 + \frac{1}{2}\gamma(U)T^2 - TS_0, \tag{3.36}$$

to obtain the low-temperature solution,

$$\boxed{U_c(T) = U_c^0 - \sqrt{\frac{2S_oT}{a}} + \mathcal{O}(T) \,.} \tag{3.37}$$

Consequently, the temperature derivative of the solution has a square root divergence at low $T$,

$$\frac{\mathrm{d}U_c(T)}{\mathrm{d}T} = -\sqrt{\frac{S_o}{2aT}} + \mathcal{O}(1) \,. \tag{3.38}$$

[40] Note that the temperatures under consideration are so small that exponential corrections can be safely neglected (cf. subsection 3.5.2).

Since the $U$ dependence in (3.37) is regular for $T \to 0$ (as long as $U_c^0$ is finite), the same must apply to the fit function $A(T)$ defined via (3.30) and (3.32), i.e.

$$A(T) = -\sqrt{\frac{S_o}{2a}}\, T^{-1/2} + \mathcal{O}(T^0)\,. \tag{3.39}$$

Using the value of $S_0 = \ln 2$, this form can be made quantitative by evaluating the parameter $a$ defined in (3.35) which equals the change of slope of $D$ at $T = 0$ and $U = U_c^0$ $(= U_{c2}^0)$. The PSCT yields a parameterization for the double occupancy of the metal as $D_m(T = 0) \approx 0.015 + 0.235(1 - U/U_c^0) \approx 0.25 - 0.0402\,U$, where $U_c^0 = 5.84$ [Georges et al. (1996), p. 71]. Since a similar parameterization for the insulator is not published, one might neglect the corresponding (much smaller) slope of $D_i^0(U)$ in a first approximation.[41] We found discrepancies when trying to reconcile the corresponding estimate for $a$ with our QMC estimates for $A(T)$ at higher $T$, even when realistic estimates for the slope of the double occupancy in the insulator were used. In fact, it is hardly conceivable that the above ground state estimate for $D_m$ could be correct. The problem is that it extrapolates exactly to the correct value $D = 0.25$ for $U = 0$, but not with the correct slope $-0.0417$ to be derived below. Thus, the correctness of the cited parameterization near $U_c^0$ would imply the existence of at least two turning points of $D(U)$ within the metallic phase which appears unlikely.

**Second Order Perturbation Theory and Beyond**

Before we derive an independent estimate for the coefficient $a$ on the basis of our finite-temperature QMC data, we would like to evaluate $D(U)$ within 2OPT. For a general symmetric DOS, the ground state energy at half filling is given by (van Dongen, 2000)

$$E(U) = -\frac{8}{2\pi} + \frac{U}{4} - U^2 \int_0^\infty \mathrm{d}x\, g^4(x) + \mathcal{O}(U^3); \qquad g(x) := \int_0^\infty \mathrm{d}\omega \rho(\omega) \mathrm{e}^{-\omega x}\,. \tag{3.40}$$

For the Bethe DOS, the function $g(x)$ can be expressed in terms of the modified Bessel function of the first kind and the modified Struve function,

$$g(x) = \frac{I_n(2x) - L_1(2x)}{2x} \overset{x\to\infty}{\longrightarrow} \frac{1}{\pi x}\,. \tag{3.41}$$

The asymptotic behavior of $g$ [which is most easily seen directly from (3.40)] can be used in the integral for $x_0 \gtrsim 8$.[42] The remaining finite integral is easily evaluated numerically with the total result

$$E_{2\mathrm{OPT}}(U) \approx -0.848826 + 0.25\,U - 0.020866\,U^2\,, \tag{3.42}$$

[41] This approach was apparently chosen when Joo and Oudovenko (2001) interpolated $T = 0$ ED data and rescaled IPT data for deriving an approximate fit form for $U_c(T)$.

[42] We checked the accuracy of this approach by computing estimates for the total integral with different integer intermediate points $x_0$ above which the approximation for the Struve function is used. For $x_0 = 1 \ldots 10$, the quadratic coefficient of $E(U)$ converges as $-0.022$, $-0.0209$, $-0.02087$, $-0.020868$, $-0.020867$, $-0.0208663$, $-0.0208662$, $-0.0208662$, $-0.0208662$.

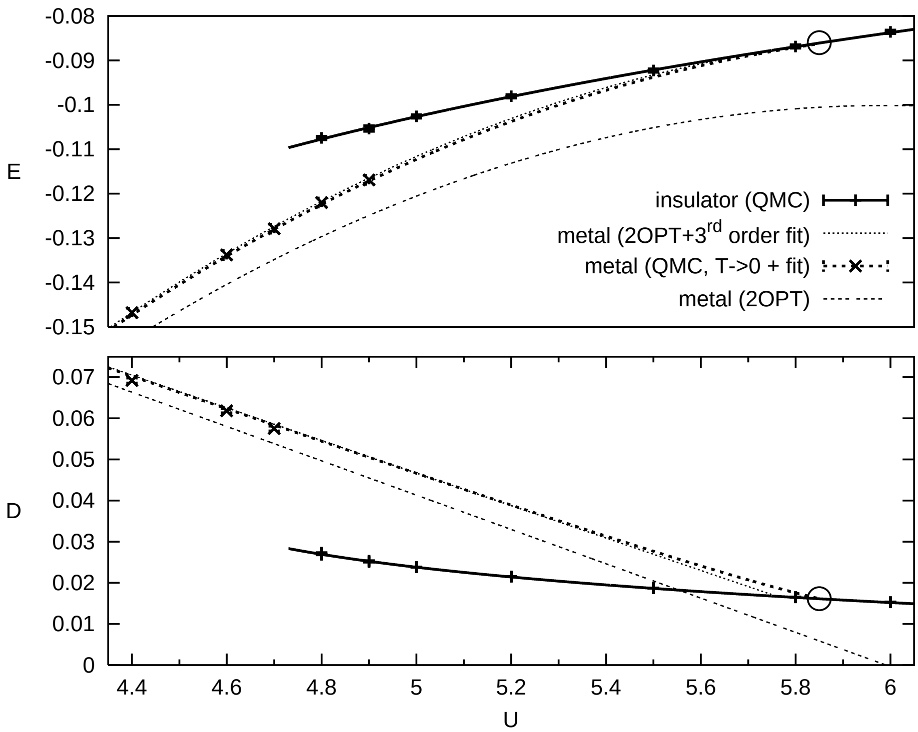



**Figure 3.45:** Ground state energy $E$ and double occupancy $D$ near the MIT. Due to the temperature-independence in the insulator, both observables can be measured using QMC without extrapolation for $T$ in the insulating phase (upright crosses and solid lines). The accuracy of second-order perturbation theory (2OPT, thin lines) for the metallic phase can be improved significantly by using the third-order fit (3.44) for $E_{\mathrm{m}}(U)$ and the consistent expression (3.45) for $D_{\mathrm{m}}(U)$ (short-dashed lines) so that $D$ and $E$ match for insulator and metal at $U = U_{\mathrm{c}}^0 \approx 5.85$ (circles) as predicted by PSCT. At these points, the fits yield $D = 0.01613$ and $E = -0.086$. These fits are supported by extrapolations of the QMC data for the metallic phase to $T = 0$ (slanted crosses). The dotted lines correspond to single-parameter plain third order fits (see text).

where all digits are significant and corrections are $\mathcal{O}(U^3)$. The corresponding estimate for the double occupancy is

$$D_{\mathrm{2OPT}}(U) \approx 0.25 - 0.041732\, U\,, \tag{3.43}$$

with corrections of $\mathcal{O}(U^2)$.

The important and somewhat surprising observation that we are heading for is that bare second-order perturbation theory gives a quantitatively reasonable description for $E$ and $D$ in the metallic phase at $T = 0$ almost up to the transition at $U_{\mathrm{c}}^0 \approx 5.84$.[43] This is seen in Fig. 3.45. Here, the solid lines represent the (temperature-independent) QMC estimates for the energy and double occupancy (upper and lower plot, respectively) discussed in the previous section. Assuming that

[43] More precisely, 2OPT would break down at $U = 5.9905$ where the its estimate for $D$ becomes negative.

the PSCT, NRG, and ED estimates for $U_c^0$ are correct (cf. Table 3.1), the MIT occurs for values of $U$, $D$, and $E$ indicated by the large circles. These points are evidently missed by the 2OPT solutions (faint thin lines); when compared to the noninteracting values, the deviations both in $D$ and $E$ are, however, quite small (about 4% and 2%, respectively). Since the curves $E(U)$ and $D(U)$ are also related nontrivially by $D(U) = \mathrm{d}E(U)/\mathrm{d}U$, the construction of a consistent, smooth fit that coincides with 2OPT for small $U$ and agrees with the insulating solutions for $U = U_c^0$ is a well-conditioned task. The correction to $E$ has to be of third order in $U$ (or higher) in order not to change the asymptotically correct 2OPT result for $U \to 0$. Using a plain 3rd order correction, the requirement that both $E$ and $D$ match for metal and insulator at a single point $U_c$ is fulfilled for $E_m(U) = E_{2OPT}(U) + 0.0000715\, U^3$, $D_m(U) = D_{2OPT}(U) + 0.0000715 \cdot 3\, U^2$ with $U_c^0 = 5.76$ (dotted line in Fig. 3.45). A smooth modification that also satisfies that $U_c^0 \approx 5.85$ (as computed within NRG, PSCT, and ED) is given by

$$E_m(U) = E_{2OPT}(U) + 0.000062\, U^3\big(1 + \frac{0.1}{6.5 - U}\big) \tag{3.44}$$

$$D_m(U) = D_{2OPT}(U) + 0.000062 \left(3\, U^2\big(1 + \frac{0.1}{6.5 - U}\big) + 0.1\, \frac{U^3}{(6.5 - U)^2}\right). \tag{3.45}$$

These fits are shown as short-dashed lines in Fig. 3.45. Both fit curves agree very well with extrapolations of QMC data for $E$ to $T \to 0$ and reasonably well with extrapolations for $D$. The change of slope of $D$ at $U_c$ implied by (3.45), (3.43), and (3.19) is $a = 0.0207$. Consequently, the asymptotic behavior of the fit function $A(T)$ defined via (3.30) and (3.32) is

$$A(T) = -4.1\, T^{-1/2} + \mathcal{O}(T^0)\,. \tag{3.46}$$

**Low-Temperature Specific Heat and Quasiparticle Weight**

Before we construct a global fit for $A(T)$ and, ultimately, for $U_c(T)$ in subsection 3.6.3, we would like to detail the extrapolation of the QMC data to $T = 0$ which was used in Fig. 3.45. The symbols in Fig. 3.46 show QMC estimates for $E$ in the metallic phase at low temperatures for various values of the interaction $U$. The solid lines represent fits of the form

$$E_m(T, U) \approx E_m^0(U) + \frac{1}{2}\gamma(U)\frac{T^2}{1 + 125\, T^2}\,, \tag{3.47}$$

i.e., with the two free parameters $E_m^0$ and $\gamma$ for each curve. The numerical parameter 125 which controls the curvature was determined for $U = 4.6$ where the QMC data is particularly precise; the resulting functional form is clearly consistent with the remaining data as well.

The accuracy of the estimate for the linear coefficient $\gamma(U)$ of the specific heat obtained by this procedure may be checked by a comparison with the quasiparticle weight $Z$ using the Fermi-liquid relation (valid for semi-elliptic DOS)

$$Z(T = 0, U) = \frac{2\pi}{3\,\gamma(U)}\,. \tag{3.48}$$

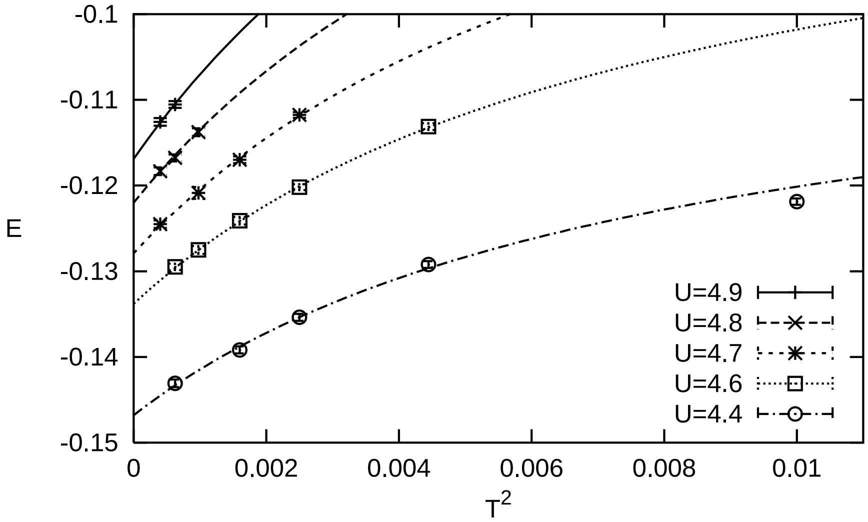



**Figure 3.46:** QMC estimates for the energy $E$ at low temperatures (at $\Delta\tau = 0$) vs. $T^2$ (symbols) and extrapolation $T \to 0$ (lines). The slope at $T = 0$ gives an estimate of one half of the linear coefficient in the specific heat, $\gamma$.

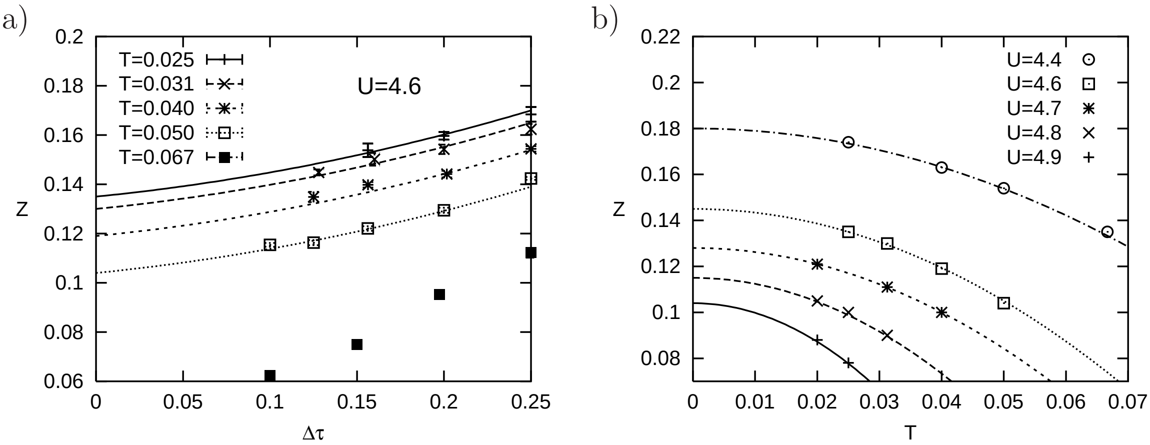



**Figure 3.47:** Extrapolation of quasiparticle weight $Z$, **a)** for $\Delta\tau \to 0$ and **b)** for $T \to 0$.

---

Since the discrete estimate (3.6) for $Z$ becomes exact for $T \to 0$, the QMC computation of $Z$ at low $T$ and its extrapolation to $T = 0$ is well-conditioned. The first step of a QMC estimate, however, the extrapolation to $\Delta\tau = 0$ has to be performed with care as we illustrate for $U = 4.6$ in Fig. 3.47 a. Clearly, the QMC data at finite $\Delta\tau$ (symbols) shows a regular dependence on $\Delta\tau$ behavior only at small enough $T$, sufficiently far below the MIT. Furthermore, numerical noise makes a reliable determination of the extrapolation law difficult, especially at small $T$ (where less accurate data for small $\Delta\tau$ is available). We found that for a large range of $U$ and $T$ a fixed extrapolation law $Z(T, U, \Delta\tau) \approx Z(T, U) + 0.07\,\Delta\tau + 0.28\,\Delta\tau^2$ is sufficiently accurate. The resulting QMC estimates for $\Delta\tau = 0$, shown in Fig. 3.47 b (symbols), may be

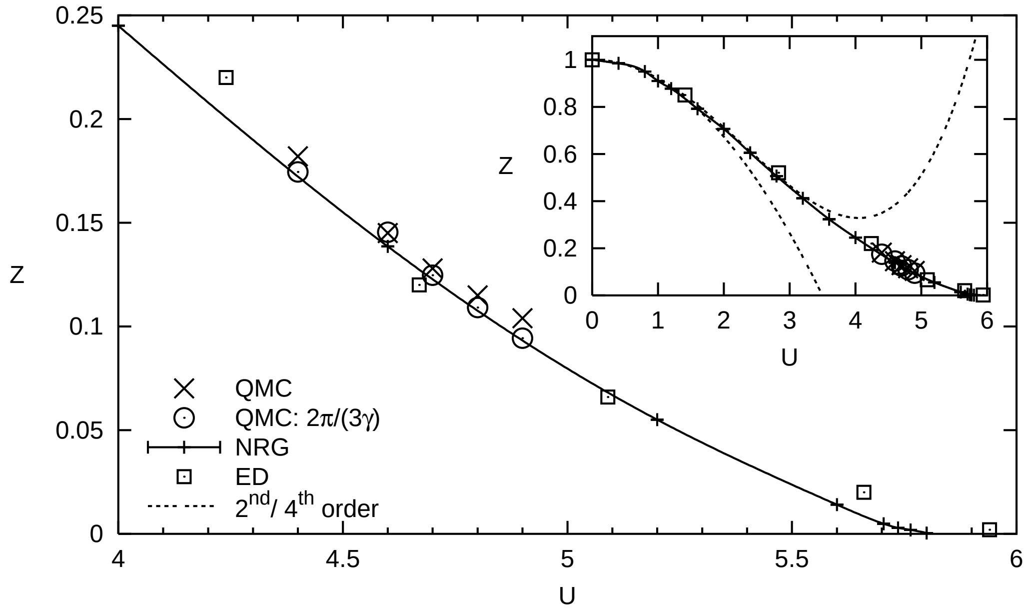



**Figure 3.48:** Quasiparticle weight $Z$ at $T = 0$: QMC Eliashberg estimate for $Z$, extrapolated for $T \to 0$ (crosses), and estimate inferred from QMC fit for $\gamma$ (cf. Fig. 3.46) in comparison with ground state results from perturbation theory (Noack and Gebhard, 1999), NRG (Bulla, 1999), and ED (Caffarel and Krauth, 1994).

extrapolated to $T = 0$ using a purely quadratic fit in $T^2$ (lines).

The actual check of internal consistency is performed in Fig. 3.48. Here, large tilted crosses denote the direct QMC estimates of $Z$ via the discrete difference expression (3.6) (in terms of the self-energy at the first Matsubara frequency) using the extrapolations illustrated in Fig. 3.47. Circles indicate the QMC estimates following from (3.48) and from the values for $\gamma$ implied by the fits shown in Fig. 3.46. The good agreement between these sets of data confirms the internal consistency of our QMC estimates. Their reliability is further supported by the good agreement with estimates for $Z$ obtained from NRG calculations (Bulla, 1999) (small crosses and solid line) and ED calculations (Caffarel and Krauth, 1994). The accuracy of these methods is, in turn, supported by perturbation theory of $2^{\mathrm{nd}}$ and $4^{\mathrm{th}}$ order [Noack and Gebhard (1999); see inset of Fig. 3.48]. The verification of the correct asymptotic slope of the fit curves shown in Fig. 3.46 implies the absence of significant systematic errors for the extrapolated ground-state energies. Therefore, the small deviations between the QMC estimates for $E$ and the corresponding fit curve shown in Fig. 3.45 strongly support the accuracy of our ground state expressions (3.44) for $E_{\mathrm{m}}(U)$ and (3.45) for $D_{\mathrm{m}}(U)$ as well as the resulting estimate for $a$.

### 3.6.3 Full Phase Diagram

So far, we have in this section derived a differential equation for the thermodynamic first-order transition line $U_{\mathrm{c}}(T)$, i.e., $\mathrm{d}U_{\mathrm{c}}(T)/\mathrm{d}T = f(T, U)$, seen that a linear

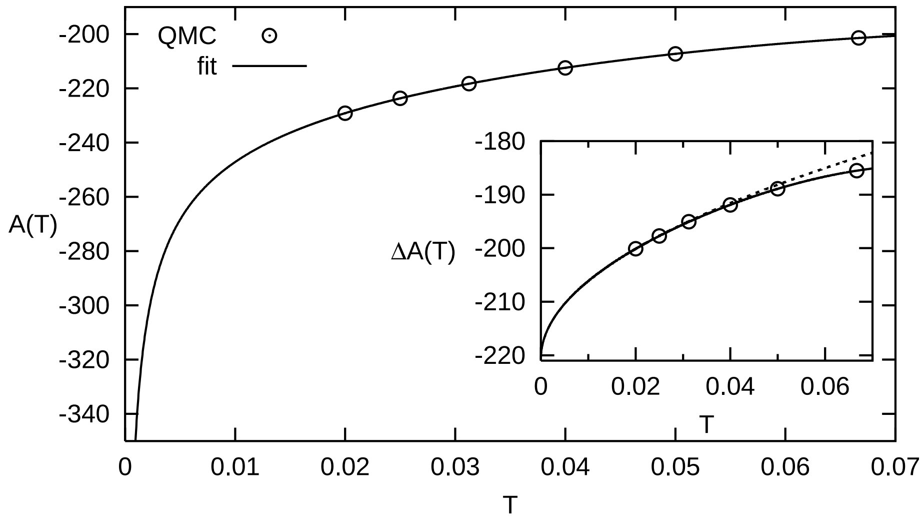



**Figure 3.49:** QMC results and fit for $T$-dependence of $A(T)$, used in the linearized differential equation for $U_c(T)$ (see text). The correction to the asymptotic low-temperature result $A(T) \sim 4.1/\sqrt{T}$ is shown in the inset. Here, the additional short-dashed line indicates the correction without $T^4$ term, i.e., the sum of the constant and the square-root term.

approximation $f(T,U) = A(T) + 43\,U$ is reasonably accurate, and determined the asymptotic low-temperature behavior for $A(T)$. The remaining task is finding an accurate interpolating fit form for $A(T)$ which can then be solved with the suitable boundary condition. Using the observation that $A(T) = \sqrt{\ln 2/(2a)}\, T^{-1/2} + \mathcal{O}(T^0)$ it might seem most natural to successively include higher and higher orders of $T^{1/2}$ in a parameterization of $A(T)$ until it matches the QMC data for the whole coexistence region. It is, however, not clear that few orders suffice for a description up to the critical end point where a perturbation series might even break down. On the other hand, we only need a smooth *inter*polation of QMC data at high temperatures; an potentially unstable *extra*polation is only needed below the lowest QMC data point at $T = 0.02$. Specifically, we have chosen the parameterization

$$A(T) = -4.1T^{-1/2} + A_1 + (A_2T)^{1/2} + (A_3T)^4 \tag{3.49}$$

with

$$A_1 = -220.68\,, \qquad A_2 = 21170\,, \qquad A_3 = 18.7\,, \tag{3.50}$$

i.e., introduced the new parameters $A_1$ and $A_2$ for the two subleading terms in a systematic expansion plus a single term of much higher order parameterized by $A_3$. This fit is shown as solid line in Fig. 3.49. Evidently, it matches the QMC data (circles) quite accurately. In the inset, the asymptotic square-root singularity has been subtracted from $A(T)$. Even at this scale no deviations of the full fit (solid line) from the QMC data is visible. In contrast, a fit including only terms up to order $T^{1/2}$ (i.e., with $A_3 = 0$) can only reproduce the low-temperature data (dashed line).

Generally, the solution of a linear differential equation can be written as a regular integral. Specifying the initial condition as $U_c(T=0) = U_c^0$, we have here

$$U_c(T) = e^{43T}\left(U_c^0 + \int_0^T dT' e^{-43T'} A(T')\right). \tag{3.51}$$

With $A(T)$ given by (3.49) and the parameters stated above, this integral can even be performed analytically. Since the derivation is both lengthy and straightforward, it is omitted here. Adjusting the initial condition so that the solution goes through the tip of the coexistence region leads to $U_c^0 \approx 5.795$. This value is smaller than the value $U_c^0 \approx 5.85$ obtained from PSCT, NRG, and ED and used in our ground state fits for $E$ and $D$. In our view, the observed discrepancies reflect the limits of the linear approximation for the dependency of $f(T, U)$ on $U$. While deviations from the linear form are small within the coexistence region, i.e., in the range $4.7 \lesssim U \lesssim 5.0$, this need not be the case for $U \to U_c^0$. Still, the deviation of about 6% (if measured from the last direct QMC data point at $U \approx 5$) is surprisingly small considering the various potential sources of errors.[44]

The full phase diagram of the half-filled Hubbard model with semi-elliptic DOS in $d \to \infty$, the main result of this chapter, is shown in Fig. 3.50. Here, crosses and thin lines indicate the boundaries of the crossover and coexistence region as already seen in Fig. 3.31. The true thermodynamic first-order phase transition line computed in this section, i.e., the solution of (3.51) for the fit defined by (3.49) and (3.50) is drawn as a thick line in Fig. 3.50. It is fully based on QMC data at least in the range $T \geq 0.02$ where direct QMC estimates for $f(T, U)$ are available. The dashed line style used below this range is a reminder that the PSCT/ED/NRG estimate of $U_c^0 \approx 5.85$ was used at some stage of the derivation of the low-temperature form of $U_c(T)$. The various numerical checks performed on the basis of extrapolated QMC data suggest, however, that the impact of possible errors in this external input on our final result is minimal.

A final test of consistency can be performed on the basis of the full solution $U_c(T)$ by computing an additional term in its low-temperature asymptotics. Then, (3.37) is refined as

$$U_c(T) = U_c^0 - \sqrt{\frac{2S_oT}{a}} + \frac{\gamma_0}{4S_0}T + \mathcal{O}(T^{3/2}), \tag{3.52}$$

where all symbols are known from subsection 3.6.2. Thus, the linear coefficient of $U_c(T)$ can be compared with asymptotic ground state results for the linear coefficient of the specific heat $\gamma(U)$ and [using $\gamma = 2\pi/(3Z)$] also with ground-state estimates of the quasiparticle weight $Z$. This comparison is shown in Fig. 3.51. In Fig. 3.51 a, the coefficients of the contributions of order $T$ and $T^{3/2}$ to $U_c(T)$ are determined as 28 and $-100$, respectively. Consequently, consistency demands that $\gamma(U) \approx 72/(U_c^0 - U)$

[44] An even better agreement at small $T$ could have been achieved by using a larger leading coefficient of the square-root divergence in $A(T)$, i.e., a smaller value of $a$. The latter would be consistent with the results of our QMC extrapolations to $T = 0$ if the curvature of the double occupancy was strongly enhanced near $U_c^0$. Such behavior is certainly possible as seen in the NRG curve for $Z(U)$ in Fig. 3.48.

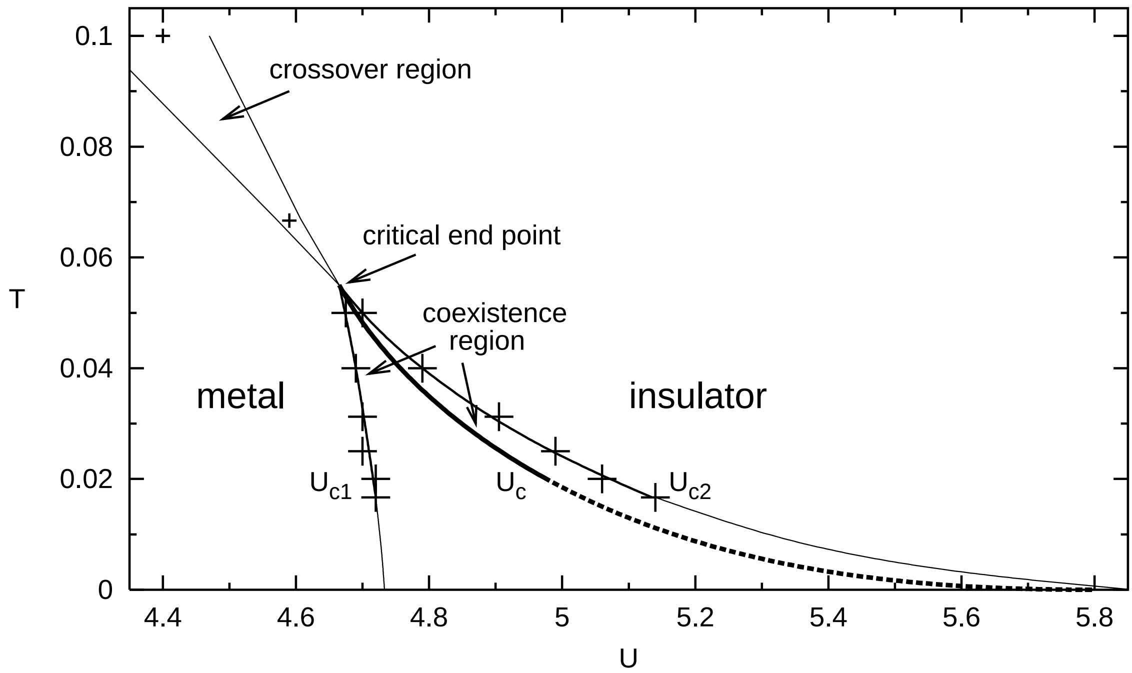



**Figure 3.50:** Final MIT phase diagram: Above the critical point (with coordinates $U^* = 4.665$ and $T^* = 0.055$), increasing interaction $U$ causes a smooth transition from a metal towards an insulator; the change is most rapid within the crossover region. For $T < T^*$, the metallic phase is thermodynamically stable for $U \leq U_\mathrm{c}$, but remains a well-defined metastable solution of the DMFT equations for $U_\mathrm{c} < U \leq U_\mathrm{c2}$. Conversely, a well-defined insulating phase solves the DMFT equations for $U \geq U_\mathrm{c1}$, but is thermodynamically stable only for $U \geq U_\mathrm{c}$. The QMC data (crosses) and thin blue lines are identical to Fig. 3.31. The thick blue line (partly dashed) represents the QMC result for $U_\mathrm{c}(T)$ obtained in this section. Faint grey lines extrapolate the QMC data for $U_\mathrm{c1}$ and $U_\mathrm{c2}$ to $T = 0$.

for $U \to U_\mathrm{c}^0$. The agreement with $T = 0$ NRG data is reasonable. Note that the curvature of $\gamma^{-1}(U)$ has to be large near $U_\mathrm{c}^0$ in order to reconcile the data in the range $U < 5.2$ with a value of $U_\mathrm{c}^0 \approx 5.85$. Therefore, the slope near $U_\mathrm{c}^0$ is necessarily much smaller than at lower $U \lesssim 5.2$. Still, the NRG predictions near $U_\mathrm{c}^0$ seem a bit low, in particular when interpreted in terms of $\gamma^{-1}(U)$ which might explain the remaining discrepancy in comparison to our linear fit. So our results have indeed passed a nontrivial test.

While our method of determining the position of the first-order phase transition line on the basis of QMC data represents the first serious successful attempt of a controlled comparison of free energies for Hubbard-like systems, the first publication of a quantitatively correct result for $U_\mathrm{c}(T)$ is due to Tong et al. (2001). Their results for $U_\mathrm{c1}$, $U_\mathrm{c}$, and $U_\mathrm{c2}$, shown as circles in Fig. 3.52 became publicly known (as preprint) in late May 2001, about half a year after our first results had been presented (cf. footnote 29 on page 111). With the exception of our lower estimates for $U_\mathrm{c1}(T)$, the agreement with our final QMC results (crosses and blue lines) appears nearly

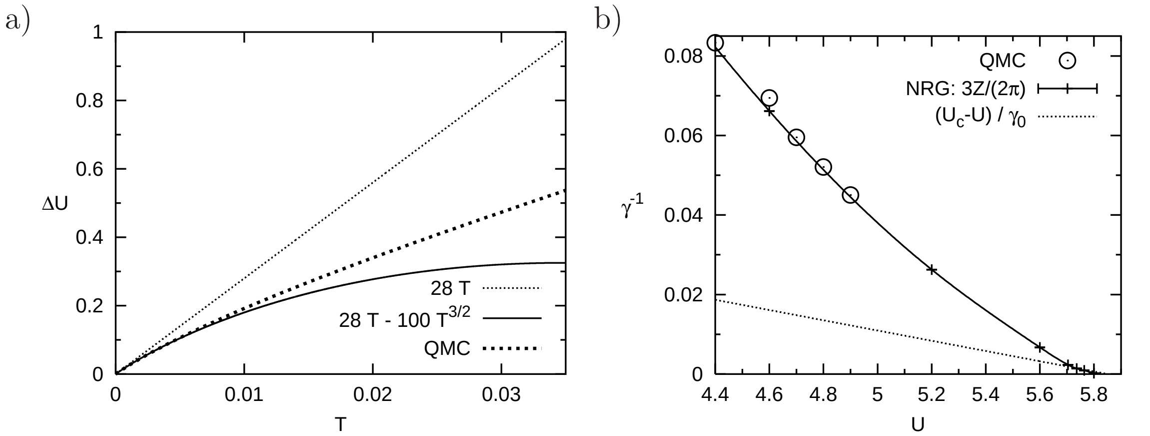



**Figure 3.51: a)** Linear and higher order contributions to the QMC estimate of the first-order phase transition line, i.e., $\Delta U(T) := U_c(T) - 5.7962 + 8.2\sqrt{T}$. **b)** The resulting asymptotic for $\gamma(U)$ (with $\gamma_0 \approx 104 S_0 \approx 72$) in comparison with more direct low-$T$ NRG and QMC estimates (cf. Fig. 3.48).

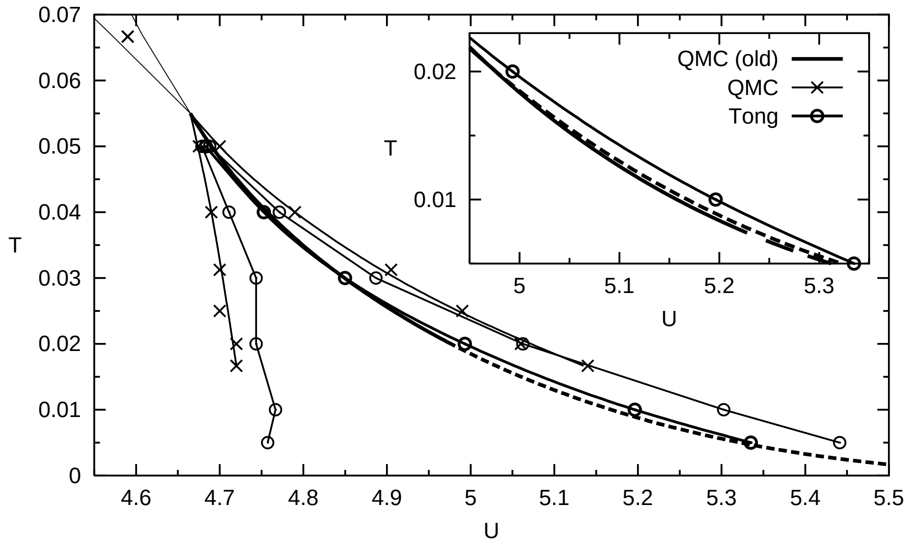



**Figure 3.52:** MIT phase diagram: comparison of QMC results (crosses) with Tong et al.'s (2001) estimates for $U_{c1}(T)$, $U_c(T)$, and $U_{c2}(T)$ (circles). In the magnified inset, also the QMC result of 2000 (red bottom line) is included (cf. Fig. 3.32 b).

perfect. In particular, the curves for $U_c(T)$ are on top of each other for $T \geq 0.025$. The discrepancy for lower temperatures is also small; still, it exceeds the uncertainty of our results. This is illustrated in the inset of Fig. 3.52 where our original estimate of October 2000 is also shown. Even though the inclusion of additional QMC simulations and a refinement of the extrapolation and fitting methods have shifted our estimates of the extent of the coexistence region and of the position of the critical point considerably, the old and new estimate of $U_c(T)$ agree within linewidths in the low-temperature range. In comparison, Tong et al.'s (2001) result deviates markedly.

The excellent overall agreement between both sets of results shown in Fig. 3.52 is even more surprising in view of the fact that the methodology employed by Tong et al. (2001) is not exact. It is based on the replacement of the original DMFT problem by a modified set of self-consistency equations. In order to obtain a result for some interaction $U$, an impurity problem is solved for some different interaction $U'$. Up to a constant offset, the shift $U' - U$ is taken proportional to the midpoint-value of the imaginary-time Green function $G(\beta/2)$ which is here regarded as order parameter (which implies a self-consistency condition for $U'$). Since the absolute value of the order parameter is larger within the metallic phase throughout the coexistence region (with decreasing difference towards the critical end point), the overlap in terms of $U'$ can be reduced until it vanishes by choosing the proportionality constant sufficiently large. The solutions of the modified self-consistency problem then becomes not only unique in terms of $U'$, but a new solution appears which continuously spans the gap in terms of $U'$ between the solutions associated with the coexisting solutions of the original problem. For the modified problem, all observables are smooth functions of $U'$. When folding back the results into the original parameter space (i.e., by plotting the results in terms of $U$ rather than $U'$), the additional solution is interpreted as a third, thermodynamically unstable, phase; observables measured for the modified problem appear as smooth s-shaped curves. This is seen in in Fig. 3.53, where full symbols represent Tong et al.'s (2001) results for the double occupancy. Apart from the square-root singularity of the metallic solution at $U_{c2}$ suggested by this method and from the additional unstable solution, the main difference in comparison to our QMC results is a shift towards larger values of $D$. This discrepancy is highly significant deep in the insulating phase (say: for $U \geq 5.6$) where the error of the QMC results is of the order $10^{-4}$, i.e., within linewidth. At the ground state transition point $U_c^0 \approx 5.85$, Tong et al. predict $D \approx 0.02$ which is far above our estimate $D \approx 0.0161$ and the PSCT result $D \approx 0.015$. Furthermore, the temperature dependence in this region is several orders of magnitude too large and is of the wrong sign. This observation already points to the main problem of the method: it cannot be exact since the impurity model is in general solved at a shifted interaction $U'$ instead of $U$. As noted in Sec. 4.6, the second moment of the self-energy is independent of the hybridization and only depends on the interaction on the impurity, equaling $U^2/4$. For the Bethe DOS, this second moment also determines the second derivative of $G(\tau)$ at $\tau = 0$. Consequently, the measurements of $G(\tau)$ performed by Tong et al. must suffer from systematic errors due to shifting the interaction, regardless of details of the self-consistency scheme (which influences the hybridization function). Still, the high accuracy of the final results cannot be accidental. In our view, the qualitative

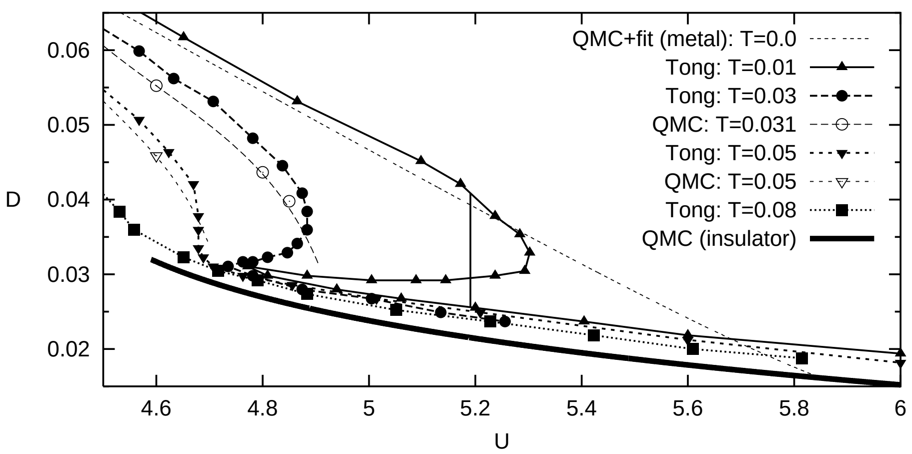



**Figure 3.53:** Double occupancy: Full symbols connected by thick lines represent results obtained using the approximate transformation technique (Tong et al., 2001). The non-unique part of the curves for $T \leq 0.05$ is interpreted as a set of metallic, unstable, and insulating solutions (from top to bottom). The first-order transition can be pinpointed by using a Maxwell-construction which includes the unstable solution (vertical line at $U \approx 5.2$ for $T = 0.01$). Empty symbols and thinner lines show the QMC results for the metallic phase; the thick black line shows the QMC result for the insulating phase (cf. Fig. 3.42).

scenario shown in Fig. 3.53 is plausible, provided that a third solution of the DMFT equations exists. Then, the knowledge of the approximate shape should suffice for a reasonably accurate estimate for $U_c(T)$ within the range spanned by $U_{c1}(T)$ and $U_{c2}(T)$. An even more direct method for the comparison of free energies will be proposed in subsection 3.7.2.

Figure 3.54 shows a comprehensive comparison of results for the MIT in the half-filled fully frustrated Hubbard model with semi-elliptic DOS in $d = \infty$. One important point that should be conveyed by this (admittedly somewhat busy) figure is the convergence of recent estimates for the boundaries $U_{c1}$ and $U_{c2}$ of the coexistence region and the significant improvement over old IPT and QMC results. Thus, the controversy stirred by Schlipf et al.'s (1999) paper has lead to secured knowledge of the properties of this model with unprecedented accuracy. Among the more recent results, those obtained by Joo and Oudovenko (2001) are least precise; in fact, their estimates for $U_{c2}$ are above upper bounds established in this work (cf. Fig. 3.31 and Fig. 3.32 b). Low-temperature information is only provided by Bulla et al. (2001) and Tong et al. (2001) with apparently about equal precision; the agreement with our QMC data for $T \leq 0.025$ is excellent. The origin of the discrepancy between NRG and the other methods near the tip of the coexistence region at $T \approx 0.05$ is unclear; possibly, the systematic NRG error (due to the finite energy discretization $\tilde{\Lambda} = 1.6$) is most significant in this temperature range. Still, the available data clearly supports the high accuracy of our QMC results.

A second point of hardly minor importance is that no qualitatively correct prediction for the first-order transition line preceded our first estimate (cf. Fig. 3.32 b);

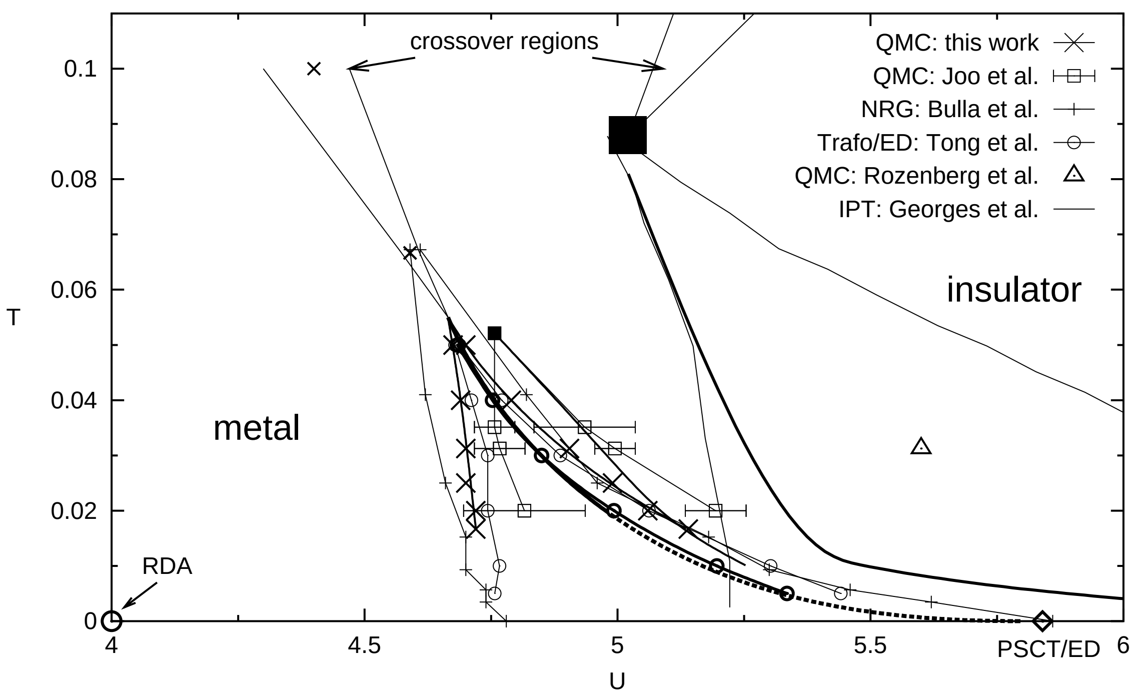



**Figure 3.54:** MIT phase diagram: recent finite-temperature results [this work, Joo and Oudovenko (2001), Bulla et al. (2001), Tong et al. (2001)] in comparison with former estimates [Rozenberg et al. (1994), Georges et al. (1996)]. In the low-temperature region, thinner lines represent $U_{c1}$ and $U_{c2}$. Thick lines show calculated first-order lines $U_c(T)$ [this work, Tong et al., Georges et al.] with the exception of Joo et al.'s estimate [with $U_c(T) \approx U_{c2}(T)$ for $T \lesssim T^*$] which is based on a rescaling of the IPT curve and a fit to ground state data. For ground state data, see Table 3.1; subsets of the finite-$T$ data are shown in Fig. 3.7 b, Fig. 3.9, Fig. 3.32 b, Fig. 3.33, Fig. 3.50, and Fig. 3.52. For Rozenberg et al.'s (1999) QMC results on coexisting phases (not shown here), see Fig. 3.32 a and Fig. 3.33.

since Tong et al.'s method is approximate (see above), our final curve shown in Fig. 3.54 is still the only controlled (and even numerically exact) result. In fact, both the original (cf. Fig. 3.7 a) and the later IPT estimates wrongly implied that $U_c(T) \approx U_{c1}(T)$ for $T \lesssim T^*$. The implication of Joo et al.'s fitting and scaling procedure that $U_c(T) \approx U_{c2}(T)$ for $T \lesssim T^*$ is closer to the truth, but still qualitatively incorrect. Only the recent independent results of this work and by Tong et al. provide significant insight with regard to the position of $U_c(T)$.

### 3.6.4 Implications of Partial Frustration

After having established the MIT scenario in the fully frustrated case (for semi-elliptic DOS), the interesting question arises what the full phase diagram would look like for partial frustration. In this context, we have to distinguish between frustration at constant DOS and the more generic case that the introduction of frustrating bonds

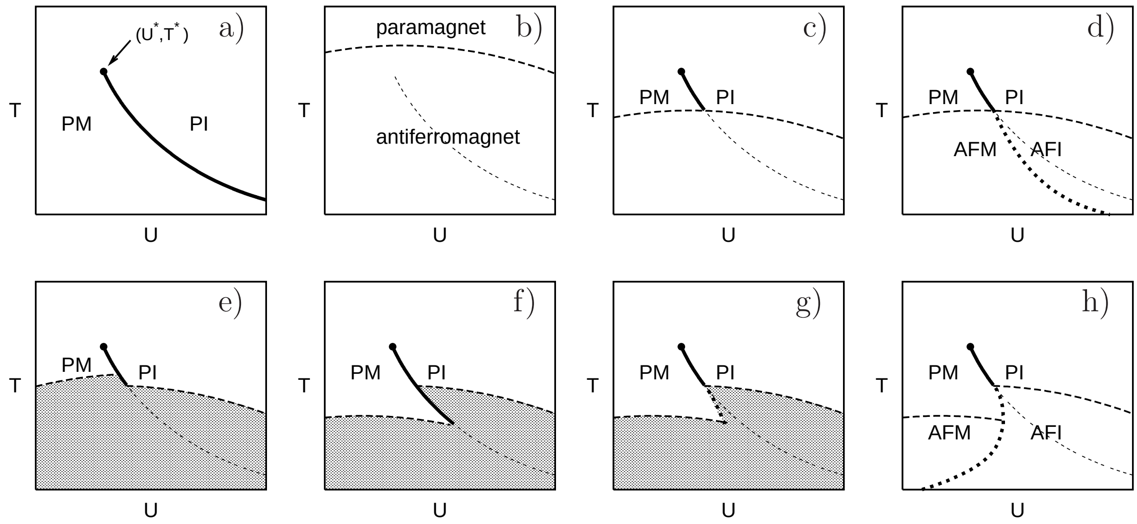



**Figure 3.55:** Schematic phase diagrams of models with semi-elliptic DOS and varying degrees of frustration. **a)** fully frustrated case with first-order paramagnetic MIT (thick solid line) for $0 \leq T < T^*$. **b)** weakly frustrated case with second-order magnetic transition (dashed line); no transitions within paramagnetic or within AF phase. **c)-h)**: stronger frustrated case with $T_N^{\max} < T^*$. **c)** first-order line ends on smooth Néel curve. **d)** first-order MIT crosses smooth Néel curve. **e)** $T_{\rm N}$ higher on metallic side of MIT line; no implications for AF phase (shaded) intended. **f)** $T_{\rm N}$ higher on insulating side of MIT line. **g)** $T_{\rm N}$ higher on insulating side of MIT line; MIT line shifted in intermediate temperature region. **h)** speculative scenario for full phase diagram: a single first-order MIT line extends from $(U^*, T^*)$ to $(0, 0)$; it agrees with the fully-frustrated line at large $T$ (solid line) and deviates below (thick short-dashed line). For the partially frustrated model, only scenarios **b)** and **g)** (with the possible completion **h)**) appear likely (see text). The MIT line of the fully frustrated model is drawn as faint thin line for comparison where it is does not represent a transition.

also leads to an asymmetry of the DOS.

### Frustration at constant semi-elliptic DOS

Within the DMFT, the lattice influences properties of homogeneous phases only via its (noninteracting) DOS. Thus, the thermodynamics of a partially frustrated model with semi-elliptic DOS agrees with that of the fully frustrated model considered so far within its paramagnetic phase.[45] In the following, we will derive the qualitative characteristics of the full phase diagram expected for such a model. For sufficiently small frustration, the first-order MIT line present in the fully frustrated model (Fig. 3.55 a) is completely hidden by the low-temperature AF phase; then, the phase diagram is qualitatively unchanged from the unfrustrated case as shown in Fig. 3.55 b. For

[45] While the claim (Rozenberg et al., 1995) that uniform NNN hopping on the Bethe lattice leaves the DOS unchanged was contradicted in subsection 2.2.2, fully disordered hopping on the NNN bonds of a Bethe lattice constitutes a realization of such a model with variable disorder.

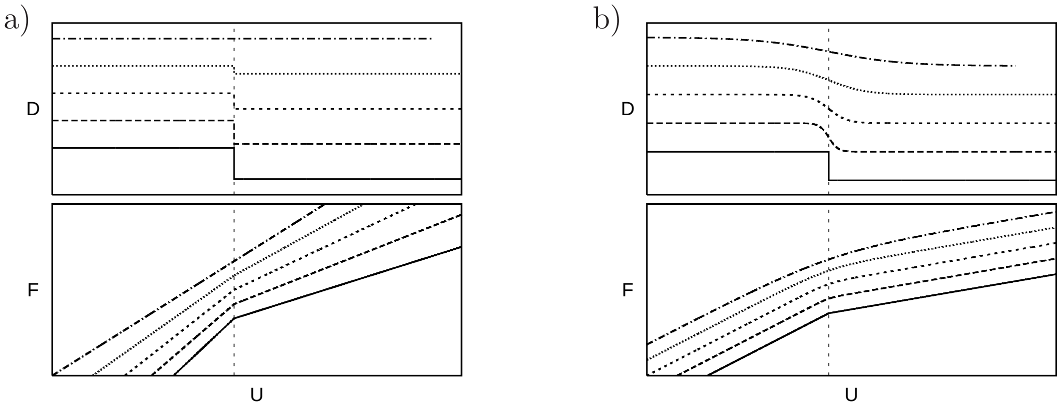



**Figure 3.56:** Scenarios for the ending of a first-order line: **a)** Upon approaching the end point (from bottom to top), the jump in the derivative of $F$ (i.e., the jump in $D$) decays to zero; at the end point, the transition is of second order. **b)** Abrupt resolution of the kink in $F$ or discontinuity in $D$ (bottom lines) into a crossover region (broadening towards the top line).

---

sufficiently strong frustration the maximum Néel temperature $T_{\mathrm{N}}^{\max}$ becomes lower than the critical temperature $T^*$ of the fully frustrated model. Then, a paramagnetic first-order MIT takes place at least for $T_{\mathrm{N}}^{\max} \leq T < T^*$. Assuming a smooth curve $T_{\mathrm{N}}(U)$ leads to a phase diagram as shown in Fig. 3.55 c. However, the energy difference on the MIT line between metallic and insulating solution is then still finite when reaching the second order line towards AF order, i.e., the Néel temperature. Thus, the first-order line cannot simply stop at this point. While the abrupt evolution of a first-order line into a crossover region (the width of which decays to zero towards the end point) as shown in Fig. 3.56 b appears to be thermodynamically consistent, such a scenario requires an abrupt change in $D$ for at least one (here: both) of the solutions at the end point. This, in turn, implies a change of order one in the corresponding solution $G(i\omega_n)$ upon an infinitesimal variation of $U$ and/or $T$ which violates the assumptions of a Ginzburg-Landau theory (cf. subsection 3.7.1). Consequently, we will discard this possibility. In contrast, the crossing of a first-order MIT line with a smooth Néel curve as shown in Fig. 3.56 d is a priori possible.

However, the Néel curve of transitions between paramagnetic and AF phases cannot be assumed to be a smooth, continuous function $T_{\mathrm{N}}(U)$ at a first metal-insulator transition. This follows from the fact that there are no precursor effects of magnetic order within the DMFT. Consequently, the antiferromagnetic susceptibility is determined by local properties of the homogeneous phase alone. Since the metallic and insulating phases are fundamentally different at the MIT and correspond to different fixed points of the self-consistency equations, it is hardly conceivable that the antiferromagnetic susceptibility diverges in both phases for exactly the same temperature $T$ [on the MIT line $U_{\mathrm{c}}(T)$]. As a result, we must expect a discontinuity (at least of the first derivative) or a non-uniqueness of the function $T_{\mathrm{N}}(U)$ at the MIT as shown in Fig. 3.56 e and Fig. 3.56 f. In the insulating phase, $D$ is smaller so that the local moment (with $\langle (n_\uparrow - n_\downarrow)^2 \rangle = 1 - 2D$ for half filling) is larger; consequently, the ordering temperature should be higher in this phase which favors Fig. 3.56 f over Fig. 3.56 e. In both cases, the first-order MIT line becomes part of the boundary of the AF phase in the intermediate temperature range.

Still, the assumption implicit in Fig. 3.56 f (and Fig. 3.56 g) that the MIT line remains unchanged from the fully frustrated case even when one of the adjacent phases is antiferromagnetic is clearly wrong. Considering that the free energy decreases in the AF phase with increasing order parameter, i.e., decreasing $T$, the phase with the higher Néel temperature must widen so that the MIT line shifts as illustrated in Fig. 3.56 g. We stress that the continuity of $D$ and $E$ within the metallic phase along this line and across the second-order Néel curve within the insulating phase implies a continuous first derivative of the MIT curve $U_c(T)$. In principle, a second-order end point of the first-order MIT line at the lower Néel temperature appears possible; however, at least a discontinuity of the order parameter would have to continue towards lower $T$ or to evolve into a crossover region analogous to the case discussed above in connection with Fig. 3.56 c; furthermore, such a coincidence appears nongeneric. Generically, one has to expect a continuation of the first-order line within the AF phase as shown in Fig. 3.56 h. This diagram contains five different phase transitions; among these, all metal-insulator transitions are of first order (i.e., PM-PI, PM-AFI, and AFM-AFI) while the remaining transitions (PM-AFM and PI-AFI) are of second order. The first-order line is continuously differentiable. The topology of this diagram agrees with the phase diagram proposed by Rozenberg, Kotliar, and Zhang [Fig. 43 in Georges et al. (1996)]; however, the details of the latter, in particular the ending of the first-order MIT in the paramagnetic phase on a (roughly perpendicular) Néel curve, are thermodynamically inconsistent.

The slope of the MIT line $U_c(T)$ at low $T$ depends on the sign of the ratio $\Delta E/\Delta D$ of differences in energy and double occupation between the AFM and AFI phases (cf. Sec. 3.6). In contrast to the PM-PI transition, both signs appear possible for the AFM-AFI transition (possibly depending on the degree of frustration); the positive slope shown in Fig. 3.56 h for $T \to 0$ reflects an arbitrary choice rather than a prediction. Note that our scenario is consistent with ground state ED calculations performed for a frustrated model with semi-elliptic DOS (Chitra and Kotliar, 1999). For increasing frustration, both Néel temperatures along the MIT line are reduced. Then, the AFM phase might disappear before the AFI phase is completely suppressed. In any case, the full paramagnetic MIT line calculated in the previous section is recovered as soon as AF order is destroyed for all $U < U_{c2b}^0$.

**Frustration by longer-range hopping**

In generic cases, longer-range hopping suitable for causing magnetic frustration also leads to an asymmetric DOS, at least in absence of disorder (cf. subsection 2.1.3 and subsection 2.2.2). In the following, let us assume that $t'^*$ parameterizes such type of hopping and, consequently, the degree of frustration and the modification of the DOS. Several scenarios come to mind:

1. The Mott transition disappears (even in a computation for the homogeneous

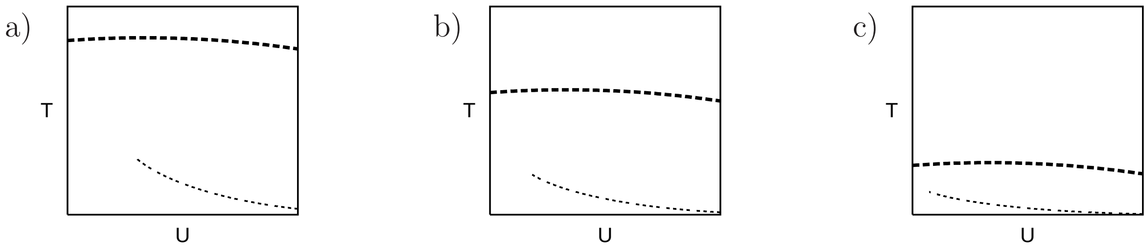



**Figure 3.57:** Schematic phase diagram for frustration by longer-range hopping. **a)** In the unfrustrated case, a paramagnetic MIT line (faint thin line) is present only for a calculation restricted to the paramagnetic phase; in an unrestricted calculation, it is hidden by the low-temperature AF phase (below the Néel curve, dashed). **b)** and **c)**: Increasing frustration suppresses both the AF phase and the MIT line (obtained in a restricted calculation) so that the latter never becomes thermodynamically relevant.

phase) and a smooth crossover develops for any $|t'^*| > 0$.

2. The first-order Mott transition survives sufficiently strong finite microscopic frustration with no or only a minor change of the critical temperature $T^*$ so that $T_\mathrm{N} < T^*$.

3. The Mott transition coexists with NNN hopping, but with $T^* < T_\mathrm{N}(U^*)$ for all hopping strengths $t^*$ as illustrated in Fig. 3.57 (with increasing frustration from left to right). The MIT line (computed in the enforced homogeneous phase) never reaches the true homogeneous phase.

Since the first suggestion implies changes of order one in solutions of the DMFT equations for an infinitesimal change of the DOS, it may be safely excluded. The second scenario leads to the same conclusions as the case of frustration at constant DOS studied above. The third scenario comprises only a single phase transition line, the Néel curve. In fact, this least interesting possibility seems to apply at least for the DMFT treatment of the cubic lattice with NNN hopping. Here, $T^*$ is strongly suppressed (from about $T^* = 1/25$ to $T^* = 1/35$) by introducing moderate frustration ($t'^*/t^* = -0.25$) according to recent QMC results (Knecht, 2002).

## 3.7 Landau Theory and Criticality

In this section, we will study some implications of Landau theory. In subsection 3.7.1, we introduce an explicit representation of the lattice contribution to the Ginzburg-Landau free-energy functional for semi-elliptic DOS. Using this formalism, we show that each step of a direct iteration of the DMFT equations proceeds downhill, i.e., opposite to the gradient (with respect to the hybridization function) of the free energy. Based on this finding, a new method for evaluating free energy differences between coexisting phases within the DMFT is proposed in subsection 3.7.2. In subsection 3.7.3, we discuss to which degree critical behavior should impact or become visible in numerical calculations as performed in this work.

### 3.7.1 Free Energy Functional for the Bethe Lattice

Within the DMFT, the free energy per particle can in the general homogeneous case be written as (Georges, Kotliar, and Si, 1992)

$$F = F_{\text{imp}} - 2T \sum_n \Big[ \ln G(i\omega_n) + \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \rho(\epsilon) \ln\big(i\omega_n + \mu - \Sigma(i\omega_n) - \epsilon\big) \Big] . \tag{3.53}$$

Here, $F_{\text{imp}}$ is the impurity part corresponding to (1.34). We wish to extract an explicit expression for the lattice contribution. For the Bethe lattice, we have $i\omega_n + \mu - \Sigma(i\omega_n) = t^{*2} G(i\omega_n) + 1/G(i\omega_n)$. Therefore, the lattice contribution can be explicitly written in terms of the Green function $G_n \equiv G(i\omega_n)$. Its derivative reads

$$-\frac{1}{2T}\frac{\partial}{\partial G_n}(F - F_{\text{imp}}) = \frac{1}{G_n} + \Big(t^{*2} - \frac{1}{G_n^2}\Big) \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \rho(\epsilon) \frac{1}{t^{*2} G_n + G_n^{-1} - \epsilon} = t^{*2} G_n . \tag{3.54}$$

Consequently, the free energy may be expressed for a semi-elliptic DOS as

$$F = F_{\text{imp}} - t^{*2} T \sum_n G_n^2 . \tag{3.55}$$

For consistency with Kotliar et al. (2000) we continue the discussion in terms of the hybridization function $\Delta_n \equiv \Delta(i\omega_n) = i\omega_n + \mu - \mathcal{G}_n^{-1}$. The impurity equation (1.31) is then equivalent (at the self-consistency) to

$$G_n = \frac{1}{2T} \frac{\partial F_{\text{imp}}[T, U, \Delta]}{\partial \Delta_n} . \tag{3.56}$$

Since for the Bethe lattice $\Delta_n = t^{*2} G_n$, the total free energy may also be written as a functional of the hybridization function

$$F[T, U, \Delta] = -T \sum_n \frac{\Delta_n^2}{t^{*2}} + F_{\text{imp}}[T, U, \Delta] . \tag{3.57}$$

The important point is that this Ginzburg-Landau functional $F_{\text{GL}} \equiv F[T, U, \Delta]$ is defined (and real valued) for an arbitrary hybridization function. While the following considerations apply for arbitrary filling, we find it convenient to introduce explicitly real-valued notation which is only valid for half filling. Using the short hand notation $O'' \equiv \operatorname{Im} O$ we then have

$$\begin{aligned} \big(\boldsymbol{\nabla}_{\Delta''} F\big)_n \equiv \frac{\partial F[T, U, \Delta'']}{\partial \Delta''_n} &= -2T\Big(G''_n[T, U, \Delta''] - \frac{\Delta''_n}{t^{*2}}\Big) && (3.58)\\ &= -2T\big(G''_{\text{new}}(i\omega_n) - G''_{\text{old}}(i\omega_n)\big) . && (3.59) \end{aligned}$$

Here, $G''_n[T, U, \Delta''] \equiv G''_{\text{new}}(i\omega_n)$ denotes the (new) solution of the impurity model for given $\Delta''_n/t^{*2} \equiv G''_{\text{old}}(i\omega_n)$. Eqn. (3.59) implies that each step of a direct iteration scheme (cf. subsection 3.2.2) advances $G$ exactly opposite to the gradient of $F$, i.e., downhill.[46] If $\nabla_\Delta F$ changes slowly on the scale of the step length, we can approximate the change in $F$ by the first term in a Taylor expansion:

$$F_{\rm new} \approx F_{\rm old} - 2\,T\,t^{*2} \sum_n \big(G''_{\rm new}(i\omega_n) - G''_{\rm old}(i\omega_n)\big)^2 < F_{\rm old}\,. \tag{3.60}$$

Consequently, the free energy of the new solution is lower for sufficiently small step size.

As a slightly alternative derivation of an equivalent result, let us consider the second derivative of the free energy functional:

$$\frac{t^{*2}}{2T}\frac{\partial^2 F[T,U,\Delta]}{\partial\Delta''_n\partial\Delta''_m} = -t^{*2}\frac{\partial G''_n}{\partial\Delta''_m} + \delta_{mn} \tag{3.61}$$

$$\frac{\partial G''_{\rm new}(i\omega_n)}{\partial G''_{\rm old}(i\omega_m)} = \delta_{mn} - \frac{t^{*2}}{2T}\left.\frac{\partial^2 F[T,U,\Delta]}{\partial\Delta''_n\partial\Delta''_m}\right|_{\Delta=G_{\rm old}}\,. \tag{3.62}$$

Since the matrix of the second derivatives on the right hand side has only positive eigenvalues[47] near a minimum of $F$, the fixed point method is contracting (as long as all eigenvalues are smaller than $4T/t^{*2}$) and is, thus, convergent to thermodynamically (locally) stable solutions. Another interesting consequence of these considerations for the iteration process is that the error decreases exponentially for the direct scheme in the neighborhood of a solution $\Delta_0$ where the free energy is quadratic in $\Delta - \Delta_0$. We have derived, but omit, the equivalent to (3.62) for the conventional iteration scheme (3.1). In this case, the iteration definitely does not proceed downhill and, therefore, coexistence regions might indeed be underestimated. After noting the connection between convergence properties of iteration schemes and the behavior of the free energy functional, we performed additional numerical simulations in the coexistence region using the direct scheme. These indicated, however, no significant increase in the resulting estimate for $U_{c2}$ in comparison to the conventional iteration scheme.

### 3.7.2 Direct Evaluation of Free Energy Differences

Another important consequence of the discussion of subsection 3.7.1 is that the gradient of the free energy with respect to the hybridization function $\Delta$ may be integrated along a line between two solutions (for fixed $U$ and $T$). The result is then nothing but their free energy difference. The integration curve is arbitrary (within a subspace compatible with analytic constraints on $\Delta(i\omega_n)$); a convenient choice is a straight line, i.e., a convex linear combination of the converged Green functions. For metallic and insulating solutions denoted by subindex "m" and "i", respectively, we define

$$\Delta_\lambda := t^{*2}G_{\rm m} + \lambda t^{*2}(G_{\rm i} - G_{\rm m}); \quad 0 \le \lambda \le 1\,. \tag{3.63}$$

[46]While this is strictly true only for an exact solution of the impurity model, the statistical noise cancels out on average due to the linear dependence for a numerically exact method like QMC. The asymptotic noise in the solution decreases with increasing accuracy, i.e., an increasing number of sweeps. Only in the case of multiple solutions, fixpoints may be lost due to too much noise.

[47]Note the advantage of our explicitly real-valued formalism at this point: written in terms of $\Delta_n = i\Delta''_n$ all eigenvalues would be negative, which we found counter-intuitive.

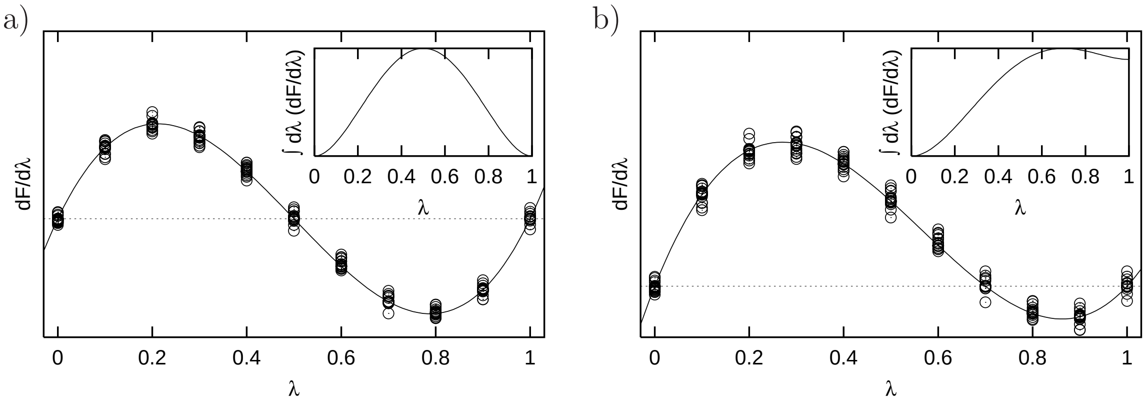



**Figure 3.58:** Illustration of an algorithm for direct evaluation of the free energy difference between two solutions A and B. Intermediate solutions are generated on a line $0 \leq \lambda \leq 1$, where $\lambda = 0$ corresponds to solution A and $\lambda = 1$ to solution B. The derivative $\mathrm{d}F/\mathrm{d}\lambda$ is computed on a grid $\lambda_i$ (circles) by solving a series of specially constructed impurity models using, e.g., QMC. Integration (insets) then yields the difference in free energy (measured from solution A). **a)** Possible result for a phase point on the first order line, where both solutions A and B have the same free energy; **b)** for a phase point where A has the higher free energy.

---

Then, the free energy difference is given by the integral

$$F_{\mathrm{i}} - F_{\mathrm{m}} = \int_0^1 \mathrm{d}\lambda \, \frac{\mathrm{d}F}{\mathrm{d}\lambda} \,, \tag{3.64}$$

where

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda} = -2T \sum_n \big(t^{*2} G''[T, U, \Delta_\lambda](i\omega_n) - \Delta''_\lambda(i\omega_n)\big)\big(G''_{\mathrm{i}}(i\omega_n) - G''_{\mathrm{m}}(i\omega_n)\big) \tag{3.65}$$

and which may be evaluated on a grid. When on each grid point multiple QMC measurements $G[T, U, \Delta_\lambda]_i$ are performed, even the error of the projection of the gradient in the "solution direction" may be evaluated so that a corresponding (statistical) error for the free energy difference is easily computed. Note that given the two converged solutions, the computation of the free energy difference within this scheme is completely parallel since no more self-consistency is involved. Another practical advantage of this method is that it checks the convergence of the existing solutions. This is illustrated in Fig. 3.58, a made-up example. Green circles indicate QMC estimates of the true derivative (solid line) of the free energy along the integration line. The limiting solutions (at $\lambda = 0$ and $\lambda = 1$) are indeed stable (gradient vanishes with positive slope), at least along the interpolation direction. The position of the intermediate fixed point only indicates a maximum of $F$ in the one-dimensional subspace under consideration; in general it will not correspond to a true maximum of $F$. Integrating up the solution (insets) shows that in Fig. 3.58 a both solutions are stable, i.e., this phase point is on the first-order line while in Fig. 3.58 b, the solution corresponding to $\lambda = 1$ is only metastable since the free energy is much higher (inset). Even though the comparatively more conventional method of pinpointing the first-order transition line used in our numerical calculations also worked well, the new method suggested here seems significantly superior since it is fully local in $U$ and $T$ and should be used in future work.[48]

Integrating the free energy in the solution space may also be useful for an absolute determination of $F$ using some reference system. In particular, the impurity model corresponding to a Lorentzian DOS can be solved analytically and may serve as a reference point. Finally, the concept can obviously be generalized for arbitrary lattices. While for these cases, the lattice contribution to the free energy will have to be computed numerically, the additional cost is still negligible in comparison with the effort for solving the impurity model.

### 3.7.3 Critical Behavior Near the MIT

Two different aspects of criticality are potentially relevant in the context of studies of the Mott metal-insulator transition: On the one hand, observables may show singular behavior (i.e., have divergent slope as a function of $T$ or $U$) near critical points or critical lines. On the other hand, critical slowing down of the iterative DMFT convergence process may arise within some regions of the parameter space.[49] In principle, "raw" QMC results (i.e., results for fixed finite $\Delta\tau$) could be negatively impacted by critical slowing down while the accuracy of extrapolations and fits could suffer from our neglect to take singular behavior explicitly into account. We will in this subsection discuss both effects and their relevance to the QMC studies presented earlier in this chapter.

As pointed out by Kotliar et al. (2000), a second-order critical end point is signaled within the DMFT by the appearance of a single zero eigenvalue $m_0$ of the fluctuation matrix (3.61) of the Ginzburg-Landau functional $F_{\mathrm{GL}}$.[50] Under the assumption that this matrix smoothly depends on the hybridization function $\Delta_n$ and on the parameters $T$ and $U$, the associated eigenvector is well-defined also in the vicinity of the critical end point where it represents a soft mode. The associated coefficient $\eta_0$ of an expansion of $\Delta_n$ in terms of the eigen basis of the fluctuation matrix then determines the critical behavior. In principle, critical exponents may be calculated from Taylor expansions in $\eta_0$. According to Kotliar et al. (2000), the parameter $\eta_0$ may be replaced by another parameter $\eta$ which fulfills a cubic equation without quadratic term, $p\eta + c\eta^3 = h$, near the critical point. As a consequence, the double occupancy is predicted to show cubic-root behavior near the critical point

[48] As pointed out in more detail in Sec. 3.6, the transformation technique used by Tong et al. (2001) for determining the thermodynamic MIT line is directly related to the concepts of this subsection. In contrast to the method suggested here, however, the transformation technique incorporates an approximation.

[49] Note that only the first consequence is strictly a property of the model; critical slowing down could in principle be overcome by more advanced iteration techniques.

[50] All other eigenvalues are, then, negative since the minimum of $F_{\mathrm{GL}}$ is unique for temperatures at or above the critical point (cf. subsection 3.7.1).

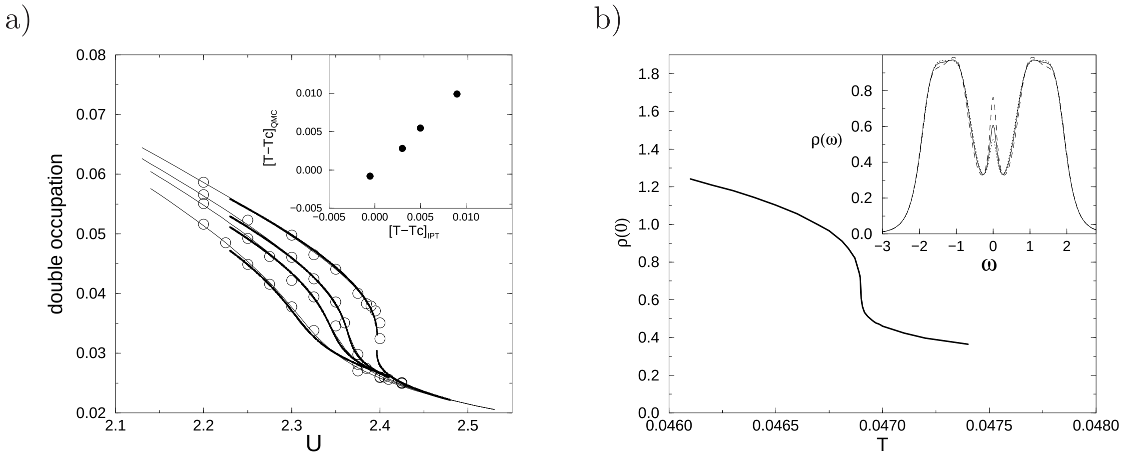



**Figure 3.59: a)** Double occupancy versus $U$ for $T$ slightly below (top set or curves/symbols) or above $T^*$ (lower 3 data sets, with increasing $T$ towards the bottom) as estimated from QMC (circles) and IPT (thin lines). Thick lines represent fits according to the Ginzburg-Landau theory. **b)** IPT density of states at the Fermi level versus $T$ for $U = 2.46316 \approx U^*_{\rm IPT}$. The insets are irrelevant for our discussion. Both figures are reproduced from the publication by Kotliar et al. (2000). Energies are in units of the half bandwidth.

---

so that the corresponding susceptibility $\chi = \partial D/\partial U$ diverges as $|U - U^*|^{-2/3}$ for $T = T^*$. In the following, we will show that the IPT data and (rescaled) QMC data presented in the cited paper fail to support this claim.

In order to simplify the discussion, we have reproduced both figures of the cited paper as Fig. 3.59. In the main panel of Fig. 3.59 a, QMC estimates for the double occupancy $D$ (circles) for the temperatures $T = 1/40$, $T = 1/35$, $T = 1/32$, and $T = 1/25$ (from top to bottom) have been scaled on IPT estimates (thin lines) for the temperatures $T = 0.0469$, $T = 0.05$, $T = 0.052$, and $T = 0.056$; here, all energies are measured in units of the half bandwidth and have to be doubled to correspond to the units used throughout the rest of this chapter. Thick lines show fits derived from the Ginzburg-Landau theory. Evidently, the only candidate for singular behavior among these data sets is that for the lowest temperature(s); a magnified view on this data is given in Fig. 3.60. Here, the QMC results (circles) are connected by lines to guide the eye. Note that the two QMC data points at $U \approx 2.40$ have been computed for the same interaction; consequently, they imply coexistence rather than a diverging slope of the metallic solution. For enhanced clarity, the IPT data is shown as crosses (instead of a thin line) in Fig. 3.60. In addition to the Ginzburg-Landau (GL) fit (solid line), we have added a pure cubic-root fit of the form $D(U) = 0.0305 + 0.035 * (2.3974 - U)^{1/3}$ for the metallic phase.[51] A first

[51] While the technique employed here for extracting the data from the postscript figures (available from the preprint archive) is exact for relative positions of symbols and line segments, a slight mismatch in overall scale and offset is in general unavoidable.

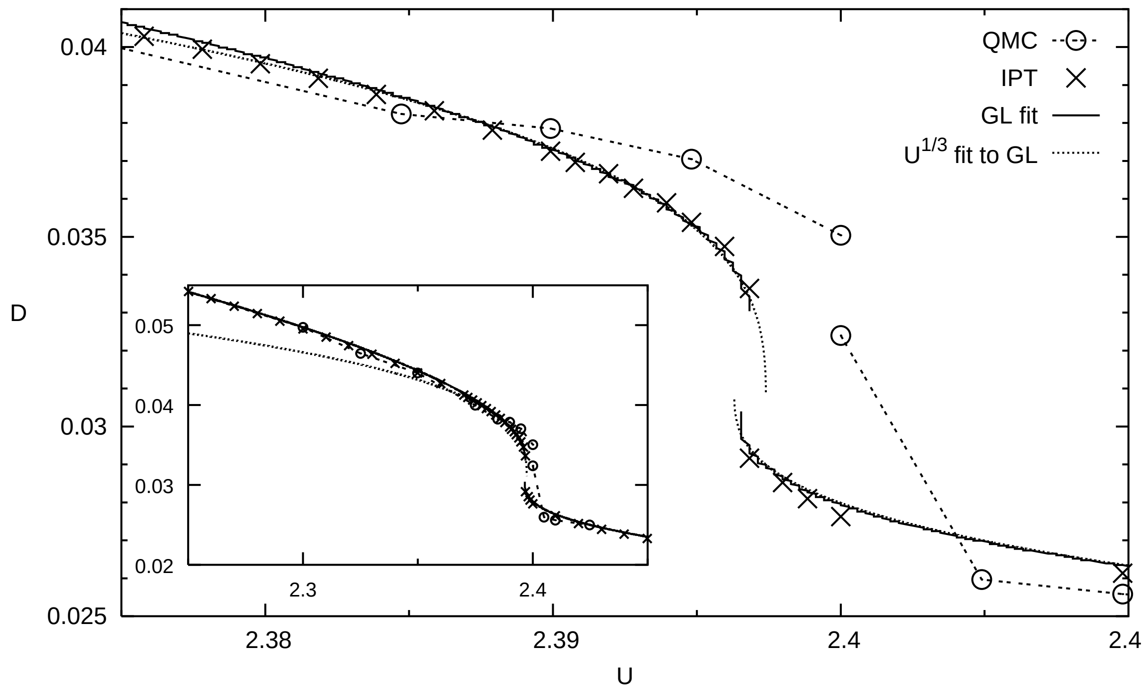



**Figure 3.60:** Double occupancy $D$ versus $U$ for $T \lesssim T^*$. Magnified view on top set of curves/circles shown in the main panel of Fig. 3.59 (Kotliar et al., 2000). In addition to the QMC data (circles) and IPT data (here: crosses) and to the Ginzburg-Landau fit (solid line), an asymptotic cubic-root fit (dotted lines) to the Ginzburg-Landau fit for the metallic IPT solution is shown. Interactions are in units of the half bandwidth.

important observation is that the slope of the QMC estimates for $D$ in the metallic phase is only slightly enhanced, but shows no signs of divergence near the transition. For the insulating phase, an enhanced slope is only suggested by a single data point in the coexistence region. The agreement of QMC data with the remaining data is poor, considering that the comparison arises from a two-parameter fit. As a second observation, we note relatively strong discrepancies between IPT data and GL fits, even close to the transition. Furthermore, strong deviations from the asymptotic cubic-root behavior (dotted line in inset of Fig. 3.60) of the GL fit appear already below about $U \approx 2.35$.

In contrast, square-root fits show excellent agreement with the IPT results, both for the metallic and the insulating phase as shown in Fig. 3.61 and Fig. 3.62, respectively. Although the shape of each of these fits (dashed lines) is determined by a single parameter (in addition to the end point coordinates $D_c$ and $U_c$) instead of at least two parameters for the GL fits, the deviation of the square-root fits from the IPT data is typically almost an order of magnitude smaller than that of the GL fits (with cubic-root asymptotics). We conclude that the critical exponent 1/3 derived by Kotliar et al. (2000) is either wrong even at the critical point or that the critical behavior at the second order end point is different from the critical behavior of coexisting solutions near their stability edge. In any case, we have established that the

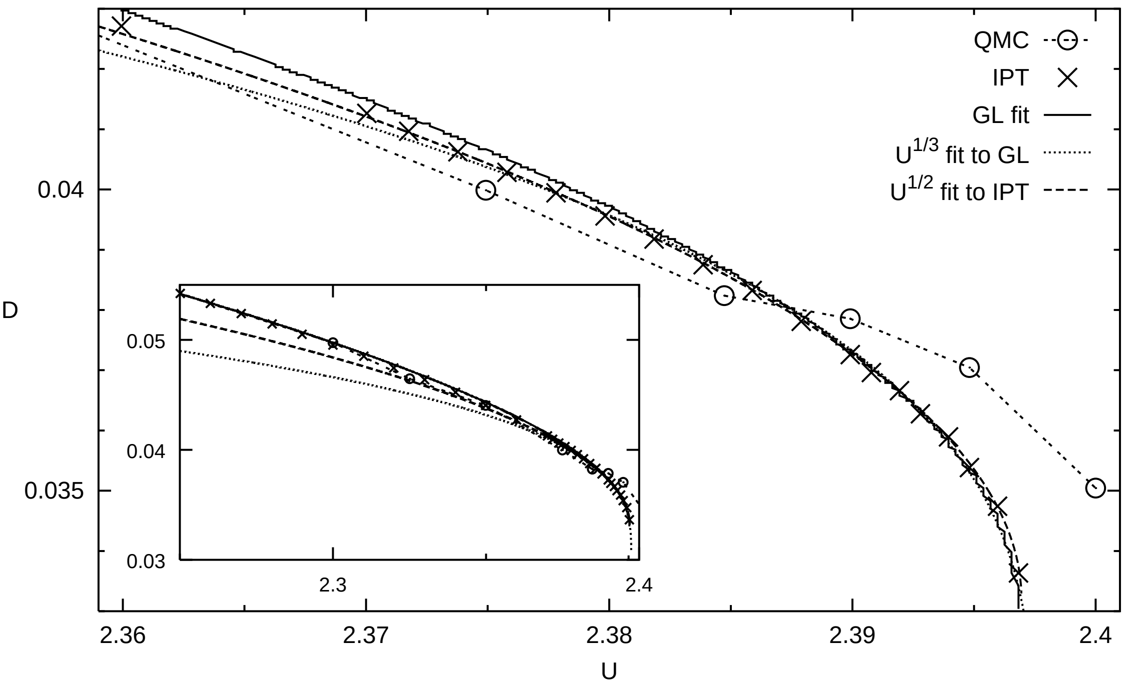



**Figure 3.61:** Double occupancy $D$ versus $U$ for $T \lesssim T^*$. Magnified view on top set of curves/circles shown in the main panel of Fig. 3.59 (Kotliar et al., 2000) for the metal. In addition to the QMC and IPT data and to the Ginzburg-Landau fit, a cubic-root fit (dotted lines) and a square-root fit of the form $D(U) = 0.0332 + 0.0488\,(2.39695 - U)^{1/2}$ (dashed lines) to the asymptotic behavior of the metallic IPT solution is shown. Interactions are in units of the half bandwidth.

latter is governed by an exponent of 1/2, at least within IPT.[52]

Given the fact that the misfit from the correct asymptotic behavior was hard to catch even for IPT data which has no statistical error and even very close to the critical end point, we may expect that critical behavior is practically undetectable using QMC at temperatures that deviate by a few percent (or more) from the critical temperature $T^*$ or when using a grid size of more than fractions of a percent in $U$. This is illustrated in Fig. 3.63 where the QMC data (symbols) presented by Kotliar et al. (2000) is reanalyzed using regular fit functions (upper lines) similar to those used in subsection 3.5.5. The bottom line shown for $U \geq 2.3$ represents our result (3.19) for the insulator (rescaled to units of the half bandwidth). The slight systematic deviation of the QMC data in the insulating phase (for $U \geq 2.4$) from this curve probably reflects a finite $\Delta\tau$ error. The overall agreement between QMC data for the metallic phase and our fit functions is very good. For $T = 1/40$, below the critical temperature, it is even much better than the agreement between QMC and either the (rescaled) IPT result or the GL fit. Only the coexistence point belonging

[52]In contrast, the IPT result for the (interacting) density of states at the Fermi level shown in Fig. 3.59 b is well described by an exponent 1/3 when $T$ is varied at constant $U = U^*_{\rm IPT}$, both on the metallic and insulating side (with different prefactors).

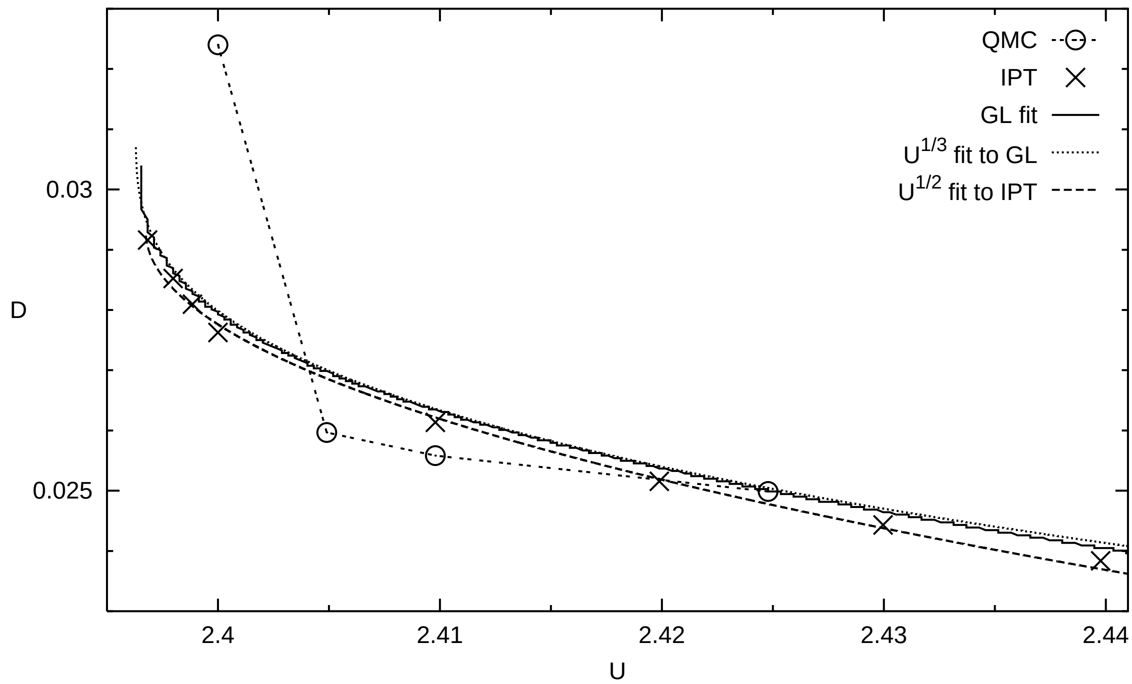



**Figure 3.62:** Double occupancy $D$ versus $U$ for $T \lesssim T^*$. Magnified view on top set of curves/circles shown in the main panel of Fig. 3.59 (Kotliar et al., 2000) for the insulator. In addition to the QMC and IPT data and to the Ginzburg-Landau fit, a cubic-root fit of the form $D(U) = 0.031 - 0.0195\,(U - 2.3963)^{1/3}$ (dotted lines) and a square-root fit of the form $D(U) = 0.0293 - 0.027\,(U - 2.39675)^{1/2}$ (dashed lines) to the asymptotic behavior of the insulating IPT solution is shown. Interactions are in units of the half bandwidth.

to the insulating phase at $U \approx 2.4$ at this temperature is incompatible with all other data or fit functions; a finer grid or an intensive convergence study would be needed to elucidate the irregular behavior of this single data point. The fit for $T = 1/32$, the highest temperature considered here, is perfect. Only for $T = 1/35$, just above the critical temperature $T^*$, a deviation in the transition region is clearly seen. The numerical significance of this discrepancy is not completely clear.[53] We conclude that potential problems arising from the fit functions used throughout this chapter are indeed limited to a very small region around the second-order end point. Since no results are reported for this region and since no signals for singular behavior are present, a negative impact on our results can be excluded.

### Critical Slowing Down

All points in parameter space (here formed by $U$ and $T$) where a new minimum of the free energy appears in the space of hybridization functions are critical in the

[53]Note that the deviation of the QMC data is of the same order of magnitude for $T = 1/32$ and $U = 2.3$ where the only viable explanation is the numerical Monte Carlo error.

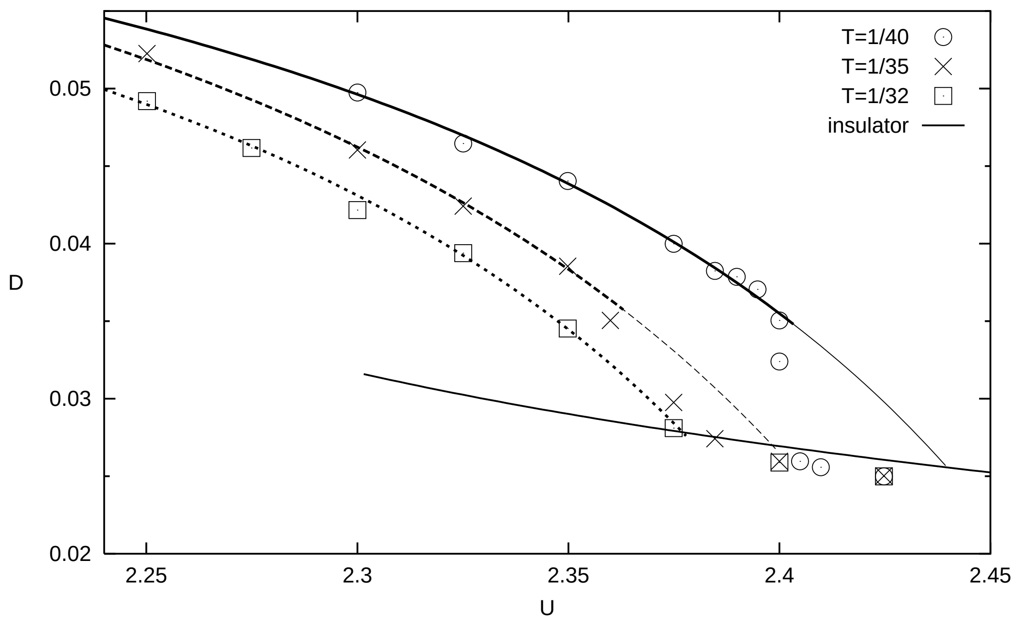



**Figure 3.63:** Double occupancy $D$ versus $U$ for $T \lesssim T^*$. QMC data for a selected set of temperatures extracted from the main panel of Fig. 3.59 (Kotliar et al., 2000) in comparison to fits with regular functions and to our result (3.19) for the double occupancy of the insulator. Thick portions of the lines indicate ranges of good agreement with the QMC data. All energies are in units of the half bandwidth.

sense that the fluctuation matrix has a zero eigenvalue.[54] Consequently, at least one component of the gradient of the free energy with respect to the hybridization function vanishes faster than linearly at its minimum. An iteration scheme that relies on linear behavior of the gradient thus slows down near the critical point; exactly at the critical point, the number of iterations necessary for decreasing the error of a solution by some factor may even diverge (in the limit of infinite target accuracy).

Therefore, difficulties in the DMFT iteration process may be expected near critical points and lines. It is not clear a priori, however, how large the problematic regions are and what accuracy can be reached at reasonable numerical cost. Joo and Oudovenko (2001) found themselves unable to present QMC estimates of the coexistence region for $T > 0.035$ due to critical slowing down.[55] Our experiences are less dramatic. We found slower convergence near critical points, in particular at higher temperatures close to the temperature of the second-order end point, both for $T \lesssim T^*$ and for $T \gtrsim T^*$. This is illustrated in Fig. 3.64 for $T = 1/15$ and $U = 4.55$ (in units of the variance of the DOS). Here, "traces" of the double occupancy $D$ are shown

[54] In the context of the MIT, these points are the second-order end point and the stability edges of solutions at the boundaries of the coexistence region.

[55] These authors even found critical slowing down for the stable solution near the boundaries of the coexistence region. This claim is rather surprising in view of the fact that the zero eigen mode only appears when the free energy is evaluated at the hybridization function corresponding to the unstable solution. It is not supported by our data at all.

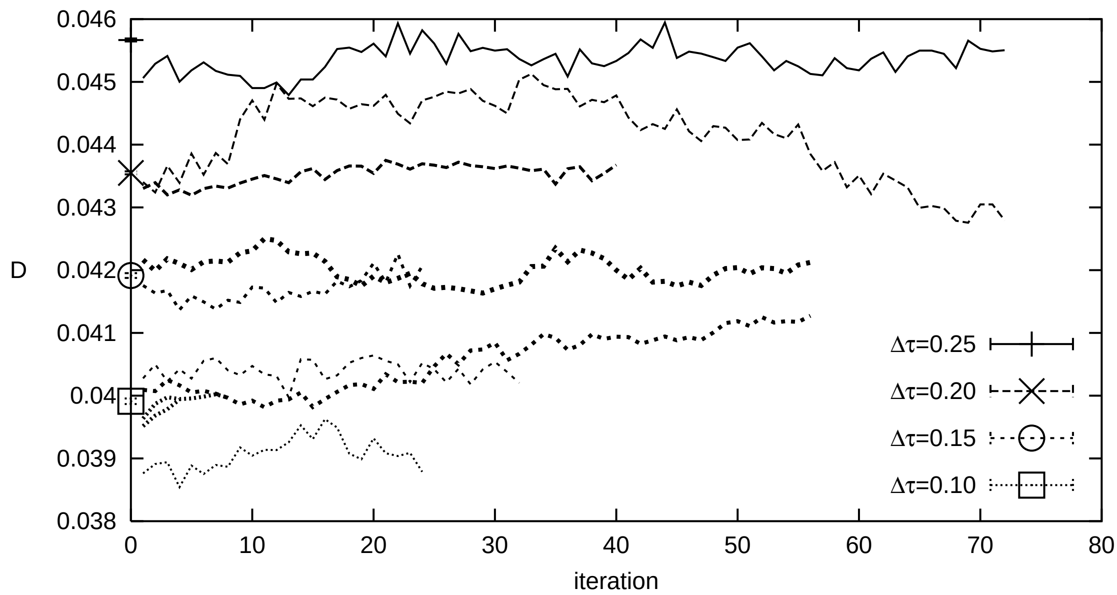



**Figure 3.64:** Convergence of double occupancy $D$ for $T = 0.067$ and $U = 4.55$. The number of sweeps is about $10^5$ for thin lines and about $10^6$ for thick lines. Symbols at the left axis indicate our best estimate of $D$ for each value of $\Delta\tau$.

versus iteration number for several independent iteration runs. Line types (solid, long-dashed, short-dashed, and dotted) denote the time discretization $\Delta\tau$ while the width of the lines give a rough indication of the numerical accuracy, i.e., the number of sweeps used in each convergence run which is larger for thicker lines. Apparently, the overall convergence is not good since not all solutions for the same value of $\Delta\tau$ tend to a common fixed point. Instead, long-scale fluctuations appear, e.g., for the run with lower accuracy for $\Delta\tau = 0.2$ (thin long-dashed line). This is exactly what we expect from Landau theory: as soon as the gradient of the free energy is smaller than statistical fluctuations, the iterated solutions follow an asymptotically free random walk (predominantly in the soft-mode direction). Two strategies for overcoming the associated problems are also apparent from Fig. 3.64: On the one hand, it is important to increase the numerical precision near critical regions in order to reduce fluctuations. On the other hand, trends can be much better recognized when comparing several independent self-consistency runs. By comparing the evolution (as a function of iteration number) of different runs with more metallic or insulating initial conditions, one can establish boundaries for the fixed point or accelerate convergence by preparing initial states near the assumed fixed point.[56] We used such strategies extensively throughout the numerical parts of this thesis. For example, the upward trend of initial solutions for $\Delta\tau = 0.15$ (two lower short-dashed lines in Fig. 3.64) prompted us to initialize runs with $D \approx 0.042$; those are seen to be stable.

In order to further illustrate these considerations and to demonstrate increased fluctuations near the critical point, we perform sets of iteration runs for $T = 0.055$,

[56]The possibilities for initialization are obviously more limited very close to a stability edge.

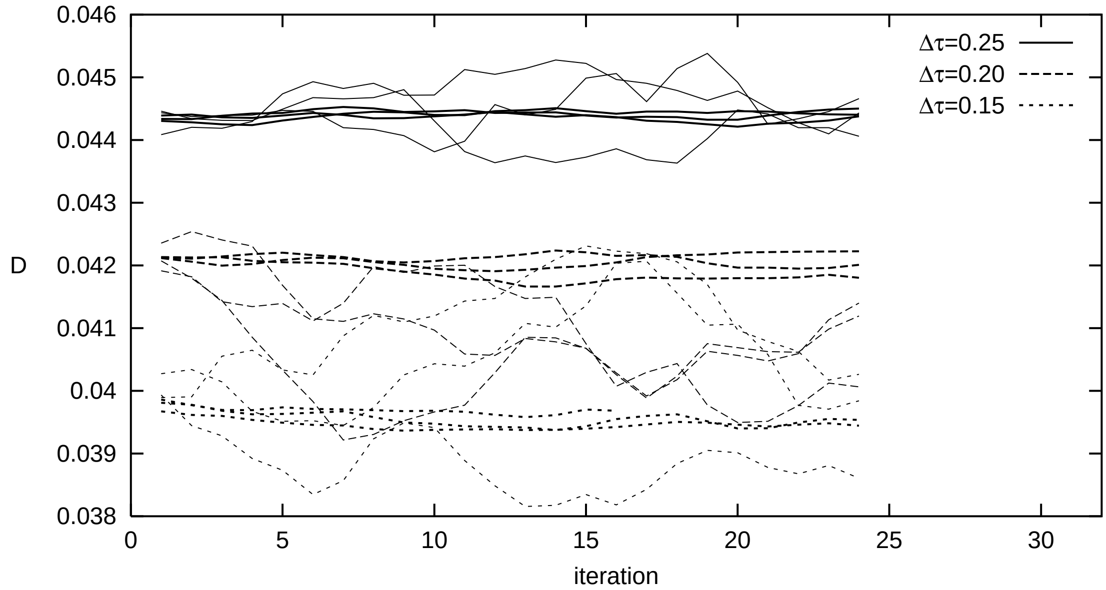



**Figure 3.65:** Convergence and fluctuations of double occupancy $D$ for $T = 0.055$ and $U = 4.65$, very close to the critical end point ($T^* = 0.055$, $U^* = 4.665$) for $\Delta\tau = 0.0$. Thin lines correspond to runs with $10^5$ sweeps per iteration, thick lines to runs with $4 \cdot 10^6$ sweeps.

---

$U = 4.65$, i.e., very close to our estimate of the critical end point (for $\Delta\tau = 0.0$). Figure 3.65 shows results for the discretizations $\Delta\tau = 0.25$, $\Delta\tau = 0.20$, and $\Delta\tau = 0.15$, with three statistically independent runs using $10^5$ sweeps ("low" accuracy) and three runs using $4 \cdot 10^6$ sweeps per iteration ("high" accuracy) each. Evidently, the amplitude of fluctuations indeed increases for decreasing $\Delta\tau$ at $T = 0.055$ and $U = 4.65$. The fluctuations are strongly reduced by a 40-fold increase in the number of sweeps, i.e., by a reduction of the statistical error of the solution of each impurity model to about 1/6. Note that the long "time" scale (i.e., the large number of iterations) associated with the fluctuations of the runs with low accuracy generically leads to an underestimation of error bars for averages obtained from a single run. One way of detecting such extremely long autocorrelation time is the comparison of several independent runs as seen in Fig. 3.65 for $\Delta\tau = 0.15$. For a finite number of runs, however, the deviations of all iteration runs may have the same sign as seen for $\Delta\tau = 0.20$ in the same figure. Thus, accurate and controlled estimates can only be obtained from runs with high accuracy, i.e., a large number of sweeps.

Corresponding data for the slightly higher temperature $T = 1/15$ is shown in Fig. 3.66. For each of the combinations of a representative sets of values of the interaction ($U = 4.4$, $U = 4.55$, $U = 4.65$, and $U = 4.8$) and of the discretization (here limited to $\Delta\tau = 0.25$, $\Delta\tau = 0.2$, and $\Delta\tau = 0.15$ by the computational resources), we used our best previous estimate of the converged solution for initializing 3 runs with high accuracy and 3 runs with low accuracy. While some initial trend is visible even in some high accuracy runs, the rapid convergence of each run and the agreement

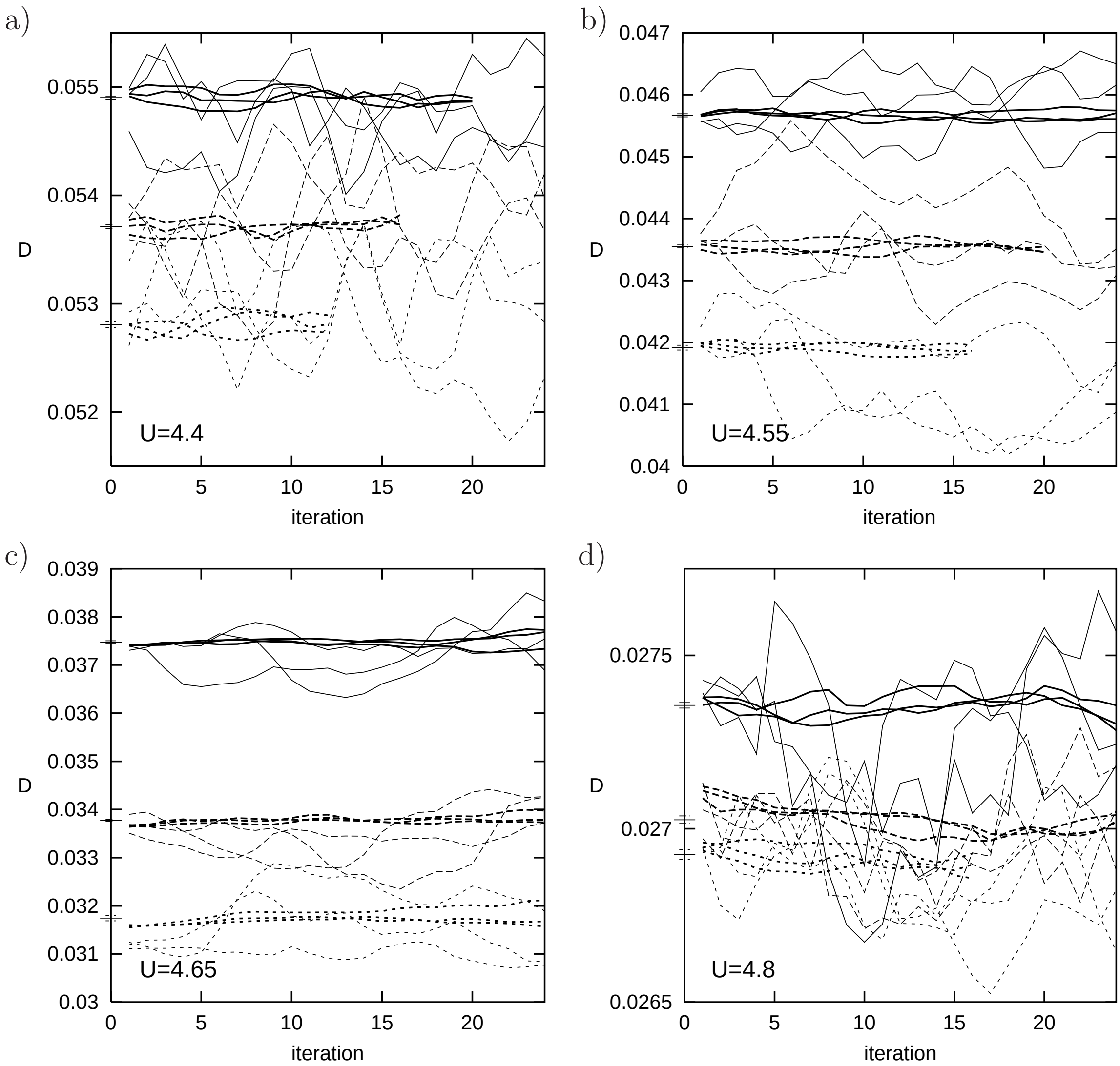



**Figure 3.66:** Fluctuations in the double occupancy for $T = 0.067$ and **a)** U=4.4, **b)** U=4.55, **c)** U=4.65, and **d)** U=4.8 using a discretization of $\Delta\tau = 0.25$ (solid lines), $\Delta\tau = 0.2$ (long-dashed lines), and $\Delta\tau = 0.15$ (short-dashed lines). Thin lines correspond to runs with $10^5$ sweeps per iteration, thick lines to runs with $4 \cdot 10^6$ sweeps. Crosses at the left axis represent final estimates for $D$ at finite $\Delta\tau$.

between the runs support the high accuracy of our final estimates.[57] The good convergence of the final solutions (i.e., an average of the converged parts of the high-accuracy runs) is further supported independently by the regular dependence of $E$ on $U$ (for each value of $\Delta\tau$) already seen in Fig. 3.23.

At lower temperatures, the fluctuations decay fast enough that runs with an order of $10^5$ sweeps suffice for reliable results. We conclude that critical slowing down poses a challenge which can be overcome by the techniques described here and by investing

[57] Note the vastly different scales for $D$ in the different parts of Fig. 3.66.

enough computer time.

## 3.8 Spectra

The objective of this section is the computation of spectra, i.e., of interacting local spectral functions $A(\omega) = -\frac{1}{\pi}\mathrm{Im}\, G(\omega)$ on the real axis. Since the QMC simulations yield results on the imaginary axis, this requires the use of the maximum entropy method (MEM). In subsection 3.8.1, we discuss advanced aspects of the maximum entropy formalism in the context of spectral functions. A specification of the algorithmic choices made in this work and numerical tests follow in subsection 3.8.2. Finally, numerical results are presented for the half-filled Hubbard model with Bethe DOS at $T = 1/20$ and $T = 1/15$ in subsection 3.8.3.

### 3.8.1 Maximum Entropy Method for Spectral Functions

Both the basic ideas of the MEM and a simple algorithm have already been introduced in Sec. 1.4. We will in this subsection discuss three potential problems in the application of the MEM for obtaining spectral functions from imaginary-time QMC data: covariance, non-Gaussian distribution of statistical errors, and systematic errors.

#### Covariance

If a series of measurements is made for a vector of observables and the variance of each component is finite,[58] the probability distribution of the average vector approaches a multivariate Gaussian distribution in the limit of an infinite number of measurements. This distribution is fully characterized by its covariance matrix the estimate of which reads in the case of the imaginary-time Green function

$$C_{kl} = \sum_{i=1}^{N_d} \frac{\left(\bar{G}^i_k - \bar{G}_k\right)\left(\bar{G}^i_l - \bar{G}_l\right)}{N_d(N_d-1)}\,. \tag{3.66}$$

As in (1.61), lower indices $l$ here correspond to time slices $\tau_l = l\Delta\tau$. The $l$ component of the result of the $i^{\text{th}}$ measurement (total number: $N_d$) is denoted by $\bar{G}^i_l$; $\bar{G}_l$ is the corresponding average. In the uncorrelated case, the matrix $C_{kl}$ reduces to its diagonal with elements $\sigma_k^2\delta_{kl}$. This is, in fact, the simplifying assumption made in Sec. 1.4 and used in a program implemented by Sandvik (Sandvik and Scalapino, 1995) which is used in parts of this work. It is the main point of this paragraph to demonstrate that this assumption does not apply in the cases of interest. The covariance matrix for a symmetric case (Bethe DOS with half filling at $T = 1/20$ and $U = 4.6$) is shown in Fig. 3.67 a. Here, crosses indicate matrix elements $C_{kl}$ for $0 \leq k \leq l \leq \Lambda/2$; lines connect elements with the same value of $k$. Diagonal

[58] The condition of finite variance is trivially met for measurements of the imaginary-time Green function since it is bounded: $0 \leq G(\tau) \leq 1$.

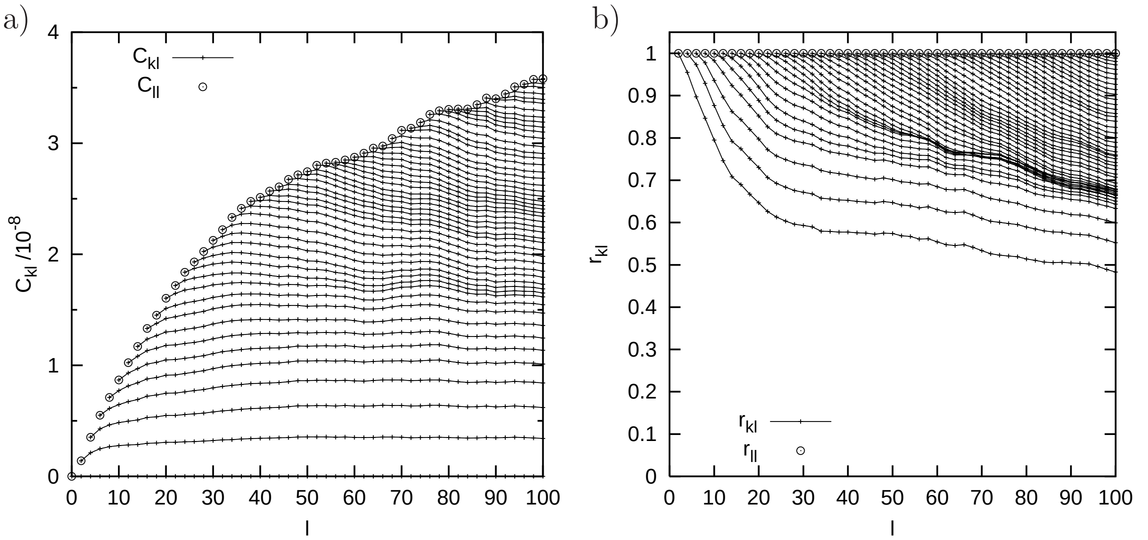



**Figure 3.67: a)** Covariance matrix $C_{kl}$ of the QMC estimate of $G(\tau)$ for Bethe DOS, $T = 1/20$, $U = 4.6$, and $\Delta\tau = 0.1$ (i.e., $\Lambda = 200$) using 300 bins with 10000 sweeps each. Circles indicate diagonal elements $C_{ll}$ as a function of $j$; offdiagonal elements with the same value of $k$ (crosses) are connected by a solid line. The indices extend over $0 \leq k \leq l \leq \Lambda/2$ (and correspond to $\tau_k = k\Delta\tau$, $\tau_l = l\Delta\tau$); the missing elements follow from the symmetries $C_{kl} = C_{lk}$ and (here) $G(\tau) = G(\beta - \tau)$. **b)** Corresponding Pearson's coefficients $r_{kl}$. For enhanced clarity, only even indices are included in the plots.

elements are additionally denoted by circles. Evidently, the offdiagonal elements are generically of the same order of magnitude as the diagonal elements. This shows that QMC measurements of the imaginary-time Green function $G$ at different time slices $\tau_l$ are not independent, but are instead highly correlated; in Bryan's terminology the full data set is oversampled. The same data is shown in a different representation in Fig. 3.67 b; here, each column and each row has been divided by the square root of the corresponding diagonal element. The result is a matrix of Pearson's coefficients $r_{kl}$, a dimensionless measure of linear correlation between pairs of stochastic variables.

Corresponding plots of the covariance matrix for an asymmetric case (LDA DOS of $La_{1-x}Sr_xTiO_3$ , $n = 0.94$, $T = 1/10$, and $U = 5.0$) are shown in Fig. 3.68. Since all three bands are degenerate and since measurements of $G$ for each band are highly correlated (e.g., the occupancy of individual bands fluctuates while the total occupancy is almost exactly constant), the most relevant covariance matrix for this case is that shown in Fig. 3.68 a for the band-averaged case; Fig. 3.68 b which applies to measurements of $G$ for one individual band (with much larger variances for $\tau \to 0$ and $\tau \to \beta$) is included for comparison and for later discussion in chapter 5 (in subsection 5.4.2). All results clearly demonstrate the large degree of covariance present in QMC estimates of $G(\tau)$.

Thus, a more general formulation of the MEM should be used which takes into account all elements of the covariance matrix. In principle, this is simple since the form (1.65) of the probability function still applies (in the limit of large bin size)

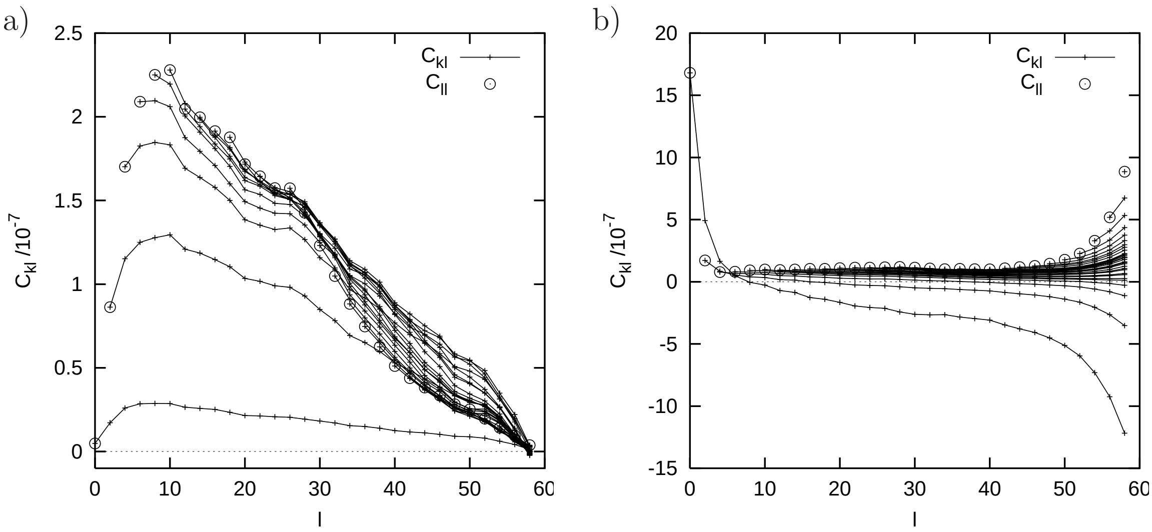



**Figure 3.68:** Covariance matrix $C_{kl}$ of the 3-band QMC estimate of $G(\tau)$ for (asymmetric) LDA DOS of $\mathrm{La}_{1\text{-}x}\mathrm{Sr}_x\mathrm{TiO}_3$ , $n = 0.94$, $T = 1/10$, $U = 5.0$, and $\Delta\tau = 0.167$ (i.e., $\Lambda = 60$) using 102 measurements with 100000 sweeps each. The indices extend over $0 \leq k \leq l \leq \Lambda - 1$; the missing elements follow from the symmetry $C_{kl} = C_{lk}$ and the constraint $G(0)+G(\beta) = 1$. **a)** Covariance matrix computed from band-averaged QMC measurements of $G$. **b)** Covariance matrix computed from QMC measurements of $G$ for individual bands.

when $\chi^2$ is generalized as

$$\chi^2 = \sum_{k,l=1}^{\Lambda} \left(\bar{G}_k - G_k\right) [C^{-1}]_{kl} \left(\bar{G}_l - G_l\right) . \tag{3.67}$$

In practice, it is necessary to apply a singular value decomposition (SVD) for diagonalizing the matrix $C$ for repeated use in the computation and to work with rotated data $\bar{G}'_l$ and $\sigma'_l$. A SVD of the kernel $K$ can also greatly accelerate the search for a solution. Another improvement realized in advanced codes is the possibility of marginalizing the spectrum over the Lagrange parameter $\alpha$ ("Bryan's method"). Jarrell's (1997) program [primarily based on Bryan (1990) and Gubernatis et al. (1991)] also used in this work realizes all of the improvements mentioned in this paragraph and also allows for a nonuniform frequency grid. While it yields more reliable spectra than the Sandvik code and even computes error bars for the integrated spectral weight in some finite regions, it requires more data sets and some additional efforts in order to produce stable results.

### Non-Gaussian Error Distribution

One problem which arises both for the Sandvik and the Jarrell code is that the assumption of a Gaussian distribution of errors breaks down for $G(\tau)$ with $\tau \approx \beta/2$ in gapped systems at low temperatures $T$. This is illustrated by the histogram of measurements of $G(\beta/2)$ for the Bethe DOS, $T = 1/20$, and $U = 5.0$ shown in

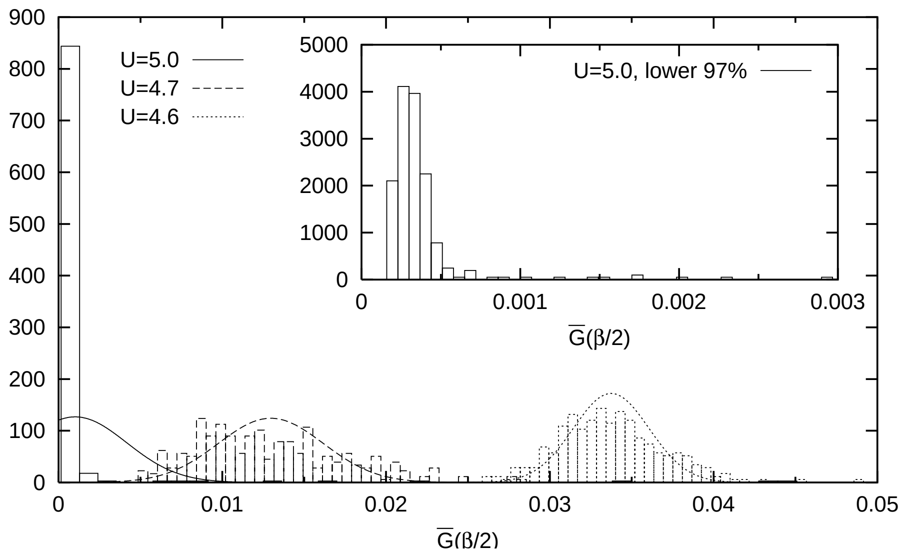



**Figure 3.69:** Histograms of $\bar{G}(\beta/2)$ for a Bethe DOS, $T = 1/20$, and $\Delta\tau = 0.1$ based on 300 measurements for each value of $U$ (boxes). Smooth lines indicate Gaussian fits reproducing both the averages and the variances of each of the measured distributions. For $U = 5.0$, the variance is dominated by a small fraction of outliers; 97 percent of all measurements are below 0.003 (inset).

Fig. 3.69. Evidently, the distribution indicated by the histogram (boxes with solid lines) strongly differs from a Gaussian distribution corresponding to the same error and variance. In particular, the true distribution is strongly asymmetric while the Gaussian fit extends into the forbidden region $G(\beta/2) < 0$. In fact, the variance is primarily associated with a small fraction of outliers [i.e., measurements corresponding to very large values of $G(\beta/2)$]; eliminating the top 3% of outliers results in a much narrower (but still asymmetric) distribution as seen in the inset of Fig. 3.69. In comparison, the histograms for the insulating phase at $U = 4.7$ (dashed lines) and for the strongly correlated metallic phase at $U = 4.6$ (dotted lines) are much better approximated by Gaussians. As a consequence, the application of the MEM formalism deep inside the insulating phase implies the use of an inappropriate likelihood function[59] unless the number and accuracy of measurements is increased (by some orders of magnitude) so that Gaussian statistics is restored for the averages.[60] MEM algorithms have been formulated, however, for non-Gaussian error statistics in other contexts (Bryan and Skilling, 1980); we are therefore confident that a practical solution could be found leading to even more reliable spectra for insulators.[61]

[59]Our attempt of solving this problem by transforming $G$ (to a new unbounded variable) lead to severe numerical instabilities so that the corresponding method could not be used in practical applications.

[60]At this point, the statistical error may be lower than systematic errors so that potentially the full error is never well approximated by a Gaussian distribution.

[61]Note that the presence of far outliers as seen for $U = 5.0$ in Fig. 3.69 poses an even more severe problem than skew and kurtosis since the characterization of the corresponding probability

**Systematic Errors**

It is important to realize that MEM can only yield reliable results when systematic errors are much smaller than statistical errors. For DMFT calculations, systematic errors can be caused by incomplete convergence of the self-consistency cycle and by a finite value of $\Delta\tau$. An accurate solution of the impurity problem and a large number of iterations (possibly using underrelaxation) can control the iteration error or effectively convert it to a statistical error by using completely independent simulation runs. A similar effect is achieved at lesser cost by performing some additional DMFT iterations at very high precision after convergence; measurements of $G(\tau)$ for each of the resulting impurity models can then be combined for analytic continuation. The spectra to be presented in this section will be based on 50 measurements for 6 impurity models (i.e., a total of 300 bins) for each parameter set; this approach brings the iteration error well under control.

In comparison, remaining $\Delta\tau$ errors in $\{\bar{G}_l\}$ are potentially more severe. Since this problem cannot be overcome by extrapolation on the level of spectra, reliable spectra require good algorithms (i.e., programs with minimal $\Delta\tau$ error) using small enough values of $\Delta\tau$. This is more important in the context of DMFT calculations than for the solution of plain SIAMs since for the DMFT even the self-consistency condition itself (which includes discretized Fourier transformations) introduces an error at finite values of $\Delta\tau$ (cf. subsection 3.4.1). For illustration of this point, a comparison of imaginary-time Green functions for the Bethe DOS in the insulating phase ($T = 1/20$ and $U = 5.0$) is shown in Fig. 3.70. In order to resolve the important region of $\tau \approx \beta/2$, we here use a strongly magnified scale (for a more complete picture, see Fig. 3.75). The thick short-dashed line represents our best estimate (for $\Delta\tau = 0.1$) of $G(\tau)$. The thinner dotted line is obtained by transforming back the MEM spectrum to be presented later in Fig. 3.80. The agreement between both curves is excellent for $\tau \lesssim 6$. Since the slight fluctuations visible in the QMC results for $6 \lesssim \tau \leq 10$ (with a local minimum for $\tau \approx 7$) violate analyticity,[62] they are filtered out by the MEM procedure which corresponds to a regularized form of $G(\tau)$ (dotted line). Both curves agree reasonably well with Jarrell's result [for $\Delta\tau = 0.167$ as used in Schlipf et al.'s (1999) paper: long-dashed line]. In contrast, Ulmke's original smoothing trick discussed in subsection 3.4.1 leads to much larger $G(\tau)$ and fails to capture the insulating character of the solution: while the former results are roughly consistent with a gap of 0.5 (here modeled using rigid semi-elliptic Hubbard bands), the Ulmke-code solution implies quite significant spectral weight near $\omega = 0$ (which even exceeds that associated with touching rigid Hubbard bands).

We conclude that the most important prerequisite for computing reliable spectra

distribution is an ill-conditioned problem. We cannot exclude the possibility that the outliers are artifacts of the QMC method, i.e., arise from the finite number of warm-up sweeps used for each measurement. Up to outliers, all data sets presented in Fig. 3.69 lead to nearly Gaussian histograms of $\log G(\beta/2)$.

[62] The positivity of all even derivatives of $G(\tau)$ can be read off from (1.55); conversely, local extrema for $0 \leq \tau \leq \beta/2$ imply negative regions of the spectrum $A(\omega)$ which is clearly unphysical. A second, weaker form of the violation of analyticity may be detected when the computation of the self energy is attempted (see Sec. 4.6).

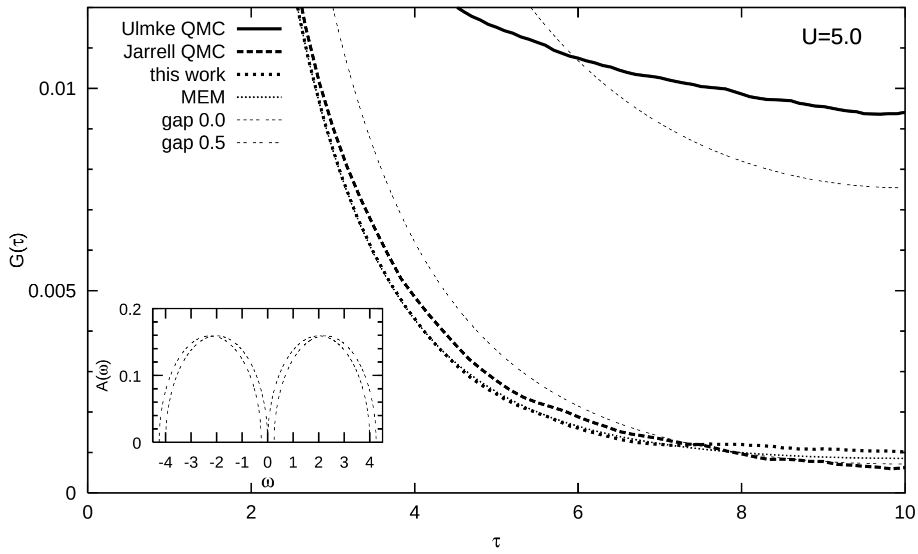



**Figure 3.70:** QMC estimates of $G(\tau)$ for Bethe DOS, $T = 1/20$, and $U = 5.0$ (thick lines) in comparison with the form of $G(\tau)$ corresponding to the MEM spectrum shown in Fig. 3.80 (dotted line) and with $G(\tau)$ corresponding to the spectra shown in the inset (thin dashed lines). Results of this work (for $\Delta\tau = 0.1$) essentially agree with results obtained by Jarrell (for $\Delta\tau = 0.167$); deviations of results obtained using Ulmke's smoothing trick (for $\Delta\tau = 0.1$) are almost two orders of magnitude larger. While the former are consistent with a spectral gap of about 0.5, the result of Ulmke's QMC method implies a gapless spectrum.

is accurate QMC data obtained using a state-of-the-art QMC algorithm. A full application of the advanced aspects of MEM can only lead to improved results once systematic errors are under control, i.e., are significantly smaller than statistical errors. As we will see in the following, this condition is not always met in this work due to the high precision of the data, i.e., the small statistical errors. By taking the systematic errors into account, we will nevertheless be able to compute quite accurate spectra.

### 3.8.2 Algorithmic Choices and Numerical Tests

In this subsection, we compare various approaches and perform numerical tests in order to establish a robust and accurate method for the analytic continuation of imaginary-time data obtained using the particular QMC code improved and used as part of this thesis.

#### Default Models

One decision that has to be made within all variants of the MEM is the choice of a default model. The simplest possibility is a constant default model $m(\omega) \equiv m_0$ (solid line in Fig. 3.71). Its only parameter is the extent of the frequency grid which should be chosen large enough that it is numerically irrelevant. Then, the usage of

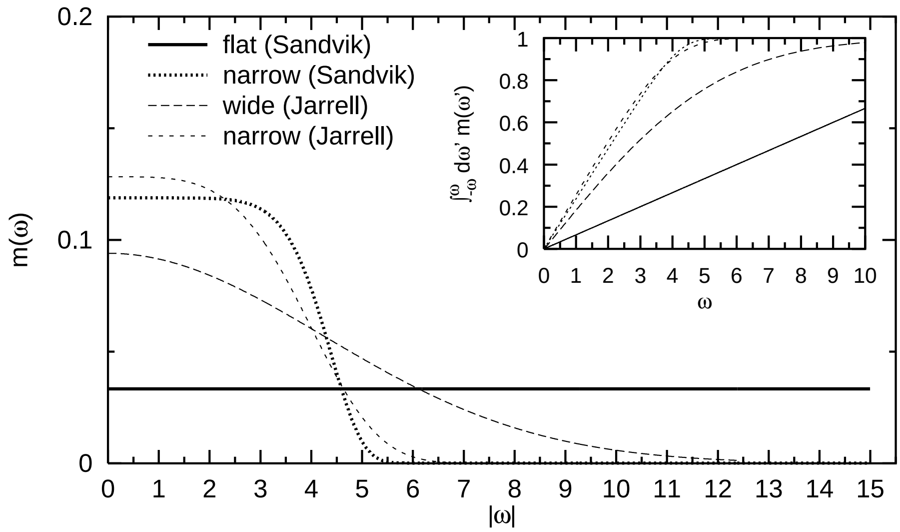



**Figure 3.71:** Default models used within this section: flat model $m(\omega) = 1/(2 * \omega_c)$ (solid line) and narrow model $m(\omega) = \big(\exp[-(0.22437\omega)^8]+0.001\big)/8.4246$ (dotted line) as used in connection with Sandvik's algorithm (both for a cutoff frequency $\omega_c = 15$) and wide model $m(\omega) = \exp[-(w/6)^2]/(6\sqrt{\pi})$ and narrow model $m(\omega) = 0.128281\ \exp[-(\omega/4.3002)^4]$ as used in connection with Jarrell's code (dashed lines). The inset shows the corresponding integrated spectral weights $\int_{-\omega}^{\omega} d\omega' m(\omega')$.

---

a constant default model should guarantee results with minimal bias. On the other hand, QMC data cannot provide sufficient information at large frequencies according to the Nyquist theorem. While the entropic ansatz prevents the MEM from producing artificial structures in this range, it often predicts too much weight at large (absolute) frequencies. This tendency can be reduced by choosing a default model which decays towards the edges of the frequency grid. Ideally, such a model should only incorporate exactly known analytic properties and should not introduce artificial features. In the case of a single-band model with sharp band edges the following observation is useful in this respect: if the total extent of the DOS is $W > 0$, the total extent of the interacting spectrum is bounded by $|U| + W$. This implies for the present symmetric case and $U > 0$:

$$A(\omega) = 0 \quad \text{for} \quad |\omega| > \frac{U+W}{2}\,. \tag{3.68}$$

Since most spectra will be computed for $U \leq 5$ and since $W = 4$ for the Bethe DOS, a good default model should essentially cut off the spectrum at $U \approx 4.5$. This is achieved by the narrow model shown as a dotted line in Fig. 3.71 which is essentially proportional to $\exp(-\omega^8)$; a small offset is necessary for numerical stability.[63] Additional default models used in connection with Jarrell's MEM code

[63] As a constant default model was initially hard-coded in Sandvik's code (as previously used within our group), the usage of nonconstant default model required a slightly generalized version of the program. In our version, important parameters of the default model can be chosen at runtime.

are shown as dashed lines in Fig. 3.71 and are specified in its caption.

It is clearly possible to devise default models which represent more accurately expectations about the true spectrum for the given parameter set, i.e., have features like upper and lower Hubbard bands, a quasiparticle peak or a gap at the Fermi energy etc. Such a default model can be obtained, e.g., as a result of approximate theories, by fitting moments, or by using the MEM result of a related parameter set. In particular, Jarrell (1997) describes an annealing method which works at a hierarchy of temperatures: here, an analytic continuation of QMC data is first performed at high temperatures where a default model obtained from perturbation theory is still quite accurate. The MEM result obtained at each temperature step is then used as a default model at the next lower temperature. While such a procedure can be justified from the observation that spectra usually depend smoothly on temperature and show more structure at lower temperatures, it has no Bayesian foundation and may break down near phase transitions. In our view, the use of the annealing method should be limited to cases where the MEM procedure cannot be stabilized by other means, possibly due to lack of sufficient computing resources. As we will see in the following, such a need does not arise within the context of this thesis.

We also will not need to apply generalizations of the MEM for enforcing analytic constraints (i.e., for moments of the spectrum) by adding further terms to the logarithm of the prior probability defined in (1.63) (i.e., in connection with a Lagrange multiplier). While such a procedure does not pose the danger of introducing artifacts, each enforced constraint removes one possible check of the accuracy of the result. For this reason and for the sake of simplicity, we prefer a plain implementation of the MEM (at least for comparison) in connection with sufficiently accurate data.[64] Note that the full band width can also be constrained by an appropriate choice of the default model.

### Error Analysis

A first point that we would like to stress is the importance of performing the error analysis in the paramagnetic phase of interest for the spin-averaged data $G(\tau) = (G_\uparrow(\tau) + G_\downarrow(\tau))/2$. By taking this average, the majority of statistical errors is eliminated since in each QMC solution of the impurity problem the estimates of $G_\uparrow(\tau)$ and $G_\downarrow(\tau)$ are highly correlated. Due to the correlations, the alternative procedure of treating both estimates as independent measurements (as previously used in our group) leads to incorrect error statistics; in fact, it grossly overestimates errors for small $\Delta\tau$ (or small $\beta - \Delta\tau$) which are here small due to the symmetry of the half-filled model. This is shown in the inset of Fig. 3.72 for $T = 1/20$ and $U = 4.6$. Here, the lower black solid line represents the true measured statistical error. Treating both

In Jarrell's code, the default model (and even the grid) is completely arbitrary, since it is read from a specified file in tabulated form.

[64] In early stages of this work, we found violations of analytic behavior in MEM results when trying to compute the real-frequency self-energy in antiferromagnetic phases. Here, the numerical MEM procedure can probably only be stabilized when some constraints are enforced. The following necessary conditions for analyticity valid for a Bethe lattice ($W = 4$) may be useful for future work: $\operatorname{Im} G_\alpha(\omega)\, ImG_{\bar\alpha}(\omega) \leq 1$ and $\operatorname{Im} G_\alpha(\omega)/\operatorname{Im} G_{\bar\alpha}(\omega) \geq \operatorname{Re} G_\alpha^2(\omega)$. Here, $\alpha$ denotes the sublattice index.

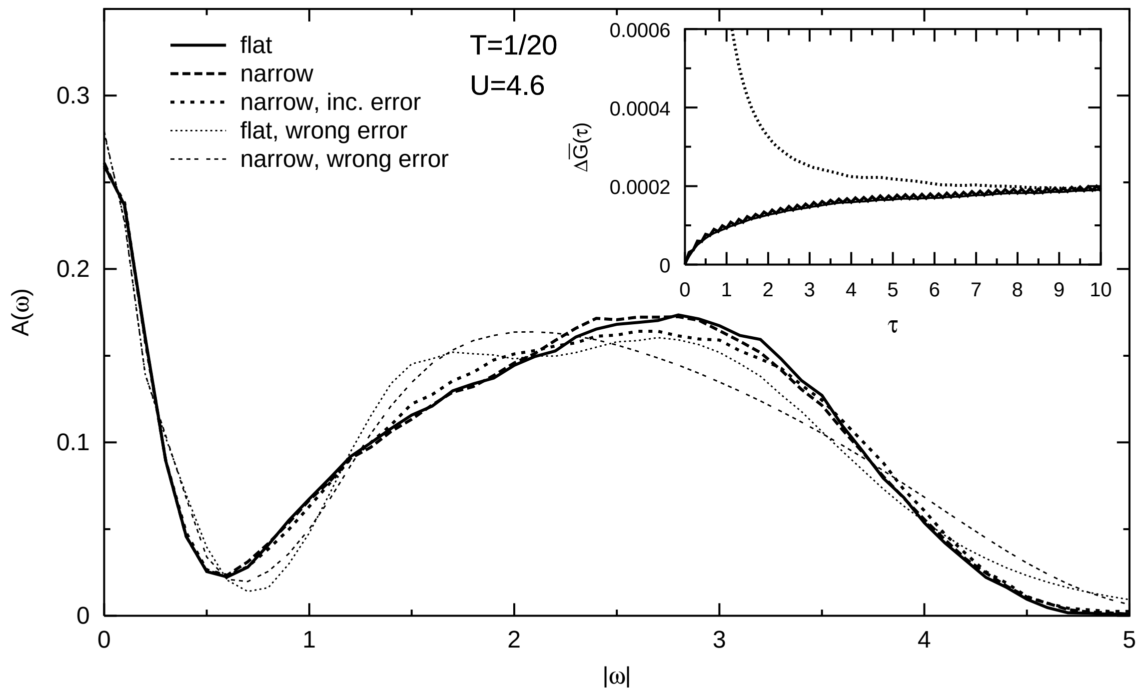



**Figure 3.72:** Local spectral function $A(\omega)$ for $T = 1/20$ and $U = 4.6$: MEM results obtained using the Sandvik scheme based on statistical errors of spin-averaged QMC measurements (thick lines) or wrongly exaggerated errors (thin lines). Solid and dotted lines correspond to a flat default model; dashed lines correspond to a narrow default model (cf. Fig. 3.71). Inset: errors estimated from spin-averaged measurements (lower lines) and exaggerated error (dotted line).

spin species as independent implies the error estimate shown as a dotted line which increases rapidly for $\tau \to 0$ (and extends to about 0.004 at $\tau = 0$). The zig-zag curve also shown in the inset of Fig. 3.72 takes additional systematic errors into account as we will discuss below. The overestimation of errors at small $\tau$ implies that this region is essentially disregarded in the MEM continuation which suppresses the most important information: on the one hand, higher-order moments of the spectrum are directly associated with derivatives $G^{(n)}(\tau)|_{\tau=0}$ via (1.59); on the other hand, the assumption of Gaussian error distribution is very accurate in this region.[65] These considerations are clearly supported by the MEM results obtained using the Sandvik algorithm shown in the main part of Fig. 3.72. Here, the solid and long-dashed lines correspond to calculations based on the adjusted error (red solid line in the inset) for the flat and the narrow default model detailed in Fig. 3.71, respectively. The excellent agreement between both results supports the good quality of the QMC data; the

[65]In fact, Jarrell recently suggested that QMC estimates of $G(\tau)$ should be disregarded in the problematic region of $\tau \approx \beta/2$ in gapped phases in order to avoid negative impact of non-Gaussian error statistics on the MEM results. In the symmetric case at hand, this suggestion would amount to using **only** QMC data at small $\tau$ (which is essentially disregarded in the scheme conventionally employed in our group).

validity of both results is also supported by the fact that only little spectral weight is predicted in the region $|\omega| > 4.3$ forbidden by (3.68). An additional result for the narrow default model that takes into account the possibility of larger systematic errors (short-dashed line) yields a somewhat more semi-elliptic shape of the Hubbard bands, but is overall consistent with the results discussed so far. In contrast, MEM results based on the exaggerated error derived from unsymmetrized measurements (thinner curves) differ clearly: both the much stronger model dependence and the large weight at high frequencies indicate the poor quality of these results.

As we have pointed out in the beginning of this section, Jarrell's program represents a careful implementation of the Bayesian MEM approach and is in general superior over simplified algorithms as implemented in the code originating from Sandvik. Consequently, it would appear natural to only use the more advanced code for this thesis. However, numerical instabilities were encountered in a direct application of this method to our best QMC data (for $\Delta\tau = 0.1$, i.e., $\Lambda = 200$ time slices), in particular for very accurate data sets with small statistical errors. While some progress could be made when the impact of the systematic iteration error was reduced by combining measurements performed for different impurity models (cf. subsection 3.8.1), consistent results could only be obtained by skipping time slices as illustrated in Fig. 3.73. Here, best results (thick solid and dotted lines) were obtained by using only every $4^{\text{th}}$ time slice (i.e., $\tau_0 = 0$, $\tau_4 = 0.4$ etc.) as can be seen from the relatively small dependence on the default model and from the comparison with the "Sandvik" result (double-dashed line) reproduced from Fig. 3.72. Skipping only every second time slice results in spectra (dashed lines) with additional features or too much weight at large frequencies; for a wide default model, even the analytic constraint $A(\omega = 0) < 1/\pi$ is violated. When using the full QMC data as input, the MEM scheme did not even converge.[66] A welcome side effect of skipping time slices is that the overall impact of covariance is reduced which stabilizes the SVD.

The primary reason for these problems becomes apparent in the main panel of Fig. 3.74. Here, the imaginary-time Green function corresponding to the MEM spectrum obtained from Sandvik's method for a narrow default model (cf. Fig. 3.72) has been subtracted from the measured QMC data. On the resulting magnified scale, an oscillatory behavior with deviations from a smooth curve which alternate on neighboring time slices becomes clearly visible. Since no such short-time oscillations appear in corresponding differences of Green functions associated [via (1.55)] with different MEM spectra (not shown), the phenomenon is recognized as a systematic error of the QMC data. Most likely, this error is associated with remaining deficiencies of the Fourier transformation scheme employed in this work and could be removed by employing a more sophisticated scheme as presented in App. C. While the individual errors are roughly in line with estimated error bars (of about $10^{-4}$) for each time slice, the pattern is so strong (in particular for small $\tau$) that the associated unphysical degree of freedom can cause a failure of the full MEM procedure (in Jarrell's code). On the other hand, the short-time oscillations are removed when only every second or every fourth time slice is included in the analysis which explains why the stability

[66] In tests for other parameter sets where numerical convergence could be obtained, additional separated peaks and/or a central dip in the quasiparticle peak appeared in the spectra.

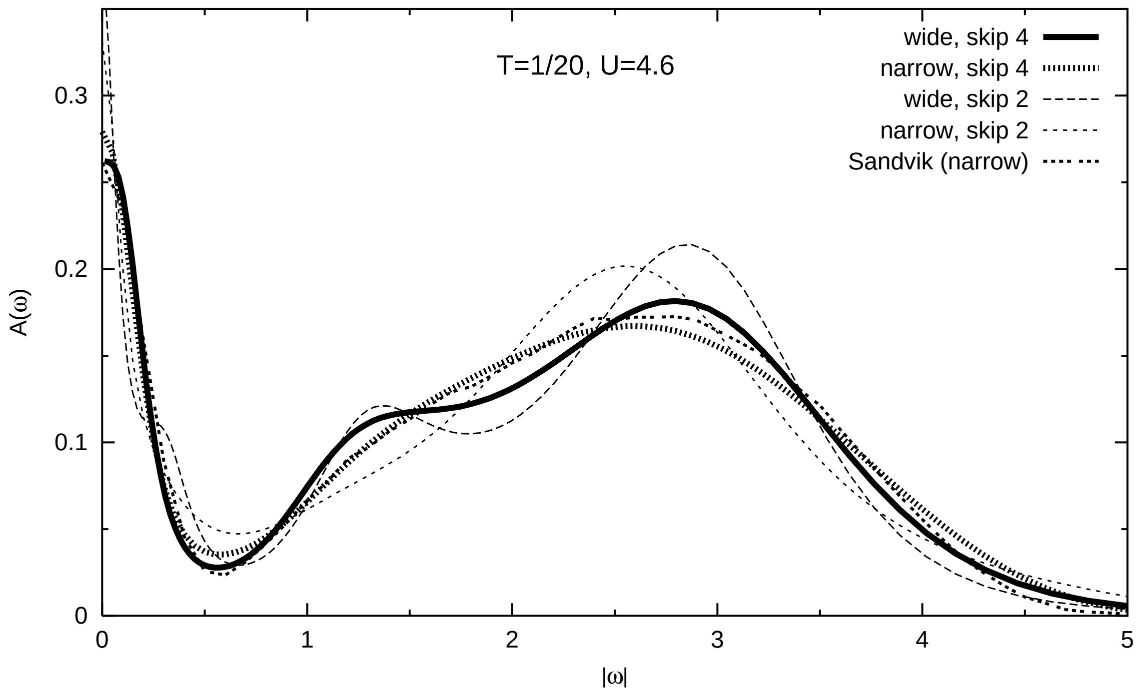



**Figure 3.73:** Local spectral function $A(\omega)$ for $T = 1/20$ and $U = 4.6$: MEM results obtained using Jarrell's program in connection with a wide or a narrow default model (cf. Fig. 3.71). In order to eliminate unphysical oscillations in the imaginary time, only every $4^{\text{th}}$ (thick lines) or every $2^{\text{nd}}$ time slice (thin dashed lines) is used for the analysis, the rest is skipped. A result of Sandvik's method (double-dashed line; cf. Fig. 3.72) is included for comparison.

of the MEM procedure is then reestablished. The longer-range oscillations visible in the difference shown in the main panel of Fig. 3.74 is also unphysical; this is seen from the presence of local maxima and minima of the (smoothed) second derivative $\mathrm{d}^2\bar{G}(\tau)/\mathrm{d}\tau^2$ shown in the inset of the same figure.

Since Sandvik's MEM scheme treats the time slices as being independent, it is not directly susceptible to the short-range oscillation. Still, the associated underestimation of errors at small $\tau$ (where the statistical errors are very small in the symmetric case) may negatively impact the results. For this reason, the computed statistical error is adjusted throughout this section by adding $2 \cdot 10^{-5}$ to $\Delta\bar{G}_l$ for all odd time slices $l$; the resulting zig-zag shape of the error estimate was already shown as a red line in the inset of Fig. 3.72. Using a larger error increment for the time slices with odd index (putting also a lower bound on the relative error estimate) (short-dashed line in the same figure) only leads to minor modification of the spectrum which demonstrates the robustness of our approach.

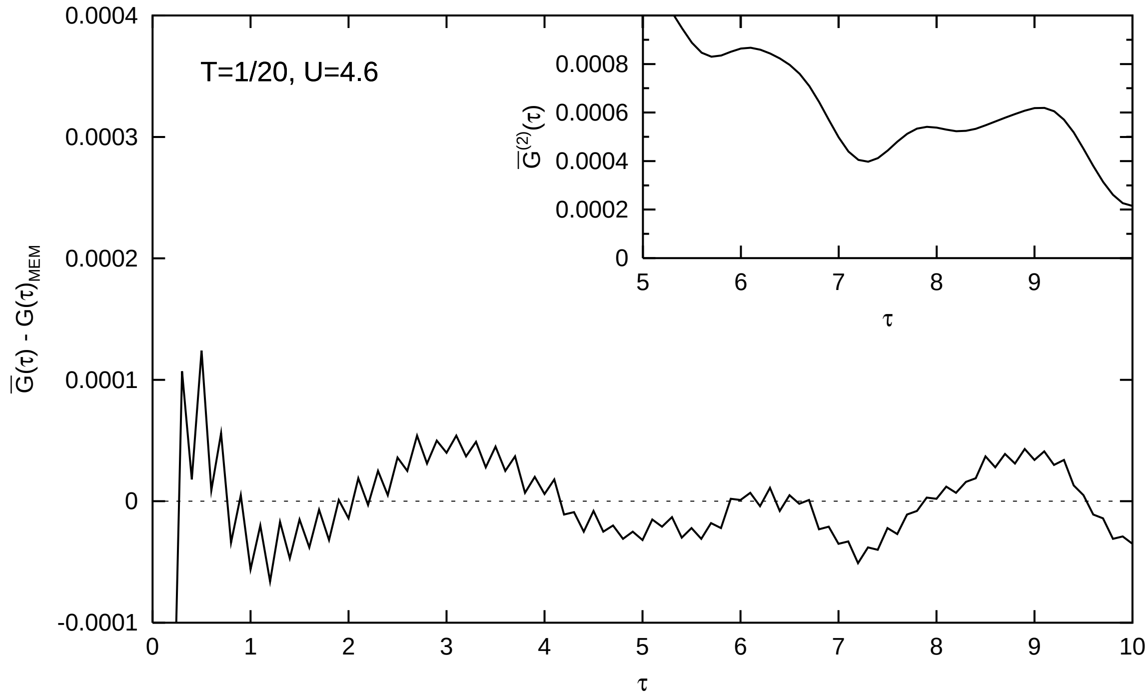



**Figure 3.74:** Difference between imaginary-time Green function $\bar{G}(\tau)$ estimated by QMC and $G(\tau)$ corresponding to the MEM spectrum obtained using Sandvik's scheme and a narrow default model (solid line in Fig. 3.72). Inset: second derivative $\mathrm{d}^2\bar{G}(\tau)/\mathrm{d}\tau^2$, smoothed with Gaussian of width 0.3.

### 3.8.3 Numerical Results for the Bethe DOS

In this subsection, we present numerical results for the local spectral function of the half-filled Hubbard model with semi-elliptic DOS at moderately low temperatures ($T = 0.05$ and $T = 0.067$). These results are based on the MEM procedure discussed above (using Sandvik's scheme with properly adjusted errors) and on the QMC estimates of the imaginary-time Green function shown in Fig. 3.75 for $T = 0.05$ and in Fig. 3.76 for $T = 0.067$. Using the symmetry $G(\beta - \tau) = G(\tau)$, both figures are split for easier comparison of metallic solutions (left panel, only shown for $\tau \leq \beta/2$) and insulating solutions (right panel, only shown for $\tau \geq \beta/2$). Furthermore, the noninteracting Green function ($U = 0.0$) is included for reference.

The spectra for the metallic phase at $T = 0.05$, i.e., slightly below the critical temperature $T^* \approx 0.055$, are shown in Fig. 3.77 and Fig. 3.78 as obtained from MEM using a flat and a narrow default model, respectively. The corresponding results for the insulating phase are depicted in Fig. 3.79 and Fig. 3.80. Evidently, the dependence of the spectra on the default model is negligible for all interactions $U$ presented in the graphs. In the metallic phase, the spectral density at the Fermi level ($\omega = 0$) is approximately pinned at the noninteracting value $\rho(0) = 1/\pi \approx 0.32$ for $U \lesssim 4.4$. Since Luttinger's theorem is exact only at $T = 0$, the slight deviation in this Fermi liquid regime can be attributed to the finite temperature. The larger deviation for

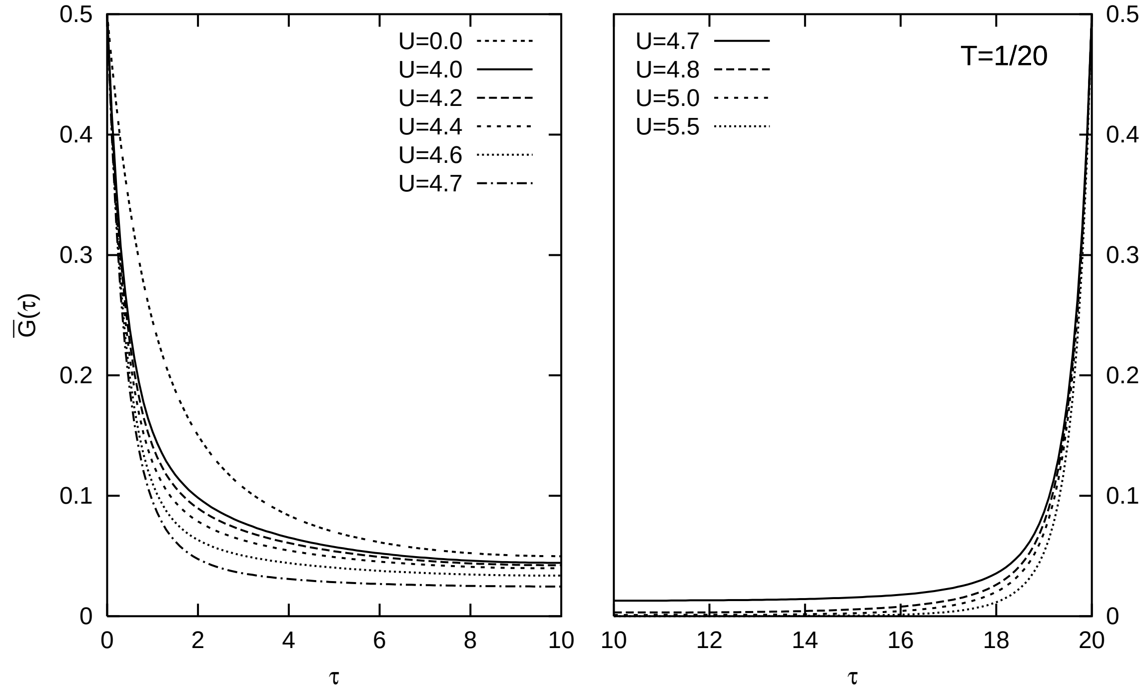



**Figure 3.75:** QMC estimates of the imaginary-time Green function for $T = 1/20$ and $\Delta\tau = 0.1$.

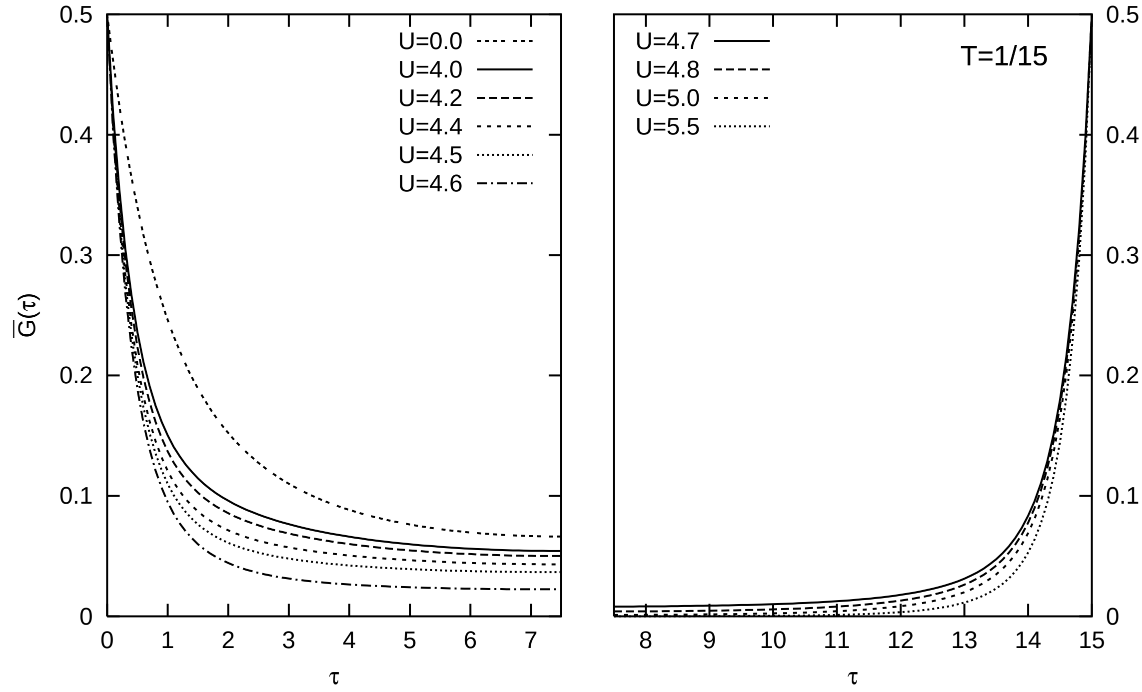



**Figure 3.76:** QMC estimates of the imaginary-time Green function for $T = 1/15$ and $\Delta\tau = 0.1$.

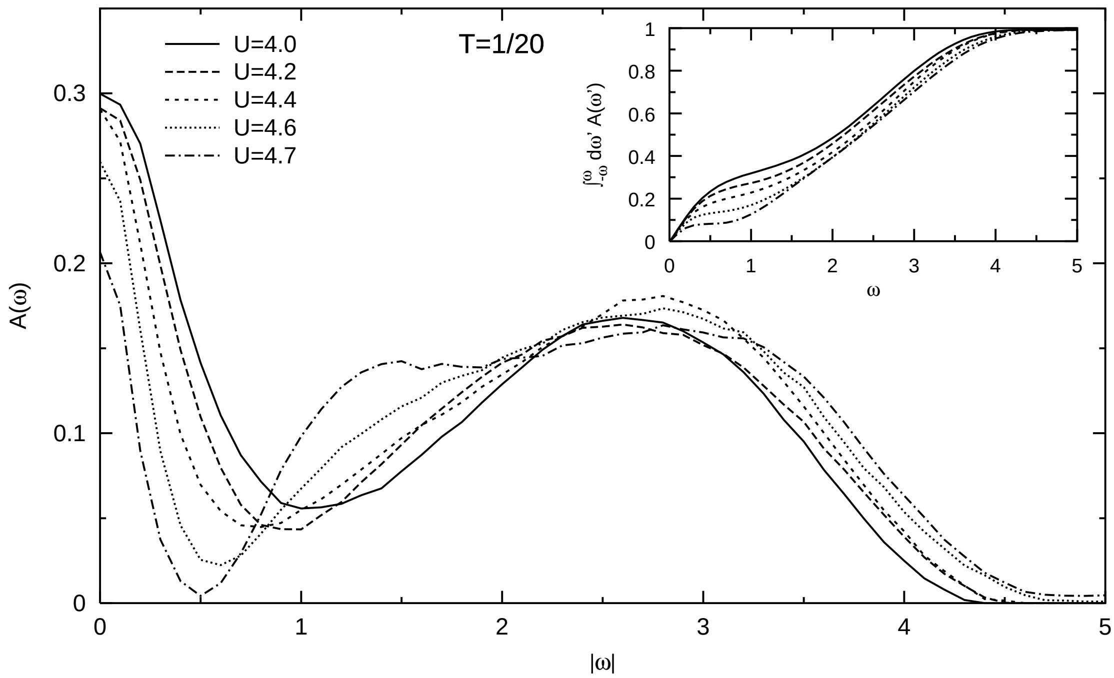



**Figure 3.77:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic metallic phase at $T = 0.05$: QMC/MEM result for $\Delta\tau = 0.1$ using a flat default model. Inset: integrated spectral weights.

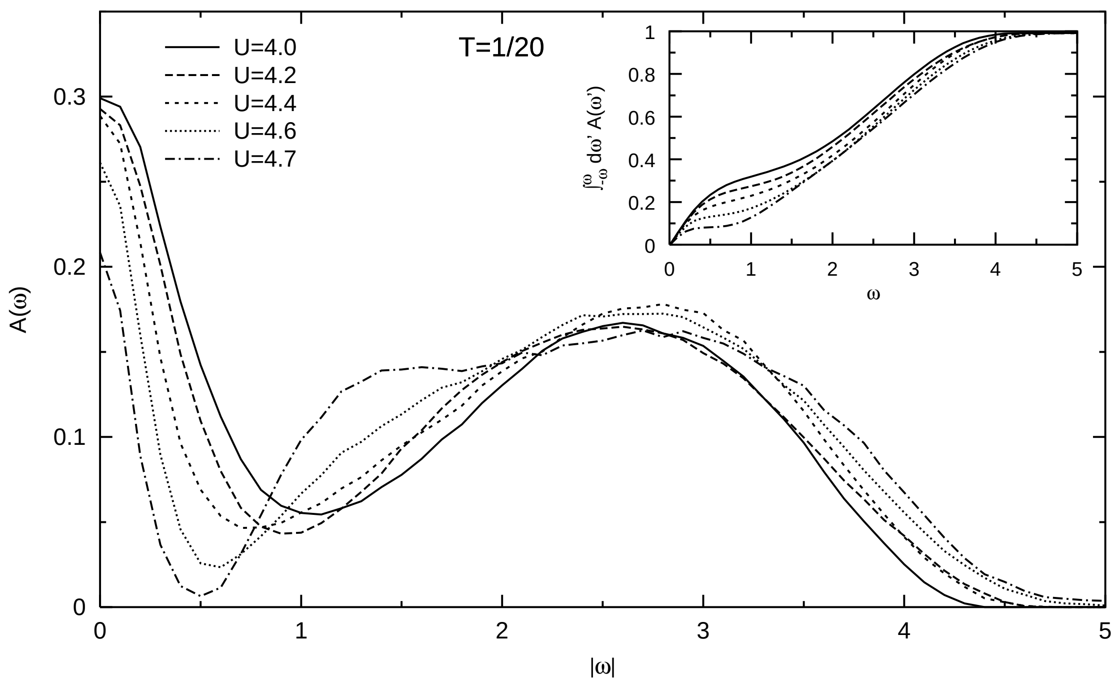



**Figure 3.78:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic metallic phase at $T = 0.05$: QMC/MEM result for $\Delta\tau = 0.1$ using a narrow default model. Inset: integrated spectral weights.

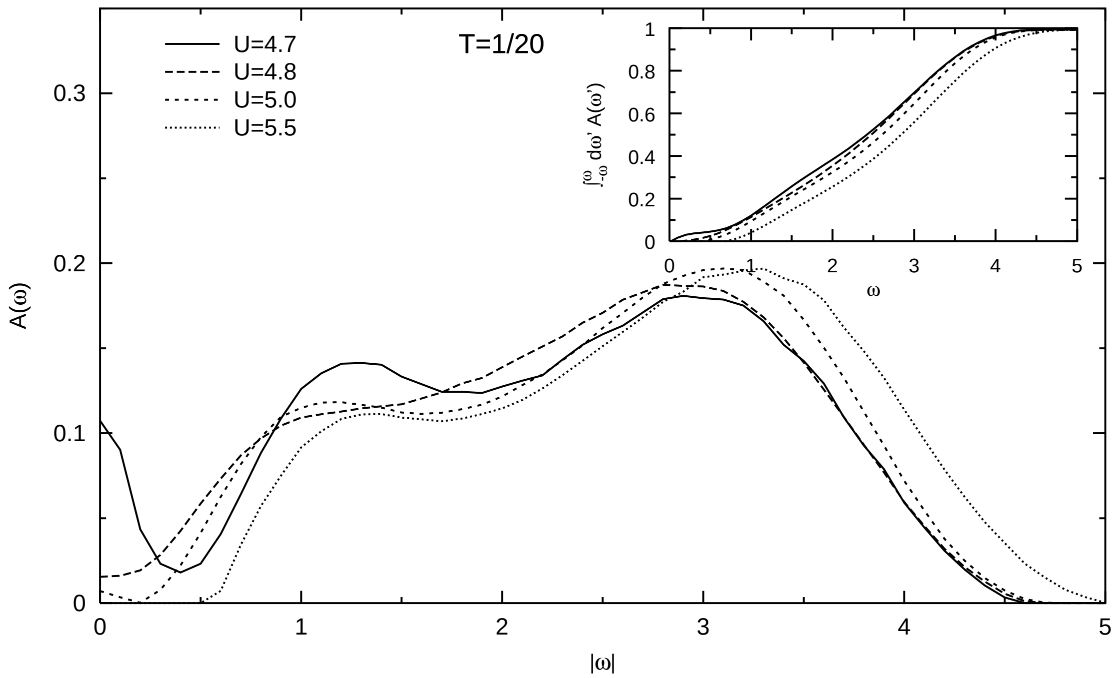



**Figure 3.79:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic insulating phase at $T = 0.05$: QMC/MEM result for $\Delta\tau = 0.1$ using a flat default model. Inset: integrated spectral weights.

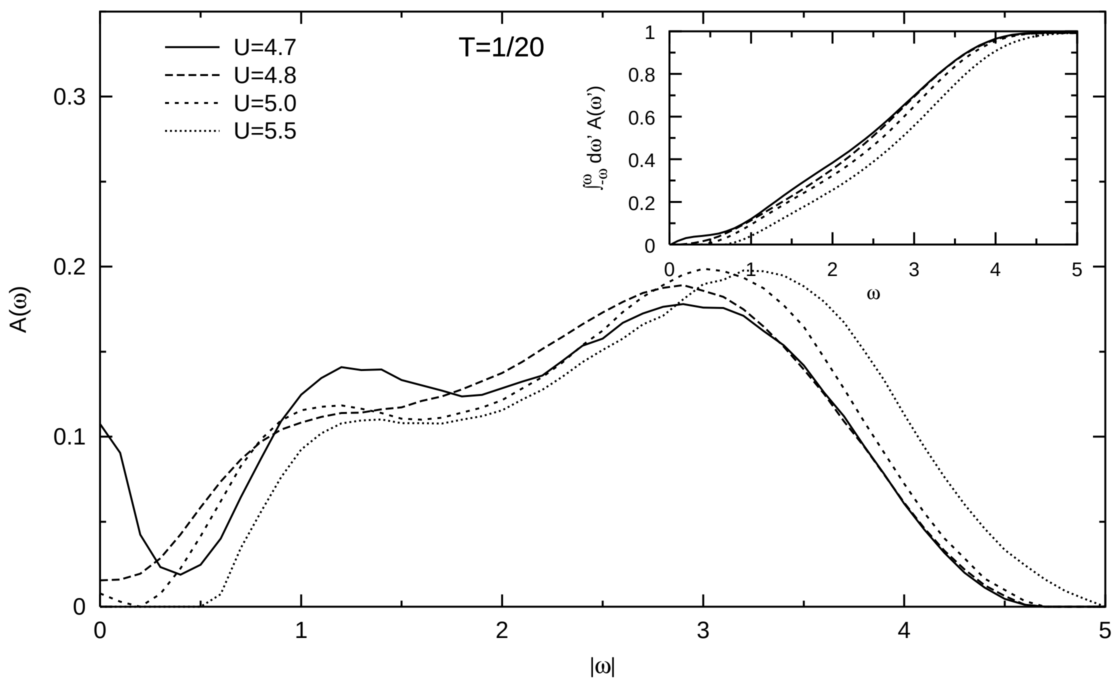



**Figure 3.80:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic insulating phase at $T = 0.05$: QMC/MEM result for $\Delta\tau = 0.1$ using a narrow default model. Inset: integrated spectral weights.

$U = 4.6$ may be exaggerated due to the finite frequency grid $\Delta\omega = 0.1$ (which is here already of the same order of magnitude as the width of the quasiparticle peak). The clearly reduced peak hight at $U = 4.7$, however, marks the destruction of the Fermi liquid. For this value of the interaction, an additional shoulder or peak appears in each Hubbard band. As seen from the integrated spectral weights shown in the insets of Fig. 3.77 and Fig. 3.78, the main difference between the spectra for $U = 4.6$ and $U = 4.7$ is a shift of spectral weight of about 0.05 from $|\omega| \lesssim 0.5$ to $|\omega| \gtrsim 0.5$.

The spectra corresponding to the insulating phase at $U = 4.7$ shown in Fig. 3.79 and Fig. 3.80 do not follow the expected form. In fact, the remnant of a quasiparticle peak points to a more metallic character. Since $U = 4.7$ is right at the edge of the stability region for $T = 0.05$ and $\Delta\tau = 0.1$ as previously seen in Fig. 3.24, this fact may be due to numerical errors (although the corresponding QMC solution has been found to be stable for some 20 iterations). Such ambiguities are absent for $U \geq 4.8$: a pseudogap with very low residual spectral weight appears for $U = 4.8$. Within the numerical accuracy, the spectrum for $U = 5.0$ is already gapped with a gap size of about 0.5 which is in line with the estimate obtained in Fig. 3.70. The gap size increases to about 1.2 for $U = 5.5$. A shoulder or two-peak structure of the Hubbard band is seen in all results for the highly correlated metallic phase and for the insulating phase which supports the genuineness of this feature. It does not become more pronounced with increasing interaction (in the insulating phase), but is slightly reduced from $U = 5.0$ to $U = 5.5$. Such a behavior is expected since rigid semi-elliptic Hubbard bands (of widths $W$ and centered around $\omega = \pm U/2$) have to be recovered for $U \to \infty$ (Gebhard, 1997).

Results for $T = 0.067$, i.e., above the critical temperature, are shown in Fig. 3.81 for the more metallic phase and in Fig. 3.82 for the more insulating phase. We caution the reader that these results have not been checked as carefully as the preceding results (for $T = 0.05$). We still expect the following observations to be genuine. It is seen in Fig. 3.81 that the height of the quasiparticle peak decreases significantly with increasing $U$ already for $4.0 \lesssim U \lesssim 4.5$. Thus, the Fermi liquid phase breaks down before the crossover towards the insulator (defined in terms of the energy dependence; cf. Fig. 3.23) takes place at $U \approx 4.62$. Again, the appearance of a second peak in the Hubbard band is associated with the disappearance of the quasiparticle peak, here for $U = 4.6$. On the insulating side, the gap opens very slowly; significant spectral weight is still present at $|\omega| \approx 0$ for $U = 4.7$ and $U = 4.8$. Deep inside the insulating phase at $U = 5.0$ and $U = 5.5$, however, the spectra are quite similar to the spectra computed for $T = 0.05$ with gap sizes of about 0.5 and 1.2, respectively. The temperature independence of the spectra in the insulating phase is in line with the corresponding observation of such independence made for the energy and the double occupancy in subsection 3.5.2.

Finally, let us compare the spectra computed within this thesis with previously published work. Fig. 3.83 a reproduces spectra for $T = 0.05$ computed by Jarrell and included in our joint publication (Schlipf et al., 1999). Considering the slight difference in the imaginary-time discretization used by Jarrell ($\Delta\tau = 0.167$ versus $\Delta\tau = 0.1$ for our results), the overall agreement is reasonable. The better agreement with Luttinger's theorem for $U = 4.0$ may be due to the finer frequency grid used by

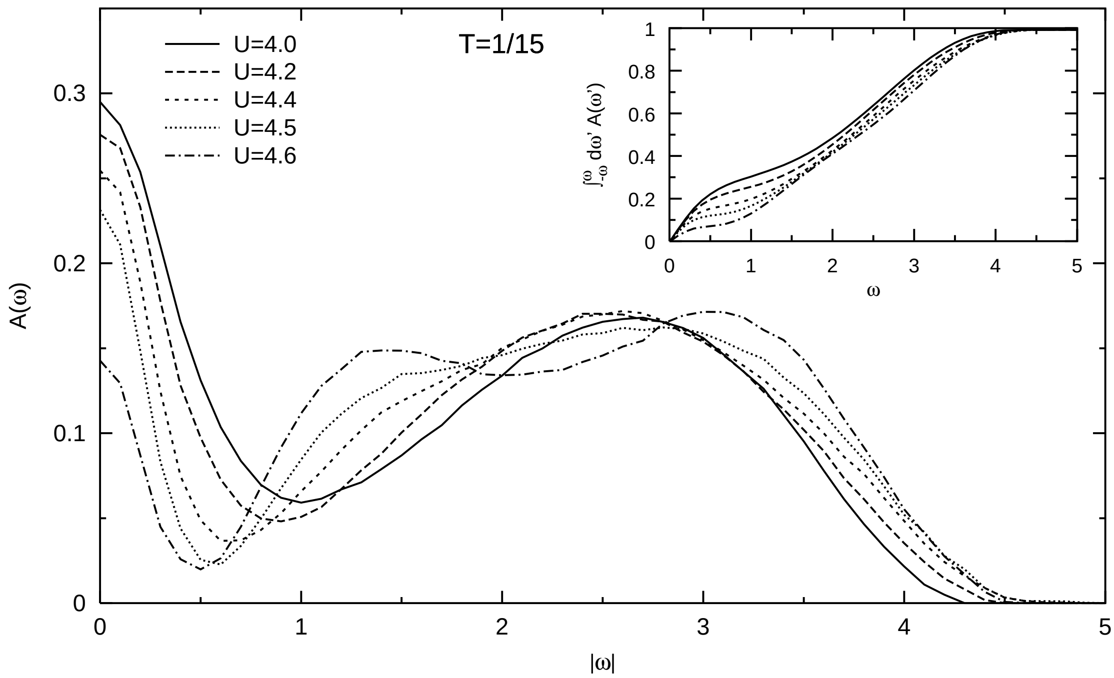



**Figure 3.81:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic metallic phase at $T = 0.067$: QMC/MEM result for $\Delta\tau = 0.1$ using a flat default model. Inset: integrated spectral weights.

---

Jarrell near $\omega \approx 0$. However, significant weight is predicted in the "forbidden" region $|\omega| > (U+W)/2 = 2+U/2$ for all interactions $U$. Furthermore, we regard the height of the peaks seen for $U = 5.2$ at $|\omega| \approx 1$ and the degree of their separation from the main peaks at $|\omega| \approx 3$ as artifacts. We attribute the higher reliability of our results primarily to a much larger investment in computer time.

NRG spectra for the significantly lower temperature $T = 0.041$ are reproduced in Fig. 3.83 b (Bulla et al., 2001). According to this method, the quasiparticle peak disappears quite abruptly at the MIT. Note, however, the significant deviations from Luttinger pinning (at $4\,\rho(0) = 1.27$) close to the MIT even at this lower temperature. A slight tendency towards formation of a shoulder is seen for $U/W = 1.22$; due to the increasingly bad resolution of the NRG at large frequency, however, this feature is not significant. The extended high-frequency tails are clearly artifacts of the logarithmic discretization (which requires corresponding broadening of all results).

The main results of this section are the construction and test of a robust scheme for computing MEM spectra of the $d = \infty$ Hubbard model and the spectra obtained for $T = 0.05$ shown in Figs. 3.77–3.80. The reliability of these results has been successfully tested for internal consistency, consistency with analytic bounds and consistency with previously published results. The spectra will pass further consistency checks when the self-energy is computed on the real axis in Sec. 4.6. An adapted MEM scheme will be applied to a 3-band model for $La_{1\text{-}x}Sr_xTiO_3$ in Sec. 5.4.

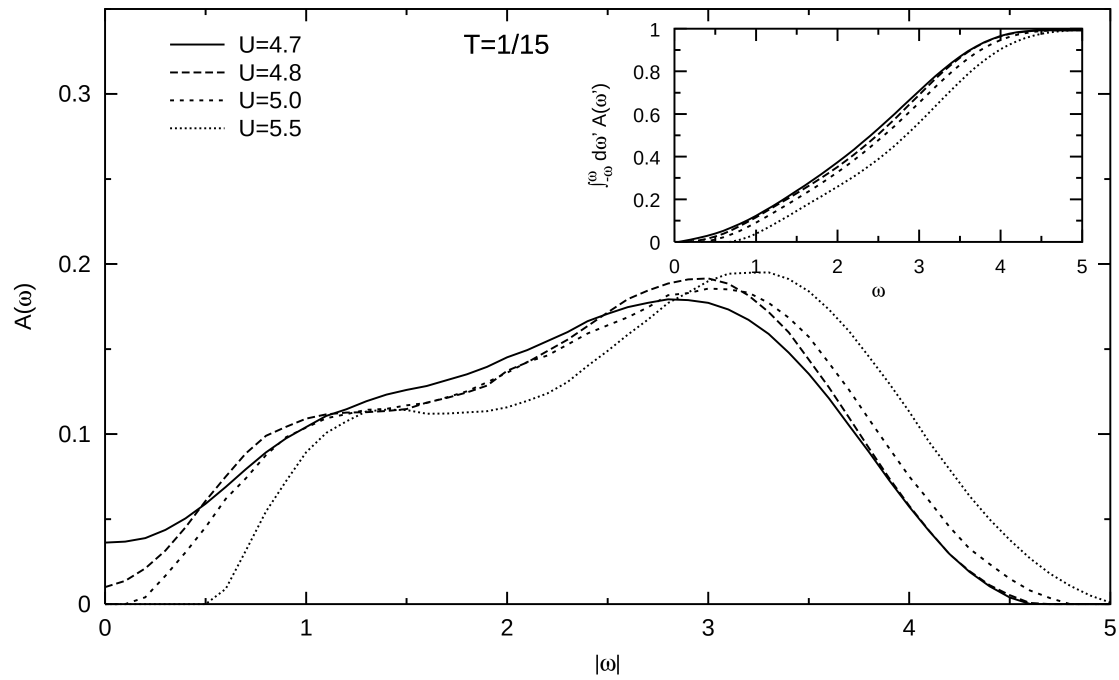



**Figure 3.82:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic insulating phase at $T = 0.067$: QMC/MEM result for $\Delta\tau = 0.1$ using a flat default model. Inset: integrated spectral weights.

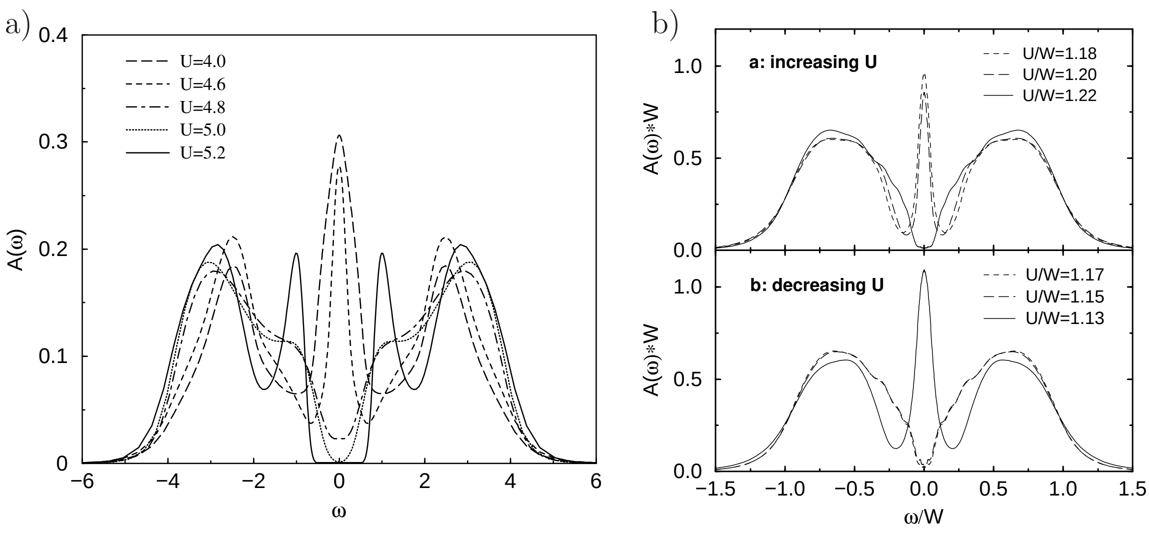



**Figure 3.83:** Local spectral function of the half-filled Hubbard model with semi-elliptic DOS for $d \to \infty$ in the paramagnetic phase: **a)** QMC/MEM results for $T = 0.05$ (and $\Delta\tau = 0.167$) obtained by Jarrell [figure adapted from Schlipf et al. (1999)]. **b)** NRG results for $T = 0.041$ (using the scale $W = 4$ of this work) published by Bulla et al. (2001).

## 3.9 Conclusion

In this chapter, we have studied the fully frustrated half-filled Hubbard model with a semi-elliptic noninteracting density of states in the limit of infinite dimensionality near the Mott metal-insulator transition (MIT).

We have computed the coexistence region of metallic and insulating phases with significantly higher accuracy than achieved previously in the literature. In the course of this study, we have discussed the signatures of a first-order transition within the DMFT and for the first time established a rigorous relation between the stability of fixed points in a DMFT iteration scheme and the thermodynamic (meta-) stability of phases. Using an improved method for the quantum Monte Carlo (QMC) calculation of the kinetic energy, we have presented the first accurate estimates of the full energy (per electron) and have shown that this observable is far better suited for the QMC determination of phase boundaries than other observables considered so far. We have explained the initial contradiction between QMC results regarding the possible extent of the coexistence region by identifying strengths and weaknesses of the QMC codes and strategies used by the major groups in the field. In particular, we have found and corrected a major flaw in the QMC code used previously by our group; while the original code was already correct in the limit of vanishing discretization $\Delta\tau \to 0$, it had introduced an unnecessary error at small frequencies which suppressed the coexisting insulating phase at attainable values of $\Delta\tau$. We have also parallelized the QMC code in order to efficiently use all available computing resources.

While the coexistence of metallic and insulating solutions indicates the presence of a first-order transition, the phase diagram is only complete with the determination of the thermodynamic first-order transition line $U_{\rm c}(T)$. This task was previously regarded as too difficult for numerically exact methods that cannot directly compute the free energy of the impurity model (such as QMC and NRG); even the prediction of the iterative perturbation theory was initially unclear. We have derived a first-order differential equation for the phase boundary; its solution is more stable than a more direct comparison of free energies due to the cancellation of errors. On the basis of highly accurate measurements for the energy and the double occupancy, we have established that the differential equation may be linearized without loss of accuracy. We have derived the leading low-temperature asymptotics of $U_{\rm c}(T)$ from Fermi liquid theory and have fixed a numerical coefficient by calculating the ground state properties of the metallic phase. This was possible by performing an interpolating fit of new (zero-temperature) second-order perturbation theory estimates with results for $U_{\rm c}(T=0)$ previously obtained by ground state methods. The quality of this fit has been stabilized by demanding thermodynamic consistency and has been successfully tested by comparison with extrapolated QMC data. As a byproduct, we have calculated the coefficient $\gamma(U)$ of the low-temperature specific heat and checked the consistency with extrapolated estimates for the quasiparticle weight $Z$. Due to the high precision of the QMC data and the numerous independent checks, the accuracy of the resulting first controlled estimate of the thermodynamic first-order phase transition line is well established.

We have analyzed Tong et al.'s (2001) recently published independent estimates for the first-order phase transition line based on a scheme using a modified set of self-consistency equations. While the agreement on the level of $U_c(T)$ is impressive, we have found significant discrepancies with respect to the measured double occupancies which this group had used in a Maxwell construction for $U_c(T)$. In addition, we have demonstrated the approximate character of this scheme. We have also studied signatures of critical behavior and have found that its impact on QMC studies is limited to the close surroundings of the (finite-temperature) critical end point of the first-order line. In addition, we have reanalyzed IPT data for the double occupancy in the critical region and have demonstrated that it is governed by a critical exponent of 1/2 rather than the value of 1/3 obtained by Kotliar et al. (2000).

Finally, we have discussed the maximum entropy method (MEM) and its application to the analytic continuation of imaginary-time Green functions computed within QMC. We have demonstrated the strong improvement resulting from our correction of the QMC code and devised a robust MEM scheme that takes the remaining systematic errors into account. By applying this method to large sets of precise QMC measurements, we have calculated the first low-temperature MEM spectra in the MIT region which do not depend on specially chosen ("annealed") default models. The overall high precision of the results is indicated by the good agreement with analytical bounds for the total bandwidth and by the excellent agreement with a sum rule for the derived real-frequency self-energy (to be discussed in Sec. 4.6).

In view of the excellent internal consistency of our results and the good agreement with data obtained using other methods (within their respective error bars), our final phase diagram Fig. 3.50 and other results presented in this chapter should represent a fairly precise and complete picture of the local properties of the fully frustrated model. However, they clearly do not represent a complete scenario for any material featuring a MIT. This follows already from the finite ground-state entropy predicted in the insulating phase which violates Nernst's law. A more realistic and consistent description is possible when the low-temperature phase is ordered already in $d = \infty$, i.e., for partial frustration. Analytical considerations for the consequences in the case of varying partial frustration at constant semi-elliptic DOS have been presented in this chapter. Investigations whether a phase diagram with a topology analogous to that seen in $V_2O_3$ can also be observed in the generic case that frustration induces an asymmetry of the DOS have recently been initiated under our supervision (Knecht, 2002); at present, this question is still open.

A further significant improvement of the QMC code based on ideas developed in this thesis and implemented by Knecht (2002) is briefly discussed in App. C; numerical comparisons shown in this appendix also demonstrate the high accuracy of the (less general) approach employed within this work. Future DMFT studies could also greatly profit from a scheme for the direct evaluation of differences in the free energy of coexisting solutions developed in this chapter.

# Chapter 4

# Optical Conductivity

In this chapter, we will discuss electronic transport, in particular the ($\boldsymbol{q} = 0$) optical conductivity $\sigma(\omega)$. Transport properties are not only very important technologically, but can also be probed accurately by a variety of techniques. As we will see below, measurements, e.g., of reflectivity, electronic energy loss, and of the direct-current (dc) conductivity are, thus, valuable tools for determining the electronic structure of materials. Specifically, a metal-insulator transition as discussed in chapter 3 is signaled by a vanishing Drude weight, i.e., a suppression of $\sigma(\omega)$ at low frequencies (and specifically for $\omega = 0$).[1]

We will here focus on the exact theory of transport strictly in the limit of infinite dimensionality ($d \to \infty$) or, more generally, infinite lattice coordination number ($Z \to \infty$), but also mention implications for an approximate treatment of the finite-dimensional case. Since all directional transport vanishes in infinite dimensions,[2] e.g.,

$$\sigma_{xx}(\omega) \stackrel{d\to\infty}{\longrightarrow} 0 + \mathcal{O}(1/d) \tag{4.1}$$

for the optical conductivity in $x$ direction, we will consider leading corrections, e.g., $d\sigma_{xx}(\omega)$ or $\sum_\alpha \sigma_{\alpha\alpha}(\omega)$. While the fact that most unscaled transport properties vanish in high dimensions might be viewed as a conceptual complication, their computation is much simplified in this limit: due to the effective locality of the vertex function (for $\boldsymbol{q} = \boldsymbol{0}$), their diagrammatic expressions reduce to the bubble diagram. Consequently, $\sigma(\omega)$ can be evaluated from the full interacting Green function and the unrenormalized current vertex. It is important to note that the current vertex introduces a dependence on lattice structure and hopping parameters beyond the noninteracting DOS so that models which are equivalent in the DMFT with respect to local properties may differ significantly in terms of transport properties. In fact, we will show in this chapter that $\sigma(\omega)$ is not uniquely defined on "the Bethe lattice", i.e., on

[1] In principle, we could have used results for $\sigma(\omega)$ in order to determine the MIT transition lines $U_{c1}$ and $U_{c2}$ in chapter 3. The attainable accuracy is, however, much better for the energy $E$ and for the double occupancy $D$ since these observables do not require an analytic continuation via MEM.

[2] Strictly speaking, this statement only applies to total or partial weights, i.e., integrals over finite frequency ranges and only when the lattice is essentially isotropic, i.e., when the kinetic energy is not dominated by a subset of directions. In particular, the limits $\omega \to 0$ and $d \to \infty$ do not necessarily commute.

lattices having a semi-elliptic DOS for $Z \to \infty$ (cf. Sec. 2.2) and present formalism and results for several possible choices. We will demonstrate why some approaches fail to describe coherent transport and that the extended-hopping general-dispersion approach proposed in Sec. 2.3 produces the physically most desirable results.

As indicated above, the numerical results of this chapter correspond to a consistent theory which is exact in $d = \infty$. We will also point out how one may alternatively use the formalism as a local approximation in finite dimensions, e.g., in connection with LDA calculations (cf. subsection 5.4.3). We will not, however, discuss vertex corrections which may change the picture completely in $d = 3$ or even lower dimensions.

The remainder of this chapter is structured as follows: In Sec. 4.1, we introduce $\sigma(\omega)$ and other dielectric quantities on a phenomenological level, collect relevant sum rules, point out the relation to experimental results, and discuss the impact of model abstractions. The Kubo formalism is reviewed in Sec. 4.2 both for continuum and lattice models. The simplifications arising in $d \to \infty$ are discussed in Sec. 4.3 where we will also point out that the formalism widely used in the literature so far is very specific to the hypercubic lattice and must be modified in the general case. Variants of the general formalism applicable to tight-binding models with a semi-elliptic density of states are studied in Sec. 4.4. Generalizations to models with varying degree of coherence and to finite dimensions as well as the theory of transport properties for frustrated lattices with $t - t'$ hopping are discussed in Sec. 4.5. Finally, we present numerical results for the Hubbard model with semi-elliptic DOS in Sec. 4.6.

## 4.1 Definition and General Properties of the Optical Conductivity

If a system is subjected to an external electric field then, in general, a redistribution of charges occurs and currents are induced. For small enough fields, the induced polarization and the induced currents are proportional to the inducing field. If the Hamiltonian of the unperturbed system is not explicitly time-dependent, we have specifically for the total current density in linear response,

$$J_\alpha(\boldsymbol{r}, t) = \sum_\beta \int \mathrm{d}\boldsymbol{r}' \int_{-\infty}^{t} \mathrm{d}t' \, \sigma_{\alpha\beta}(\boldsymbol{r}, \boldsymbol{r}', t - t') E_\beta(\boldsymbol{r}', t'), \tag{4.2}$$

where $\alpha, \beta$ label space directions and $\boldsymbol{E}$ is the (total) electric field. If the system is homogeneous above some length scale $a$ (e.g., for a lattice spacing $a$), then the response to perturbations which are slowly varying on this length scale is also homogeneous,

$$J_\alpha(\boldsymbol{r}, t) = \sum_\beta \int \mathrm{d}\boldsymbol{r}' \int_{-\infty}^{t} \mathrm{d}t' \, \sigma_{\alpha\beta}(\boldsymbol{r} - \boldsymbol{r}', t - t') E_\beta(\boldsymbol{r}', t'), \tag{4.3}$$

which can be Fourier-transformed to yield

$$J_\alpha(\boldsymbol{q},\omega) = \sum_\beta \sigma_{\alpha\beta}(\boldsymbol{q},\omega) E_\beta(\boldsymbol{q},\omega). \tag{4.4}$$

Since photons have a much steeper (linear) energy dispersion than electrons within the Brillouin zone, the transmitted momentum is essentially zero for not too high energies. Thus, in the optical frequency regime and below we can take the long-wavelength limit $\boldsymbol{q} \to 0$. For cubic symmetry, the response is then isotropic, i.e., the conductivity tensor becomes diagonal:[3]

$$\boldsymbol{J}(\omega) = \sigma(\omega)\boldsymbol{E}(\omega). \tag{4.5}$$

Note that causality, i.e., the restriction to $t' < t$ in (4.2) and (4.3), implies a retarded conductivity $\sigma(\omega)$ which is analytic in the upper half plane.

### 4.1.1 Connection between Conductivity and Reflectivity

Under the conditions used in the derivation of (4.4), the dielectric function $\epsilon(\omega)$ is defined as relating the (coarse-grained) electrical displacement $\boldsymbol{D}$ to the (coarse-grained) electric field $\boldsymbol{E}$,[4]

$$\boldsymbol{D}(\omega) = \epsilon(\omega)\boldsymbol{E}(\omega). \tag{4.6}$$

The dielectric function can be directly expressed in terms of the optical conductivity,

$$\epsilon(\omega) = 1 + \frac{4\pi i \sigma(\omega)}{\omega}. \tag{4.7}$$

Note that here $\sigma(\omega)$ here has the dimension of a frequency (or energy) which is only valid in the Gaussian system of units. Also note the importance of the absolute scale of $\sigma(\omega)$ which determines qualitative features of $\epsilon(\omega)$ as, e.g., the zeros of $\epsilon_1(\omega)$. For later use, we also state the equivalent equation valid in the SI system [$\epsilon \to \epsilon/\epsilon_0$, $\sigma \to \sigma/(4\pi\epsilon_0)$]

$$\epsilon_{\mathrm{SI}}(\omega) = \epsilon_0 + \frac{i\sigma(\omega)}{\omega}; \quad \epsilon_0 = 8.85 \times 10^{-12}\,\mathrm{Fm}^{-1} = 8.85 \times 10^{-12}\,\frac{\mathrm{s}}{\Omega\mathrm{m}}. \tag{4.8}$$

In some cases, solids can be prepared in thin enough slices that absorption and phase shift can be directly measured in transmission experiments in the energy ranges of interest. Then, the index of diffraction $n(\omega)$ determines the phase velocity $c_{\mathrm{medium}}(\omega) = c/n(\omega)$, i.e., the phase shift in a sample of finite thickness, while $k(\omega)$ quantifies the absorption per wavelength [see, e.g., Hecht (1997) or Landau and Lifschitz (1990)]. The functions $\epsilon(\omega)$ and $\sigma(\omega)$ are obtained from such measurements via

$$\kappa(\omega) \equiv \sqrt{\epsilon(\omega)} = n(\omega) + ik(\omega)\,, \tag{4.9}$$

[3]This is not true at finite $\boldsymbol{q}$, even in an homogeneous system. Later, in Sec. 4.4, we will also need to consider the more general case where $\sigma(\omega)$ is still diagonal, but not isotropic.

[4]Unless stated otherwise, we will here and in the following use the Gaussian system of units (cgs).

where the square root is taken such that $n$ is positive, and via (4.7).

If transmission experiments are not possible, but at least good (i.e., smooth and clean) surfaces can be prepared, then optical properties can still be probed in reflection mode. The reflectivity for normally incident light is

$$r(\omega) = \left|\frac{1-\kappa(\omega)}{1+\kappa(\omega)}\right|^2 = \frac{\big(1-n(\omega)\big)^2 + k^2(\omega)}{\big(1+n(\omega)\big)^2 + k^2(\omega)}, \tag{4.10}$$

Since $\kappa(\omega)$ [like $\sigma(\omega)$ and $\epsilon(\omega)$] is analytic in the upper half plane,[5] its real part $n(\omega)$ and its imaginary part $k(\omega)$ are related by Kramers-Kronig transformations. Using information about $r(\omega)$ at all frequencies, one can extract $n$ and $k$ by Kramers-Kronig analysis which yields $\epsilon(\omega)$ and finally $\sigma(\omega)$. This procedure and the implicit extrapolations to frequencies outside the range of measurements can be avoided by measurements at variable angle $\Theta$ and/or polarization of the incident light. In this case, the reflectivity is given (for permeability $\mu = 0$, i.e., nonmagnetic materials) by (Stratton, 1941)

$$r_\perp(\Theta,\omega) = \frac{\big(q(\Theta,\omega)-\cos\Theta\big)^2 + p^2(\Theta,\omega)}{\big(q(\Theta,\omega)+\cos\Theta\big)^2 + p^2(\Theta,\omega)}, \tag{4.11}$$

$$r_\parallel(\Theta,\omega) = r_\perp(\Theta,\omega)\frac{\big(q(\Theta,\omega)-\sin\Theta\tan\Theta\big)^2 + p^2(\Theta,\omega)}{\big(q(\Theta,\omega)+\sin\Theta\tan\Theta\big)^2 + p^2(\Theta,\omega)}, \tag{4.12}$$

where all quantities are real and [$n \equiv n(\omega)$, $k \equiv k(\omega)$ as above]

$$p^2(\omega) = \frac{1}{2}\left[k^2 - n^2 + \sin^2\Theta + \sqrt{4n^2k^2 + (n^2-k^2+\sin^2\Theta)^2}\right], \tag{4.13}$$

$$q^2(\omega) = \frac{1}{2}\left[n^2 - k^2 - \sin^2\Theta + \sqrt{4n^2k^2 + (n^2-k^2+\sin^2\Theta)^2}\right]. \tag{4.14}$$

Since $p^2$ and $q^2$ reduce to $k^2$ and $n^2$, respectively, for normal incidence ($\Theta = 0$) it is easily checked that (4.11) and (4.12) indeed reduce to (4.10) in this limit.

Transmission experiments, if possible, are not only preferable over reflection experiments since the Kramers-Kronig analysis or mathematical complications are avoided, but also, since it is more certain that bulk properties are measured (as opposed to surface properties).

For illustration, the dielectric functions of a single-mode Lorentz oscillator model[6] are shown in Fig. 4.1. Here, we have chosen a nonzero oscillator frequency with a comparatively small damping term so that the model describes an insulator. As we

[5] At finite $\boldsymbol{q}$, this is rigorously true only for the transverse part $\epsilon_t(\boldsymbol{q},\omega)$; also analytic is the inverse of the longitudinal part, $\epsilon_l^{-1}(\boldsymbol{q},\omega)$ (v. Baltz, 1997). This distinction vanishes, however, for $\boldsymbol{q} = 0$.

[6] This model is based on damped harmonic oscillator modes with weight $f_j$, resonance frequency $\omega_j$, and damping term $\gamma_j$, leading to $\epsilon(\omega) = 1+4\pi e^2/m\sum_j f_j/((w_j^2-w^2)-2i\gamma_j w)$; see, e.g., Altarelli, Dexter, Nussenzveig, and Smith (1972). We use it here as a convenient fit formula producing an analytic function with a strongly peaked imaginary part. On a phenomenological level, more realistic curves are easily obtained by using several oscillator levels.

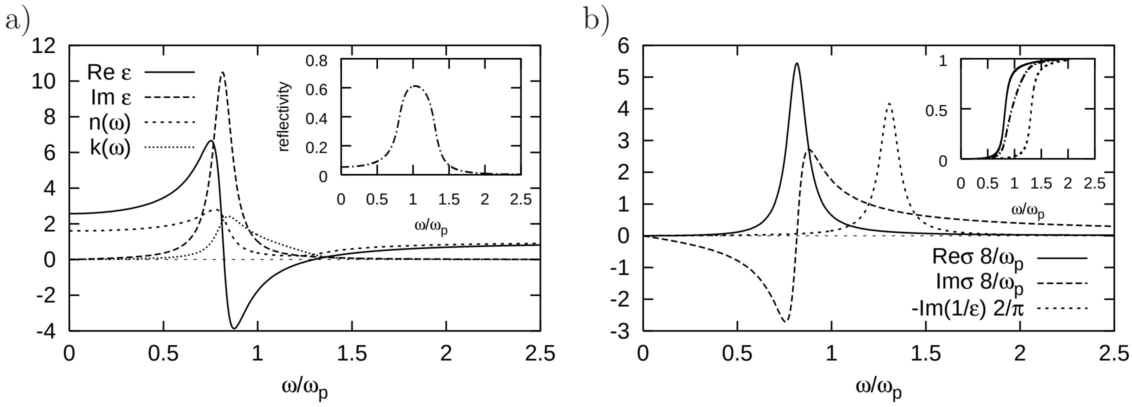



**Figure 4.1:** Scheme: dielectric functions in insulator (Lorentz oscillator model). **a)** dielectric function $\epsilon(\omega)$, refractive index $n(\omega)$, and absorption coefficient $k(\omega)$. Inset: reflectivity for normal incidence. **b)** Scaled optical conductivity $\sigma(\omega)$ and electronic energy loss function $\mathrm{Im}\,\frac{-1}{\epsilon(\omega)}$. Inset: integrals $\int \frac{\mathrm{d}\omega}{\omega_\mathrm{p}} \frac{8\sigma(\omega)}{\omega_\mathrm{p}}$ (solid line), $\int \frac{\mathrm{d}\omega}{\omega_\mathrm{p}} \frac{\omega}{\omega_\mathrm{p}} \frac{4}{\pi} k(\omega)$ (dash-dotted line), and $\int \frac{\mathrm{d}\omega}{\omega_\mathrm{p}} \frac{\omega}{\omega_\mathrm{p}} \frac{2}{\pi} \mathrm{Im}\,\frac{-1}{\epsilon(\omega)}$ (dotted line) related to the sum rules (4.18)-(4.20).

will discuss in the following subsection 4.1.2, $\epsilon(\omega) \stackrel{\omega\to\infty}{\longrightarrow} 1 - \omega_\mathrm{p}^2/\omega^2$ for some plasma frequency $\omega_\mathrm{p}$ which we used here to rescale $\omega$ and $\sigma(\omega)$ so that the resulting plots are dimensionless. Near the oscillator frequency $\omega_0 \approx 0.82\,\omega_\mathrm{p}$, the rapid changes in $\mathrm{Re}\,\epsilon(\omega)$ and $n(\omega)$ with negative slope show anomalous dispersion. Here, the usual expression $u = c/(1 + \mathrm{d}n/\mathrm{d}\omega)$ for the group velocity of light would imply $u > c$. It is not applicable in this frequency range, however, due to the strong absorption indicated by peaks in $\mathrm{Im}\,\epsilon(\omega)$, $\mathrm{Re}\,\sigma(\omega)$, and $k(\omega)$. The peak of the reflectivity is at slightly higher frequencies $\omega \approx \omega_\mathrm{p}$ (inset of Fig. 4.1 a). In our example, the absorption is so strong that $\mathrm{Re}\,\epsilon(\omega)$ changes sign at intermediate frequencies. When it vanishes outside the absorption region (here at $\omega_{p,\mathrm{bulk}} \approx 1.3\,\omega_\mathrm{p}$), the system is susceptible to a bulk plasma oscillation since (4.6) then allows for finite electric field without external field, $D \equiv 0$. This oscillator mode can be excited by injecting high-energy electrons which then lose energy in quanta of $\hbar\omega_{p,\mathrm{bulk}}$. More precisely, the distribution of energy losses for single scattering (transmission through thin slices) is approximately proportional to the electron energy-loss function $-\mathrm{Im}\,\{1/\epsilon(\omega)\}$ depicted in Fig. 4.1 b.

### 4.1.2 Optical $f$-sum Rules

The so-called $f$-sum rules are based on the Thomas-Reiche-Kuhn oscillator sum rule (Thomas, 1925; Kuhn, 1925; Reiche and Thomas, 1925) for atomic spectra which was later generalized to the momentum-dependent case (Bethe, 1930). On a phenomenological level, the sum rules follow from analytic properties of the dielectric functions and the very general assumption that absorptive processes only extend over a finite frequency range. Their common unknown parameter, i.e., the plasma frequency $\omega_\mathrm{p}$, can only be determined on a microscopic level.

The Kramers-Kronig equation for the real part of $\epsilon(\omega)$ reads

$$\begin{aligned}\mathrm{Re}\,\epsilon(\omega) &= 1+\frac{1}{\pi}P\int_{-\infty}^{\infty}\mathrm{d}\omega'\,\mathrm{Im}\,\epsilon(\omega')\frac{1}{\omega'-\omega} && (4.15)\\ &= 1+\frac{2}{\pi}\int_{0}^{\infty}\mathrm{d}\omega'\,\mathrm{Im}\,\epsilon(\omega')\frac{\omega'}{\omega'^2-\omega^2}, && (4.16)\end{aligned}$$

where in the second line we used the fact that $\epsilon(-\omega)=\epsilon^*(\omega)$ since $\boldsymbol{E}$ and $\boldsymbol{D}$ are real. Under the assumption that excitations (at $\boldsymbol{q}=\boldsymbol{0}$) are restricted to a finite energy range $0<\omega<\omega_{\max}$, we can take the $\omega\to\infty$ limit of (4.16),

$$\begin{aligned}\lim_{\omega\to\infty}\omega^2\big(1-\mathrm{Re}\,\epsilon(\omega)\big) &= \frac{2}{\pi}\lim_{\omega\to\infty}\int_{o}^{\omega_{\max}}\mathrm{d}\omega'\,\mathrm{Im}\,\epsilon(\omega')\,\omega'\left(1+\frac{\omega'^2}{\omega^2}-\dots\right)\\ &= \frac{2}{\pi}\int_{0}^{\infty}\mathrm{d}\omega'\,\mathrm{Im}\,\epsilon(\omega')\omega' \;=:\; \omega_{\mathrm{p}}^2, && (4.17)\end{aligned}$$

where $\omega_{\mathrm{p}}$ is the plasma frequency. Using (4.17) and (4.7) we arrive at the $f$-sum rule for the optical conductivity,

$$\int_{0}^{\infty}\mathrm{d}\omega\,\mathrm{Re}\,\sigma(\omega)=\frac{\omega_{\mathrm{p}}^2}{8}\,. \tag{4.18}$$

Other useful $f$-sum rules which follow from (4.17) are

$$\begin{aligned}\int_{0}^{\infty}\mathrm{d}\omega\,\omega k(\omega) &= \frac{\pi}{4}\omega_{\mathrm{p}}^2 && (4.19)\\ \int_{0}^{\infty}\mathrm{d}\omega\,\omega\,\mathrm{Im}\frac{-1}{\epsilon(\omega)} &= \frac{\pi}{2}\omega_{\mathrm{p}}^2\,. && (4.20)\end{aligned}$$

Note that the limit $\omega\to\infty$ used in the above equations is problematic for several (related) reasons: most importantly, photons of high frequency have a short wavelength so that any theory which depends on the long-wavelength limit $\boldsymbol{q}\to 0$ breaks down. While the range of validity could be enlarged by considering finite $q$, ultimately the homogeneity condition is violated when $\omega\gtrsim\frac{c}{a}$, for lattice spacing $a$. In fact, the response at high frequencies, i.e., (inelastic) $x$-ray diffraction, must show strong angular dependence due to the lattice structure factor. Furthermore, the excitation of core electrons will lead to relativistic effects for systems containing heavy nuclei.

Theoretically, these problems may be avoided by considering the jellium model of interacting electrons in a neutralizing homogeneous background. In this case, the value of the plasma frequency only depends on the total electron density $n$ (for an electron charge $e$ and mass $m$),

$$\omega_{\mathrm{p}}^2 = \frac{4\pi n e^2}{m} \,, \tag{4.21}$$

and additional exact sum rules may be derived (Nozières and Pines, 1958; Nozières and Pines, 1959; Nozières and Pines, 1989). It is an interesting question to what extent the formalism and this simple result carry over to real systems. The full sum rules remain valid for real solids (Nozières and Pines, 1989); for a theoretical analysis, see subsection 4.2.1. At least for light metals like lithium or beryllium (where relativistic effects are negligible and the homogeneity condition can be met) the integrals (4.18)-(4.20) should, thus, be observable with $\omega_{\mathrm{p}}$ given by (4.21). Experimental data at soft $X$-ray frequencies $\hbar\omega \gtrsim 100$ eV is, however, scarce. Only after an extensive literature search we could find an accurate experimental verification of the full sum rule for aluminum (see next subsection).

Alternatively, one may consider partial sum rules, i.e., limit the integrals to some maximum frequency. This approach is most promising when the valence electrons are well separated energetically from core electron states. Then, the high-energy part of the valence electron contribution to dielectric functions is still well defined and corresponds to a valence electron plasma frequency

$$\omega_{p,\mathrm{valence}}^2 = \frac{4\pi n_v e^2}{m_c} \,. \tag{4.22}$$

Here, $n_v$ is the valence electron density and $m_c$ is a mass which characterizes the electron mobility and is renormalized by polarization effects. The polarizability of ion cores also causes an almost frequency-independent shift of the dielectric function so that the intermediate asymptotic form of the dielectric function reads $\epsilon(\omega) \approx \epsilon_{\mathrm{core}} + \omega_{p,\mathrm{valence}}^2/\omega^2$.

### 4.1.3 Experiments

In this subsection, we show experimental results for the dielectric functions and the optical conductivity of the good metal aluminum. This material was chosen as an example due to the completeness of the available data which allows for the illustration of partial sum rules and for checking the universality of the full $f$-sum rule. Finally, we will see that photon experiments (reflection or transmission measurements) are superior over electronic energy-loss spectroscopy (EELS) for determining $\sigma(\omega)$.

Figure 4.2 a shows reflectance data for Al in the valence electron energy range up to $\hbar\omega = 22$ eV. The reflectivity is asymptotically nearly perfect at very small frequencies (see Fig. 4.2 b), then remains large up to the far ultra violet (UV) range with a dip at about $\hbar\omega \approx 1.5$ eV and decreases quite abruptly at $\hbar\omega \approx 15$ eV. By Kramers-Kronig analysis of this data, Ehrenreich et al. (1963) computed $\epsilon(\omega)$ shown in Fig. 4.2 c and $\sigma(\omega)$ as well as the loss function shown in Fig. 4.2 d. Here, the metallic behavior is visible as a Drude peak in $\mathrm{Re}\,\sigma(\omega)$ (not fully shown) and a divergence of $\mathrm{Im}\,\epsilon(\omega)$. The dip in the reflectance finds correspondence in a two-peak structure of

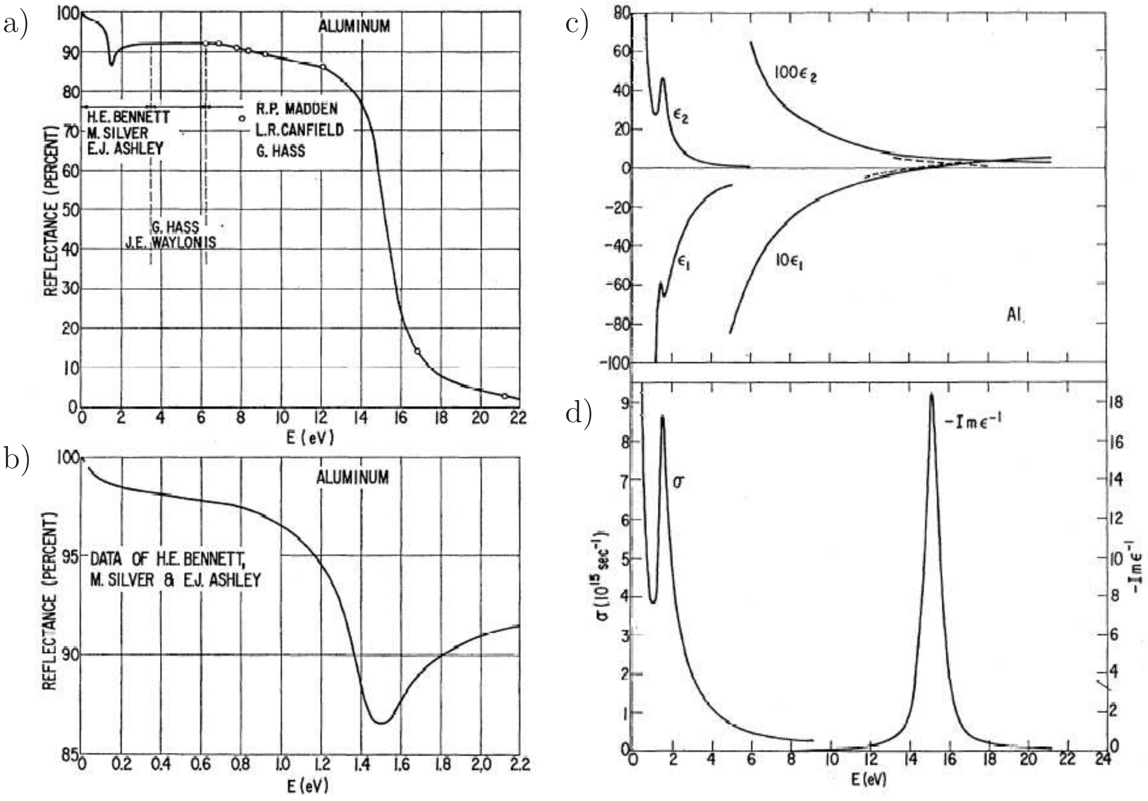



**Figure 4.2:** Low-energy optical properties of aluminum: **a)+b)** measured specular reflectivity; **c)** corresponding dielectric function $\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$. **d)** Optical conductivity $\sigma(\omega)$ and loss function $-\mathrm{Im}\,\epsilon^{-1}(\omega)$ (Ehrenreich et al., 1963).

$\sigma(\omega)$. The electron energy-loss function $-\mathrm{Im}\,\epsilon^{-1}(\omega)$ is peaked at $\hbar\omega \approx 15$ eV which is close to the free valence-electron plasma frequency $\hbar\omega^0_{p,\mathrm{valence}} \approx 15.8$ eV. Hence, (4.22) applies for three valence electrons per atom and an only weakly renormalized mass $m_c \approx 1.05m$.

As an alternative to defining renormalized electron masses for valence electrons as in (4.22) one may use the bare mass and introduce effective electron densities. This is even more physical since the total weight is always proportional to the full electron density; consequently, the polarizability of core electrons only shifts the contribution of valence electrons to higher energies without modifying their total integral. Rescaling the integrands in (4.18)-(4.20) so that the total sum equals the number of electrons per atom, the weight up to some finite frequency may then be regarded as the effective number of electrons for that frequency range. As seen in Fig. 4.3 a, the partial sums deviate strongly from each other in and near the low frequency absorption range $\hbar\omega \lesssim 15$ eV. Since Al is nearly free-electron like, the effective number corresponding to (4.18) (indicated as "$\epsilon_2$") saturates very close to 3, the true valence electron number, while those corresponding to (4.19) and (4.20) are reduced by a factor of $1/\sqrt{\epsilon_{\mathrm{core}}}$ and $1/\epsilon_{\mathrm{core}}$, respectively. At $\hbar\omega \approx 1$ keV, the effective numbers reach 11 while above 10 keV the sum rules correctly measure a total of 13

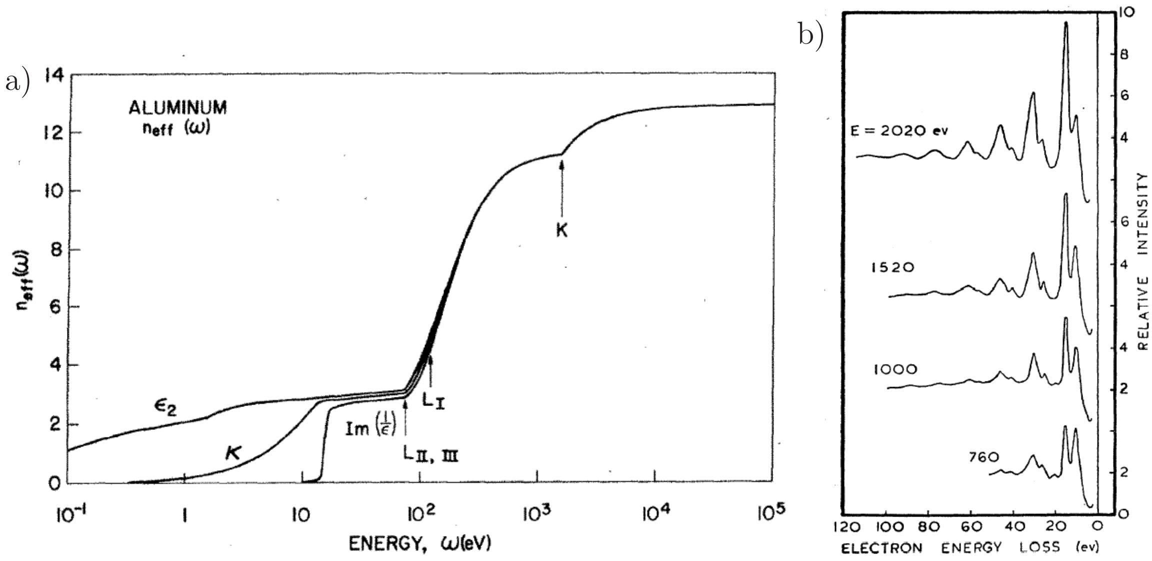



**Figure 4.3:** Optical properties of aluminum: **a)** Effective number of valence electrons defined via partial sum rules (Smith and Shiles, 1978). **b)** Experimental electron energy loss spectrum (reflection mode) with different incident electron energies $E$ (Powell and Swan, 1959).

electrons per atom. A corresponding plot is also shown in the inset of Fig. 4.1b for the Lorentz oscillator model.

Figure 4.3b shows EELSdata obtained in reflection mode (Powell and Swan, 1959). One can clearly distinguish a peak at $\hbar\omega \approx 15$ eV corresponding to the peak of the electronic energy loss function derived from reflectance data shown earlier in Fig. 4.2d. Additional peaks at multiples of this energy are associated with multiple energy losses. Additionally, a peak at $\hbar\omega \approx 10$ eV is seen which signals the excitation of a surface plasmon with $\omega_{p,\text{surface}} \approx \omega_{p,\text{bulk}}/\sqrt{2}$. Evidently, the relative weights of these peaks change upon changing the incident electron energy: with decreasing beam energy, the relative probability for multiple scattering decreases while the importance of surface effect increases. In general, it is difficult to extract even the shape of the linear-response electronic energy loss function from EELS; an unknown prefactor then has to be estimated from the sum rule (4.20). More reliable results may be obtained in transmission EELS experiments, where the contribution from multiple scattering vanishes with decreasing sample thickness and also momentum-dependent dielectric functions can be measured (Batson and Silcox, 1983).

### 4.1.4 Impact of Electronic Model Abstractions

In the previous subsections we have seen that the dielectric functions have important universal properties and relations. Even the full plasma frequency $\omega_\text{p}$ was found to be universal, i.e., only dependent on the total electron density, at least in the non-relativistic limit (this point will be further investigated in subsection 4.2.1). Model

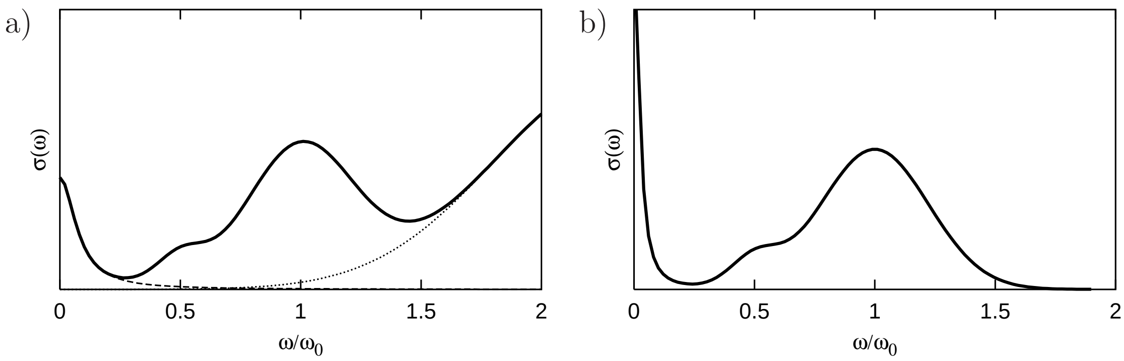



**Figure 4.4:** Scheme for $\sigma(\omega)$: Effects of disorder/phonons as well as interband excitations which would be seen in experiment (**a)**, solid line) have to be identified (dashed and dotted lines) and subtracted when comparing to **b)** results of a "clean" valence electron calculation.

abstractions and approximations can violate the $f$-sum rules; then, the plasma frequency may become dependent on parameters of the model or even on temperature. Before we treat transport properties and the conductivity sum rule on a microscopic level in the next sections, we want to discuss the qualitative impact of the neglect of lattice degrees of freedom and of a reduction to a few valence bands. Furthermore, we sketch the generic overall behavior of $\sigma(\omega)$ for the Hubbard model.

In Fig. 4.4 a, we show schematically the optical conductivity of a typical strongly correlated metal (solid line). One can distinguish a relatively broad Drude peak near $\omega = 0$ and an incoherent peak at finite frequencies $\omega \approx \omega_0$ with a shoulder (here at $\omega \approx \omega_0/2$). In a single-band picture, $\hbar\omega_0$ roughly corresponds to the local Coulomb interaction $U$. Then, the contributions at $\omega \gg \omega_0$ have to be interpreted as interband contributions. Therefore, they have to be quantified (dotted line) and subtracted in order to define an effective valence electron contribution as shown in Fig. 4.4 b. Often, the width of the Drude peak (dashed line in Fig. 4.4 b) is dominated by lattice degrees of freedom which are not taken into account in purely electronic models. Therefore, this contribution has been replaced by a much narrower Drude peak with the same weight in Fig. 4.4 b. Clearly, this mapping is somewhat ad-hoc and not generally reliable.[7] Still, the example given here illustrates that results obtained for an electronic valence-electron model will not in general yield meaningful estimates for the plasma frequency (not even in the sense of partial sum rules) and that even the shape of the low-frequency part of $\sigma(\omega)$ may be unrealistic.

In the remainder of this section, we will classify the qualitative behavior of $\sigma(\omega)$ for electronic models which are defined in the continuum or on a Bravais lattice. Then, electronic states may be labeled by their (crystal) momentum $\boldsymbol{k}$, band index $\nu$, and spin $\sigma$. The associated paramagnetic current density operator $\hat{\boldsymbol{\jmath}}$ and Fermi

[7] Our example is very similar to an analysis of experimental data for $CaVO_3$ by Rozenberg, Inoue, Makino, Iga, and Nishihara (1996).

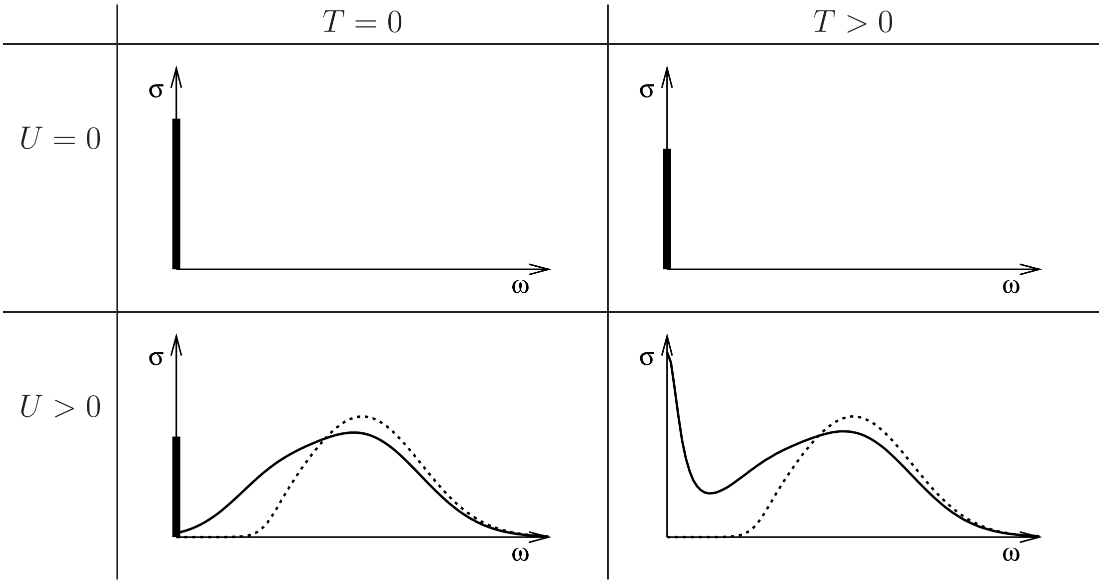



**Figure 4.5:** Scheme: qualitative behavior of the optical conductivity in Hubbard type models. For $U > 0$, one may distinguish metals (solid lines) and insulators (dashed lines). The thick vertical bars at $\omega = 0$ indicate $\delta$-functions.

velocity $\boldsymbol{v}_{\boldsymbol{k},\nu}$ read[8] for a noninteracting dispersion $\epsilon_{\boldsymbol{k},\nu}$

$$\hat{\boldsymbol{\jmath}} = \sum_{\boldsymbol{k},\nu,\sigma} e\, \boldsymbol{v}_{\boldsymbol{k},\nu}\, \hat{n}_{\boldsymbol{k},\nu,\sigma}; \quad \boldsymbol{v}_{\boldsymbol{k},\nu} = \frac{1}{\hbar} \boldsymbol{\nabla}_{\boldsymbol{k}}\, \epsilon_{\boldsymbol{k},\nu}\,. \tag{4.23}$$

The important observation which we are heading for is that due to the absence of dissipation no finite-frequency response arises in the noninteracting limit: In absence of electron-electron interaction, $\hat{H} = \hat{H}_0 = \sum_{\boldsymbol{k},\nu,\sigma} \epsilon_{\boldsymbol{k},\nu} \hat{n}_{\boldsymbol{k},\nu,\sigma}$. We thus have

$$[\hat{\boldsymbol{\jmath}}, \hat{H}] = e \sum_{\boldsymbol{k},\nu,\sigma} \boldsymbol{v}_{\boldsymbol{k},\nu} [\hat{n}_{\boldsymbol{k},\nu,\sigma}, \hat{H}_0] = 0 \tag{4.24}$$

$$\Rightarrow \quad \hat{\boldsymbol{\jmath}}(t) = \mathrm{e}^{i\hbar\hat{H}t}\, \hat{\boldsymbol{\jmath}}(0)\, \mathrm{e}^{-i\hbar\hat{H}t} = \hat{\boldsymbol{\jmath}}(0) \tag{4.25}$$

$$\Rightarrow \quad \sigma(\omega) = D\delta(\omega)\,, \tag{4.26}$$

where $D > 0$ is the Drude weight. The diamagnetic current could be neglected for this argument, since its contribution at finite frequencies is purely imaginary. Equation (4.26) remains valid for static mean-field theories of the density type, i.e., the Hartree approximation as long as translational invariance is retained. Thus, even in this case, $\sigma(\omega)$ is reduced to a Drude peak (of zero width). This general behavior is illustrated in Fig. 4.5, where we also distinguish between zero and finite temperature. Insulating behavior (dashed lines) is possible only for $U > 0$. Note that the total weight is in general temperature-dependent. Later we will see that some models corresponding to a semi-elliptic Bethe DOS do not yield a Drude peak in the noninteracting limit since their current operator and their kinetic energy operator do not commute.

[8] A more thorough introduction of the formalism follows in Sec. 4.2.

## 4.2 Kubo Formalism

In this section, we will discuss the formalism leading to expressions for $\sigma(\omega)$ both in the continuum and on the lattice in enough detail to expose the microscopic origin and ranges of validity of sum rules.

### 4.2.1 Kubo Formalism in the Continuum

As reviewed in App. D.1, the interaction of a particle at position $\boldsymbol{r}$ with charge $q$ with an external time varying electric field may be fully described in terms of a shift of its momentum by $q\boldsymbol{A}(\boldsymbol{r})/c$, where $\boldsymbol{A}(\boldsymbol{r})$ is the vector potential. Furthermore, within the Coulomb gauge, the electrical interaction between charged particles is instantaneous. In the nonrelativistic limit, we can, thus, assume the following first-quantized electronic Hamiltonian for a number of $N_e$ electrons,[9]

$$H = H_{\text{int}}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_{N_e}) + \sum_{i=1}^{N_e} \frac{1}{2m}\left(\boldsymbol{p}_i - \frac{e}{c}\boldsymbol{A}(\boldsymbol{r}_i, t)\right)^2 . \tag{4.27}$$

Here, the interaction Hamiltonian $H_{\text{int}}$ may contain arbitrary interactions of electrons with a (static) lattice potential as well as electron-electron interactions. This term will enter the results only implicitly via its influence on the eigenstates of the unperturbed system and their occupation in the thermal ensemble. From (4.27) and $\boldsymbol{v}_i = \nabla_{\boldsymbol{p}_i} H$ we can read off the total current,

$$\int \mathrm{d}\boldsymbol{r}\, \boldsymbol{J}(\boldsymbol{r}, t) = \sum_{i=1}^{N_e} e\, \boldsymbol{v}_i = \sum_{i=1}^{N_e} \frac{e}{m}\left(\boldsymbol{p}_i - \frac{e}{c}\boldsymbol{A}(\boldsymbol{r}_i)\right) . \tag{4.28}$$

In (4.28), we may distinguish a diamagnetic contribution $-\sum_{i=1}^{N_e} \frac{e^2}{mc}\boldsymbol{A}(\boldsymbol{r}_i)$ from the paramagnetic current, the density of which reads in symmetrized form,[10]

$$\boldsymbol{j}(\boldsymbol{r}, t) = \frac{e}{2m}\sum_{i=1}^{N_e}\big(\boldsymbol{p}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) + \delta(\boldsymbol{r} - \boldsymbol{r}_i)\boldsymbol{p}_i\big) . \tag{4.29}$$

Denoting the thermal expectation value[11] of some observable $\mathcal{O}$ in the presence of the vector potential by $\langle \mathcal{O} \rangle_A$, the current density is, thus, exactly given by

$$\langle \boldsymbol{J}(\boldsymbol{r}, t)\rangle_A = \langle \boldsymbol{j}(\boldsymbol{r}, t)\rangle_A - \frac{e^2}{mc}\boldsymbol{A}(\boldsymbol{r}, t)\langle n(\boldsymbol{r}, t)\rangle_A . \tag{4.30}$$

Since we only need the linear response current, i.e., contributions to $\boldsymbol{J}$ which are at most linear in $\boldsymbol{A}$, we can take the density expectation value in (4.30) in zeroth order. In the long-wavelength limit, the diamagnetic contribution is then $-\frac{e^2 n_0}{mc}\boldsymbol{A}(\boldsymbol{r}, t)$,

[9] Note that following the usual convention the electron charge is here denoted as $e$, not $-e$.

[10] Here and in the following, we use the upper-case symbol $\boldsymbol{J}$ for the total current and the lower-case symbol $\boldsymbol{j}$ for the paramagnetic current or its density.

[11] Due to the explicit time dependence of $H$, we here have to use a generalized definition of the thermal average in terms of a time-dependent density operator $\hat{\rho}(t)$; see App. D.2.

where $n_0$ is the average density. The remaining problem is the calculation of $\langle \boldsymbol{j}(\boldsymbol{r},t)\rangle_A$ in linear order in $\boldsymbol{A}$. For this purpose, it suffices to consider the linear perturbing part of the Hamiltonian[12] which may be expressed as

$$H'(t) = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\, \boldsymbol{j}(\boldsymbol{r},t)\cdot \boldsymbol{A}(\boldsymbol{r},t)\,. \tag{4.31}$$

From (D.2) we have for periodic field $\boldsymbol{E}(\boldsymbol{r},t) = \boldsymbol{E}_{\boldsymbol{q}}\mathrm{e}^{i(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}$

$$\boldsymbol{A}(\boldsymbol{r},t) = -\frac{ic}{\omega}\boldsymbol{E}(\boldsymbol{r},t)\,. \tag{4.32}$$

Thus, we can write, switching to second quantized notation (with $V$ denoting the total volume),

$$\hat{H}'(t) = -\frac{V}{c}\,\hat{\boldsymbol{\jmath}}_{-\boldsymbol{q}}(t)\cdot \boldsymbol{A}_{\boldsymbol{q}}(t) = \frac{iV}{\omega}\,\hat{\boldsymbol{\jmath}}_{-\boldsymbol{q}}(t)\cdot \boldsymbol{E}_{\boldsymbol{q}}\mathrm{e}^{-i\omega t}\,. \tag{4.33}$$

Using (D.21) and assuming that the response is restricted to the same wave vector as the perturbing field, i.e., for not too short wavelengths, the paramagnetic linear response current reads (denoting directions as $\alpha$ or $\beta$),

$$\langle \hat{\jmath}_{\boldsymbol{q},\alpha}(\omega)\rangle_A = -\frac{iV}{\omega+i0^+}E_{\boldsymbol{q},\beta}(\omega)\,\langle\langle \hat{\jmath}_{\boldsymbol{q},\alpha}, \hat{\jmath}_{-\boldsymbol{q},\beta}\rangle\rangle\,(\omega+i0^+) \tag{4.34}$$

From (4.4), (4.30), and (D.22) we get the general expression for continuum systems[13]

$$\sigma_{\alpha\beta}(\boldsymbol{q},\omega) = \frac{V}{\hbar(\omega+i0^+)}\int_0^\infty \mathrm{d}t\,\mathrm{e}^{i(\omega+i0^+)t}\langle[\hat{\jmath}^\dagger_{\boldsymbol{q},\alpha}(t),\hat{\jmath}_{\boldsymbol{q},\beta}(0)]\rangle + i\frac{n_0e^2}{m(\omega+i0^+)}\delta_{\alpha\beta} \tag{4.35}$$

Here, the time dependence of the operator is in the interaction picture (see Sec. D.2) and the long wavelength limit has been used to replace $n(\boldsymbol{r},t)$ by the average density $n_0$. In the optical limit $\boldsymbol{q}\to 0$, the current density operator $\hat{\boldsymbol{\jmath}} = \hat{\boldsymbol{\jmath}}^\dagger$ is hermitian. For systems with cubic symmetry, the current flow in perpendicular directions is then uncorrelated and $\sigma_{\alpha\beta}(\boldsymbol{0},\omega)$ reduces to the scalar $\sigma(\omega)$. Equation (4.35) is correct both for $T=0$ (then the outer brackets indicate the ground state expectation value) and for $T>0$ (when the brackets imply thermal averaging) and for longitudinal as well as transverse electric fields.

Since the susceptibility, i.e., the integral in (4.35) falls off as $1/\omega$ or faster [cf. (D.25)], the high-frequency limit of $\sigma(\omega)$ is determined by the second, diamagnetic term which is proportional to the squared free plasma frequency $\omega_\mathrm{p}^2 = 4\pi n_0 e^2/m$. Thus, the frequency sum (4.19) is unchanged by the interactions of electrons, both with the lattice ions and among themselves,

$$\int_0^\infty \mathrm{d}\omega\,\mathrm{Re}\,\sigma_{\alpha\beta}(\omega) = \frac{1}{2}\int_{-\infty}^\infty \mathrm{d}\omega\frac{n_0e^2}{m}\pi\delta(\omega)\,\delta_{\alpha\beta} = \frac{\omega_\mathrm{p}^2}{8}\,\delta_{\alpha\beta} \tag{4.36}$$

[12] Note that the quadratic term in $H$ had to be considered for deriving the correct expression for the full current.

[13] Note that (4.35) is valid in Gaussian units. For SI units, its right hand side has to be divided by $4\pi\epsilon_0$.

In the noninteracting limit, where correlations vanish, $\mathrm{Re}\,\sigma(\omega)$ reduces to the diamagnetic $\delta$-function contributions. Interactions shift weight from $\omega = 0$ to finite frequencies without changing the total weight. Thus, the full universal $f$-sum rule is indeed found to be valid, at least within the (zeroth order) Born-Oppenheimer approximation and in the limit that optical absorption decays below a cutoff energy $E$ with $\frac{E}{\hbar c} \ll \frac{\pi}{a}$.

### 4.2.2 Kubo Formalism on a Lattice

In Wannier representation, the vector potential can be treated (Kohn, 1964; Shastry and Sutherland, 1990; Scalapino, White, and Zhang, 1992; Dagotto, 1994) by a gauge transformation of the Wannier wave functions $\Phi_l(\boldsymbol{r})$ centered at site $\boldsymbol{R}_l$,

$$\Phi_l(\boldsymbol{r}) \mapsto \tilde{\Phi}_l(\boldsymbol{r}, t) = \left[\exp\left(i\frac{e}{c\hbar}\int\limits_{\boldsymbol{R}_0}^{\boldsymbol{R}_l} \mathrm{d}\boldsymbol{r}\cdot\boldsymbol{A}(\boldsymbol{r}', t)\right)\right]\Phi_i(\boldsymbol{r}), \tag{4.37}$$

where the free choice of $\boldsymbol{R}_0$ determines the arbitrary overall phase factor. In the long-wavelength limit $\boldsymbol{q} \to 0$, the vector potential can be taken out of the integral. Then, the change in the Hamiltonian consists of multiplying the hopping elements defined in (1.10) by so-called Peierls phase factors,[14]

$$t_{ij} = t_{ij}^0 \exp\left(-i\frac{e}{c\hbar}(\boldsymbol{R}_i - \boldsymbol{R}_j)\cdot\boldsymbol{A}\right). \tag{4.38}$$

Note that only hopping elements with components parallel to the vector field $\boldsymbol{A}$ are modified. In comparing (4.38) with (D.3) we identify the additional factor as $\exp(-i\Delta S/\hbar)$, where $\Delta S$ is the field contribution to the action for the world line connecting the lattice sites in absence of a scalar potential $\phi$.

Before considering the general case, we specialize on the more commonly treated case of the hypercubic tight-binding model with nearest-neighbor hopping, where the unperturbed kinetic energy is given by $\hat{K}^0 = -t\sum_{\langle ij\rangle,\sigma}\left(\hat{c}^\dagger_{i\sigma}\hat{c}_{j\sigma} + \hat{c}^\dagger_{j\sigma}\hat{c}_{i\sigma}\right)$. Expanding the phase factors to lowest orders in $\boldsymbol{A}$,[15] we obtain the modified kinetic energy:

$$\hat{K} = \hat{K}^0 + \sum_{\alpha=1}^{d}\big(-\frac{V}{c}\hat{\jmath}_\alpha A_\alpha + \frac{e^2a^2}{2\hbar^2c^2}\hat{K}^0_\alpha A_\alpha^2 + \ldots\big) \tag{4.39}$$

where the lattice paramagnetic current density operator and the kinetic energy in

[14] While, in general, the interaction matrix elements defined in (1.11) also acquire Peierls phases, on-site elements and intersite contributions of the density-density type remain unchanged.

[15] We need quadratic order in order to retain the linear contribution to the current operator, $\hat{J}_\alpha$.

direction $\alpha$ are given by (lattice constant $a$)[16]

$$\hat{j}_\alpha = -\frac{ieta}{V\hbar}\sum_{i,\sigma}\left(\hat{c}^\dagger_{i\sigma}\hat{c}_{i+\boldsymbol{e}_\alpha\sigma} - \hat{c}^\dagger_{i+\boldsymbol{e}_\alpha\sigma}\hat{c}_{i\sigma}\right), \tag{4.40}$$

$$\hat{K}^0_\alpha = -t\sum_{i,\sigma}\left(\hat{c}^\dagger_{i\sigma}\hat{c}_{i+\boldsymbol{e}_\alpha\sigma} + \hat{c}^\dagger_{i+\boldsymbol{e}_\alpha\sigma}\hat{c}_{i\sigma}\right). \tag{4.41}$$

The full current density operator is obtained by variation of $\hat{K}$ with respect to the vector potential (Scalapino et al., 1992),

$$\hat{J}_\alpha = -\frac{c}{V}\frac{\delta\hat{K}}{\delta A_\alpha} = \hat{j}_\alpha + \frac{e^2a^2}{Vc\hbar^2}\hat{K}^0_\alpha A_\alpha. \tag{4.42}$$

As in the continuum case, the conductivity can be expressed in terms of a correlation function and of the diamagnetic contribution,

$$\sigma_{\alpha\alpha}(\omega) = \frac{iV}{\omega}\langle\langle\hat{j}_\alpha,\hat{j}_\alpha\rangle\rangle\,(\omega+i0^+) - i\frac{e^2a^2}{V\hbar^2(\omega+i0^+)}\langle\hat{K}^0_\alpha\rangle \tag{4.43}$$

Evidently, the reduction to a single-band model has changed the diamagnetic contribution and, consequently, the sum-rule to a nonuniversal form. For the assumption used in deriving (4.43) of nearest-neighbor hopping on a hypercubic lattice, the frequency sum is seen to be proportional to the kinetic energy,

$$\int_0^\infty \mathrm{d}\omega\,\sigma_{\alpha\alpha}(\omega) = -\frac{\pi}{2}\frac{e^2a^2}{V\hbar^2}\langle\hat{K}^0_\alpha\rangle = -\frac{\pi e^2}{2d\hbar^2}\,a^{2-d}\,\frac{\langle\hat{K}^0\rangle}{N}\,. \tag{4.44}$$

For the more general case of arbitrary hopping on a Bravais lattice, we have to consider that the phase shift due to the vector potential $\boldsymbol{A}$ increases with hopping distance. For later use, we want to obtain expressions in momentum representation. Thus, we assume for the unperturbed Hamiltonian (with band index $\nu$),

$$\hat{H} = \sum_{\boldsymbol{k},\nu,\sigma}\epsilon_{\boldsymbol{k},\nu}\,\hat{n}_{\boldsymbol{k},\nu,\sigma} + \hat{H}_{\mathrm{ee}}\{\hat{n}_{i,\nu,\sigma}\}\,, \tag{4.45}$$

where the interaction term is of the density-density type.[17] Due to inversion symmetry we can always write for a selected lattice direction $x$,

$$\epsilon_{\boldsymbol{k},\nu} = \sum_{n=1}^\infty f_{\nu,n}(\boldsymbol{k}_\perp)\cos(na_xk_x) \tag{4.46}$$

[16]Note that for this particular case and up to prefactors, the kinetic energy and the current operator are sums over the real and the imaginary part, respectively, of the same bond hopping operator.

[17]Note that, in contrast to the continuum formulation, the restriction to density-density type interaction terms is in general a serious approximation in the Wannier representation; see subsection 1.1.3.

where $\boldsymbol{k}_\perp$ is the projection of $\boldsymbol{k}$ on directions perpendicular to $\boldsymbol{e}_x$. Since the $n^{\text{th}}$ term in (4.46) contains hopping to $n^{\text{th}}$-nearest neighbors, we can use straightforward generalizations of (4.39) and (4.42) to find for the paramagnetic current density operator $\hat{\jmath}_x$ and diamagnetic current density operator $\hat{\jmath}_x^{\text{dia}}$, respectively,

$$\begin{aligned} \hat{\jmath}_x &= -\frac{ea_x}{V\hbar}\sum_{\boldsymbol{k},\nu,\sigma}\hat{n}_{\boldsymbol{k},\nu,\sigma}\sum_{n=1}^{\infty} n f_{\nu,n}(\boldsymbol{k}_\perp)\sin(nk_x) = \frac{e}{V\hbar}\sum_{\boldsymbol{k},\nu,\sigma}\hat{n}_{\boldsymbol{k},\nu,\sigma}\frac{\partial}{\partial k_x}\epsilon_{\boldsymbol{k},\nu} &(4.47)\\ \hat{\jmath}_x^{\text{dia}} &= \frac{e^2A}{Vc\hbar^2}\sum_{\boldsymbol{k},\nu,\sigma}\hat{n}_{\boldsymbol{k},\nu,\sigma}\frac{\partial^2}{\partial k_x^2}\epsilon_{\boldsymbol{k},\nu} &(4.48)\end{aligned}$$

We, thus, obtain for the conductivity,

$$\boxed{\begin{aligned}\sigma_{\alpha\alpha}(\omega) &= \frac{ie^2}{V\hbar^2(\omega+i0^+)}\sum_{\nu,\nu',\sigma,\sigma'}\sum_{\boldsymbol{k},\boldsymbol{k}'} v_{\boldsymbol{k},\nu,\alpha}v_{\boldsymbol{k}',\nu',\alpha}\langle\langle\hat{n}_{\boldsymbol{k},\nu,\sigma},\hat{n}_{\boldsymbol{k}',\nu',\sigma'}\rangle\rangle(\omega+i0^+)\\ &\quad+\frac{ie^2}{V\hbar^2(\omega+i0^+)}\sum_{\boldsymbol{k},\nu,\sigma}\langle\hat{n}_{\boldsymbol{k},\nu,\sigma}\rangle\frac{\partial^2}{\partial k_\alpha^2}\epsilon_{\boldsymbol{k},\nu}\,,\end{aligned}} \qquad (4.49)$$

where $v_{\boldsymbol{k},\nu,\alpha} = \frac{\partial}{\partial k_\alpha}\epsilon_{\boldsymbol{k},\nu}$ is the $\alpha$ component of the Fermi velocity of band $\nu$ for momentum $\boldsymbol{k}$. Consequently, the $f$-sum rule reads

$$\boxed{\int_0^\infty \mathrm{d}\omega\,\sigma_{\alpha\alpha}(\omega) = \frac{\pi e^2}{2V\hbar^2}\sum_{\boldsymbol{k},\nu,\sigma}\langle\hat{n}_{\boldsymbol{k},\nu,\sigma}\rangle\frac{\partial^2}{\partial k_\alpha^2}\epsilon_{\boldsymbol{k},\nu}\,.} \qquad (4.50)$$

For a full electronic model, we can use the free dispersion for a single band with $\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k^2} = \frac{\hbar^2}{m}$ which leads to the free diamagnetic term and, thus, to the free sum rule (4.21), i.e., $\int\mathrm{d}\omega\,\sigma_{\alpha\alpha}(\omega) = \frac{\omega_{\text{p}}^2}{8}$ with $\omega_{\text{p}}^2 = \frac{4\pi ne^2}{m}$.

### 4.2.3 General Confirmation of the $f$-sum Rule

While the fact that the paramagnetic current will not contribute to the $f$-sum rule follows on quite general grounds from (D.25), we can drop the condition of a bounded spectrum used in Sec. D.2. In the course of re-deriving the sum rule (4.50), we will also identify a contribution to the paramagnetic current which cancels the diamagnetic current. In this subsection, only the single-band case is considered explicitly.

For the Hamiltonian (4.45), the polarization operator $\hat{\boldsymbol{P}} = \sum_i \boldsymbol{r}_i\hat{n}_i$ commutes with the interaction term

$$[\hat{H}_{\text{ee}},\hat{\boldsymbol{P}}] = \sum_i \boldsymbol{r}_i[\hat{H}_{\text{ee}},\hat{n}_i] = 0\,. \qquad (4.51)$$

and can be used for expressing the current density operators:

$$\begin{aligned}\hat{\jmath}_\alpha &= \frac{e}{V}\frac{\mathrm{d}}{\mathrm{d}t}\hat{P}_\alpha = -\frac{ie}{V\hbar}[\hat{P}_\alpha,\hat{K}] &(4.52)\\ \hat{\jmath}_\alpha^{\text{dia}} &= \frac{e^2A}{Vc\hbar^2}[\hat{P}_\alpha,[\hat{P}_\alpha,\hat{K}]]\,. &(4.53)\end{aligned}$$

Thus, we can rewrite the current-current correlation function,

$$\langle\langle \hat{\jmath}_\alpha, \hat{\jmath}_\alpha \rangle\rangle(\omega) = \frac{ie}{V\hbar}\int_0^\infty \mathrm{d}t\, \mathrm{e}^{i\omega t} \langle [\frac{\mathrm{d}}{\mathrm{d}t}\hat{P}_\alpha(t), \hat{\jmath}_\alpha(0)]\rangle \tag{4.54}$$

$$= \frac{ie}{V\hbar}\left( \mathrm{e}^{i\omega t}\langle [\hat{P}_\alpha(t), \hat{\jmath}_\alpha(0)]\rangle\big|_0^\infty - \int_0^\infty \mathrm{d}t\,(i\omega)\mathrm{e}^{i\omega t}\langle [\hat{P}_\alpha(t), \hat{\jmath}_\alpha(0)]\rangle \right) \tag{4.55}$$

$$= -\frac{ie}{V\hbar}\langle [\hat{P}_\alpha, \hat{\jmath}_\alpha]\rangle + \frac{e\,\omega}{V\hbar}\int_0^\infty \mathrm{d}t\,\mathrm{e}^{i\omega t}\langle [\hat{P}_\alpha(t), \hat{\jmath}_\alpha]\rangle \tag{4.56}$$

The operator in the first term, $-\frac{ie}{V\hbar}[\hat{P}_\alpha, \hat{\jmath}_\alpha] = -\frac{e^2}{V^2\hbar^2}[\hat{P}_\alpha, [\hat{P}_\alpha, \hat{K}]]$, cancels the diamagnetic contribution to $\sigma_{\alpha\alpha}(\omega)$. Thus, the $f$-sum reads,[18]

$$\int_0^\infty \mathrm{d}\omega\, \mathrm{Re}\,\sigma_{\alpha\alpha}(\omega) = \frac{1}{2}\frac{ie}{\hbar}\int_0^\infty \mathrm{d}t\, 2\pi\delta(t)\langle [\hat{P}_\alpha(t), \hat{\jmath}_\alpha]\rangle \tag{4.57}$$

$$= \frac{\pi e^2}{2V\hbar^2}\langle [\hat{P}_\alpha, [\hat{P}_\alpha, \hat{K}]]\rangle\,, \tag{4.58}$$

which reduces to (4.50) since $[\hat{P}_\alpha, [\hat{P}_\alpha, \hat{K}]] = \sum_{\boldsymbol{k},\sigma} \hat{n}_{\boldsymbol{k},\sigma}\frac{d^2}{dk_\alpha^2}\epsilon_{\boldsymbol{k}}$.

From (4.56), we can also express the total optical conductivity as

$$\sigma_{\alpha\alpha}(\omega) = e\langle\langle \hat{P}_\alpha, \hat{\jmath}_\alpha \rangle\rangle(\omega)\,. \tag{4.59}$$

This form makes the qualitative behavior seen in Fig. 4.5 physically plausible: In the presence of (small) interactions, quasiparticles acquire a finite lifetime and, thus, correlations generically decay in real time which leads to $\sigma(\omega)$ being finite (i.e., not infinite) at all $\omega$. At $T = 0$, however, quasiparticles directly at the Fermi surface are still stable and may give rise to a (reduced) Drude $\delta$-peak. Within Fermi liquid theory, the reduction factor is $Z$, the quasiparticle weight. At finite $T$, all lifetimes are finite (but still large near the Fermi surface) so that the Drude peak broadens with increasing interaction until both Fermi liquid theory breaks down and the Drude peak may disappear in the incoherent background or vanish when a gap opens. Assuming, on the other hand, that $\sigma(\omega)$ remains finite at all $\omega$, one can as well work with the real part of the paramagnetic contribution for $\omega > 0$ as given by (4.49) and obtain $\sigma_{\mathrm{dc}}$ as $\lim_{\omega\to 0}\sigma(\omega)$. In fact, in numerical calculations at $T = 0$ the Drude weight is often obtained via the $f$-sum rule (Dagotto, 1994).

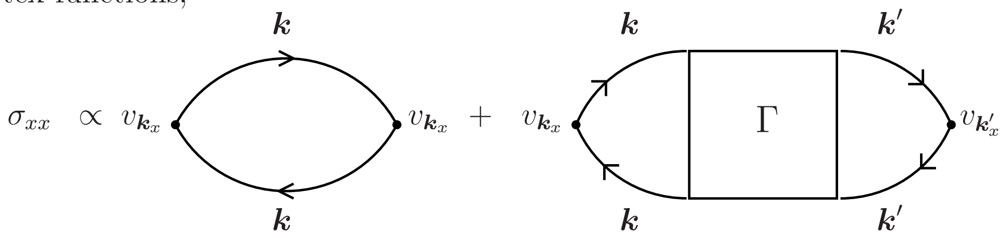
[18] Strictly speaking, the integral on the right hand side of (4.57) is undefined. For a more rigorous proof, one needs to keep a "small" real $\eta > 0$, use $\mathrm{Re}\int_0^\infty \mathrm{d}\omega \mathrm{e}^{i\omega(t+i\eta)} = \eta/(\eta^2 + t^2)$ and the fact that $\mathrm{Im}\,\langle i[\hat{A}, \hat{B}]\rangle = 0$ for Hermitian operators $\hat{A}$, $\hat{B}$. Then, (4.58) can be derived from $\lim_{\eta\to 0}\mathrm{Re}\int_0^\infty \mathrm{d}\omega\int_0^\infty \mathrm{d}t \mathrm{e}^{i\omega(t+i\eta)} f(t) = f(0^+)\lim_{\eta\to 0}\int_0^\infty \mathrm{d}t \eta/(\eta^2+t^2) = f(0^+)\pi/2$ for a real function $f(t)$.

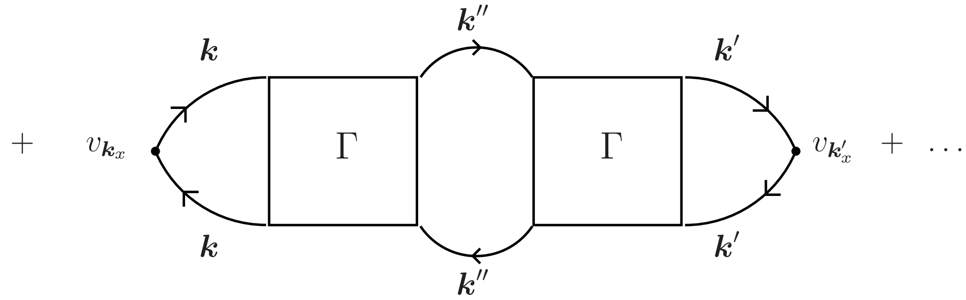
## 4.3 Optical Conductivity in the Limit $d \to \infty$

In this section, we will initially follow the conventional treatment used for the derivation of $\sigma(\omega)$ for the hypercubic lattice with NN hopping for $d \to \infty$ (Pruschke et al., 1993), but keep the treatment general for a single-band model defined on a lattice which is periodic and exhibits inversion symmetry in current direction $x$.[19] In this case, (4.49) implies for the paramagnetic contribution to the optical conductivity[20]

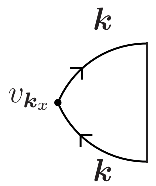
$$\sigma_{xx}(\omega) = \frac{ie^2}{V\hbar^2(\omega + i0^+)} \sum_{\boldsymbol{k},\boldsymbol{k}'} \sum_{\sigma\sigma'} v_{\boldsymbol{k}_x} v_{\boldsymbol{k}'_x} \langle\langle \hat{n}_{\boldsymbol{k}\sigma}, \hat{n}_{\boldsymbol{k}'\sigma'} \rangle\rangle (\omega + i0^+) \,. \tag{4.60}$$

For simplicity, we have here introduced the notation $v_{\boldsymbol{k}_x}$ for the $x$ component of the vector $\boldsymbol{v_k}$. In a diagrammatic series expansion, the correlation function can be expressed in terms of interacting Green function lines and particle-hole irreducible vertex functions,




The particle-hole irreducible vertex function $\Gamma$ becomes effectively local in $d \to \infty$, i.e., becomes independent of $\boldsymbol{k}$, $\boldsymbol{k}'$. Thus, the summations over $\boldsymbol{k}$, $\boldsymbol{k}'$ can be performed independently (Khurana, 1990; Pruschke et al., 1993; Georges et al., 1996). Furthermore, the dispersion has even parity with respect to $k_x$, i.e., $\epsilon_{(-k_x,\boldsymbol{k}_\perp)} = \epsilon_{(k_x,\boldsymbol{k}_\perp)}$. Therefore, $G_{\boldsymbol{k}\sigma}(\omega) = 1/(\omega - \epsilon_{\boldsymbol{k}} - \Sigma_\sigma(\omega))$ also has even parity while $v_{\boldsymbol{k}_x}$ has odd parity with respect to $k_x$. Consequently, each individual $\boldsymbol{k}$ summation involving only one current vertex $v_{\boldsymbol{k}_x}$ vanishes:




$= 0 \,.$

[19] For lattices which are not periodic in all directions, one would have to use the replacement $\boldsymbol{k} \to (\epsilon, k_x)$ in the following. Thus, we would label eigenstates of the noninteracting system by their momentum in current direction and total energy instead of the full momentum.

[20] Here, the diamagnetic contribution is neglected so that resulting expressions will be complete only for $\mathrm{Re}\,\sigma(\omega)$ for $\omega \neq 0$.

Thus, vertex corrections do not contribute in $d \to \infty$ and (4.60) reduces to the bubble contribution (with inverse temperature $\beta$ and Matsubara frequencies $i\omega_n$)

$$\sigma_{xx}(i\nu) = \frac{e^2}{V\hbar^2\nu} T \sum_{\boldsymbol{k},\sigma,\omega_n} v_{\boldsymbol{k}_x}^2 G_{\boldsymbol{k}\sigma}(i\omega_n)\, G_{\boldsymbol{k}\sigma}(i\omega_n + i\nu)\,. \tag{4.61}$$

The momentum independence[21] of $\Sigma(\omega)$ can now be used for replacing the momentum sum by an energy integral (where $N$ denotes number of sites)

$$\sigma_{xx}(i\nu) = 2\frac{e^2}{\hbar^2\nu}\frac{N}{V}\int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \tilde{\rho}_{xx}(\epsilon)\, T \sum_n G_\epsilon(i\omega_n) G_\epsilon(i\omega_n + i\nu)\,. \tag{4.62}$$

Here, we have introduced [cf. (3.11)]

$$G_\epsilon(i\omega_n) \equiv G_{\epsilon_{\boldsymbol{k}},\sigma}(i\omega_n) = \frac{1}{i\omega_n - \epsilon_{\boldsymbol{k}} - \Sigma_\sigma(i\omega_n)} \tag{4.63}$$

and

$$\tilde{\rho}_{xx}(\epsilon) := \frac{1}{N}\sum_{\boldsymbol{k}} v_{\boldsymbol{k}_x}^2 \delta(\epsilon - \epsilon_{\boldsymbol{k}})\,. \tag{4.64}$$

and specialized on the paramagnetic case. The factor of 2 in (4.62) is for spin degeneracy. Note that $\tilde{\rho}_{xx}(\epsilon) = \mathcal{O}(1/d)$; in the isotropic case, we have $d\tilde{\rho}_{xx}(\epsilon) = \tilde{\rho}(\epsilon) := \frac{1}{N}\sum_{\boldsymbol{k}} |\boldsymbol{v}_{\boldsymbol{k}}|^2 \delta(\epsilon - \epsilon_{\boldsymbol{k}})$. Analytic continuation of (4.62) yields

$$\boxed{\mathrm{Re}\,\sigma_{xx}(\omega) = \sigma_0 \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \tilde{\rho}_{xx}(\epsilon) \int_{-\infty}^{\infty} \mathrm{d}\omega'\, A_\epsilon(\omega') A_\epsilon(\omega' + \omega) \frac{n_\mathrm{f}(\omega') - n_\mathrm{f}(\omega + \omega')}{\omega}\,,} \tag{4.65}$$

where

$$A_\epsilon(\omega) = -\frac{1}{\pi}\mathrm{Im}\, G_\epsilon(\omega)\,, \qquad n_\mathrm{f}(\omega) = \left(1 + \mathrm{e}^{\beta(\omega-\mu)}\right)^{-1}\,, \tag{4.66}$$

and

$$\sigma_0 := \frac{2\pi e^2}{\hbar^2}\frac{N}{V}\,. \tag{4.67}$$

We note that the limit of infinite dimensionality enters (4.65) only via the locality of the self-energy and of the irreducible vertex function. Using both properties as an approximation, (4.65) may therefore also be applied in finite dimensions. For numerical results, we will set $e$, $\hbar$, and the lattice spacing $a$ to unity which corresponds to $\sigma_0 = 2\pi$.

[21] If $\Sigma(\omega)$ was momentum-dependent, the momentum sum in the general bubble-sum expression (4.61) could not be replaced by an energy integral. For piecewise constant $\Sigma(\omega)$ like in the DCA, however, a generalization of (4.152) could be applied on each patch. The individual contributions for each value of the coarse-grained momentum would then have to be summed up for the total optical conductivity.

### 4.3.1 Optical Conductivity for the Hypercubic Lattice

For a hypercubic lattice with NN hopping, $\tilde{\rho}_{xx}(\epsilon)$ is proportional to the DOS,

$$\tilde{\rho}^{\text{hc}}_{xx}(\epsilon) = \frac{t^{*2}a^2}{d}\rho^{\text{hc}}(\epsilon) \tag{4.68}$$

We will here first reproduce Pruschke et al.'s (1993) proof using the same Fourier transform technique that had been used before for calculating the (Gaussian) DOS of the hc lattice (Müller-Hartmann, 1989*a*) and then derive a shorter version which is more general. Using $\epsilon_{\boldsymbol{k}} = -2t\sum_{\alpha=1}^{d}\cos(k_\alpha)$ and $\frac{\partial \epsilon_{\boldsymbol{k}}}{\partial k_\alpha} = 2t\sin(k_\alpha)$, which corresponds to lattice spacing $a = 1$, $\tilde{\rho}^{\text{hc}}_{xx}(\epsilon)$ transforms as follows,

$$\begin{aligned}
\int\limits_{-\infty}^{\infty} \mathrm{d}\epsilon\, \tilde{\rho}_{xx}(\epsilon)\, \mathrm{e}^{is\epsilon} &= \left[\int\limits_{-\pi}^{\pi}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{-2ist\cos(k)}\right]^{d-1}\left[\int\limits_{-\pi}^{\pi}\frac{\mathrm{d}k}{2\pi}4t^2\sin^2(k)\mathrm{e}^{-2ist\cos(k)}\right] &(4.69)\\
&= 2t^2\left[J_0(2st)\right]^d\left(1+\frac{J_2(2st)}{J_0(2st)}\right) &(4.70)\\
&= 2t^2\exp\left[d\ln\left(1-\frac{(2st)^2}{4}+\cdots\right)+\ln\left(1+\frac{(2st)^2}{8}+\cdots\right)\right]\\
&= 2t^2\exp\left[-\frac{(st^*)^2}{2}+\mathcal{O}(\frac{1}{d})\right]. &(4.71)
\end{aligned}$$

Keeping only leading terms, transforming back, and determining the dependence on $a$ from scaling, one obtains (using $Z = 2d$)

$$\tilde{\rho}_{xx}(\epsilon) = 2t^2a^2\rho(\epsilon)\,, \tag{4.72}$$

which confirms (4.68). Thus, the conductivity reads,

$$\boxed{\begin{aligned}
\mathrm{Re}\,\sigma^{\text{hc}}_{xx}(\omega) &= \frac{2\pi e^2a^2t^{*2}}{d\hbar^2}\frac{N}{V}\int\limits_{-\infty}^{\infty}\mathrm{d}\epsilon\,\rho(\epsilon)\\
&\times \int\limits_{-\infty}^{\infty}\mathrm{d}\omega'\,A_\epsilon(\omega')A_\epsilon(\omega'+\omega)\,\frac{n_{\text{f}}(\omega')-n_{\text{f}}(\omega+\omega')}{\omega}\,.
\end{aligned}} \tag{4.73}$$

While the above approach using Bessel functions $J_\nu(x)$ has the advantage that it allows for the calculation of corrections terms of $\mathcal{O}(1/d)$ and beyond, the $d \to \infty$ result can be obtained more directly: Let us assume that the total noninteracting kinetic energy separates into a contribution depending only on hopping in the current direction and a contribution which does not involve hopping in that direction and that the latter contribution dominates the bandwidth,

$$\epsilon_{\boldsymbol{k}} = \epsilon^{\perp}_{\boldsymbol{k}_\perp} + \epsilon^{x}_{k_x}\,, \quad \frac{\sqrt{\langle(\epsilon^x_{k_x} - \langle\epsilon^x_{k_x}\rangle)^2\rangle}}{\sqrt{\langle(\epsilon_{\boldsymbol{k}} - \langle\epsilon_{\boldsymbol{k}}\rangle)^2\rangle}} = \mathcal{O}(1/d)\,. \tag{4.74}$$

Then, to leading order,

$$\tilde{\rho}_{xx}(\epsilon) = \frac{1}{N}\sum_{\boldsymbol{k}} v_{k_x}^2 \delta(\epsilon - \epsilon_{\boldsymbol{k}_\perp}^{\perp}) = \rho(\epsilon)\,\frac{a_x}{2\pi}\int\limits_{-\pi/a_x}^{\pi/a_x} \mathrm{d}k_x\, v_{k_x}^2\,. \tag{4.75}$$

Thus, expression (4.73) directly applies to all stacked lattices in $d \to \infty$ when hopping in current direction (lattice spacing $a_x$) is restricted to nearest neighbors with amplitude $t^*/\sqrt{Z}$. The suitable generalization for arbitrary-range hopping in current direction follows from (4.75). Note that the definition of hopping elements in current direction does not impact any local properties and that the frequency sum over $\sigma_{xx}(\omega)$ will be directly related to local properties (such as the kinetic energy) only for isotropic models.

### 4.3.2 $f$-sum Rule within the DMFT

In this subsection, we will derive a very general new form of the $f$-sum rule for the optical conductivity which applies to all cases where expression (4.65) for $\sigma(\omega)$ is used, in particular to $d = \infty$. We do not need to make any assumptions about the functional form of $\tilde{\rho}_{xx}(\epsilon)$ other than that it has to be piecewise differentiable. In particular, we do not need to know the dispersion $\epsilon_{\boldsymbol{k}}$; thus, our results can even be applied to lattices without (full) $\boldsymbol{k}$-space. Furthermore, we can show that the sum rule is fulfilled also for approximate DMFT solutions as long as the self-energy is causal. Finally, we present an efficient computation scheme for the frequency sum directly on the imaginary axis.

We start with the noninteracting limit. Then the spectral function becomes a $\delta$-function, $A_\epsilon(\omega) \to \delta(\omega - \epsilon)$, and we obtain from (4.65) [with $\sigma_0$ defined in (4.67)],

$$\begin{aligned}
\int\limits_0^\infty \mathrm{d}\omega\, \sigma_{xx}(\omega) &= \frac{\sigma_0}{2}\int\limits_{-\infty}^{\infty}\mathrm{d}\epsilon\, \tilde{\rho}_{xx}(\epsilon)\int\limits_{-\infty}^{\infty}\mathrm{d}\omega\int\limits_{-\infty}^{\infty}\mathrm{d}\omega'\,\delta(\omega'-\epsilon)\delta(\omega+\omega'-\epsilon)\frac{n_{\mathrm{f}}(\omega')-n_{\mathrm{f}}(\omega'+\omega)}{\omega}\\
&= -\frac{\sigma_0}{2}\int\limits_{-\infty}^{\infty}\mathrm{d}\epsilon\,\tilde{\rho}_{xx}(\epsilon)\frac{\mathrm{d}}{\mathrm{d}\epsilon}n_{\mathrm{f}}(\epsilon) \;=\; \frac{\sigma_0}{2}\int\limits_{-\infty}^{\infty}\mathrm{d}\epsilon\,\tilde{\rho}'_{xx}(\epsilon)\,n_{\mathrm{f}}(\epsilon) &(4.76)\\
&= \frac{\sigma_0}{4}\Big\langle\frac{\tilde{\rho}'_{xx}(\epsilon)}{\rho(\epsilon)}\Big\rangle\,. &(4.77)
\end{aligned}$$

Here and in the following, the prime denotes the derivative with respect to the argument, $\tilde{\rho}'_{xx}(\epsilon) := \frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\rho}_{xx}(\epsilon)$. No boundary terms arise in the partial integration in (4.76) since $\tilde{\rho}(\epsilon)$ vanishes at the band edges. We further introduced the notation

$$\langle f(\epsilon)\rangle := \int\limits_{-\infty}^{\infty}\mathrm{d}\epsilon\,\rho(\epsilon)\,(n_{\epsilon,\uparrow}+n_{\epsilon,\downarrow})\,f(\epsilon)\,, \tag{4.78}$$

where the expectation value of the generalized momentum distribution function is for arbitrary interaction given in terms of Fermi distribution function $n_{\mathrm{f}}(\omega)$, chemical potential $\mu$, and spectral function $A_{\epsilon,\sigma}(\omega)$ by

$$n_{\epsilon,\sigma} := \langle \hat{n}_{\epsilon,\sigma} \rangle = \int_{-\infty}^{\infty} \mathrm{d}\omega \, n_{\mathrm{f}}(\omega - \mu) A_{\epsilon,\sigma}(\omega) \, . \tag{4.79}$$

In the paramagnetic case, $n_\epsilon \equiv n_{\epsilon,\sigma}$. While the expression (4.77) was derived for the noninteracting limit, and, thus, explicitly only determines the weight of the Drude peak, one might suspect that it remains valid in connection with (4.79) in the interacting case. This would amount to a generalization of the hc sum rule (4.44), where (for $t^* = 1$ and $a = 1$), $\tilde{\rho}(\epsilon) = \rho(\epsilon) = \mathrm{e}^{-\epsilon^2/2}$ and $\langle \tilde{\rho}'(\epsilon)/\rho(\epsilon) \rangle = \langle -\epsilon \rangle$ is minus the kinetic energy. For a proof of (4.77) in the general case, we will show that the frequency dependent part in (4.65), denoted as $I(\epsilon)$, sums up to $-\mathrm{d}n_\epsilon/\mathrm{d}\epsilon$,

$$\begin{aligned}
I(\epsilon) &= \int_{-\infty}^{\infty} \mathrm{d}\omega \int_{-\infty}^{\infty} \mathrm{d}\omega' A_\epsilon(\omega') A_\epsilon(\omega' + \omega) \frac{n_{\mathrm{f}}(\omega') - n_{\mathrm{f}}(\omega' + \omega)}{\omega} && (4.80)\\
&= \int_{-\infty}^{\infty} \mathrm{d}\omega'' \int_{-\infty}^{\infty} \mathrm{d}\omega' \, A_\epsilon(\omega') A_\epsilon(\omega'') \frac{n_{\mathrm{f}}(\omega') - n_{\mathrm{f}}(\omega'')}{\omega'' - \omega'} && (4.81)\\
&= \int_{-\infty}^{\infty} \mathrm{d}\omega' \, n_{\mathrm{f}}(\omega') A_\epsilon(\omega') \int_{-\infty}^{\infty} \mathrm{d}\omega'' \, \frac{A_\epsilon(\omega'')}{\omega'' - \omega'} \\
&\quad + \int_{-\infty}^{\infty} \mathrm{d}\omega'' \, n_{\mathrm{f}}(\omega'') A_\epsilon(\omega'') \int_{-\infty}^{\infty} \mathrm{d}\omega' \, \frac{A_\epsilon(\omega')}{\omega' - \omega''} && (4.82)\\
&= \frac{2}{\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega \, n_{\mathrm{f}}(\omega) \, \mathrm{Re}\, G_\epsilon(\omega) \, \mathrm{Im}\, G_\epsilon(\omega) && (4.83)\\
&= -\frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega \, n_{\mathrm{f}}(\omega) \, \frac{\mathrm{d}}{\mathrm{d}\epsilon} \, \frac{\mathrm{Im}\, \Sigma(\omega)}{(\omega - \epsilon - \mathrm{Re}\, \Sigma(\omega))^2 + \mathrm{Im}\, \Sigma(\omega)^2} && (4.84)\\
&= -\frac{\mathrm{d}}{\mathrm{d}\epsilon} \, n_\epsilon \, . && (4.85)
\end{aligned}$$

Thus, the $f$-sum rule for the optical conductivity within the DMFT, i.e., within the bubble approximation with a $\boldsymbol{k}$-independent self-energy reads

$$\begin{aligned}
\int_0^{\infty} \mathrm{d}\omega \, \sigma_{xx}(\omega) &= \frac{\sigma_0}{2} \int_{-\infty}^{\infty} \mathrm{d}\epsilon \, \tilde{\rho}_{xx}(\epsilon) \Big( -\frac{dn_\epsilon}{d\epsilon} \Big) && (4.86)\\
&= \frac{\sigma_0}{2} \int_{-\infty}^{\infty} \mathrm{d}\epsilon \, \tilde{\rho}'_{xx}(\epsilon) \, n_\epsilon = \frac{\sigma_0}{4} \, \langle \frac{\tilde{\rho}'_{xx}(\epsilon)}{\rho(\epsilon)} \rangle \, . && (4.87)
\end{aligned}$$

This sum rule is of great practical value as a check of numerical results. Note that it cannot only be evaluated from real-frequency data, but also directly on the imaginary axis, since

$$
\begin{aligned}
-\frac{dn_\epsilon}{d\epsilon} &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty} \mathrm{d}\omega \,\mathrm{Im}\,\big(G_\epsilon^2(\omega)\big) n_\mathrm{f}(\omega) \;=\; T\sum_n G_\epsilon^2(i\omega_n) \\
&= T\sum_n \frac{(\epsilon+\mathrm{Re}\,\Sigma_n)^2-(\omega_n-\mathrm{Im}\,\Sigma_n)^2}{\big((\epsilon+\mathrm{Re}\,\Sigma_n)^2+(\omega_n-\mathrm{Im}\,\Sigma_n)^2\big)^2}\,. \qquad (4.88)
\end{aligned}
$$

Since the terms in (4.88) decay as $1/\omega_n^2$, the sum over $n$ does not require additional convergence factors. Furthermore, due to the fast convergence, the sum is insensitive to the large-frequency behavior of the self-energy. Since, on the other hand, QMC methods in general and the particular Fourier-transform techniques used in this work do not yield good estimates for $\Sigma_n$ at large frequencies, the estimate for (4.88) will be even much better conditioned than, e.g., the estimate (3.10) for the kinetic energy. In fact, we found that computing $n_\epsilon$ by integrating over $\mathrm{d}n_\epsilon/\mathrm{d}\epsilon$ as given by (4.88) reduces the $\Delta\tau$ error for the kinetic energy considerably when compared with the direct computation (3.10).

### 4.3.3 *f*-sum Rule and General Dispersion Formalism

Within the framework of the general dispersion method developed in Sec. 2.3, the full noninteracting dispersion $\epsilon_{\boldsymbol{k}}$ can be written in terms of the transformation function $\mathcal{F}$ and the dispersion $\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}$ of the hypercubic lattice with NN hopping only. Then, the derivatives appearing in (4.47) through (4.50) read

$$
\frac{\partial\epsilon_{\boldsymbol{k}}}{\partial k_x} = \frac{\partial}{\partial k_x}\mathcal{F}(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}) = \mathcal{F}'(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}})\frac{\partial\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}}{\partial k_x} \qquad (4.89)
$$

and

$$
\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k_x^2} = \mathcal{F}'(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}})\frac{\partial^2\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}}{\partial k_x^2} + \mathcal{F}''(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}})\left(\frac{\partial\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}}{\partial k_x}\right)^2 . \qquad (4.90)
$$

Using (for $a=1$)

$$
\epsilon_{\boldsymbol{k}}^{\mathrm{hc}} = \sqrt{\frac{2}{d}}\sum_{\alpha=1}^{d}\cos(k_\alpha); \quad \frac{\partial\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}}{\partial k_x} = -\sqrt{\frac{2}{d}}\sin(k_x); \quad \frac{\partial^2\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}}{\partial k_x^2} = -\sqrt{\frac{2}{d}}\cos(k_x) \qquad (4.91)
$$

we can carry out the sum in (4.64)

$$
\begin{aligned}
\tilde\rho(\epsilon) &= \lim_{d\to\infty} d\tilde\rho_{xx}(\epsilon) \\
&= \frac{1}{N}\sum_{\boldsymbol{k}}\lim_{d\to\infty} d\Big(\frac{\partial\epsilon_{\boldsymbol{k}}}{\partial k_x}\Big)^2\delta(\epsilon-\epsilon_{\boldsymbol{k}}) \\
&= \lim_{d\to\infty}\frac{d}{(2\pi)^d}\int\!\mathrm{d}k_x\int\!\mathrm{d}\boldsymbol{k}_\perp\big(\mathcal{F}'(\epsilon_\perp^{\mathrm{hc}}(\boldsymbol{k}_\perp)+\epsilon_\parallel(k_x))\big)^2\,\frac{2}{d}\sin^2(k_x) \\
&\quad\times\delta\left(\epsilon-\mathcal{F}\left(\epsilon_\perp^{\mathrm{hc}}(\boldsymbol{k}_\perp)+\epsilon_\parallel(k_x)\right)\right) \\
&= \left[\lim_{d\to\infty}\frac{1}{(2\pi)^{d-1}}\int\!\mathrm{d}\boldsymbol{k}_\perp\left(\mathcal{F}'\left(\epsilon_\perp^{\mathrm{hc}}(\boldsymbol{k}_\perp)\right)\right)^2\delta\left(\epsilon-\mathcal{F}\left(\epsilon_\perp^{\mathrm{hc}}\right)\right)\right]\frac{1}{2\pi}\int\!\mathrm{d}k_x\,2\sin^2(k_x) \\
&= \rho(\epsilon)\left(\mathcal{F}'(\mathcal{F}^{-1}(\epsilon)\right)^2 , \qquad (4.92)
\end{aligned}
$$

where we have exploited the fact that the energy of the hypercubic lattice separates into contributions depending on $k_x$ and on components of $\boldsymbol{k}$ perpendicular to the $x$ direction, respectively. Since the former is of order $1/d$ it could be neglected in the fourth line of (4.92). In an analogous fashion one can compute the sum in (4.50)

$$\lim_{d\to\infty}\sum_{\boldsymbol{k},\sigma}\langle\hat{n}_{\boldsymbol{k}}\rangle\, d\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k_x^2} = 2\int_{-\infty}^{\infty} \mathrm{d}\epsilon\,\rho(\epsilon)\left[\mathcal{F}''(\mathcal{F}^{-1}(\epsilon)) - \mathcal{F}'(\mathcal{F}^{-1}(\epsilon))\mathcal{F}^{-1}(\epsilon)\right]\langle\hat{n}_\epsilon\rangle \tag{4.93}$$

Finally, since $\tilde{\rho}'(\epsilon)$ is related with the second derivative of the dispersion,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\rho}(\epsilon) &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big(\big(\mathcal{F}'\big(\mathcal{F}^{-1}(\epsilon)\big)\big)^2\rho(\epsilon)\Big)\\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\big(\mathcal{F}'\big(\mathcal{F}^{-1}(\epsilon)\big)\,\rho^{\mathrm{hc}}\big(\mathcal{F}^{-1}(\epsilon)\big)\big)\\
&= \mathcal{F}'\big(\mathcal{F}^{-1}(\epsilon)\big)\,\big(\rho^{\mathrm{hc}}\big)'\big(\mathcal{F}^{-1}(\epsilon)\big)\,\big(\mathcal{F}^{-1}(\epsilon)\big)'\\
&\quad+\mathcal{F}''\big(\mathcal{F}^{-1}(\epsilon)\big)\,\big(\mathcal{F}^{-1}(\epsilon)\big)'\,\rho^{\mathrm{hc}}\big(\mathcal{F}^{-1}(\epsilon)\big)\\
&= \big(\rho^{\mathrm{hc}}\big)'\big(\mathcal{F}^{-1}(\epsilon)\big) + \frac{\mathcal{F}''\big(\mathcal{F}^{-1}(\epsilon)\big)}{\mathcal{F}'\big(\mathcal{F}^{-1}(\epsilon)\big)}\rho^{\mathrm{hc}}\big(\mathcal{F}^{-1}(\epsilon)\big)\\
&= \big[\mathcal{F}''\big(\mathcal{F}^{-1}(\epsilon)\big) - \mathcal{F}^{-1}(\epsilon)\mathcal{F}'\big(\mathcal{F}^{-1}(\epsilon)\big)\big]\,\rho(\epsilon)
\end{aligned} \tag{4.94}$$

we can reformulate the general DMFT sum rule (4.87) in terms of the dispersion:

$$\boxed{\begin{aligned}
\int_0^\infty \mathrm{d}\omega\,\sigma_{xx}(\omega) &\overset{(4.87)}{=} \frac{\sigma_0}{4d}\Big\langle\frac{\tilde{\rho}'(\epsilon)}{\rho(\epsilon)}\Big\rangle\\
&\overset{(4.94)}{=} \frac{\sigma_0}{4d}\,\big\langle\mathcal{F}''\big(\mathcal{F}^{-1}(\epsilon)\big) - \mathcal{F}^{-1}(\epsilon)\mathcal{F}'\big(\mathcal{F}^{-1}(\epsilon)\big)\big\rangle\\
&\overset{(4.93)}{=} \frac{\sigma_0}{4}\sum_{\boldsymbol{k},\sigma}\langle\hat{n}_{\boldsymbol{k}\sigma}\rangle\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k_x^2}\,.
\end{aligned}} \tag{4.95}$$

Since (4.95) is equivalent to (4.50) for a single band, we have demonstrated the consistency of our formalism.

## 4.4 Optical Conductivity for the Bethe Lattice

As discussed in Sec. 2.2, the Bethe "lattice" is more accurately labeled as a tree (a special planar graph) since it is does not constitute a regular lattice. In fact, the periodicity of a regular lattice and the associated graph formed by, e.g., its nearest-neighbor bonds generally imply the existence of an infinite number of inequivalent loops for $Z > 2$ while the Bethe "lattice" is defined by the complete absence of loops (and a constant number of nearest neighbors relative to each site). Due to its definition in topological terms only, concepts like directions, angles, and distances

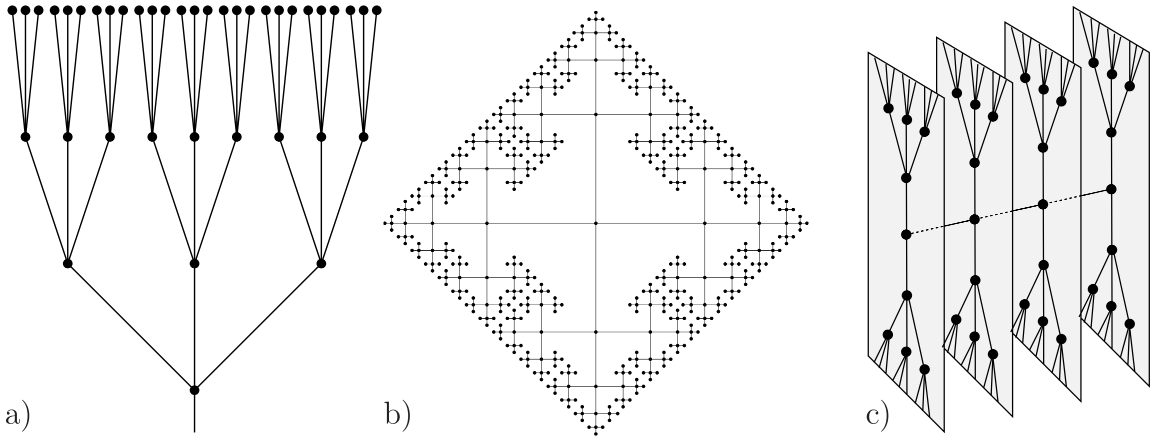



**Figure 4.6:** Bethe lattice: equivalent schematic representations of portions of a Bethe lattice for $Z = 4$.

of bonds or between sites are a priori undefined for the Bethe lattice.[22] It is not possible to embed the Bethe lattice in any dimension so that it is locally isotropic with respect to NN bond lengths and angles. Thus, no $\boldsymbol{k}$ space can be associated with the Bethe lattice. These facts are illustrated by the three equivalent schematic representations of finite portions of the Bethe lattice for $Z = 4$ in Fig. 4.6. In the conventional tree-like drawing Fig. 4.6 a, the sites are arranged in levels. Each site is connected by one bond to the lower level while $K = Z - 1$ bonds extend to the upper level. In the planar drawing Fig. 4.6 b, we have chosen to arrange the bonds with constant angle but geometrically decreasing bond length counted from some central site so that the total spatial extent of the full graph is finite (with both outer and inner surfaces). Drawing Fig. 4.6 c, finally, illustrates a layout of the lattice on stacked planes where one central site on each plane is connected to the neighboring planes. Since the different layouts leave the topology invariant, they correspond to the same (unperturbed) lattice Hamiltonian and, therefore, share all local properties for arbitrary coordination number $Z$.

Consideration of directional transport such as the optical conductivity $\sigma(\omega)$ obviously requires at least one direction to be defined; a finite conductivity can only arise if some bonds are nonorthogonal with respect to this direction. In this section, we will first explore the possibilities strictly for the Bethe lattice, i.e., for fixed topology, and then also in the DMFT context, i.e., for fixed local properties in the limit $Z \to \infty$. Figure 4.6 a suggests applying the fields in the vertical direction. For equidistant levels and in the long-wavelength limit this definition would at first sight resemble the situation in current direction in regular lattices. The uniform current per bond expected from linear response theory within this picture (cf. subsection 4.4.1) would, however, change the charge on each bond due to the lack of inversion symmetry and the broken translational invariance. Then, the system would act more like a capacitor

[22] Therefore, hopping matrix elements cannot be computed from Wannier functions, but have to be regarded as independent parameters of an abstract model.

than as a usual conductor. For a uniform current flow, one would have to balance incoming and outgoing current on each site; in the long-wavelength limit this would amount to geometrically reducing the level spacing upward by a factor of $1/(Z-1)$ and, consequently, imply a semi-finite lattice. In fact, the latter situation is very similar to that depicted in Fig. 4.6 b for which any current definition would also seem artificial. In contrast, the layout depicted in Fig. 4.6 c is translationally invariant in the stacking direction so that fields and currents can be unambiguously defined in this direction. The main drawbacks of this interpretation are that the lattice sites within a plane are then no longer equivalent and that the current flow is limited to a single chain of bonds and, thus, does not scale to the thermodynamic limit. As we will show below, there is also no Drude peak in the noninteracting limit for this model.

Clearly, the pathologies of the Bethe lattice are even worse for transport phenomena than for local properties.[23] Still, the availability of DMFT data has led many researchers to trying to extract meaningful results for transport properties, in particular $\sigma(\omega)$, as well. Rozenberg et al. (1995) just ignored the difficulties, used the formalism valid for the hypercubic lattice, and wrongly stated that the hc $f$-sum rule (4.44) was still observed (cf. subsection 4.4.3). Freericks found a derivation within the "tree level picture" Fig. 4.6 a (Freericks and Jarrell, 1995*a*; Freericks, 2000). Stumpf (1999) discussed $\sigma(\omega)$ for the single-link stacked picture Fig. 4.6 c. In contrast to above approaches, Uhrig and Vlaming (1993) explored the freedom within the DMFT of changing terms of $\mathcal{O}(1/d)$ in the Hamiltonian and computed $\sigma(\omega)$ for stacked Bethe lattices.

For the remainder of this section, we attempt a systematic and exhaustive discussion of ways of defining and computing $\sigma(\omega)$ consistently on the basis of DMFT calculations using a semi-elliptic DOS. For each choice, we characterize the general behavior of $\sigma(\omega)$ and the specific form of the $f$-sum rule. Specifically, the derivations strictly for the Bethe lattice by Freericks and by Stumpf are discussed in subsection 4.4.1 and subsection 4.4.2, respectively. Going beyond the cited work, we point out an inconsistency in the former approach and explain the lack of coherent transport in the latter. Conductivity for stacked lattices is reviewed in a broad context in subsection 4.4.3. In subsection 4.4.4, we show that full off-diagonal disorder generally leads to a very similar expression for $\sigma(\omega)$ as the single-link stacked approach by Stumpf. Finally, we show that our generalized-dispersion approach for the first time yields definition and expressions for an isotropic conductivity corresponding to the semi-elliptic Bethe DOS in subsection 4.4.5.

### 4.4.1 Treelike Layout of the Bethe Lattice

In this subsection, we will discuss a formalism developed by Freericks which is based on the treelike layout of the Bethe lattice shown in Fig. 4.6 a. Initially, Freericks and coworkers had followed Rozenberg's approach of computing $\sigma(\omega)$ for the Bethe lattice by application of the expression (4.73) derived for the hc lattice (Chung and

[23]For a discussion of the importance of surface effects, see Sec. 2.2.

Freericks, 1998*b*). After noting a numerical violation of the corresponding $f$-sum rule (4.44), i.e., a lack of proportionality between the integral over $\sigma(\omega)$ and the full kinetic energy, they realized the need for a careful treatment of the linear response formalism applicable to the Bethe lattice. They proposed (Chung and Freericks, 1998*a*)[24] an application of (4.73), but using a DOS multiplied by a nontrivial (squared) Fermi velocity prefactor of the form $\langle v_{\boldsymbol{k}_x}^2\rangle(\epsilon) = 4-\epsilon^2$, i.e., $\tilde{\rho}(\epsilon) = (4-\epsilon^2)\rho(\epsilon)$. This prefactor vanishes at the band edges (at $\epsilon=\pm 2$) as expected on general physical grounds. Furthermore, the proportionality between the $f$-sum and the total kinetic energy was seen to be reestablished, albeit with an unexpected prefactor 3. No derivations were published; Freericks did, however, make a sketch of his treatment available privately (Freericks, 2000) which forms the basis of our derivation given below.

Let us first show analytically which sum rule obtains from using the hc decoupling $\tilde{\rho}_{xx}(\epsilon) = \langle v_{\boldsymbol{k}_x}^2\rangle\rho(\epsilon)$ with $\rho(\epsilon)$ being the semi-elliptic Bethe DOS (and effectively constant $\langle v_{\boldsymbol{k}_x}^2\rangle$). This is easily accomplished using the formalism developed in subsection 4.3.2,

$$\tilde{\rho}_{xx}(\epsilon) = \langle v_{\boldsymbol{k}_x}^2\rangle\rho(\epsilon); \quad \rho(\epsilon) = \frac{1}{2\pi}\sqrt{4-\epsilon^2}$$

$$\Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\rho}_{xx}(\epsilon) = -\langle v_{\boldsymbol{k}_x}^2\rangle\frac{\epsilon}{2\pi}\big(4-\epsilon^2\big)^{-1/2} \tag{4.96}$$

$$\Rightarrow \quad \int_0^\infty \mathrm{d}\omega\,\sigma_{xx}(\omega) = \langle v_{\boldsymbol{k}_x}^2\rangle\frac{\sigma_0}{4}\langle\frac{-\epsilon}{4-\epsilon^2}\rangle\,. \tag{4.97}$$

Evidently, the frequency sum is indeed not proportional to the expectation value of the kinetic energy. This fact is not surprising since the proportionality of the $f$-sum to the kinetic energy is very specific to the hc case as shown in subsection 4.3.1.

The central assumption made by Freericks is that kinetic energy eigenstates can be written in the following form,[25]

$$|\epsilon\rangle = \sum_x (\gamma_\epsilon)^x \sum_{\alpha_x} |x,\alpha_x\rangle\,. \tag{4.98}$$

Here, the integer $x$ counts the level number (i.e., is the vertical coordinate in Fig. 4.6 a) and $\alpha_x$ is an intra-level coordinate, i.e., labels lattice sites within level $x$. The latter is written with index $x$ since the set of values that the intra-level coordinate can take depends on $x$ for $Z>2$. The factor[26] $\gamma_\epsilon$ takes the role of a phase shift between levels although, as we will see, its absolute value does not in general equal 1. An expression for $\gamma_\epsilon$ in terms of $\epsilon$ is obtained from the eigenvalue equation for the kinetic energy.

[24] Note that this reference is the published version of the preprint (Chung and Freericks, 1998*b*).

[25] Our reproduction of the derivation differs considerably in form from the sketch made available to us (Freericks, 2000). In particular, we found it necessary to introduce indices within the planes. Unless noted, however, the treatments are equivalent.

[26] In the following, we will also write $\gamma_\epsilon^x \equiv (\gamma_\epsilon)^x$ where $x$ is still an exponent rather than an index.

For the kinetic energy, we can write

$$\hat{T} = -2t \sum_{y} \sum_{\alpha_y} \hat{c}^{\dagger}_{y,\alpha_y} \Big( \hat{c}_{y-1,f(\alpha_y)} + \sum_{i=1}^{Z-1} \hat{c}_{y+1,g(\alpha_y,i)} \Big) , \tag{4.99}$$

where $f(\alpha_y)$ maps onto the intra-level coordinate of the "root" (on level $y-1$) of site $(y, \alpha_y)$ . Conversely, $g(\alpha_y, i)$ maps onto the intra-level coordinate of one of the $(Z-1)$ "children" on the higher level of the same site.[27] We obtain for the eigenvalue equation

$$\begin{aligned} \hat{T}|\epsilon\rangle &= -2t \sum_{x} \Big( \gamma_\epsilon^{x-1} + (Z-1)\gamma_\epsilon^{x+1} \Big) \sum_{\alpha_x} |x, \alpha_x\rangle & (4.100) \\ &= -2t\Big(\gamma_\epsilon^{-1} + (Z-1)\gamma_\epsilon\Big)|\epsilon\rangle . & (4.101) \end{aligned}$$

By definition, the coefficient if front of $|\epsilon\rangle$ must equal $\epsilon$; inverting this equation, we obtain

$$\gamma_\epsilon = \frac{-\epsilon \pm \sqrt{\epsilon^2 - 4(Z-1)t^2}}{2(Z-1)t} . \tag{4.102}$$

Expressing the current operator[28] in a form analogous to the kinetic energy,

$$\hat{\jmath} = iet \sum_{y} \sum_{\alpha_y} \hat{c}^{\dagger}_{y,\alpha_y} \Big( - \hat{c}_{y-1,f(\alpha_y)} + \sum_{i=1}^{Z-1} \hat{c}_{y+1,g(\alpha_y,i)} \Big) , \tag{4.103}$$

the associated eigenvalue equation reads,

$$\begin{aligned} \hat{\jmath}|\epsilon\rangle &= iet\Big( - \gamma_\epsilon^{-1} + (Z-1)\gamma_\epsilon\Big)|\epsilon\rangle & (4.104) \\ &= \pm e\sqrt{4(Z-1)t^2 - \epsilon^2}|\epsilon\rangle . & (4.105) \end{aligned}$$

The requirement that the eigenvalues of $\hat{\jmath}$ have to be real fixes the band edges, $|\epsilon| \leq 2t\sqrt{Z-1}$.[29] The square root in (4.105) is to be interpreted as $\langle v_{\boldsymbol{k}_x}\rangle(\epsilon)$ which implies that in the limit of $Z \to \infty$,

$$\tilde{\rho}_{xx}(\epsilon) = (4t^{*2} - \epsilon^2)\rho(\epsilon) . \tag{4.106}$$

Applying the general expression (2.24) specifically for the semi-elliptic DOS (for $t^* = 1$) then leads to

$$\begin{aligned} \tilde{\rho}_{xx}(\epsilon) &= \frac{1}{2\pi}\big(4 - \epsilon^2\big)^{3/2} & (4.107) \\ \int_0^\infty \mathrm{d}\omega\, \sigma_{xx}(\omega) &= \frac{\sigma_0}{4}\langle\frac{\tilde{\rho}'_{xx}(\epsilon)}{\rho(\epsilon)}\rangle = \frac{\sigma_0}{4}\langle -3\epsilon\rangle , & (4.108) \end{aligned}$$

[27] More formally, we can demand that the mapping $(y, \alpha_y, i) \to (y+1, \alpha_{y+1}) \equiv (y+1, g(\alpha_y, i))$ is injective and that $f(g(\alpha_y, i)) = \alpha_y$ for all $y$, $\alpha_y$, and $i \in \{1, \ldots, Z-1\}$.

[28] Since no volume can be defined, we consider here the full current instead of the current density. Furthermore, we set $a = 1$ and $\hbar = 1$.

[29] Freericks links the restriction of $\epsilon$ to be within the band edges to the requirement that the states are extended. See also the discussion below.

where for the sum rule we have used (4.87). Since only one direction is well defined, we can identify $\sigma_{xx}(\omega)$ with $\mathrm{Tr}\left\{\sigma_{xx}(\omega)\right\}$ and find that the corresponding sum rule exceeds the "usual" hypercubic case by a factor of 3. While, in view of the preceding section, we are not surprised to find different sum rules (in terms of specific expectation values) for different lattices, one might view the appearance of the factor of 3 as a clue to some inconsistency in this particular case.

We will show in the following that the preceding derivation (4.98) - (4.108) is incorrect for the Bethe lattice and instead applies to some effective one-dimensional lattice. In order to establish operator identities or, in this case, that the eigenvalue of the current operator can always be expressed in terms of the eigenvalue of the kinetic energy [cf. (4.104)], one needs to demonstrate such a relationship for a complete basis. However, the class of wave functions represented by the ansatz (4.98) is not complete for $Z > 2$. This is easily seen from the fact that, according to this ansatz, the occupation is constant within each level, $\langle \hat{n}_{x,\alpha_x}\rangle = \langle \hat{n}_{x,\alpha'_x}\rangle$ while each single hopping process induces an intraplanar variation of the occupancy.[30] Therefore, Freericks' derivation at best applies to some subspace of the full Hilbert space only.[31] Let us characterize this reduced Hilbert space by finding an orthogonal basis for a section of the Bethe lattice with an arbitrary but finite number L of levels. Setting

$$\gamma_\epsilon = \frac{\tilde{\gamma}_\epsilon}{\sqrt{Z-1}}, \tag{4.109}$$

we can write

$$\begin{aligned} \langle\epsilon|\epsilon'\rangle &= N_0 \sum_{x=1}^{L} (Z-1)^x \gamma_\epsilon^x \gamma_{\epsilon'}^{*\,x} = N_0 \sum_{x=1}^{L} \left(\tilde{\gamma}_\epsilon \tilde{\gamma}_{\epsilon'}^*\right)^x & (4.110) \\ &= N_0 \sum_{x=1}^{L} \mathrm{e}^{i(\varphi_\epsilon - \varphi_{\epsilon'})x} . & (4.111) \end{aligned}$$

Here, $N_0$ is the number of sites in level 0. In (4.111), we have used the fact that $\tilde{\gamma}_\epsilon$ is of unit absolute value for physical values of $\epsilon$ and have introduced its (real) phase $\varphi_\epsilon = \arg(\tilde{\gamma}_\epsilon)$. For orthogonality, we have to require that the phases differ by integer multiples of $2\pi/L$ which fixes the allowed values of $\tilde{\gamma}_l \equiv \tilde{\gamma}_{\epsilon_l}$ up to some arbitrary phase chosen as zero,

$$\begin{aligned} \tilde{\gamma}_l &= \mathrm{e}^{2\pi i l/L}; \quad l \in \{0, ..., L-1\} & (4.112) \\ \epsilon_l &= -2t\sqrt{Z-1}\,\cos\left(\frac{2\pi l}{L}\right) . & (4.113) \end{aligned}$$

[30] While for any energy eigenstate the total occupation of each level is also constant (which is clearly unphysical since it implies infinite occupation per lattice site for level number $x \to -\infty$), the total occupancy per level can be adjusted independently for each level by choosing a suitable linear combination of eigenfunctions (4.98).

[31] Note that the situation is completely different for wave functions than it is for Green functions. In the latter case, the position where a particle is destroyed singles out a central site. The distance from this site then uniquely defines levels and bond directions (for the unperturbed Hamiltonian). In contrast, the Hilbert space defined in terms of wave functions has to restore the equivalence of all lattice sites.

With these settings, we have $\langle \epsilon_l | \epsilon_{l'} \rangle = N_0 L \delta_{ll'}$. Up to the prefactor $\sqrt{Z-1}$, (4.113) is identical to an enumeration of states of the usual $d = 1$ lattice. Therefore, in the limit $L \to \infty$, the density of states corresponding to the selected set of wave function reads,[32]

$$\rho(\epsilon) = \frac{1}{\pi\sqrt{4t^2(Z-1)-\epsilon^2}} \quad \xrightarrow{Z\to\infty} \quad \frac{1}{\pi\sqrt{4t^{*2}-\epsilon^2}} \tag{4.114}$$

Setting $t^* = 1$, we obtain for the sum rule,

$$\tilde{\rho}_{xx}(\epsilon) = (4-\epsilon^2)\rho(\epsilon);\;\ \rho(\epsilon) = \frac{1}{\pi\sqrt{4-\epsilon^2}} \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}\epsilon}\tilde{\rho}_{xx}(\epsilon) = -\frac{\epsilon}{\pi}\big(4-\epsilon^2\big)^{-1/2} \tag{4.115}$$

$$\Rightarrow \quad \int_0^\infty \mathrm{d}\omega\, \sigma_{xx}(\omega) = \frac{\sigma_0}{4}\langle -\epsilon \rangle\,, \tag{4.116}$$

as for the hypercubic lattice. So the factor of 3 was indeed incorrect as the derivation really applied to some effective one-dimensional model.[33] In our view, the inadvertent reduction of the Hilbert space cannot be healed within Freerick's approach since this would involve considering arbitrary level ordering of the lattice sites for a given energy. Even if such an enumeration was possible it seems extremely unlikely that linear combinations can be found which are also eigenstates of the current operator.

Finally, we note that Millis and collaborators also derived an expression for $\sigma(\omega)$ for the Bethe lattice by assuming that the "usual" hc sum rule (4.44) holds (Chattopadhyay, Millis, and Das Sarma, 2000). Consequently, their result agrees with (4.107) and (4.108) up to a factor of 1/3. Since this assumption is also unjustified, there is at present no valid indication at all that an expression of the form $\langle v^2_{\boldsymbol{k}_x}\rangle(\epsilon) \propto (4-\epsilon^2)$ can be derived for any layout of the Bethe lattice or that the treelike layout of the Bethe lattice (cf. Fig. 4.6 a) is useful for deriving an expression for $\langle v^2_{\boldsymbol{k}_x}\rangle(\epsilon)$. Therefore, we have to turn to other approaches which will be discussed in the following subsections.

## 4.4.2 Single-Chain Stacked Bethe Lattice

In this section, we will use an alternative real-space picture where the Bethe lattice is laid out in stacked equivalent plains which are connected by a single chain of bonds as shown in Fig. 4.6 c. Once the layout is chosen, the derivation by Logan and his student Stumpf (1999), which we will discuss in the next paragraph, corresponds to a straightforward application of a general real-space expression for the (paramagnetic contribution to the) optical conductivity in high dimensions (Metzner, Schmit, and Vollhardt, 1992). The main virtue of this method is that it leaves the topology of the Bethe lattice intact and is therefore as specific to the Bethe lattice as possible. Its

[32] Note that this expression cannot be correct for any lattice model with hopping parameter $t$ for $Z \neq 2$ since as shown in Sec. 2.2 the variance for NN hopping must always equal $Zt^2$ while from (4.114) it computes to $2(Z-1)t^2$.

[33] Interestingly, the total expression for $\tilde{\rho}_{xx}(\epsilon)$ equals two times the Bethe DOS $\rho(\epsilon)$. So an application of the results of this subsection as an approximation would imply using $\langle v^2_{\boldsymbol{k}_x}\rangle(\epsilon) = 4$ which up to constant prefactors is equivalent to using the hc formula for the Bethe lattice.

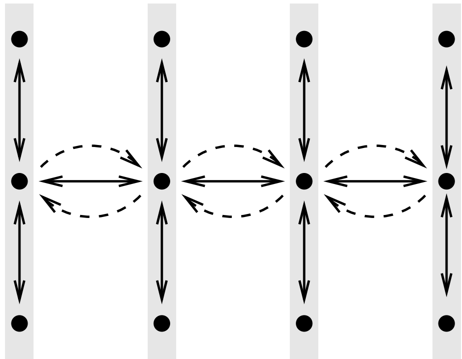


**Figure 4.7:** Schematic representation of hopping processes for a lattice consisting of stacked hyperplanes connected by a single chain of bonds in current direction (e.g., Stumpf's view on the Bethe lattice). Contributions to the current operator are restricted to this chain (dashed lines) while contributions to the kinetic energy (full lines) arise both from hopping along the chain and within the hyperplanes.

---

main characteristics are that it describes only incoherent transport and that the total current is not extensive.[34] The lack of coherence is clearly a shortcoming with respect to a realistic description of Fermi liquids, but very much in line with properties of the unperturbed Bethe lattice (i.e., with $\boldsymbol{A} = 0$). In fact, we have pointed out already in subsection 2.2.2 that all lattices with Bethe DOS have an effective mean free path of one lattice spacing. As we will see in this section, the same statement can be made about all lattices to which the expression for $\sigma(\omega)$ resulting from Stumpf's treatment applies. Specifically, the optical conductivity of periodically stacked lattices where the bonds in current direction are fully disordered agrees with Stumpf's result except that it is extensive.

Let us consider a model as depicted in Fig. 4.7, where a current can flow in $x$ direction along a single chain of bonds which connect central sites of equivalent sublattices. The remaining hopping bonds of each sublattice are confined in stacked hyperplanes orthogonal to the current direction. For full correspondence to Fig. 4.6, each sublattice would have to consist of a Bethe lattice with two bonds removed at the central site (since these are assumed to connect neighboring planes); for $Z = \infty$, however, the removal of a finite number of bonds per site is irrelevant so that each hyperplane can be regarded as a full Bethe lattice. More generally, any DMFT lattice type can be assumed to be replicated on each hyperplane with the same resulting expressions for the conductivity $\sigma_{xx}(\omega)$ in stacking direction.

Labeling sites by the linear coordinate $x$ and some intraplanar coordinate $\alpha$, where $\alpha = 0$ denotes the central sites, the paramagnetic current operator (4.47) takes the form

$$\hat{j}_x = \frac{ieta}{\hbar} \sum_{i,\sigma} \left( \hat{c}^{0\dagger}_{i+1,\sigma} \hat{c}^{0}_{i,\sigma} - \hat{c}^{0\dagger}_{i,\sigma} \hat{c}^{0}_{i+1,\sigma} \right) . \tag{4.117}$$

[34]The latter directly follows from the restriction of the current to the single chain of bonds lying along the current direction.

a) b) c) d)

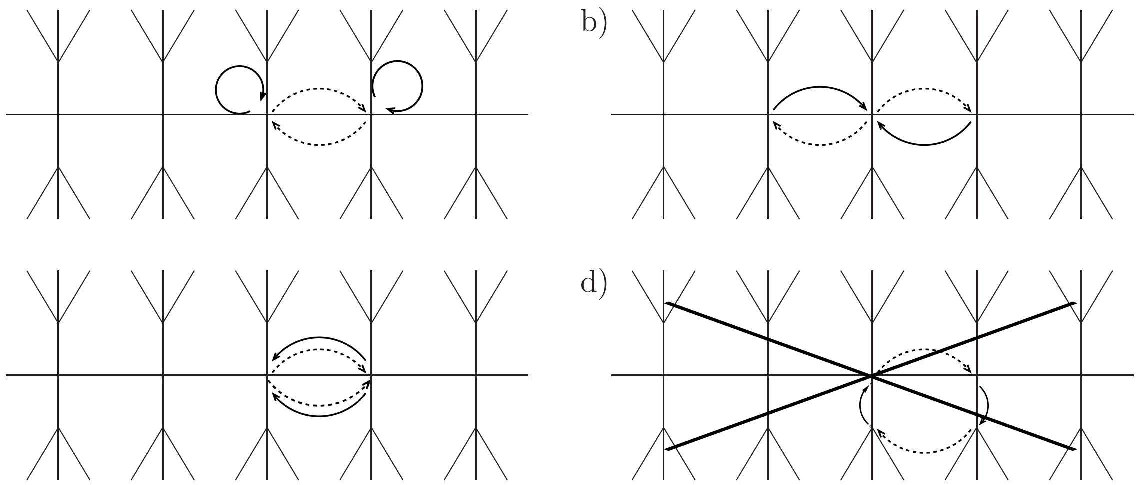


**Figure 4.8:** Diagrammatic representation of terms of the form (4.119) entering the computation of the optical conductivity: **a)** Leading contribution scaling as $\mathcal{O}(1/Z)$. **b)** and **c)** subleading contributions being $\mathcal{O}(1/Z^2)$. **d)** Contributions which include loops do not exist for the Bethe lattice (but are among the leading diagrams for regular lattices).

Here, the upper index 0 of operators stands for $\alpha = 0$, the central site within a sublattice. Consequently, the paramagnetic contribution to the optical conductivity

$$\sigma_{xx}(\omega) = -\frac{e^2a^2t^2}{i\hbar^2(\omega+i0^+)N}\sum_{i,j,\sigma}\langle\langle \hat{c}^{0\dagger}_{i+1,\sigma}\hat{c}^{0}_{i,\sigma} - \hat{c}^{0\dagger}_{i,\sigma}\hat{c}^{0}_{i+1,\sigma}, \hat{c}^{0\dagger}_{j+1,\sigma}\hat{c}^{0}_{j,\sigma} - \hat{c}^{0\dagger}_{j,\sigma}\hat{c}^{0}_{j+1,\sigma}\rangle\rangle\,(\omega+i0^+) \tag{4.118}$$

contains terms of the form (with suppressed $\omega$-dependence)

$$\Gamma_{ijkm} := \langle\langle \hat{c}^{0\dagger}_{i\sigma}\hat{c}^{0}_{k\sigma}, \hat{c}^{0\dagger}_{m\sigma}\hat{c}^{0}_{j\sigma}\rangle\rangle\,\delta_{i,k\pm1}\delta_{m,j\pm1}\,. \tag{4.119}$$

In a diagrammatic evaluation of the corresponding real-time correlation function (cf. Sec. D.2), the exponentials carrying the time dependence in the Heisenberg picture introduce additional paths between pairs of sites; the resulting full diagram only contributes to the implicit trace when it represents a closed path. A selection of such diagrams is shown in Fig. 4.8 where the current paths[35] (dashed lines) are closed by time-dependent propagators (solid lines). In the DMFT formalism, each path of taxi-cab length $l$ carries a factor of $Z^{-l/2}$ so that the diagrams in Fig. 4.8 a, Fig. 4.8 b, and Fig. 4.8 c, have prefactors of $1/Z$, $1/Z^2$, and $1/Z^2$, respectively. Due to the absence of loops, the additional powers of $1/Z$ introduced by nonlocal propagators cannot be compensated by factors of $Z$ arising from a sum over equivalent diagrams. Thus, up to a shift along the current direction, diagram Fig. 4.8 a is the only (leading) contribution for $Z \to \infty$. For comparison, a diagram typical of regular lattices is shown in Fig. 4.8 d; such diagrams usually contribute in leading order $Z/Z^2$, but do not exist on the Bethe lattice. The two time-dependent propagators closing the path do not interfere for $Z \to \infty$ on any lattice (see subsection 4.4.4) which implies that

[35] In momentum representation, these "current vertices" contract to a pair of dots.

the expectation values decouple and lead to a pair of single-particle Green functions, i.e., to the usual bubble sum. In the specific case under consideration, i.e., without loops, both Green functions become local, so that the conductivity reads

$$\sigma_{xx}(\omega) = \frac{\pi e^2 a^2 t^2}{\hbar^2} \frac{L_x}{L} \sum_\sigma \int_{-\infty}^{\infty} \mathrm{d}\omega' A_{A\sigma}(\omega') A_{B\sigma}(\omega'+\omega) \frac{n_\mathrm{f}(\omega') - n_\mathrm{f}(\omega'+\omega)}{\omega} . \quad (4.120)$$

Here, $A(\omega)$ denotes the full local spectral function, $L_x$ the linear extent in current direction, and $L$ the number of sites for some finite part of the system. As indicated by the sublattice indices $A, B$ and spin index $\sigma$, the result is valid also in phases with broken symmetry, i.e., in an antiferromagnetic phase. Note that (4.120) is not directly dependent on the type of the underlying sublattice (confined in each hyperplane), not even via a noninteracting DOS.[36]

The most striking feature of the present derivation and its result (4.120) is that the dc conductivity always remains finite (or vanishes) in the paramagnetic phase, even in the noninteracting limit. This follows from the fact that in the paramagnetic phase, the local spectral function is bounded by the maximum of the noninteracting DOS. For a symmetric DOS which is peaked at zero frequency, the integral in (4.120) is therefore bounded by $\rho^2(0)$.[37] Consequently, there is no qualitative difference between the conductivities for noninteracting and interacting systems: only the energy eigenvalues of each pair of states determines its contribution to $\sigma_{xx}(\omega)$; the associated momenta are unimportant. This behavior is in striking contrast to the Bravais lattice case discussed in subsection 4.1.4 where the dc conductivity diverges in the noninteracting limit. In this sense, (4.120) describes incoherent transport. As we will show in subsection 4.5.1, the expression is generally valid within the DMFT for $\boldsymbol{k}_\perp$-nonconserving hopping between planes.

### Commutators and Loops

As discussed in subsection 4.1.4, the presence of optical absorption at finite frequencies in the noninteracting limit implies that the paramagnetic current operator and the kinetic energy do not commute. In the following, we will demonstrate that indeed $[\hat{H}_0, \hat{\jmath}] \neq 0$ for the single-chain layout discussed in this subsection. For this purpose, we focus first on a single contribution to the current operator and study separately and in turn its commutator with kinetic energy contributions (anti)parallel or perpendicular to its direction. We will see that the former contributions vanish quite generically (in absence of disorder) in summations along the current chains. In contrast, the latter contributions can only cancel in the presence of certain loops of length four which do not exist on the Bethe lattice.

For a single hopping term contributing to the current operator, the commutator

[36] In subsection 4.5.1, we will present a unified treatment of incoherent and coherent transport where $\boldsymbol{k}$-sums and noninteracting DOSs reappear for the case of regular sublattices.

[37] In contrast, momentum-dependent spectral functions diverge in the noninteracting limit.

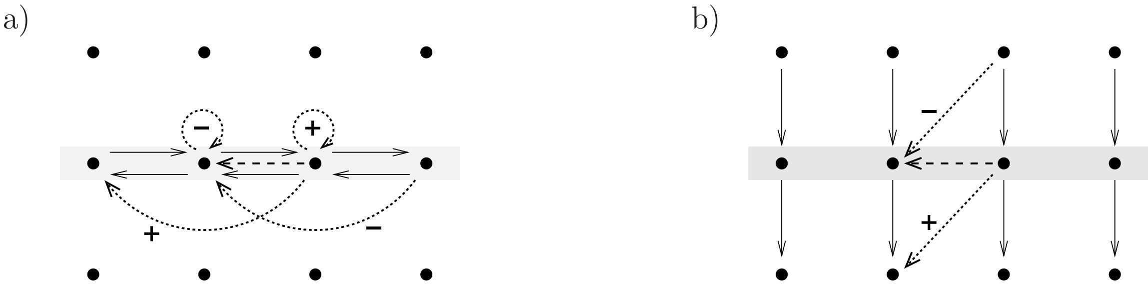



**Figure 4.9:** Commutator of kinetic energy (solid arrows) and one paramagnetic current term (dashed arrow). **a)** Contributions from hopping along the current lead to on-site terms and next-nearest neighbor hopping terms in the commutator (dotted arrows) vanish in sums along the current chain. **b)** Perpendicular hopping induces NNN hopping terms (dotted arrows) away from and towards the current chain (shaded).

$[\hat{H}_0, \hat{\jmath}]$ contains terms from hopping in current direction,

$$\begin{aligned}\sum_i \left[\hat{c}_i^{0\dagger}\hat{c}_{i+1}^{0}, \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\right] &= \sum_i \left[\hat{c}_i^{0\dagger}\big(\delta_{i+1,j}^{0} - \hat{c}_j^{0\dagger}\hat{c}_{i+1}^{0}\big)\hat{c}_{j+1} - \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\,\hat{c}_i^{0\dagger}\hat{c}_{i+1}^{0}\right] \\ &= \hat{c}_{j-1}^{0\dagger}\hat{c}_{j+1}^{0} - \sum_i \left[\hat{c}_j^{0\dagger}\hat{c}_i^{0\dagger}\hat{c}_{j+1}^{0}\hat{c}_{i+1}^{0} + \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\,\hat{c}_i^{0\dagger}\hat{c}_{i+1}^{0}\right] \\ &= \hat{c}_{j-1}^{0\dagger}\hat{c}_{j+1}^{0} - \hat{c}_j^{0\dagger}\hat{c}_{j+2}^{0} \qquad (4.121)\end{aligned}$$

and from hopping opposite to the current direction,

$$\sum_i \left[\hat{c}_{i+1}^{0\dagger}\hat{c}_i^{0}\ , \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\right] = \hat{c}_{j+1}^{0\dagger}\hat{c}_{j+1}^{0} - \hat{c}_j^{0\dagger}\hat{c}_j^{0}\ . \qquad (4.122)$$

Both the generating and the generated terms for the parallel and antiparallel cases are shown schematically in Fig. 4.9 a. Evidently, they both cancel in a sum over the full current,

$$\sum_{i,j} \left[\hat{c}_i^{0\dagger}\hat{c}_{i+1}^{0}, \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\right] = 0 = \sum_{i,j} \left[\hat{c}_{i+1}^{0\dagger}\hat{c}_i^{0}\ , \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\right]. \qquad (4.123)$$

In contrast, hopping processes perpendicular to the current introduce "diagonal" terms in the commutator as depicted in Fig. 4.9 b,[38]

$$\sum_{\substack{i,\alpha \\ \alpha \text{ NN of } 0}} \left[\hat{c}_i^{\alpha\dagger}\hat{c}_i^{0}\ , \hat{c}_j^{0\dagger}\hat{c}_{j+1}^{0}\right] = \sum_{\substack{\alpha \\ \alpha \text{ NN of } 0}} \hat{c}_j^{\alpha\dagger}\hat{c}_{j+1}^{0}\,, \qquad (4.124)$$

which can only vanish in summations over parallelogram loops of length 4 which are not present in the model under consideration. For the hypercubic case of a regular lattice, the cancellation is diagrammatically represented in Fig. 4.10. Consequently,

[38] Within the current notation, also terms of the form $\hat{c}_i^{0\dagger}\hat{c}_i^{\alpha}$ contribute to the kinetic energy. Fig. 4.9 b shows a selection of both contributions.

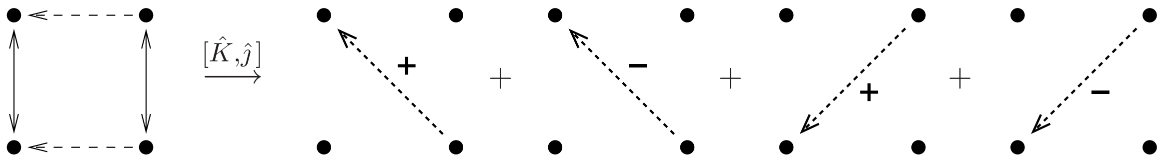



**Figure 4.10:** Commutator of kinetic energy and paramagnetic current: in a hypercubic lattice, contributions of the form Fig. 4.9 b cancel in summations over square loops.

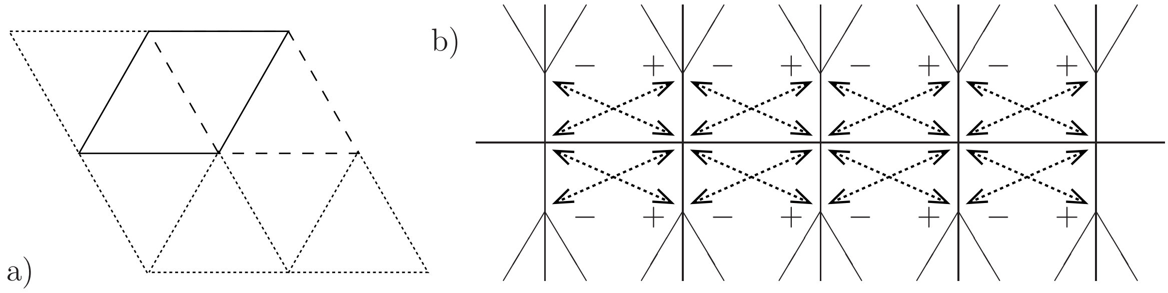



**Figure 4.11: a)** Sums over parallelogram loops (solid lines, dashed lines) lead to a cancellation of all contributions of $[\hat{K}, \hat{\jmath}]$ in an fcc lattice. **b)** Nonvanishing full commutator $[\hat{K}, \hat{\jmath}]$ (dotted arrows) for the Bethe lattice.

the existence of loops of length four as shown in Fig. 4.10 is a necessary condition for the absence of absorption in the noninteracting limit. Note that with suitable modifications (allowing for several current chains) the present discussion generally applies to regular Bravais lattices.[39] Then, the loops along which the commutator cancels consist of pairs of parallel edges as indicated for the case of the fcc lattice in Fig. 4.11 a. Obviously, the properties of a Bravais lattice guarantee that each pair of hopping bonds associated with a common lattice site can be closed to such a parallelogram. Thus, the real-space argument leads to the same result for Bravais lattices as the $\boldsymbol{k}$-space derivation in subsection 4.1.4: the reduction of $\sigma(\omega)$ to a Drude peak in the noninteracting limit.

In contrast, the failure of coherence in the present view on the Bethe lattice can be unambiguously attributed to the nonvanishing terms in the commutator of kinetic energy and current operator represented in Fig. 4.11 b.

### 4.4.3 Periodically Stacked Lattices

In this subsection, we will leave behind approaches for computing $\sigma(\omega)$ strictly for the Bethe lattice and instead discuss a very general concept which can be used for establishing expressions for the conductivity for one particular current direction in the limit of $Z \to \infty$. While this method can be applied to any underlying lattice

[39]The restriction to lattices of the Bravais type is necessary; in fact, our argument correctly implies that absorption at finite frequencies takes place for lattices with several atoms per unit cell even in the noninteracting limit just like for the Bethe lattice.

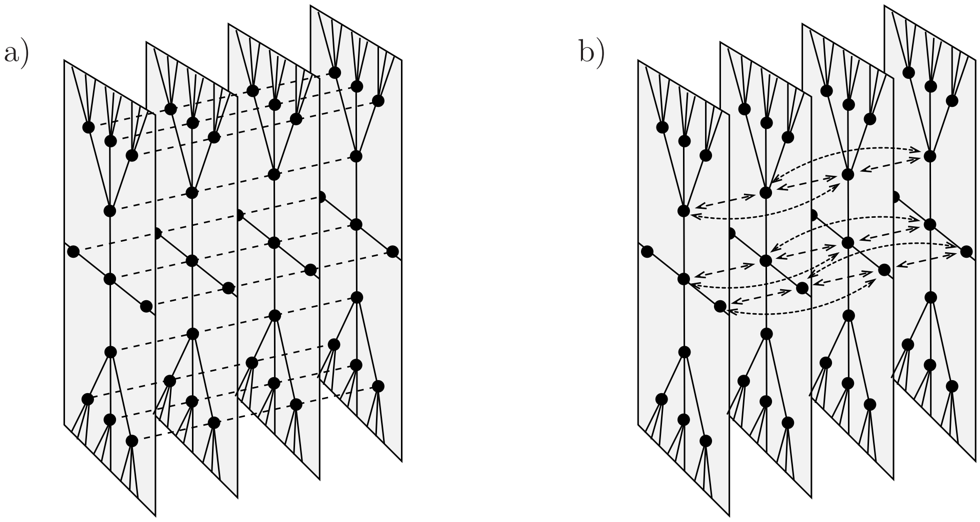



**Figure 4.12:** Stacked lattices with **a)** NN hopping in stacking direction and **b)** simultaneous NN and NNN hopping, here shown for Bethe lattice hyperplanes with $Z = 4$. For clarity, only a selection of bonds is shown in b).

or pseudolattice, it is important to note that it in general describes anisotropic ("*c*-axis") transport. The basic idea of the approach is that the influence of adding a finite number of additional hopping bonds per site on the kinetic energy and on all other local properties will scale to zero in the limit $Z \to \infty$. Thus, the unperturbed Hamiltonian (i.e., without coupling to external fields) can be changed in order to make the transport problem well-defined (or to redefine it), while leaving the local DMFT problem unchanged.

The additional bonds are defined by extending the vector space in which the original lattice lives[40] by one extra dimension. Copies of the hyperplane (i.e., the original vector space) can then be stacked along the new dimension, connecting sites with the same intraplanar coordinates by hopping bonds. We will in the following assume that the additional hopping matrix elements scale like $1/\sqrt{Z}$ so that their contribution to the variance of the total kinetic energy is $\mathcal{O}(1/Z)$.[41] This construction is illustrated in Fig. 4.12 a for nearest-neighbor hopping in current direction $x$. Obviously, the resulting lattice is periodic (at least) in this direction. Furthermore, all lattice sites remain fully equivalent.[42] Therefore, the component $k_x$ of the wave vector is well-defined which makes the general results of the beginning of this section

[40] Since the Bethe lattice is a planar graph, it can be arranged in 2 dimensions. Within the present approach, the particular layout within the plane and even the embedding dimension are irrelevant.

[41] While this assumption is most natural in the DMFT sense it suffices for the theory to apply that the additional terms decay for $Z \to \infty$. In fact, there is no compelling reason in this anisotropic case that hopping terms in the anisotropy direction should have the same scaling law $\propto 1/Z$ as those within the hyperplanes.

[42] Note the contrast to the alternative approaches discussed in subsection 4.4.1 and subsection 4.4.2 where the equivalence of lattice sites was broken by the assignment of level numbers to sites or by the selection of current carrying bonds, respectively.

applicable. For NN hopping in current direction, the conductivity is thus given by

$$\operatorname{Re}\sigma_{xx}(\omega) = \frac{2\pi e^2 a^2 t^2}{\hbar^2}\frac{N}{V}\int\limits_{-\infty}^{\infty} \mathrm{d}\epsilon\rho(\epsilon)\int\limits_{-\infty}^{\infty} \mathrm{d}\omega' A_\epsilon(\omega')A_\epsilon(\omega'+\omega)\frac{n_\mathrm{f}(\omega')-n_\mathrm{f}(\omega+\omega')}{\omega}\,, \tag{4.125}$$

which agrees with (4.73) for a uniform hopping amplitude $t = t^*/\sqrt{d}$. For the particular case of the stacked Bethe lattice, this fact was observed and used by Uhrig and Vlaming (1993). Longer-range hopping in stacking direction as illustrated in Fig. 4.12 b only leads to a modified constant prefactor. For this general case, an expression for $\sigma_{xx}(\omega)$ is found from (4.65) in conjunction with (4.75).

As verified by Uhrig and Vlaming (1993) within their locator formalism, the resulting sum rule for NN hopping in current direction is indeed of the form characteristic for the hypercubic lattice [with $\sigma_0 = 2\pi e^2 N/(V\hbar^2)$ as defined in (4.67)],

$$\int\limits_0^\infty \mathrm{d}\omega\,\sigma_{xx}(\omega) = -\frac{\sigma_0 t^2 a^2}{4}\langle \hat{T}_x\rangle\,. \tag{4.126}$$

This observation is, however, not very useful in the DMFT context since the applicability of (4.125) depended on the fact that $\langle \hat{T}_x\rangle$ is negligible with respect to the full kinetic energy. Thus, (4.126) alone gives no relation between the frequency sum and any expectation value surviving the DMFT limit. Using instead the general DMFT $f$-sum rule (4.87), we obtain for the general stacked case,

$$\int\limits_0^\infty \mathrm{d}\omega\,\sigma_{xx}(\omega) = \frac{\sigma_0}{4}\langle\frac{\frac{\mathrm{d}}{\mathrm{d}\epsilon}\rho(\epsilon)}{\rho(\epsilon)}\rangle\int\limits_0^{2\pi/a_x}\mathrm{d}k_x v_{k_x}^2\,, \tag{4.127}$$

where the integral equals $t^2a^2$ for NN hopping in current direction. Thus, we have not only identified a form of the $f$-sum rule which is generally applicable for anisotropic lattices in high dimensions, but also found an expression for the expectation value of the kinetic energy in the stacking direction. For the particular case of the stacked Bethe lattice with NN hopping, the sum rule reads [cf. (4.97)]

$$\int\limits_0^\infty \mathrm{d}\omega\,\sigma_{xx}(\omega) = \frac{t^2a^2\sigma_0}{4}\langle\frac{-\epsilon}{4-\epsilon^2}\rangle\,. \tag{4.128}$$

### 4.4.4 Offdiagonal Disorder

In the DMFT limit, diagrammatic expressions for directional transport are always dominated by diagrams with minimal extent in current direction. Specifically, for lattices where all lattice sites are arranged in planes perpendicular to the current and where the interplane contribution to the kinetic energy has zero weight for $Z \to \infty$, the surviving diagrams are confined to two neighboring planes with only two connections, the current vertices, between the planes as illustrated in Fig. 4.13: here

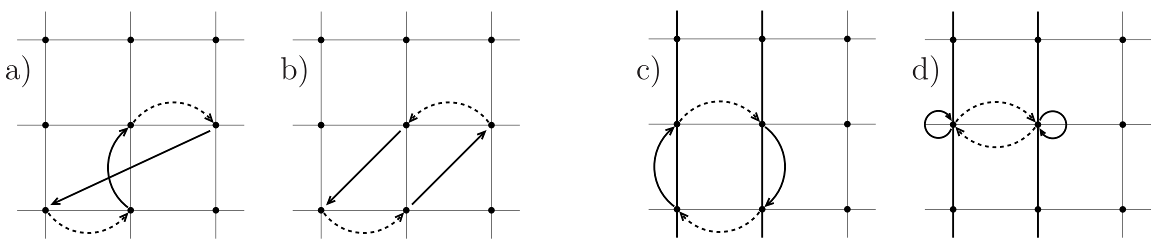



**Figure 4.13:** Diagrammatic contributions to the current-current correlation function (cf. Fig. 4.8). Here, the vertical axis symbolizes all directions perpendicular to the current direction (horizontal axis). In finite dimensions, contributions arise for **a)** parallel and **b)** antiparallel current (dashed arrows) with arbitrary offset. In $d \to \infty$, however, only the antiparallel contribution survives with no offset in current direction, **c)**+**d)**. The connecting Green functions (solid arrows) are, then, restricted to two neighboring planes.

the horizontal axis is the current direction whereas the vertical axis represents all other (perpendicular) directions. Diagrams as shown in Fig. 4.13 a and Fig. 4.13 b which include sites at more than 2 planes are suppressed in high dimensions where only diagrams like shown in Fig. 4.13 c and Fig. 4.13 d survive. The reason is that each hopping in current direction contributes a factor that scales to zero in the DMFT limit (and is proportional to $1/\sqrt{Z}$ in the isotropic case) while all other directions are inequivalent by definition so that these factors cannot cancel from a sum over equivalent diagrams. As a consequence, the time evolution paths (solid lines) between the current vertices are each restricted to a separate hyperplane and cannot interfere which implies that the full correlation function can be expressed in terms of a pair of Green functions (Metzner et al., 1992). In more formal terms, when we label the involved sites as $A_1$ and $A_2$ for one (hyper)plane and $B_1$, $B_2$ for the second plane then the expectation value of the time-dependent current-current correlation function decomposes in the following way:

$$\begin{aligned}
\langle \mathrm{e}^{-i\hat{H}t} c^\dagger_{A_1} c_{B_1} \mathrm{e}^{i\hat{H}t} c^\dagger_{B_2} c_{A_2} \rangle &= \langle \mathrm{e}^{-i\hat{H}t} c_{B_1} \mathrm{e}^{i\hat{H}t} c^\dagger_{B_2}\; c_{A_2} \mathrm{e}^{-i\hat{H}t} c^\dagger_{A_1} \mathrm{e}^{i\hat{H}t} \rangle && (4.129)\\
&= \langle \mathrm{e}^{-i\hat{H}t} c_{B_1} \mathrm{e}^{i\hat{H}t} c^\dagger_{B_2} \rangle\; \langle c_{A_2} \mathrm{e}^{-i\hat{H}t} c^\dagger_{A_1} \mathrm{e}^{i\hat{H}t} \rangle && (4.130)\\
&= G_{B_1B_2}(t) G_{A_2A_1}(-t)\,. && (4.131)
\end{aligned}$$

Here, we have used the invariance of the implicit traces under ring permutations and the time invariance of the Hamiltonian. Thus, this real-space argument implies a re-derivation of the bubble formula (4.62) for $\sigma(\omega)$ in $d \to \infty$. In contrast to the $\boldsymbol{k}$-space formulation presented in Sec. 4.3 for regular lattices, the real-space formulation applies to arbitrary (pseudo-) lattices and even in the presence of disorder.

Full diagonal disorder is realized within the DMFT when all hopping matrix elements are chosen independently from a random distribution with $\langle t_{ij} \rangle = 0$ (cf. subsection 2.2.2). Then, only diagrams which are self-retracing in real space have finite expectation values. Taken together with the restriction to diagrams living on two neighboring planes, this means that the current-current correlation function becomes fully local, i.e., reduces to contributions of the form $\langle\langle \hat{c}^\dagger_{\boldsymbol{r}} \hat{c}_{\boldsymbol{r}+\boldsymbol{e}_x}, \hat{c}^\dagger_{\boldsymbol{r}+\boldsymbol{e}_x} \hat{c}_{\boldsymbol{r}} \rangle\rangle$ .

Consequently, the optical conductivity can be expressed in terms of local spectral functions (Dobrosavljević and Kotliar, 1993),

$$\sigma_{xx}(\omega) = \frac{\pi e^2 a^2 t^2}{\hbar^2} \frac{N}{V} \sum_{\sigma} \int_{-\infty}^{\infty} \mathrm{d}\omega' A_{A\sigma}(\omega') A_{B\sigma}(\omega' + \omega) \frac{n_{\mathrm{f}}(\omega') - n_{\mathrm{f}}(\omega' + \omega)}{\omega}, \quad (4.132)$$

where $t^2 = \langle t^2_{\boldsymbol{r},\boldsymbol{r}+\boldsymbol{e}_x} \rangle$. Note that this expression essentially agrees with (4.120), but does not contain the factor $L_x/L$ which scales to zero in the thermodynamic limit. We also observe that (4.132) is valid as long as the hopping terms in current direction are fully disordered in the sense defined above regardless if the bonds perpendicular to the current (i.e., within the hyperplanes) are ordered or disordered. This result can, therefore, be applied in connection with any underlying lattice and is, then, characteristic of the stacked disordered anisotropic model.

### 4.4.5 General Dispersion Method

In the preceding subsections, we have seen that all previous approaches of defining and computing an optical conductivity compatible with DMFT calculations for a semi-elliptic DOS had an inherently anisotropic character. Using the general formalism developed in Sec. 2.3 and applied to the Bethe DOS in Sec. 2.4, we can here for the first time consider the isotropic case. We remind the reader that in the general dispersion formalism hopping matrix elements only depend on taxi-cab distance $D$ which, for $d \to \infty$, leads to the dispersion $\epsilon_{\boldsymbol{k}} = \sum_{D=1}^{\infty} \mathrm{He}_D(\epsilon^{\mathrm{hc}}_{\boldsymbol{k}})\, t^*_D/\sqrt{D!}$, where $\mathrm{He}_D(x)$ is the Hermite polynomial of order $D$ and $\epsilon^{\mathrm{hc}}_{\boldsymbol{k}}$ is the dispersion of the hypercubic lattice with NN hopping [cf. (2.52)]. Due to the completeness of the Hermite polynomials, an arbitrary transformation function $f$ with $\epsilon_{\boldsymbol{k}} = f(\epsilon^{\mathrm{hc}}_{\boldsymbol{k}})$ can be realized by an appropriate choice of scaled hopping matrix elements $t^*_D$. In particular, any bounded function $f$ leads to corresponding band edges in the resulting DOS. Using (2.61), a suitable (monotonous) transformation function $f$ may be determined for an arbitrary "target" DOS which then allows for the computation of transport properties, e.g., via (2.63). Furthermore, the hopping matrix elements defining the corresponding microscopic model can be computed according to (2.55).

For the "redefined Bethe lattice", a regular isotropic hc lattice with extended hopping along the diagonals and a semi-elliptic DOS, we have from (2.70),

$$\tilde{\rho}(\epsilon) = \frac{1}{2\sqrt{1-\epsilon^2/4}} \exp\left[-2\left(\mathrm{erf}^{-1}\left(\frac{\epsilon\sqrt{1-\epsilon^2/4} + 2\sin^{-1}(\epsilon/2)}{\pi}\right)\right)^2\right] \quad (4.133)$$

which defines $\sigma_{xx}(\omega)$ in conjunction with the general expression (4.65) and is shown in Fig. 4.14. Also shown in the figure is the function $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ which determines the sum rule (4.95). It is a straightforward but lengthy exercise to derive the explicit analytic expression from the chain rule. Let us, instead, only point out that it approaches $-\pi\epsilon/2$ at $\epsilon \approx 0$, but clearly deviates from this asymptotic form near the band edges. Thus, the $f$-sum cannot be proportional to the kinetic energy in any limit. The

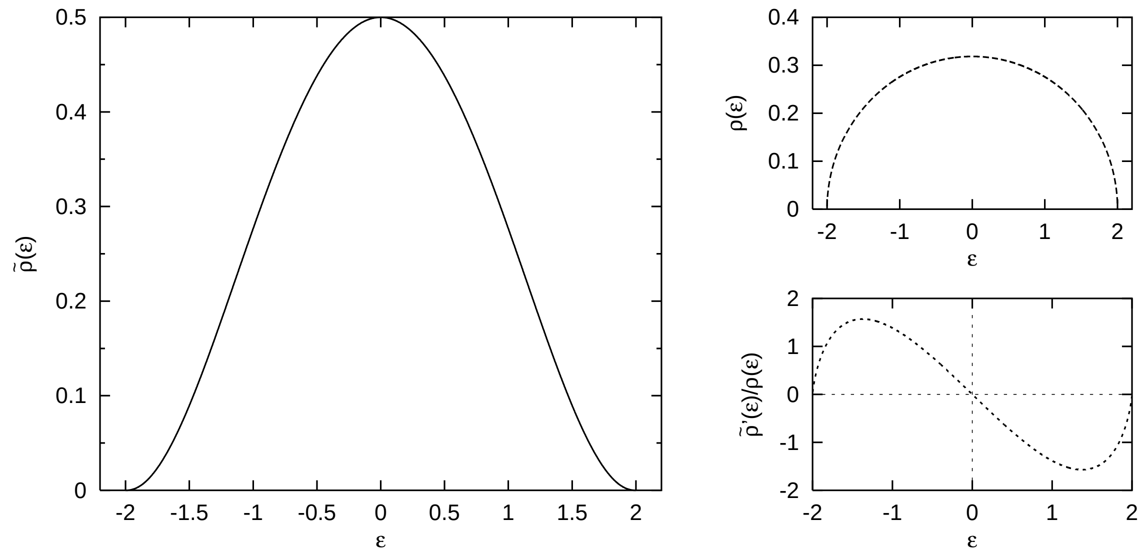



**Figure 4.14:** DMFT transport characteristics of redefined Bethe lattice in $d = \infty$: for this isotropic, regular lattice, the optical conductivity is well defined by (4.65) using the input $\tilde{\rho}(\epsilon)$ (left) while its semi-elliptic DOS (top right) guarantees local single-particle properties to be identical to that of the Bethe lattice. The $f$-sum (4.87) can be computed from measurements of the generalized momentum distribution function $n_\epsilon$ using the input $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ (bottom right). At small frequencies, the latter function deviates from the hc form $-\epsilon$ only quantitatively by a factor $\pi/2$ while the decay at the band edges marks a very important qualitative difference.

only similarity one can expect to the hc case is that, for small $U$ and $T$, changes in the frequency sum will be proportional to changes in the kinetic energy. Let us emphasize that the information collected in Fig. 4.14 gives a complete description of the lattice within the DMFT with respect to local and transport properties. There are several possible choices for the transport input $\tilde{\rho}(\epsilon)$ consistent with the same DOS (which by itself completely determines local properties within the DMFT). The particular choice (4.133) depicted here is the most natural since it is consistent with isotropic hopping matrix elements and since the noninteracting energy increases monotonously from the center of the Brillouin zone. Furthermore, the scaled hopping matrix elements decay very fast for this choice, cf. Table 2.2. Thus, we have found a minimal extension of the formalism developed for the hypercubic lattice which is consistent with a semi-elliptic "Bethe" DOS.

As discussed in chapter 2, one should check whenever possible whether results obtained for $d = \infty$ survive as a reasonable approximation in finite (and not too high) dimensions. Within the framework of the general dispersion method and, thus, for the model at hand, finite dimensionality $d$ implies a finite maximal hopping distance $D_{\max} \leq d$; the selection of hopping matrix elements even relied on $D_{\max} \ll d$. We have already demonstrated in subsection 2.4.2 and subsection 2.4.3 that the effects of a truncation of the hopping range and of finite dimensionality on the DOS, i.e.,

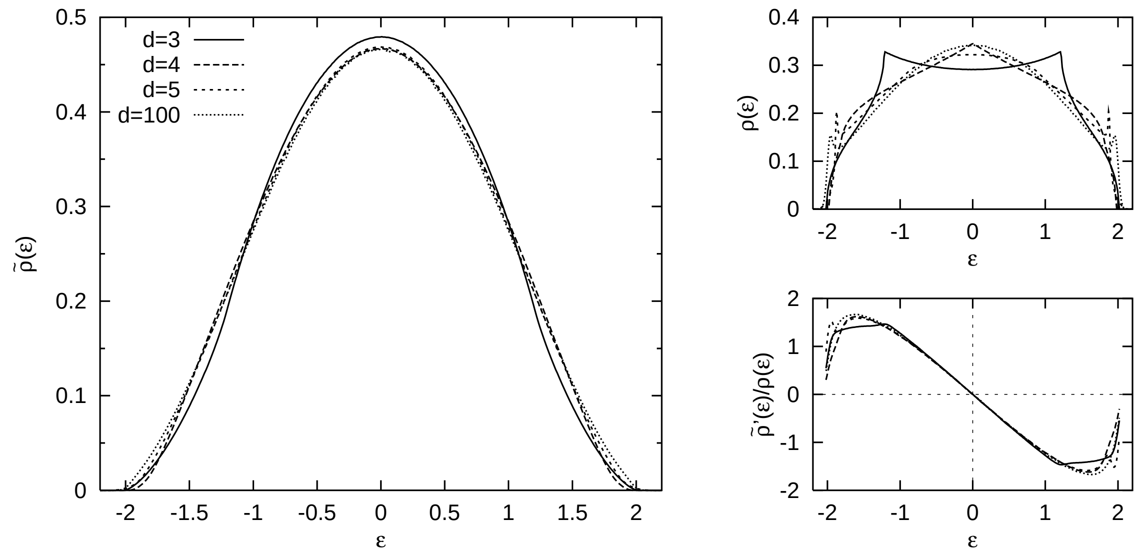



**Figure 4.15:** DMFT transport characteristics of the hc lattice with 3rd NN hopping (analogous to Fig. 4.14). $\tilde{\rho}(\epsilon)$ changes much less with dimension $d$ that the DOS, in particular for $d \geq 4$.

---

on local properties, are small even for $d$, $D_{\max} \lesssim 10$. Thus, the model relies much less on extremely high dimensionality than, e.g., the $d = \infty$ fcc lattice. We will now extend the discussion to transport related properties.

In Fig. 4.15, the analogue of Fig. 4.14 is shown for the redefined Bethe lattice with the hopping range truncated at $D_{\max} = 3$ and for a range of dimensions $3 \leq d \leq 100$. Evidently, $\tilde{\rho}(\epsilon)$, the lattice input for $\sigma(\omega)$, is hardly affected by a change of dimension, in particular for $d > 3$. The differences are even much smaller for $\tilde{\rho}(\epsilon)$ than for the DOS $\rho(\epsilon)$ since Van-Hove singularities in the DOS for some energy $\epsilon^*$ are cancelled by a corresponding dip in the average squared Fermi velocity $\langle |\boldsymbol{v_k}|^2 \rangle(\epsilon^*)$. As we will see below, this behavior is very general so that results for $\sigma(\omega)$ of a local theory in finite dimensions will depend on $d$ predominantly via $A(\omega)$, the interacting DOS, and only very little via $\tilde{\rho}(\epsilon)$. From comparing the main panels of Fig. 4.14 and Fig. 4.15, we also see that the impact of the truncation to $D_{\max} = 3$, i.e., only two nonvanishing scaled hopping matrix elements, is typically much smaller than 10%. As seen in the lower right part of Fig. 4.15, the function $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ entering the $f$-sum rule (4.95) converges fast for $d > 3$; the spikes near the band edges are due to numerical errors. With the inclusion of 5th-nearest-neighbor hopping, effects of finite dimensionality (for $d \geq 5$) on $\tilde{\rho}(\epsilon)$ become almost invisible as seen in Fig. 4.16. Also the effect of truncation is further reduced.

In conclusion, the results of this subsection demonstrate that the transport properties derived for the redefined Bethe lattice within the general dispersion formalism for $d = \infty$ are not specific to very high dimensions or to infinite hopping range. Instead, a local theory of transport, i.e., usage of a local self-energy and of the bubble approximation, applied in finite dimensions for models with similar hopping elements

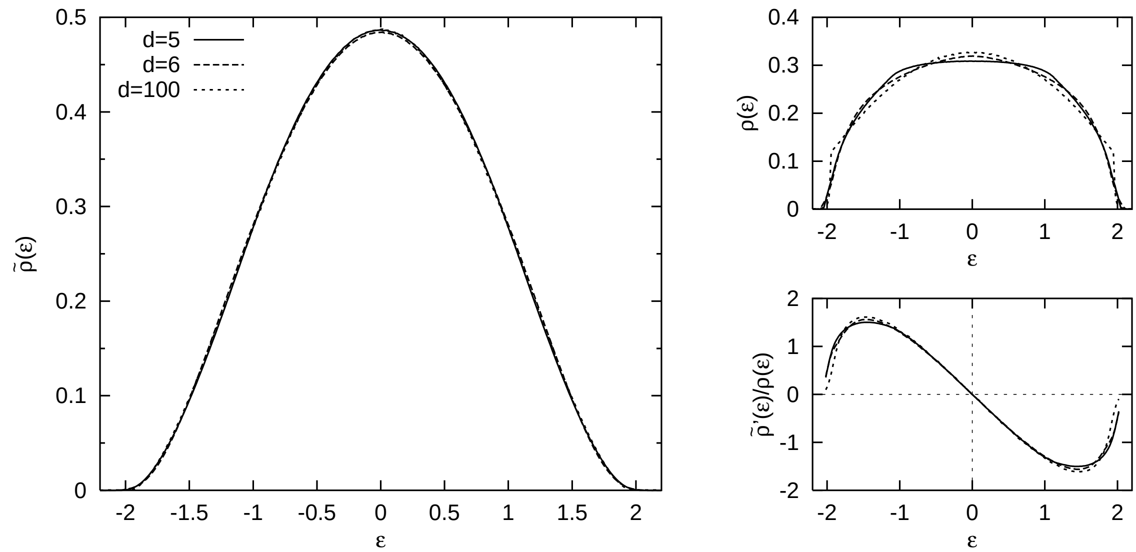



**Figure 4.16:** DMFT transport characteristics of hc lattice with 5$^{\text{th}}$ NN hopping (analogous to Fig. 4.14). $\tilde{\rho}(\epsilon)$ is practically independent of dimension.

leads to similar expressions for $\sigma(\omega)$ and would, thus, also lead to similar results. In this sense, the final numerical results which we will present in Sec. 4.6 are not specific to the limit $d = \infty$, but apply at least qualitatively also for a local theory of conductivity in finite $d$, as long as the noninteracting DOS is symmetric and roughly semi-elliptic.

## 4.5 Generalizations

In the previous section, we have focused on the question of how to define and compute $\sigma(\omega)$ in a way consistent with DMFT calculations for a semi-elliptic noninteracting DOS and within the paramagnetic phase. Before showing numerical results based on these studies in the next subsection, we use this section to present generalizing concepts. In subsection 4.5.1, we discuss various cases of stacked lattices which lead to various degrees of coherence in the transport along the stacking direction. Properties entering a local theory of $\sigma(\omega)$ in finite dimensions are collected for various lattice types in subsection 4.5.2. Finally, the effect of frustration on transport is studied analytically using the $t - t'$ lattice, i.e., a hc lattice with NN and NNN hopping in subsection 4.5.3.

### 4.5.1 Coherent versus Incoherent Transport in High Dimensions

As discussed in subsection 4.4.4, diagrammatic contributions to the optical conductivity along a symmetry direction generically only involve two neighboring hyperplanes.

Therefore, the Green functions appearing in the bubble formula for $\sigma(\omega)$ are local in the current direction. Let us for the moment specialize in the general stacked case where the hyperplanes are connected by nearest-neighbor bonds in current direction which is assumed to be perpendicular to the hyperplanes. If we consider the homogeneous case and denote the real space Green function connecting site $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ as $G_{\boldsymbol{r}_2-\boldsymbol{r}_1}(\omega)$, the resulting expression contributing to the conductivity is of the form

$$\sum_{\boldsymbol{r}_\perp} G_{(0,\boldsymbol{r}_\perp)}(\omega)\; G_{(0,-\boldsymbol{r}_\perp)}(\omega')\,. \tag{4.134}$$

The important point is that the Green functions are local in current direction; $\boldsymbol{r}_\perp$ is the perpendicular component of $\boldsymbol{r}$. If the hyperplanes are regular lattices, this translates in Fourier space into a contribution

$$\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'} G_{\boldsymbol{k}}(\omega)\, G_{\boldsymbol{k}'}(\omega')\, \delta(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp)\,. \tag{4.135}$$

This simple form applies to the hypercubic lattice and all "full" stacked lattices, where each lattice site is connected to neighboring sites by bonds in current direction. Compared with the finite-dimensional case where the bubble summation would have to include diagrams as shown in Fig. 4.13 c and which would give rise to the full $\delta$-function $\delta(\boldsymbol{k}-\boldsymbol{k}')$, the restriction of the $\delta$-function to perpendicular directions might already be viewed as a reduction of coherence. Obviously the fully incoherent case is realized when there is only a single chain of links in current direction as discussed in subsection 4.4.2 or if the hopping in current direction is fully disordered as discussed in subsection 4.4.4: in both cases, the $\delta$-function is replaced by unity.

The main point of this subsection is that the degree of coherence in $\boldsymbol{k}_\perp$ may be tuned for an anisotropic model by attaching bonds in current direction only to a sublattice of sites within each hyperplane (with the extreme cases where sublattice and full lattice within the hyperplane are identical or when only a single site per plane is connected). For example, when every second site (i.e., any site with coordinates $\boldsymbol{r}$ for which $(-1)^{||\boldsymbol{r}_\perp||} = 1$) is connected, the transport is coherent up to the AF wave vector which leads to a contribution[43]

$$\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'} G_{\boldsymbol{k}}(\omega)\, G_{\boldsymbol{k}'}(\omega')\, \frac{1}{2}\big(\delta(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp) + \delta(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp - \boldsymbol{Q}_\perp)\big)\,. \tag{4.136}$$

Finally, for a hc lattice (with arbitrary hopping range in the perpendicular direction) where every $n_\alpha^{\text{th}}$ site in direction $\alpha$ is part of the "current sublattice", the $\boldsymbol{k}$ summation reads

$$\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'} G_{\boldsymbol{k}}(\omega)\, G_{\boldsymbol{k}'}(\omega') \prod_{\alpha=2}^{d} \left( \frac{1}{n_\alpha} \sum_{l=0}^{n_\alpha-1} \delta(k_\alpha - k'_\alpha - \frac{l}{n_\alpha}\frac{2\pi}{a_\alpha}) \right)\,. \tag{4.137}$$

Here, $a_\alpha$ is the lattice spacing in direction $\alpha$. The completely incoherent case is recovered for $n_\alpha \to \infty$ when all $\delta$-functions approach unity.

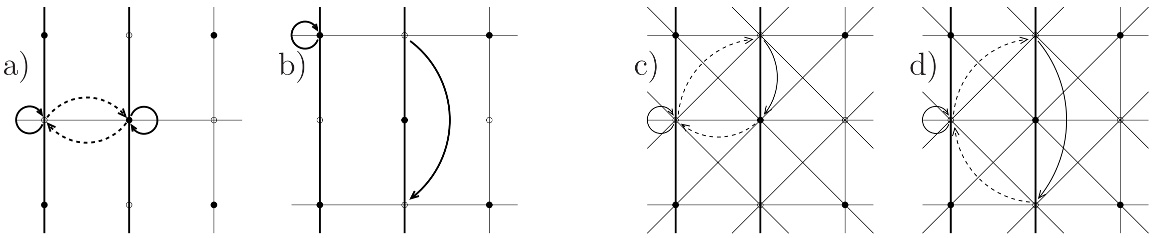



**Figure 4.17: a)** For stacked lattices with only one current chain (here in horizontal direction), the current-current correlation function reduces to its (antiparallel) on-bond part; the Green functions become local. **b)** For stacked lattices with interplane connections only for every second site, the connecting Green functions (solid arrows) are diagonal in sublattice index $\alpha \in \{A, B\}$. **c)+d)** Generalized stacked lattices with additional hopping along diagonals (see text).

A complementary view is taken in Fig. 4.17. While the completely incoherent case with a single link in current direction leads to diagrams involving only local Green functions as shown in Fig. 4.17 a, the case where current bonds touch every second site implies diagrams where all Green functions are local with respect to the sublattice index as shown in Fig. 4.17 b. Thus, the lack of coherence with respect to $\boldsymbol{Q}$ as demonstrated in (4.136) translates into a locality of the Green functions with respect to the sublattice index.

In the more general case that current may also flow along diagonals, the two Green functions do no longer only connect equivalent pairs of sites so that nontrivial factors appear in the momentum summation. For example, NNN bonds imply diagrams as shown in Fig. 4.17 c which lead to contributions of the form

$$2t_{\parallel}\, t'_{\parallel} \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}'} G_{\boldsymbol{k}}(\omega) G_{\boldsymbol{k}'}(\omega') \Big( \sum_{\alpha} \cos(k_\alpha) \Big) \delta(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp) \tag{4.138}$$

and diagrams as shown in Fig. 4.17 d with contributions[44]

$$4t'^2_{\parallel} \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}'} G_{\boldsymbol{k}}(\omega) G_{\boldsymbol{k}'}(\omega') \Big( \sum_{\alpha} \cos(k_\alpha) \Big) \Big( \sum_{\alpha'} \cos(k_{\alpha'}) \Big) \delta(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp)\,. \tag{4.139}$$

Here, $t_{\parallel}$ and $t'_{\parallel}$ denote hopping matrix elements which contribute to the current (and may differ from matrix elements for transport within the planes). Using these and analogous considerations it is easily possible to derive expressions for $\sigma(\omega)$ for arbitrary stacked versions of arbitrary underlying regular lattices. One can even carry out the momentum sums and, thus, arrive at expressions $\tilde{\rho}_{xx}(\epsilon)$ for any such combination. As one example, we remark that pure fcc-like hopping between planes as

[43]Note that this form immediately generalizes to the case of stacked AF planes (both with FM or AF order in current direction).

[44]It is easy to check that the special case $\alpha = \alpha'$ is properly accounted for in this equation. Indeed the terms $2\cos(2k_\alpha)$ (corresponding to the actual diagram shown in Fig. 4.17 d) and 2 (corresponding to parallel current vertices) sum up to $4\cos^2(k_\alpha)$.

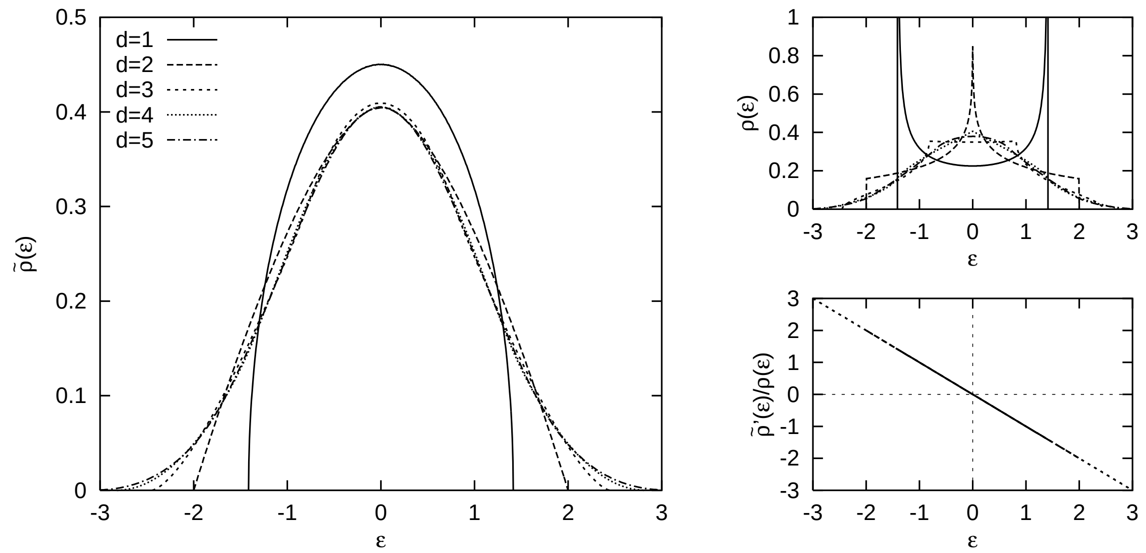



**Figure 4.18:** DMFT transport characteristics of the hypercubic (hc) lattice with NN hopping (cf. Fig. 4.14). The approach of the limiting Gaussian form (for $d \to \infty$) is much faster for $\tilde{\rho}(\epsilon)$ than for the DOS $\rho(\epsilon)$. The exact relation $\tilde{\rho}'(\epsilon)/\rho(\epsilon) = -\epsilon$ implies the proportionality of the $f$-sum to the kinetic energy.

described by (4.139) leads to $v_{\boldsymbol{k},x}^2 = 2dt_{\parallel}'^2 \epsilon_{\boldsymbol{k}}^{hc\,2}$. If the hyperplanes are, e.g., hypercubic, we get $\langle v_{\boldsymbol{k},x}^2 \rangle(\epsilon) = 2dt_{\parallel}'^2 \epsilon^2$. In contrast, if the hyperplanes are also of the fcc-type with $t_\perp' = t_\parallel' = \frac{t'^*}{\sqrt{2d(d-1)}}$, the resulting expression is $\langle v_{\boldsymbol{k},x}^2 \rangle(\epsilon) = 2\frac{t'^{*2}}{d}(1 + \frac{\sqrt{2}}{t'^*}\epsilon)$. In the latter case we find that the squared Fermi velocity vanishes at the band edge as always for isotropic systems.

## 4.5.2 Optical Conductivity in Finite Dimensions

In this subsection we continue the discussion of transport related lattice properties in finite dimensions for a series of commonly treated lattice types. With the exception of the dimension-independent analytic forms for $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ available for the hc and fcc cases, all curves have been obtained by Monte Carlo sampling of the appropriate $\boldsymbol{k}$-space expressions. Apart from possible uses of this data as input for computations of $\sigma(\omega)$ in a local theory in finite dimensions, the study is also useful as support for the results of the general-dispersion method which may be used as a black-box algorithm for generating $\tilde{\rho}(\epsilon)$ from the noninteracting DOS $\rho(\epsilon)$ alone.

Fig. 4.18 shows that $\tilde{\rho}(\epsilon)$ converges much faster to its Gaussian $d = \infty$ form than the DOS on the hc lattice. This is possible since van-Hove singularities appearing in the DOS are cancelled by corresponding dips in the squared Fermi velocity. As visualized in Fig. 4.19 a, the Fermi velocity becomes more and more effectively constant and approaches unity for $d \to \infty$. In Fig. 4.19 b corresponding curves are shown for the fcc lattice which approach a tilted straight line vanishing at the band edge for

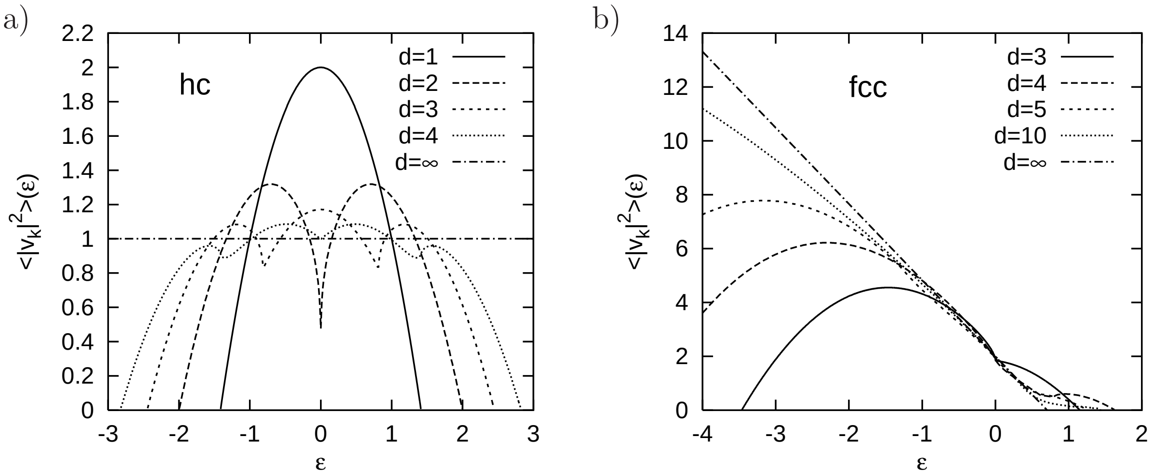



**Figure 4.19:** DMFT transport characteristics: $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon) = \tilde{\rho}(\epsilon)/\rho(\epsilon)$ **a)** for the hc lattice and **b)** for the fcc lattice. In the former case, the graph approaches unity for $d \to \infty$; in the latter, the asymptotic form is a straight line which vanishes at the band edge $\epsilon = 1/\sqrt{2}$.

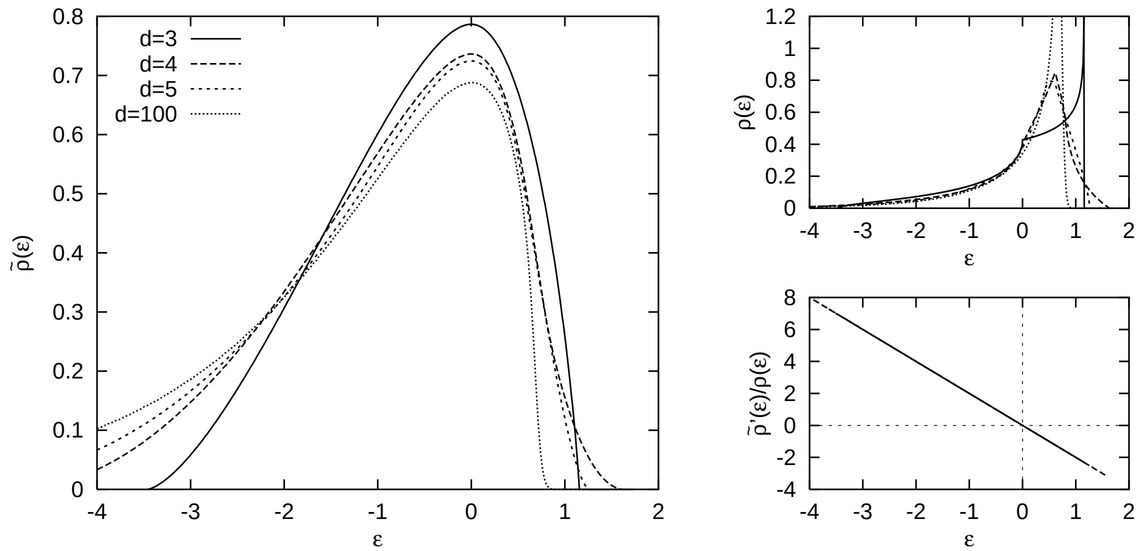



**Figure 4.20:** DMFT transport characteristics of fcc lattice with NN hopping (cf. Fig. 4.14). As in the hc case, the $f$-sum is proportional to the kinetic energy; here, $\tilde{\rho}'(\epsilon)/\rho(\epsilon) = -2\epsilon$.

---

$d \to \infty$. In particular, neither the finite-dimensional nor the limiting forms are constant. This fact which is clearly important for quantitative studies of the influence of frustration on transport properties was overlooked in a recent publication (Merino and McKenzie, 2000) as we will discuss in the next subsection. In Fig. 4.20, the $f$-sum rule is still seen to be proportional to the kinetic energy, but with an extra factor of 2 compared to the hc case. This proportionality is lost in the simultaneous presence of hopping terms to nearest and next-nearest neighbors as studied in Fig. 4.21. Here,

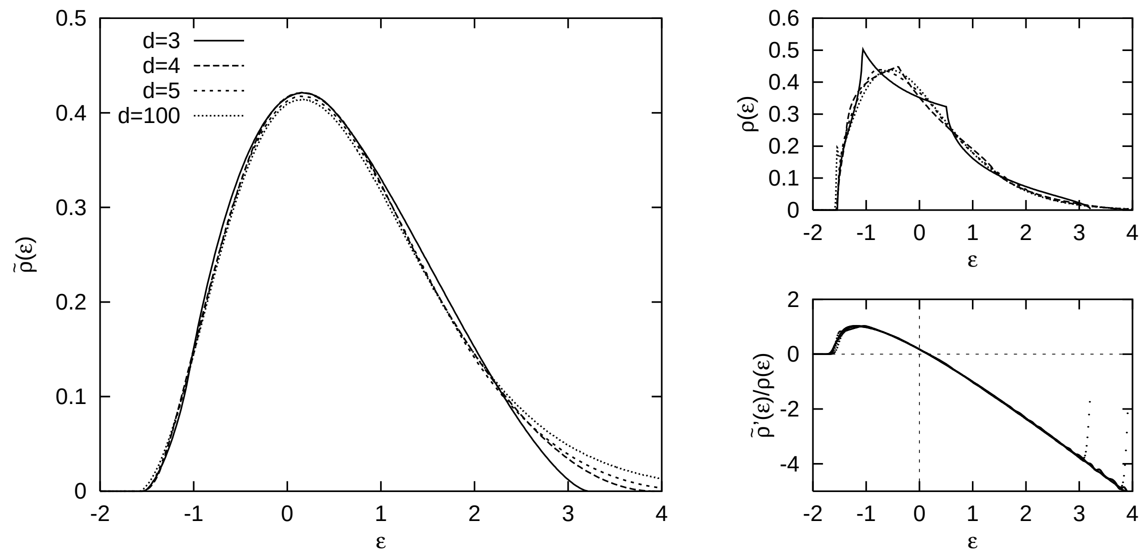



**Figure 4.21:** DMFT transport characteristics of the $t-t'$ hc lattice with NN and NNN hopping ($a^* = -0.25$). In contrast to Fig. 4.18 and Fig. 4.20, $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ is no longer proportional to the kinetic energy $\epsilon$.

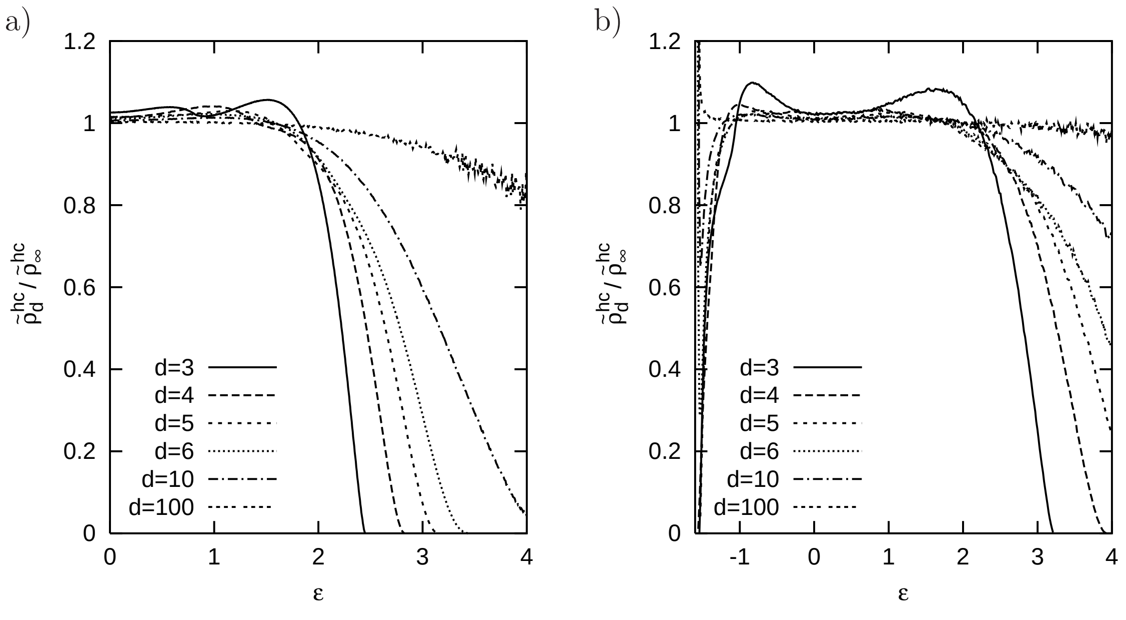



**Figure 4.22:** Dimensional dependence of DMFT transport characteristics of the hc lattice **a)** with only NN hopping and **b)** with both NN and NNN hopping ($a^* = -0.25$).

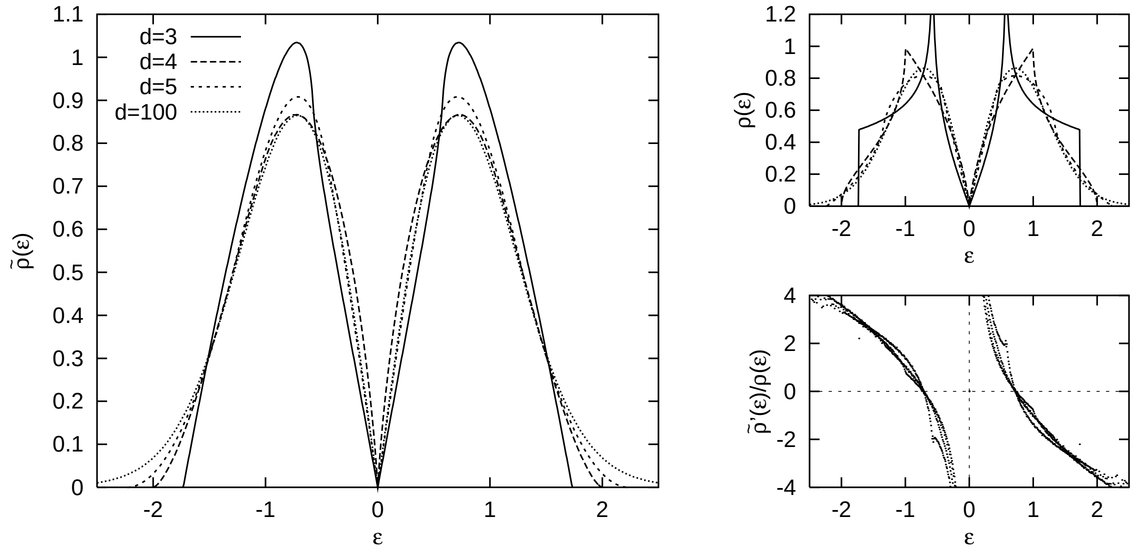



**Figure 4.23:** DMFT transport characteristics of the hyperdiamond lattice.

$\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ acquires a nontrivial shape the limiting form of which will be derived in the next subsection. Again, the deviations of $\tilde{\rho}(\epsilon)$ and of $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ in finite dimensions from the asymptotic form (for $d \to \infty$) are seen to be small. This fact is further illustrated for $\tilde{\rho}(\epsilon)$ in Fig. 4.22 a and Fig. 4.22 b, where the $d$-dimensional results have been divided by the asymptotic results, for the hc lattice with only NN hopping and for the t-t' lattice with $a^* = -0.25$, respectively.

As a final example, Fig. 4.23 shows results for the hyperdiamond lattice which is a semimetal at half filling (cf. App. B). At the pseudogap, the function $\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ entering the $f$-sum rule is not even continuous.

### 4.5.3 Impact of Frustration by $t - t'$ Hopping

Throughout large parts of this thesis we have studied models with a symmetric DOS. While this is a convenient special case it is certainly not generic. As discussed in subsection 2.1.3, true magnetic frustration which is an essential microscopic ingredient for models with nontrivial low-temperature phases is generically associated with an asymmetry of the DOS. Furthermore, for extensions of the hc lattice, it is associated with even-range hopping (rather than only odd-range hopping). Thus, at least in finite dimensions, the hc lattice with nearest neighbor (NN) and next-nearest neighbor (NNN) hopping is the prototype lattice for studying magnetic frustration.

Our interest for performing actual calculations was triggered by a recent DMFT paper on the impact of frustration on the optical conductivity, on the dc Hall Resistance, and on the thermopower of correlated electron systems (Merino and McKenzie, 2000). In this work, the expressions valid for the transport properties of the hc lattice were used also in the case of finite frustration $a^* = 0.1$ and $a^* = 0.3$.[45] We

[45] In the publication, the authors explicitly point out that this simplification is used "in order

will here derive the correct expressions, i.e., perform the necessary momentum sums and compare the results to an evaluation of the hc expressions.

In the low-frequency limit of the homogeneous phase, the transverse conductivity reduces to (Shastry, Shraiman, and Singh, 1993; Merino and McKenzie, 2000)[46]

$$\sigma_{xy} = \sigma_0^{\mathrm{H}} \int_{-\infty}^{\infty} \mathrm{d}\omega \, \frac{\partial n_f(\omega)}{\partial \omega} \frac{1}{N} \sum_{\boldsymbol{k}} \Big(\frac{\partial \epsilon_{\boldsymbol{k}}}{\partial k_x}\Big)^2 \frac{\partial^2 \epsilon_{\boldsymbol{k}}}{\partial k_y^2} A^3_{\epsilon_{\boldsymbol{k}}}(\omega)\,. \tag{4.140}$$

Here, $\sigma_0^{\mathrm{H}} = 4\pi^2|e|^3 aB/3\hbar^2$ and $B$ is the magnetic field in $z$ direction. The spectral function $A_{\epsilon_{\boldsymbol{k}}}(\omega)$ depends on momentum $\boldsymbol{k}$ only via the free dispersion $\epsilon_{\boldsymbol{k}}$. The dc Hall coefficient

$$R_{\mathrm{H}} = \frac{\sigma_{xy}}{B\sigma_{xx}^2(\omega=0)} \tag{4.141}$$

follows from (4.140) and the limit $\omega \to 0$ of (4.65). The thermopower is given by the expression (Schweitzer and Czycholl, 1991; Pruschke, Jarrell, and Freericks, 1995; Merino and McKenzie, 2000)

$$S = -\frac{k_B}{|e|T} \frac{L_{12}}{L_{11}} \tag{4.142}$$

where the transport integrals reduce for $d \to \infty$ to[47]

$$L_{ij} = -\int_{-\infty}^{\infty} \mathrm{d}\omega \, \frac{\partial n_f(\omega)}{\partial \omega} \Big[\frac{2}{N} \sum_{\boldsymbol{k}} \big(\frac{\partial \epsilon_{\boldsymbol{k}}}{\partial k_x}\big)^2 A^2_{\epsilon_{\boldsymbol{k}}}(\omega)\Big]^i \omega^{j-1} \tag{4.143}$$

In order to perform the sums over momenta, we will use the fact that the dispersion of the $t-t'$ lattice may be expressed in terms of the dispersion $\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}$ of the hypercubic lattice with unit scaled hopping amplitude $t^* = 1$,

$$\epsilon(\boldsymbol{k}, t^*, a^*) \equiv \epsilon_{\boldsymbol{k}} = t^* \epsilon_{\boldsymbol{k}}^{\mathrm{hc}} + \frac{a^* t^*}{\sqrt{2}} \Big(1 - \big(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}}\big)^2\Big) \equiv f(\epsilon_{\boldsymbol{k}}^{\mathrm{hc}})\,. \tag{4.144}$$

Consequently, we may write

$$v_{\boldsymbol{k},x} \equiv \frac{\partial \epsilon_{\boldsymbol{k}}}{\partial k_x} = t^* \big(1 - \sqrt{2}\, a^* \epsilon_{\boldsymbol{k}}^{\mathrm{hc}}\big) \frac{\partial \epsilon_{\boldsymbol{k}}^{\mathrm{hc}}}{\partial k_x} \tag{4.145}$$

where $\partial \epsilon_{\boldsymbol{k}}^{\mathrm{hc}}/\partial k_x = \sqrt{\frac{2}{d}} \sin(k_x)$. Averaging over $k_x$, this yields for the squared Fermi velocity

$$\begin{aligned} \langle v_{\boldsymbol{k},x}^2 \rangle(\epsilon) &= \frac{t^{*2}}{d} \Big(1 - \sqrt{2} a^* \tilde{\epsilon}\Big)^2 \Big|_{\epsilon = f(\tilde{\epsilon})} && (4.146) \\ &= \frac{1 + 2a^{*2} - 2\sqrt{2} a^* \sqrt{1 + a^{*2}\epsilon}}{1 + a^{*2}}\,. && (4.147) \end{aligned}$$

---

to avoid the cumbersome sums over momentum". They claim it to be justified by their focus "on many-body effects and not on how different band structures may change the results slightly".

[46]In finite dimensions, the second derivative in $x$ direction is supplemented by a term $-\big(\frac{\partial \epsilon_{\boldsymbol{k}}}{\partial k_y}\big)^2 \frac{\partial^2 \epsilon_{\boldsymbol{k}}}{\partial k_x \partial k_y}$ (which carries an additional factor $1/d$ and, therefore, vanishes for $d \to \infty$).

[47]Note that this expression is only approximate since it does not take contributions due to the flow of doubly occupied sites into account (Georges et al., 1996).

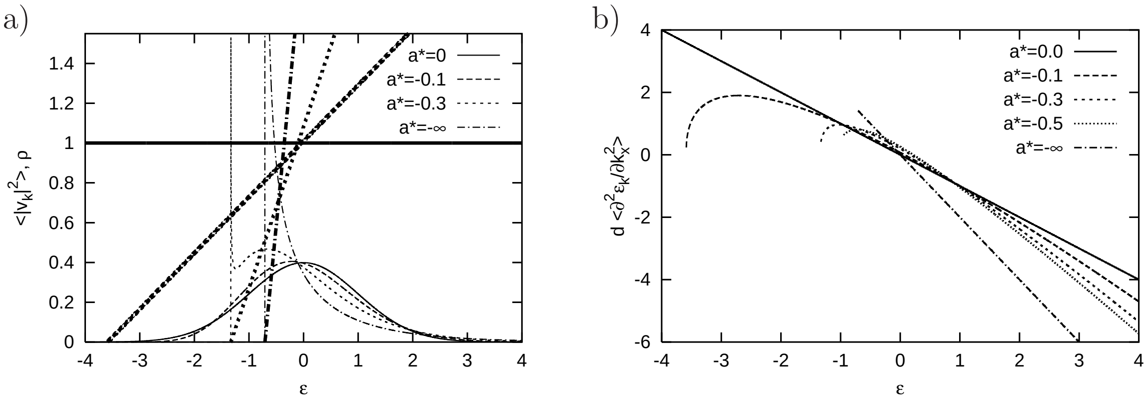



**Figure 4.24:** DMFT transport characteristics of the $t-t'$ lattice: **a)** For finite frustration, the squared Fermi velocity $\langle|v_{\boldsymbol{k}}|^2\rangle(\epsilon)$ (thick lines) is not constant, but vanishes linearly at the band edge of the noninteracting DOS (thin lines). **b)** $d\langle\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k_x^2}\rangle(\epsilon)=\tilde{\rho}'(\epsilon)/\rho(\epsilon)$ needed for computing the Hall conductivity and the $f$-sum. Deviations from the linear hc form ($a^*=0$) are most significant near the band edge.

---

In (4.147) we have already set $t^*=\sqrt{1/(1+a^{*2})}$ in order to guarantee unit variance of the corresponding DOS. Note that the inverse transformation function $\mathcal{F}^{-1}$ is not unique for the $t-t'$ lattice; this didn't matter here because both roots corresponding to the same $\epsilon_{\boldsymbol{k}}$ also yield the same $v^2_{\boldsymbol{k},x}$. As expected on physical grounds, the Fermi velocity vanishes at the band edge. In fact, $\langle v^2_{\boldsymbol{k},x}\rangle(\epsilon)$ is proportional the distance from the band edge as illustrated in Fig. 4.24 (thick lines). Evidently, the assumption of $\langle v^2_{\boldsymbol{k},x}\rangle(\epsilon)$ being constant (as implied by the formulas derived for the hc lattice) is bad already for $|a^*|=0.1$ and fails completely for stronger frustration $|a^*|=0.3$ and beyond. In particular, the dependence of transport properties on the degree of frustration may even be dominated by the change in $\langle v^2_{\boldsymbol{k},x}\rangle(\epsilon)$.

The evaluation of the second derivative is a bit more complicated since here both roots of $\mathcal{F}(\tilde{\epsilon})=\epsilon$ contribute differently. Therefore, we split the total DOS of the $t-t'$ lattice

$$\rho(\epsilon,a^*)=\begin{cases}\sqrt{\frac{2}{\pi}}\frac{1+a^{*2}}{r(\epsilon,a^*)}\exp\left(-\frac{1+r^2(\epsilon,a^*)}{4a^{*2}}\right)\cosh\left(\frac{r(\epsilon,a^*)}{2a^{*2}}\right) & \text{for }\operatorname{sign}(a^*)\epsilon<r_0\\ 0 & \text{otherwise}\,,\end{cases}\tag{4.148}$$

where $r(\epsilon,a^*)=\sqrt{1+2a^*(a^*-\sqrt{2}\epsilon\sqrt{1+a^{*2}})}$ and where the band edge is given by $r_0=(a^*+1/(2a^*))/\sqrt{2}$, into two contributions (with the same boundaries):

$$\rho^{\pm}(\epsilon,a^*)=\sqrt{\frac{1}{2\pi}}\frac{1+a^{*2}}{r(\epsilon,a^*)}\exp\left(-\frac{\big(1\pm r(\epsilon,a^*)\big)^2}{4a^{*2}}\right).\tag{4.149}$$

Using the notation

$$h_0(\epsilon,a^*)=\frac{1}{2a^{*2}}-\sqrt{2}(a^*+\frac{1}{a^*});\quad h^{\pm}(\epsilon,a^*)=\frac{\sqrt{2}a^*h_0(\epsilon,a^*)\pm\sqrt{1+h_0(\epsilon,a^*)}}{1+a^{*2}}\tag{4.150}$$

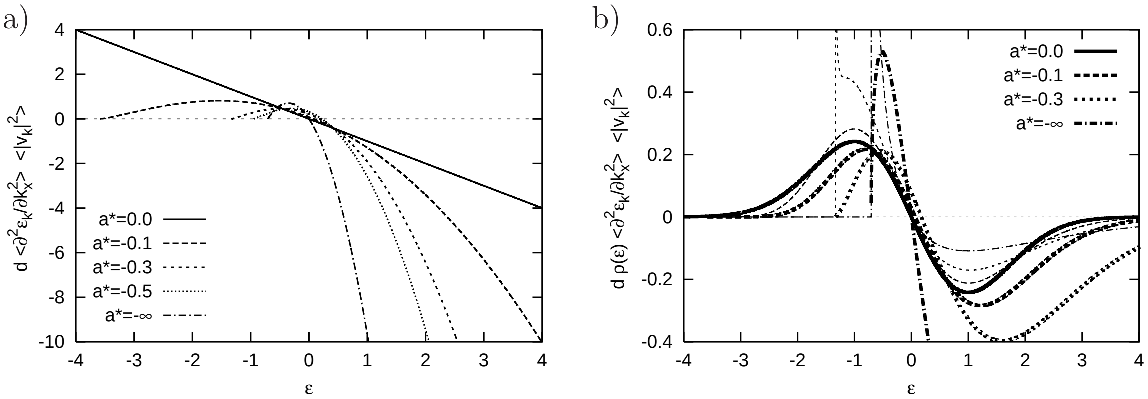



**Figure 4.25:** DMFT transport characteristics of the $t-t'$ lattice: Lattice-dependent terms entering the expression (4.140) for the transverse conductivity. **a)** Intermediate result: $d\langle\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k_x^2}\rangle(\epsilon) = \tilde{\rho}'(\epsilon)/\rho(\epsilon)$ multiplied by $\langle|v_{\boldsymbol{k}}|^2\rangle(\epsilon) = \tilde{\rho}(\epsilon)/\rho(\epsilon)$. **b)** Multiplication by the DOS $\rho(\epsilon)$ yields the correct lattice dependence (thick lines). Thin lines: corresponding results of an (unjustified) application of expressions valid for the pure hc lattice as used by Merino and McKenzie (2000). The limit $a^* = -\infty$ is equivalent to the fcc lattice.

the effective second derivative can be derived to read (for $a^* \leq 0$)

$$d\langle\frac{\partial^2\epsilon_{\boldsymbol{k}}}{\partial k_x^2}\rangle(\epsilon, a^*) = \frac{h^+(\epsilon, a^*)\rho^+(\epsilon, a^*) + h^-(\epsilon, a^*)\rho^-(\epsilon, a^*)}{\rho(\epsilon, a^*)}\,. \tag{4.151}$$

Evaluations of these expressions for moderate frustration are shown in Fig. 4.25 a. Note the strong deviations compared to the hc result (straight solid line). Figure 4.25 b depicts expectation values multiplied by the noninteracting DOS as they enter the computation of the transverse conductivity. Clearly, the correct curves (thick lines) bear little resemblance to an evaluation of the hc results for the $t - t'$ DOS (thin lines). Consequently, the study by Merino and McKenzie (2000) should be redone using the results presented in this subsection.

## 4.6 QMC Results for the Bethe Lattice

In this section, we discuss important practical aspects for numerical DMFT calculations of the optical conductivity $\sigma(\omega)$ and present results for the half-filled one-band Hubbard model with semi-elliptic DOS near the metal-insulator transition. Building upon the framework developed in Sec. 4.4, we do not only present conductivity data for the stacked Bethe lattice considered previously in the literature, but also for the new isotropic model with longer-range hopping and for the fully disordered model. While it is clear a priori that all results corresponding to the same local spectra share some common features, the differences will be seen to be significant, in particular in the metallic phase. Properties like the Hall conductance and the thermopower, which we discussed in the previous section, need not be computed since they vanish identically for the particle-hole symmetric systems under consideration.

### 4.6.1 Numerical Procedure for QMC Data

Within the DMFT framework, the optical conductivity may be expressed in the coherent case[48] as [see (4.64) - (4.67)]

$$\mathrm{Re}\,\sigma_{xx}(\omega) = \sigma_0 \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \tilde{\rho}_{xx}(\epsilon) \int_{-\infty}^{\infty} \mathrm{d}\omega'\, A_\epsilon(\omega') A_\epsilon(\omega'+\omega)\, \frac{n_\mathrm{f}(\omega') - n_\mathrm{f}(\omega+\omega')}{\omega}\,, \qquad (4.152)$$

where $\sigma_0 = \frac{2\pi e^2}{\hbar^2}\frac{N}{V}$ and $n_\mathrm{f}(\omega) = \big(1 + \mathrm{e}^{\beta(\omega-\mu)}\big)^{-1}$. Evidently, practical calculations require two basic ingredients: $\tilde{\rho}_{xx}(\epsilon) = \frac{1}{N}\sum_{\boldsymbol{k}} v^2_{\boldsymbol{k}_x}\delta(\epsilon - \epsilon_{\boldsymbol{k}})$ measures the noninteracting density of states weighted with the squared Fermi velocity and thus contains all necessary information about the lattice. For a given model (characterized by its lattice topology and hopping parameters), this function can be determined once and for all, independent of temperature and interaction. For lattices with semi-elliptic Bethe DOS, this task was solved in Sec. 4.4. The remaining problem is the computation of the "momentum-dependent" (i.e., dependent on the noninteracting dispersion $\epsilon$) interacting spectral function

$$A_\epsilon(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\, G_\epsilon(\omega)\,, \qquad (4.153)$$

where

$$G_\epsilon(\omega) = \frac{1}{\omega - \epsilon - \Sigma(\omega)}\,. \qquad (4.154)$$

An evaluation for real frequencies $\omega$ yields

$$A_\epsilon(\omega) = -\frac{1}{\pi}\,\frac{\mathrm{Im}\,\Sigma(\omega)}{\big(\omega - \epsilon - \mathrm{Re}\,\Sigma(\omega)\big)^2 + \big(\mathrm{Im}\,\Sigma(\omega)\big)^2}\,. \qquad (4.155)$$

Once $\Sigma(\omega)$ is known on the real axis, the computation of $\sigma(\omega)$ reduces to the evaluation of a two-dimensional integral (over the variables $\epsilon$ and $\omega'$). Note that the "momentum"-dependent spectral functions $A_\epsilon(\omega)$ are strongly peaked in the metallic phase (where $|\mathrm{Im}\,\Sigma(\omega)|$ is small for $\omega \approx 0$; cf. Fig. 4.29) which can in principle cause numerical instabilities; still, this step is straightforward.

In the context of QMC calculations, the real-frequency self-energy $\Sigma(\omega)$ has to be extracted in a multi-step procedure[49] which is summarized in Fig. 4.26 and will be discussed in the remainder of this subsection. All steps will be illustrated on the basis of the spectral function of the Hubbard model with semi-elliptic DOS at $T = 0.05$ and $U = 4.0$ shown as solid line in Fig. 3.77 in Sec. 3.8 (resulting from QMC using $\Delta\tau = 0.1$ and MEM using a flat default model). The corresponding imaginary part of the real-frequency Green function is indicated by circles in Fig. 4.27. It will become apparent later that the frequency grid $\Delta\omega = 0.1$ used in the MEM procedure

[48] For the cases of disorder or a single-link topology which imply incoherent transport, see subsection 4.4.2 and the results presented in subsection 4.6.3.

[49] The QMC method itself and the MEM determination of the local spectral function $A(\omega) = -\frac{1}{\pi}\mathrm{Im}\, G(\omega)$ have already been discussed in Sec. 1.3, Sec. 1.4, and Sec. 3.8.

$$A(\omega) = -\frac{1}{\pi}\mathrm{Im}\, G(\omega)$$

Kramers–Kronig $\Big\downarrow$ $\quad \mathrm{Re}\, G(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} \mathrm{d}\omega' \, \frac{\mathrm{Im}\, G(\omega')}{\omega - \omega'}$

$$G(\omega)$$

inverted Dyson Eq. $\Big\downarrow\Big\uparrow$ $\quad G(\omega) = \int_{-\infty}^{\infty} \mathrm{d}\epsilon \, \frac{\rho(\epsilon)}{\omega - \epsilon - \Sigma(\omega)}$

$$\Sigma(\omega), A_\epsilon(\omega)$$

numerical integration $\Big\downarrow$ $\quad \longleftarrow \quad \tilde{\rho}(\epsilon)$

$$\sigma(\omega)$$

**Figure 4.26:** Scheme for the computation of $\sigma(\omega)$ on the basis of the local spectral function $A(\omega)$ (cf. Sec. 1.4 and Sec. 3.8), of $\rho(\epsilon)$, and of $\tilde{\rho}(\omega)$. A Kramers-Kronig transformation produces the full Green function from which the self-energy $\Sigma(\omega)$ can be obtained via an inversion of the momentum integrated lattice Dyson equation. Finally, the numerical evaluation of the integrals in (4.152) yields $\sigma(\omega)$.

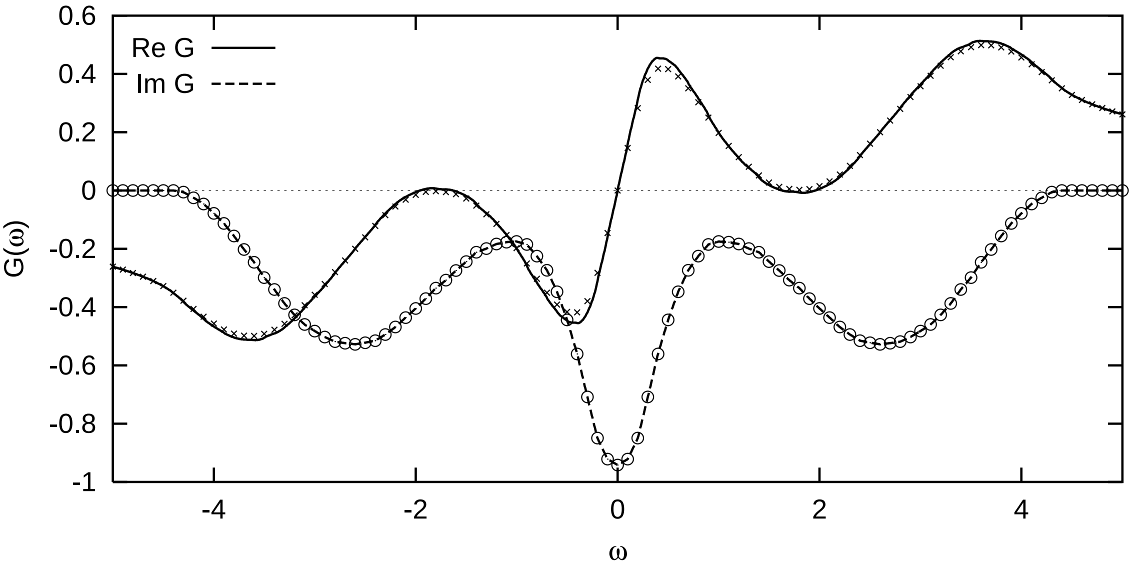



**Figure 4.27:** QMC results (for $\Delta\tau = 0.1$) at half filling, semi-elliptic DOS, $T = 0.05$, and $U = 4.0$: Imaginary part of the real-frequency Green function $\mathrm{Im}\, G(\omega)$ (circles) as obtained from MEM on a grid with $\Delta\omega = 0.1$. This function is interpolated on a fine grid with $\Delta\omega = 0.005$ (dashed line) in order to obtain $\mathrm{Re}\, G(\omega)$ (solid line) via Kramers-Kronig transformation. Small crosses indicate the corresponding result without interpolation.

is too coarse for the numerical integrations which will be performed according to the lowest-order approximation

$$\int_{x_0}^{x_n} dx\, f(x) \approx \Delta x \sum_{l=0}^{n-1} \frac{f(x_l) + f(x_{l+1})}{2}\,, \tag{4.156}$$

where $x_i = i\Delta x$. Therefore, $\mathrm{Im}\, G$ is first interpolated[50] and evaluated on a finer grid with the resolution $\Delta\omega = 0.005$ (dashed line in Fig. 4.27). The Kramers-Kronig transformation is then performed using (4.156) which yields $\mathrm{Re}\, G$ on the same fine grid (solid line). In order to demonstrate the importance of the interpolation procedure, the result of a direct Kramers-Kronig transformation is also shown in Fig. 4.27 (small crosses). The difference between both estimates for $G$ are relatively small, but clearly visible.

The next step, the numerical inversion of the $\boldsymbol{k}$ integrated Dyson equation (on the fine grid) falls into the class of complex root-finding problems (plus a trivial integration). Such problems are less trivial than real-valued one-dimensional root-finding problems since the solution cannot be bracketed. We have accomplished the solution using an adapted two-dimensional (in the Gaussian plane) Newton scheme; its Jacobian is specified by

$$\frac{\partial \mathrm{Im}\, G}{\partial \mathrm{Im}\, \Sigma} = \frac{\partial \mathrm{Re}\, G}{\partial \mathrm{Re}\, \Sigma} = \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \rho(\epsilon) \frac{(\omega - \epsilon - \mathrm{Re}\, \Sigma)^2 - (\mathrm{Im}\, \Sigma)^2}{\left[(\omega - \epsilon - \mathrm{Re}\, \Sigma)^2 + (\mathrm{Im}\, \Sigma)^2\right]^2}\,, \tag{4.157}$$

where we have suppressed the $\omega$-dependence of $G$ and $\Sigma$ for notational convenience (since Dyson equation is local in $\omega$ anyway), and

$$\frac{\partial \mathrm{Im}\, G}{\partial \mathrm{Re}\, \Sigma} = -\frac{\partial \mathrm{Re}\, G}{\partial \mathrm{Im}\, \Sigma} = \int_{-\infty}^{\infty} \mathrm{d}\epsilon\, \rho(\epsilon) \frac{2\,(\omega - \epsilon - \mathrm{Re}\, \Sigma)\, \mathrm{Im}\, \Sigma}{\left[(\omega - \epsilon - \mathrm{Re}\, \Sigma)^2 + (\mathrm{Im}\, \Sigma)^2\right]^2}\,. \tag{4.158}$$

In our program, the iterative search can be stabilized by underrelaxation (typically using a factor of 0.5). Note that the existence of an analytical solution $\Sigma(\omega)$ (with $\mathrm{Im}\, \Sigma(\omega) < 0$) of the inverse Dyson equation is by no means guaranteed for an arbitrary spectral function $A(\omega)$. For example, the maximum value of the spectral function must not exceed the maximum of the noninteracting DOS in the homogeneous case. On the other hand, numerical errors can easily produce such unphysical spectra since the (unrestricted) MEM does not incorporate model specific physical information. Then, the root-finding algorithm will fail to converge at least for some isolated frequencies $\omega$. In our implementation of the algorithm, such isolated failures can be fixed by linear interpolation of the solutions at adjacent grid points. The fact that this was necessary only for the strongest interaction $U = 5.5$ is a further

[50] We have chosen linear interpolation not only for sake of simplicity, but also for robustness. While for the metallic case a cubic spline interpolation might yield a better approximation of the smooth spectrum, such a higher order scheme is dangerous in general since it may lead to ringing (oscillations) and to nonanalytic solutions (i.e., solutions of the wrong sign) near band edges.

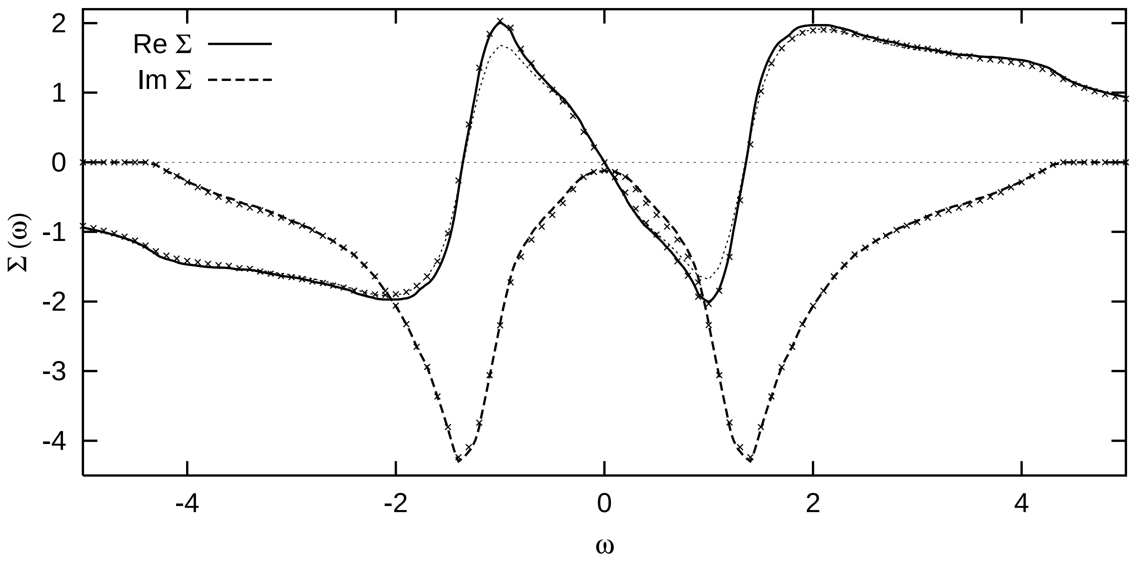



**Figure 4.28:** QMC results (for $\Delta\tau = 0.1$) at half filling, semi-elliptic DOS, $T = 0.05$, and $U = 4.0$: Real (solid line) and imaginary (dashed line) part of the real-frequency self-energy as derived from the corresponding curves for $G(\omega)$ shown in Fig. 4.27 by inversion of the Dyson equation. Results obtained using a coarse grid (crosses) are seen to violate analyticity: a Kramers-Kronig transform (dotted line) of the "coarse" result for $\operatorname{Im}\Sigma(\omega)$ deviates significantly from the direct "coarse" estimate for $\operatorname{Re} G(\omega)$.

indication of the high quality of the MEM spectra. For the particular case of the semi-elliptic Bethe DOS studied in this section, the root-finding effort could have been avoided (and was checked) by using the explicit analytic expression. The numerical approach, however, is completely general and will also be used in the LDA+DMFT context in chapter 5 (where it is unavoidable).

The resulting estimates for the self-energy (symmetrized and with a slightly coarser grid resolution $\Delta\omega = 0.025$) are shown as solid and dashed lines for $\operatorname{Re}\Sigma(\omega)$ and $\operatorname{Im}\Sigma(\omega)$, respectively, in Fig. 4.28. The corresponding results, computed on the coarse grid (crosses) are seen to deviate slightly from these curves. More importantly, the "coarse" results violate analyticity significantly: a Kramers-Kronig transformation of the "coarse" estimate for $\operatorname{Im}\Sigma(\omega)$ yields an estimate for $\operatorname{Re}\Sigma(\omega)$ (dotted line) which is clearly distinguishable from from the direct "coarse" estimate (and also from the "fine" result). Precise analyticity of $\Sigma(\omega)$, however, is absolutely crucial since otherwise spectral functions computed via the Dyson equation are not properly normalized. "Momentum" resolved spectral functions $A_\epsilon(\omega)$ are shown in the main panel of Fig. 4.29 for $U = 4.0$ and a range of values of $\epsilon \geq 0$. As seen in the inset, the normalization of $A_\epsilon(\omega)$ is correct within $10^{-3}$ when the fine grid is used, i.e., an interpolation is performed before applying the Kramers-Kronig transformation to $\operatorname{Im} G$. Using the nonanalytic self-energy obtained without interpolation, the deviations are about two orders of magnitude larger and would yield clearly visible effects on the level of the momentum distribution function or of the conductivity. Thus, the effort

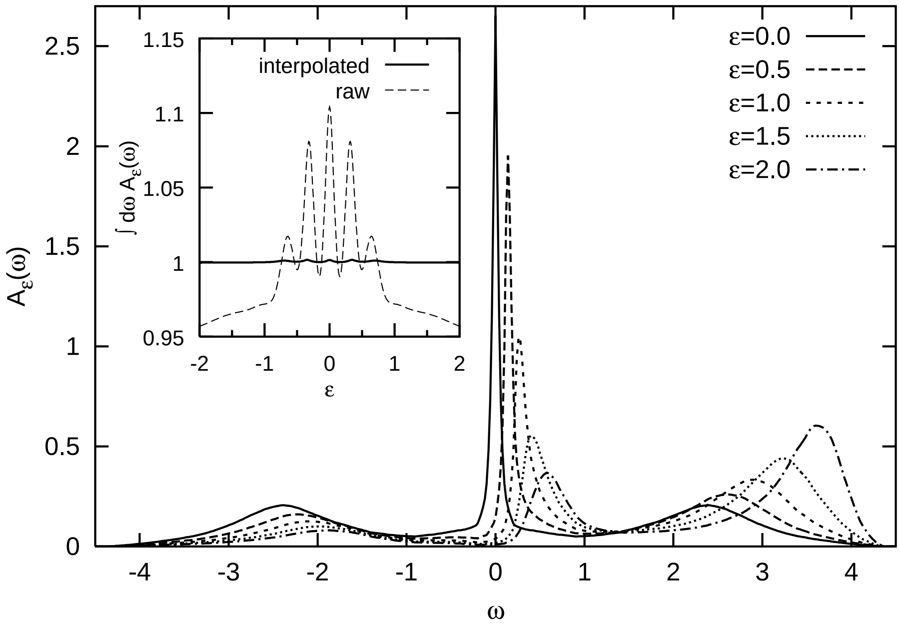



**Figure 4.29:** Momentum resolved spectral functions $A_\epsilon(\omega)$ for $T = 0.05$ and $U = 4.0$, based on $\Sigma(\omega)$ as presented in Fig. 4.28. $A_\epsilon(\omega)$ shows a three-peak structure for all $\epsilon$; the central resonance is sharpest for $\epsilon = 0$. The inset shows the numerical norm $\int_{-\infty}^{\infty} d\omega A_\epsilon(\omega)$ which should equal 1 for all $\epsilon$. The deviations from this sum rule do not exceed $10^{-3}$ when the fine grid is used; for a coarse grid, the deviations reach $10^{-1}$.

of ensuring fully analytic estimates for $\Sigma(\omega)$ is certainly warranted.[51]

## 4.6.2 Results: Self-Energy on the Real Axis

In this subsection, we present QMC estimates of the real-frequency self-energy of the half-filled Hubbard model with Bethe DOS for $T = 0.05$. All numerical results are based on the procedure outlined in the previous subsection and on the MEM spectra (corresponding to a flat default model with a frequency grid $\Delta\omega = 0.1$) presented in Fig. 3.77 and Fig. 3.79 in Sec. 3.8.

The self-energy $\Sigma(\omega)$ for the metallic phase is shown in Fig. 4.30. Qualitatively, the overall behavior meets the expectations for a strongly correlated Fermi liquid: $\mathrm{Re}\,\Sigma(\omega)$ is linear near the Fermi energy $\omega = 0$ with negative slope which becomes steeper upon approaching the transition. The imaginary part $\mathrm{Im}\,\Sigma(\omega)$ is nearly quadratic (as a function of $\omega$) in this region. While the curvature increases steadily with increasing $U$, the offset at $\omega = 0$ is initially almost constant before a relatively rapid increase at $U = 4.7$ signals the breakdown of the Fermi liquid. However, the strong features seen for $U = 4.7$ appear exaggerated: from continuity one would expect that some weight should be shifted from the peaks in $\mathrm{Im}\,\Sigma(\omega)$ to slightly larger

[51]In the following, slight remaining discrepancies in the Kramers-Kronig relations were eliminated by replacing $\mathrm{Re}\,\Sigma(\omega)$ by the Kramers-Kronig transform of $\mathrm{Im}\,\Sigma(\omega)$.

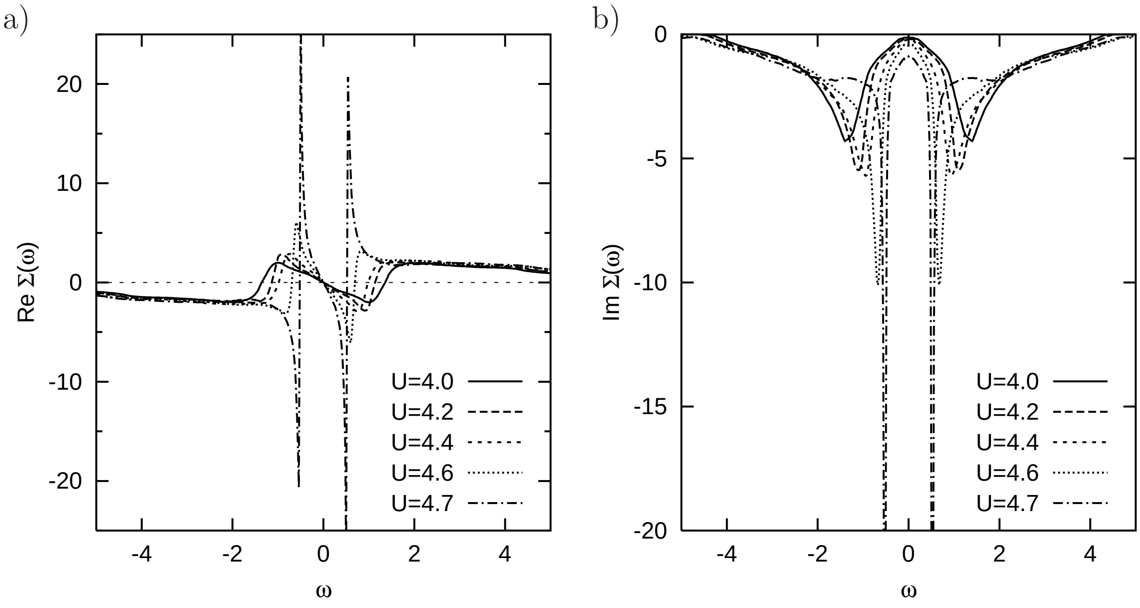



**Figure 4.30:** QMC results for $T = 0.05$: Self-energy in the metallic phase near the MIT. **a)** The width of the central linear region of $\mathrm{Re}\,\Sigma(\omega)$ decreases with increasing $U$ while the absolute slope near $\omega = 0$ increases. **b)** $\mathrm{Im}\,\Sigma(\omega)$ is nearly quadratic at small $|\omega|$. The very sharp features for $U = 4.7$ might be exaggerated due to numerical errors.

---

frequencies $|\omega|$. On the other hand, the result for $U = 4.7$ (as well as for all other values of the interaction) is robust numerically, i.e., is nearly invariant under the choice of a default model (within the MEM procedure).

Results for the insulating phase are displayed in Fig. 4.31. Here, $\mathrm{Im}\,\Sigma$ shows a single peak at $\omega = 0$ which narrows with increasing $U$. While this peak is part of a continuous spectrum [of $\Sigma(\omega)$] for $U \leq 5.0$, it appears $\delta$-function like and surrounded by a full gap for $U = 5.5$. Corresponding to the single-peak structure of $\mathrm{Im}\,\Sigma(\omega)$, the derivative of the real part $\mathrm{Re}\,\Sigma(\omega)$ at $\omega = 0$ is positive (and of large absolute value) which implies that the quasiparticle weight $Z$ as defined in (3.5) is negative.

Before turning to results for the optical conductivity, let us perform some additional checks of the results obtained so far. Since $\Sigma(\omega)$ is a retarded function which is analytic in the upper half plane, it can be expressed in terms of its positive semidefinite spectral function $A_\Sigma(\omega) := -\frac{1}{\pi}\mathrm{Im}\,\Sigma(\omega)$. The norm of this spectral function gives some measure of the total influence of the interaction on the single-particle properties. A priori, one would expect this norm to be a smooth function of $U$, except possibly for kinks or jumps at phase transitions. Thus, an evaluation of this norm can be useful for identifying outliers caused by numerical errors. This expectation is indeed correct as seen in Fig. 4.32, where circles indicate the numerical estimates based on the self energies shown in Fig. 4.30 and Fig. 4.31. In fact, all results are nearly on a straight line, even across the phase transition at $U \approx 4.7$. Under the assumption that this behavior is universal, i.e., that the true norm does not depend on the temperature and on the model DOS, the total self-energy spectral weight must be proportional

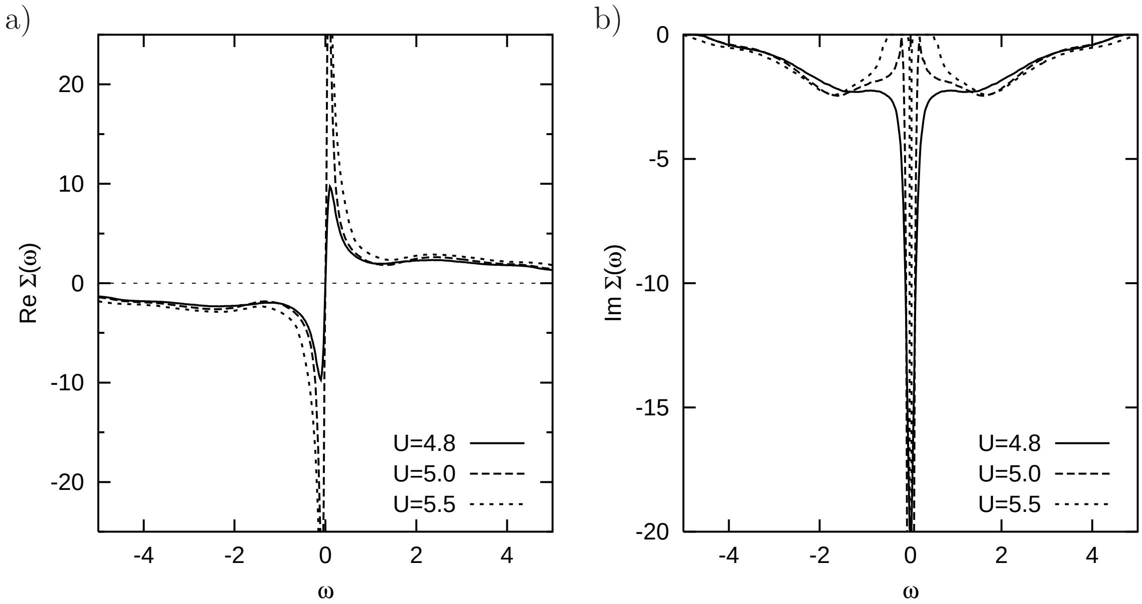



**Figure 4.31:** QMC results for $T = 0.05$: Self-energy in the insulating phase near the MIT. **a)** With increasing $U$, the real part of $\Sigma(\omega)$ approaches $1/\omega$ for not too small $|\omega|$. **b)** The imaginary part shows a sharp central peak already for $U = 4.8$ which evolves into a numerically almost singular delta function form inside a gap for $U = 5.5$.

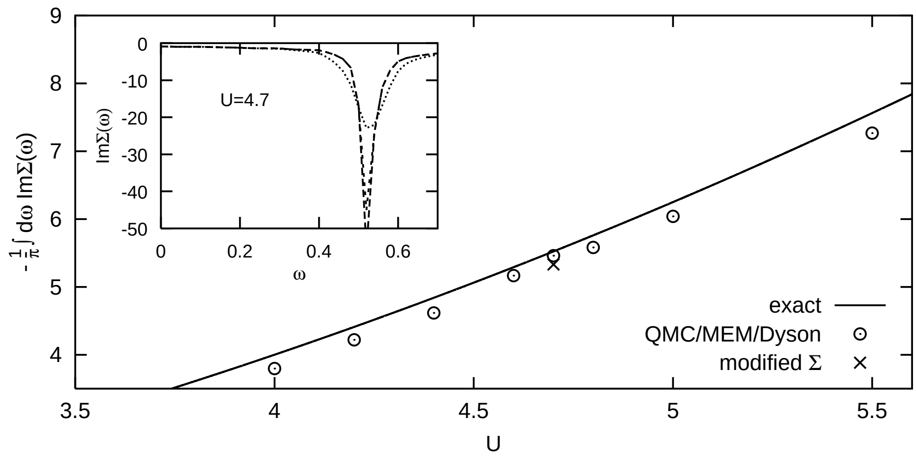



**Figure 4.32:** QMC results for $T = 0.05$: The numerical results (circles) fulfill the sum rule $-\frac{1}{\pi}\int_{-\infty}^{\infty} d\omega \, \mathrm{Im}\,\Sigma(\omega) = U^2/4$ for the self-energy within about 2%. The remaining discrepancy can be traced to using the finite discretization $\Delta\tau = 0.1$ in QMC which makes the solutions slightly more metallic. The inset gives a magnified view on $\mathrm{Im}\,\Sigma(\omega)$ for $U = 4.7$: original curve (long dashes) and modifications (short-dashed and dotted lines; see text) which reduce the frequency sum (cross in main panel).

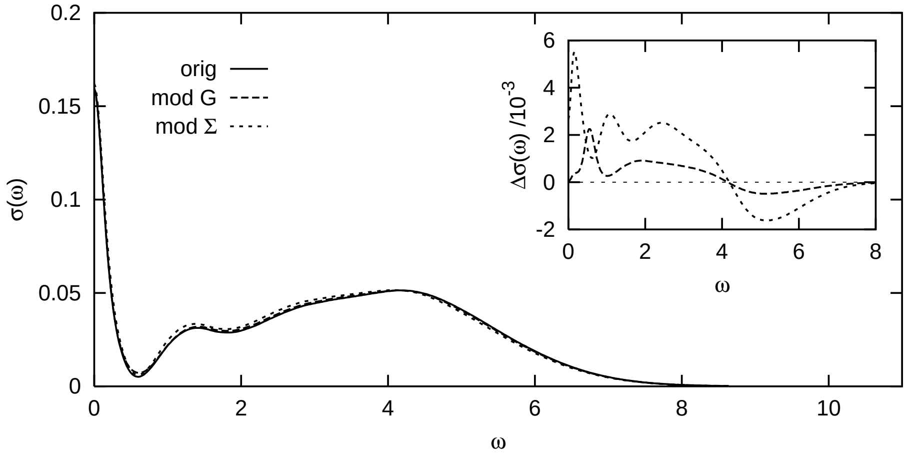



**Figure 4.33:** Optical conductivity for the isotropic "redefined Bethe lattice" for $T = 0.05$ and $U = 4.7$: The slight modifications in $\operatorname{Im}\Sigma(\omega)$ shown in the inset of Fig. 4.32 lead to almost indistinguishable results for $\sigma(\omega)$; the inset gives a magnified view of the differences.

to $U^2$ by dimensional analysis. This assumption is indeed correct as can be shown by evaluating the equations of motion of the impurity model (Potthoff, Wegner, and Nolting, 1997); in the case of interest, the norm reads $\int_{-\infty}^{\infty} d\omega A_\Sigma(\omega) = U^2/4$ (solid line in Fig. 4.32).[52] Considering the numerous potential sources of errors entering the sum rule (such as Trotter and statistical errors, incomplete convergence and the MEM procedure) the deviations of typically 2% of our numerical results from the exact values appear small; the good agreement lends further credibility to these results.

Furthermore, we can estimate if the total weight of the peaks in $\operatorname{Im}\Sigma(\omega)$ for $U = 4.7$ (dash-dotted lines in Fig. 4.30 b) is too large: Since the total numerical norm is relatively higher (though closer to the exact value) for $U = 4.7$ than for other values of $U$, one might suspect that the extra weight is due to a numerical error and should be corrected. In order to investigate this question and in particular the possible impact on the resulting estimate for $\sigma(\omega)$, we have generated two adjusted estimates for $\Sigma(\omega)$ shown in the inset of Fig. 4.32. In the first case (short-dashed line), the peak height in the original estimate for $\operatorname{Im}\Sigma(\omega)$ (long-dashed line) has been reduced so that the total is in line with the norms obtained for other values of $U$ (cross in the main panel). Alternatively, the MEM spectrum has been adjusted slightly so that it lead to a softer peak in $\operatorname{Im}\Sigma(\omega)$ (dotted line in the inset). The resulting three different estimates for $\Sigma(\omega)$, however, yield almost indistinguishable conductivities as depicted in Fig. 4.33. Thus, at least the most prominent uncertainty for $\Sigma(\omega)$ for $U = 4.7$ is qualitatively irrelevant for the results to be presented in the remainder of this section.

[52]This sum rule which (to the best of our knowledge) has not been mentioned in the DMFT-QMC community so far can also be directly applied on the imaginary axis in the context of QMC calculations. For its use in an improved Fourier scheme and its generalization to arbitrary filling and to the multi-band case, see App. C and Knecht (2002).

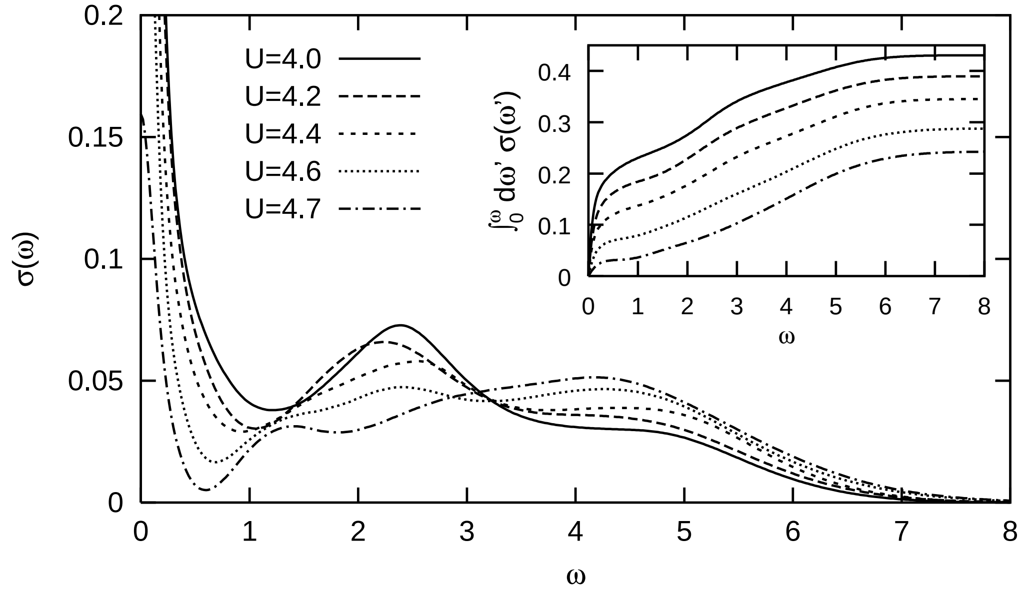



**Figure 4.34:** Optical conductivity $\sigma(\omega)$ for the isotropic "redefined Bethe lattice" at $T = 0.05$ in the metallic phase. With increasing $U$, weight is shifted from a Drude peak at $\omega \approx 0$ to a mid-infrared peak/shoulder at $\omega \approx U/2$ and, finally, to a broad incoherent peak at $\omega \approx U$. The inset shows the partial $f$-sum $\int_0^\omega d\omega' \sigma(\omega')$ which saturates for $\omega \approx 8$.

### 4.6.3 Results: Optical Conductivity

This subsection comprises the main numerical results of this chapter. Using the procedures detailed in the previous subsection, we will here present results in the paramagnetic low-temperature phase near the Mott metal-insulator transition. The optical conductivity data will in turn be shown for three different definitions, i.e., different models compatible with a semi-elliptic DOS: first, the isotropic "redefined Bethe lattice" specified in Sec. 2.4 (using the general dispersion approach developed in Sec. 2.3), then the (anisotropic) stacked Bethe lattice, and, finally, the incoherent definition which corresponds to a disordered model or to a single-chain layout of the Bethe lattice. The numerical results for the first two cases will be based on the self-energies $\Sigma(\omega)$ shown in Fig. 4.30 and Fig. 4.31. An estimate of the self-energy is not needed in the last, incoherent case; here, the optical conductivity will be computed directly from the MEM spectra $A(\omega)$ shown in Fig. 3.77 and Fig. 3.79. In all cases, the constants $e$, $\hbar$, and the lattice spacing $a$ are set to unity which corresponds to setting $\sigma_0 = 2\pi$ [for the coherent case; cf. (4.67)].

**Coherent Transport**

We start the presentation of optical conductivity data with the redefined Bethe lattice; for this model, $\tilde{\rho}(\omega)$ has the exact analytic expression (4.133). Figure 4.34 displays results for the metallic phase near the MIT. Each curve features a prominent "Drude" peak at low frequencies which is associated with excitations within

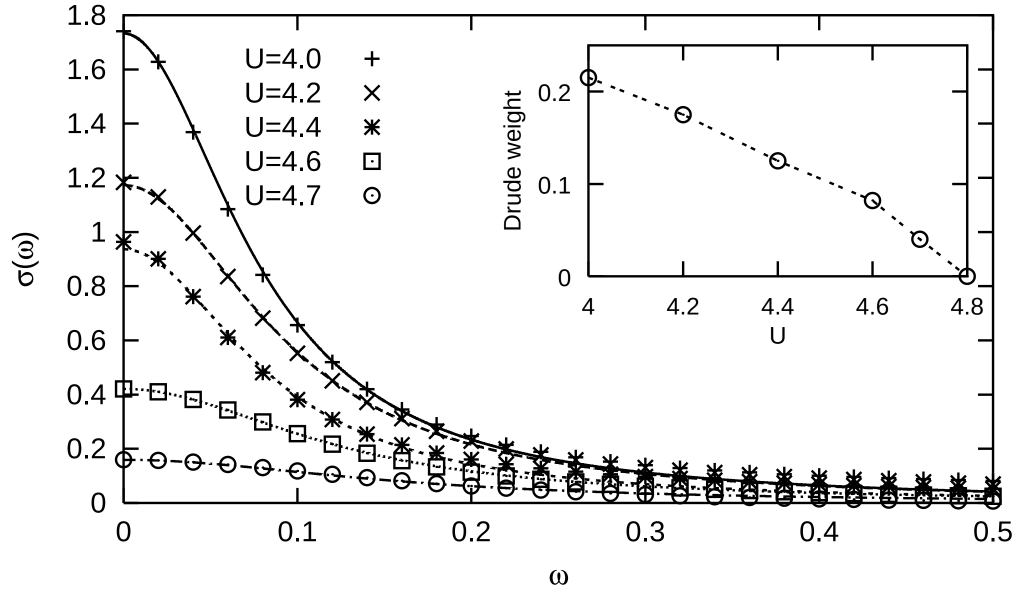



**Figure 4.35:** Drude part of the optical conductivity $\sigma(\omega)$ for the isotropic "redefined Bethe lattice" at $T = 0.05$ in the metallic phase. The numerical results (symbols) are fitted with Lorentzian curves (lines); their normalization, the Drude weight, decreases rapidly with increasing $U$ as seen in the inset, while their width increases only slightly.

---

the quasiparticle peaks (of the local spectra, cf. Fig. 3.77 and Fig. 3.79). Incoherent absorption bands, roughly in the range above $\omega \gtrsim 1$ can be subdivided into mid-infrared contributions at $1 \lesssim \omega \lesssim 3$ which are associated with excitations from the lower Hubbard band to the quasiparticle peak or from the quasiparticle peak to the upper Hubbard band while a contribution at $\omega \approx U$ is associated with excitations from the lower to the upper Hubbard band. Depending on the interaction strength $U$, this subdivision appears in the form of shoulders or genuine peaks. The inset of Fig. 4.34 shows the partial optical $f$-sum which allows for a direct comparison of the weights associated with the different contributions. Note the analogy of this representation to the experimental curves displayed in Fig. 4.3. As expected, both the contribution of the Drude peak and the total $f$-sum decrease for increasing $U$. Figure 4.35 concentrates on the coherent peaks seen at low frequencies. Here, the numerical results are indicated by symbols, while the lines represent Lorentzian fits. The good quality of the fits allows for a determination of the weight and the width of each Drude peak. The drude weights are shown as a function of $U$ in the inset of Fig. 4.35; its initially linear decay with increasing $U$ becomes more rapid near the MIT. Here, lines are guides to the eye only. The width of the Drude peak (not shown) is nearly constant (about 0.08) for $U \leq 4.4$ and increases to 0.16 for $U = 4.7$. Figure 4.36 shows $\sigma(\omega)$ and its $f$-sum for the insulating phase using the same scales as in Fig. 4.34. Just above the MIT, i.e., for $U = 4.8$, the conductivity spectrum is very broad with long tails both for small and large frequencies. Further away from the transition, weight shifts towards the central region, i.e., the width of the absorption

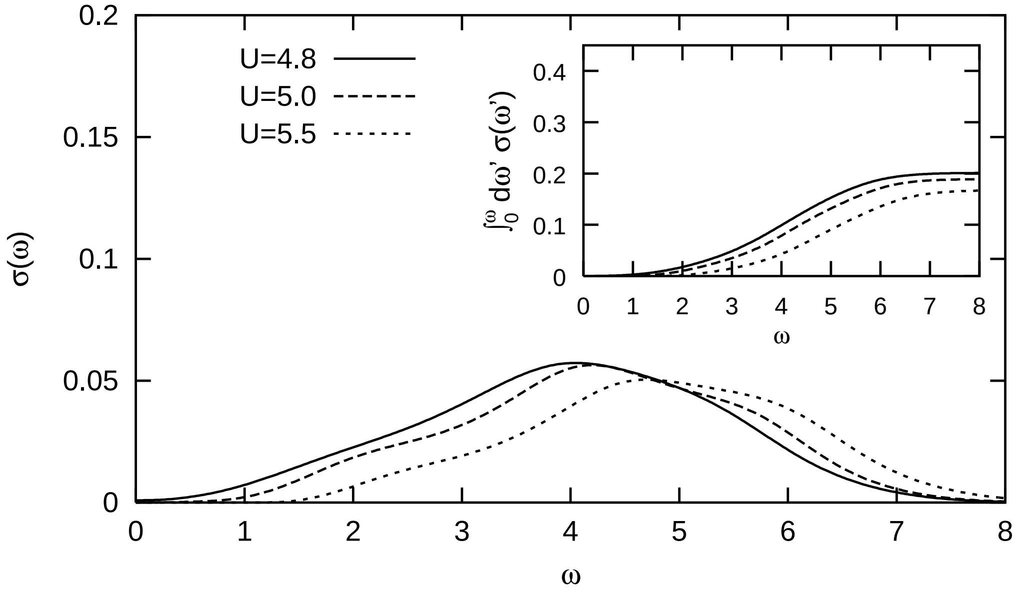



**Figure 4.36:** Optical conductivity $\sigma(\omega)$ for the isotropic "redefined Bethe lattice" at $T = 0.05$ in the insulating phase. The inset shows the partial $f$-sum (cf. Fig. 4.34).

band shrinks slightly (by about 15% from $U = 4.8$ to $U = 5.5$).

In the following, plots analogous to the last three figures will be shown (using the same scales) for the stacked Bethe lattice discussed in subsection 4.4.3. For notational simplicity, we will continue to write $\sigma(\omega)$ instead of using the notation $\sigma_{xx}(\omega)$ appropriate for an anisotropic model. We expect to find qualitatively very similar results as for the isotropic model considered before; still the differences will be seen to be significant. In fact, Fig. 4.37 and Fig. 4.38 demonstrate that within the metallic phase both the Drude weights and the total $f$-sums are reduced significantly compared with the isotropic case. The widths of the Drude peaks, however, are nearly identical. No significant changes are observed in the insulating phase in Fig. 4.39 up to a slight reduction in the total $f$-sum.

### Incoherent Transport

Completely incoherent transport follows both from Stumpf's single-chain view of the Bethe lattice detailed in subsection 4.4.2 and from offdiagonal disorder as discussed in subsection 4.4.4. We will here consider the disordered case which yields an extensive conductivity. The formal expression (4.132) reads for the paramagnetic phase and the choice of units specified above

$$\sigma_{xx}(\omega) = 2\pi \int_{-\infty}^{\infty} \mathrm{d}\omega'\, A(\omega')\, A(\omega'+\omega)\, \frac{n_{\mathrm{f}}(\omega') - n_{\mathrm{f}}(\omega'+\omega)}{\omega}\,. \tag{4.159}$$

The primary new feature seen in Fig. 4.40 and Fig. 4.41 for the metallic phase is

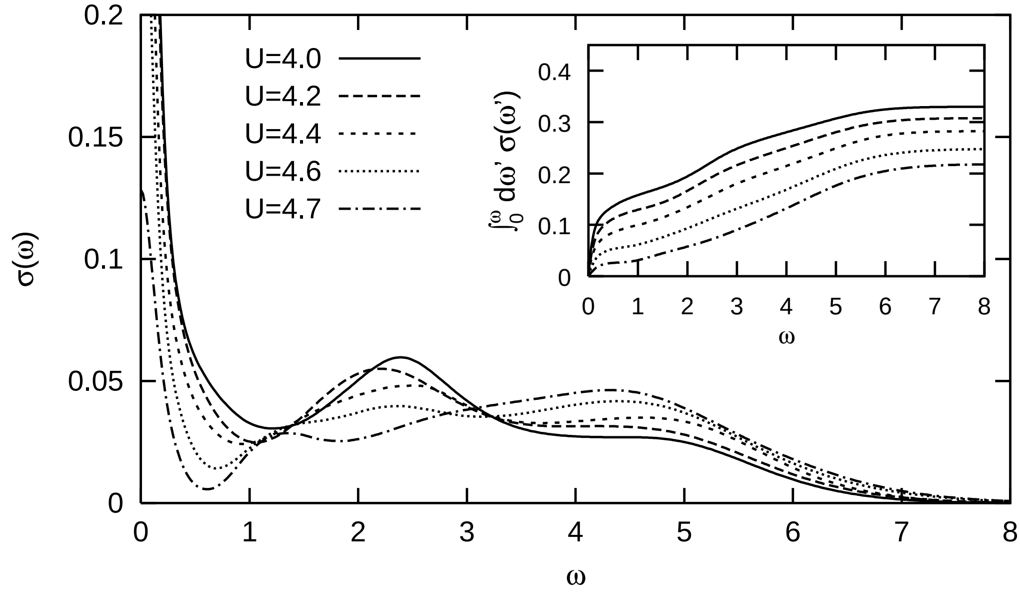



**Figure 4.37:** Optical conductivity $\sigma(\omega)$ for the anisotropic stacked Bethe lattice at $T = 0.05$ in the metallic phase; analogous to Fig. 4.34.

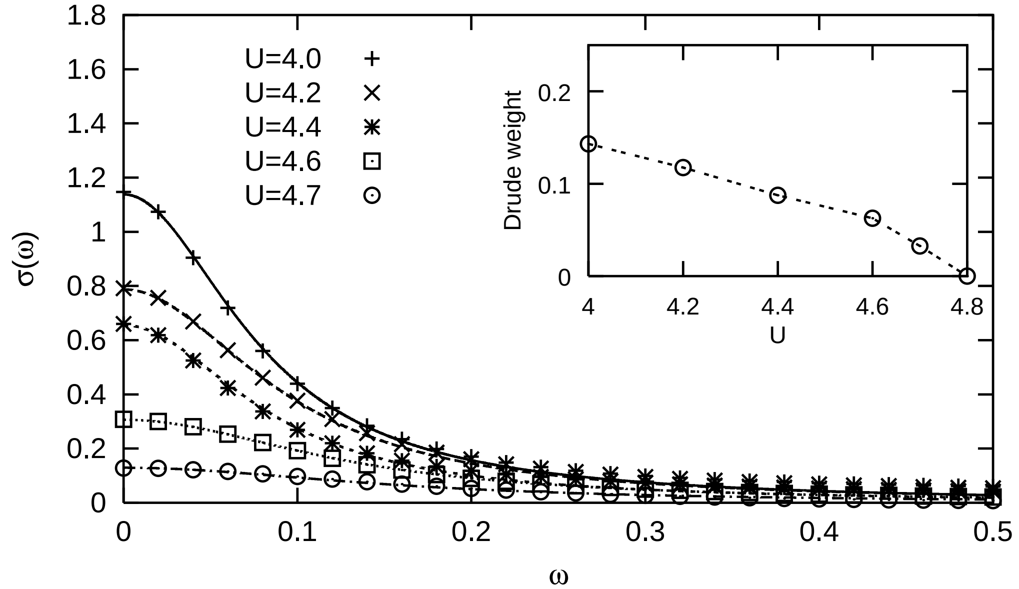



**Figure 4.38:** Drude part of the optical conductivity $\sigma(\omega)$ for the anisotropic stacked Bethe lattice at $T = 0.05$ in the metallic phase; analogous to Fig. 4.35.

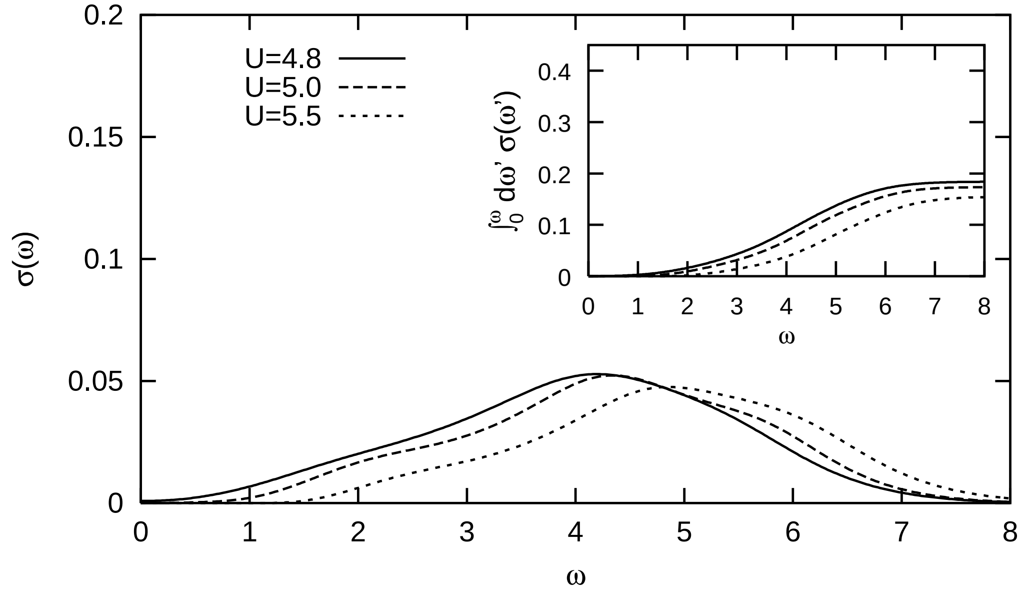



**Figure 4.39:** Optical conductivity $\sigma(\omega)$ for the anisotropic stacked Bethe lattice at $T = 0.05$ in the insulating phase; analogous to Fig. 4.36.

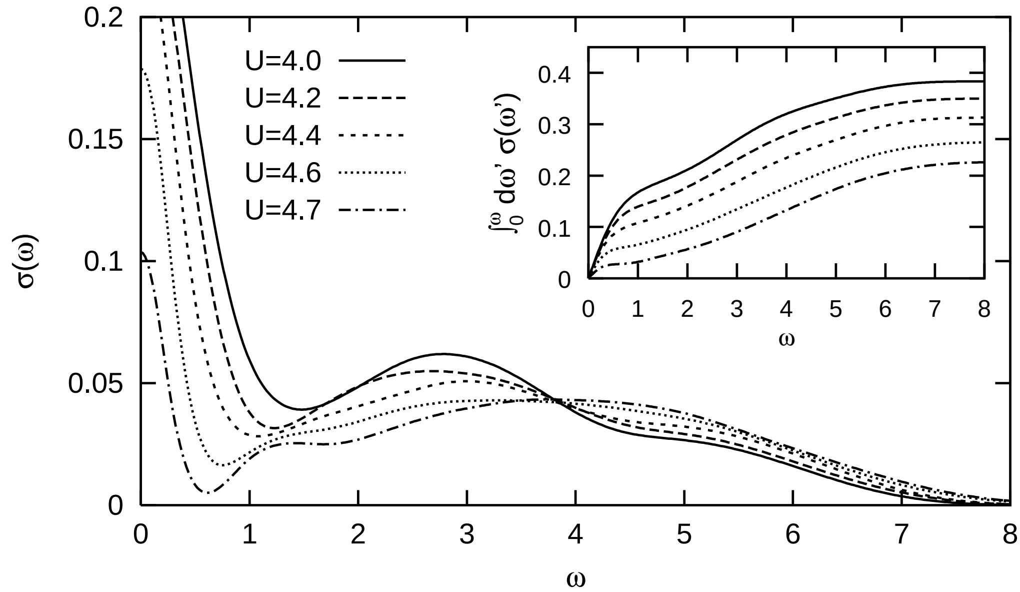



**Figure 4.40:** Optical conductivity $\sigma(\omega)$ for a fully disordered model or Stumpf's layout of the Bethe lattice at $T = 0.05$ in the insulating phase; analogous to Fig. 4.34.

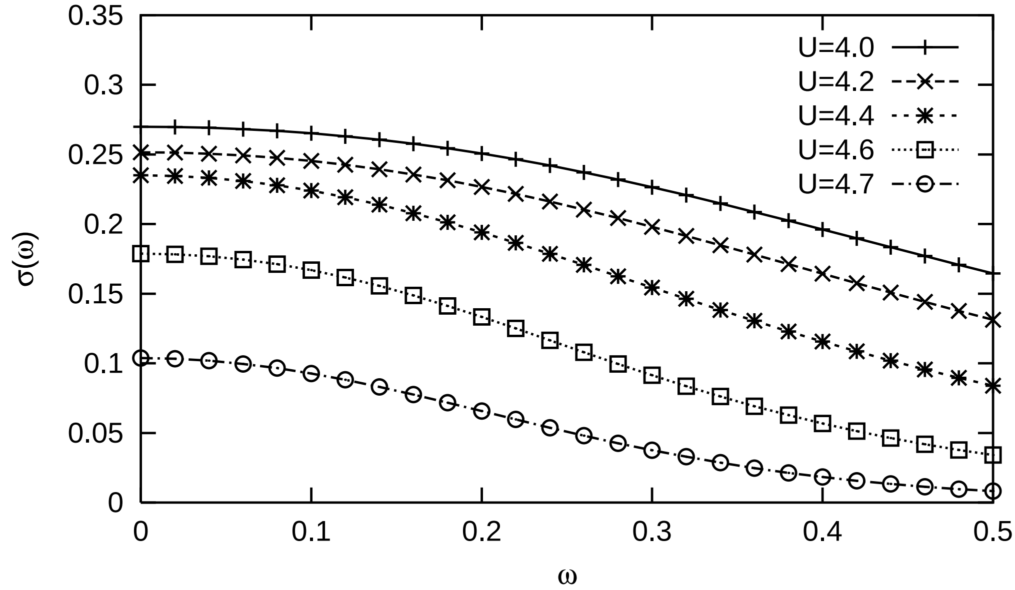



**Figure 4.41:** Low-frequency part of the optical conductivity $\sigma(\omega)$ for a fully disordered model or Stumpf's layout of the Bethe lattice at $T = 0.05$ in the metallic phase. Due to the incoherent character, no Drude peaks can be identified.

the large width of the low-frequency peaks which here depends on the width of the quasiparticle peak in the spectrum and, therefore, decreases with increasing $U$. Since the low-frequency part of $\sigma(\omega)$ does not have a Drude shape, the lines in Fig. 4.41 are guides to the eye only. The results for the insulating phase presented in Fig. 4.42 are hardly distinguishable from the corresponding results obtained for the clean stacked case (Fig. 4.39). We may conclude that the conservation of momentum (which applies for the isotropic and the stacked, but not for the disordered model) does not matter for transport in the insulator. In this phase, all relevant states are spread out so far in momentum space that the overlap for a given momentum becomes asymptotically proportional to a constant, i.e., $A_\epsilon(\omega)A_\epsilon(\omega') \propto A(\omega)A(\omega')$. In order to emphasize that offdiagonal disorder matters in general, let us take a look at the noninteracting case. Then, coherent transport is dispersionless, i.e., $\sigma(\omega) \propto \delta(\omega)$. In contrast, dissipation arises for energies up to the bandwidth in the disordered case as shown in Fig. 4.43 for various temperatures and band fillings.

### Comparison of the $f$-sums

The optical $f$-sums corresponding to the estimates of $\sigma(\omega)$ obtained in the last subsection are displayed in Fig. 4.44 a. Evidently, the $f$-sum is larger for the isotropic than for the stacked Bethe lattice, in particular in the metallic phase. This difference can be attributed to the enhanced squared Fermi velocity in the isotropic model (see Fig. 2.17): The enhancement is largest near the Fermi surface, at $\epsilon = 0$ by a factor of $\pi/2 \approx 1.57$. A corresponding increase in the $f$-sum is expected in the

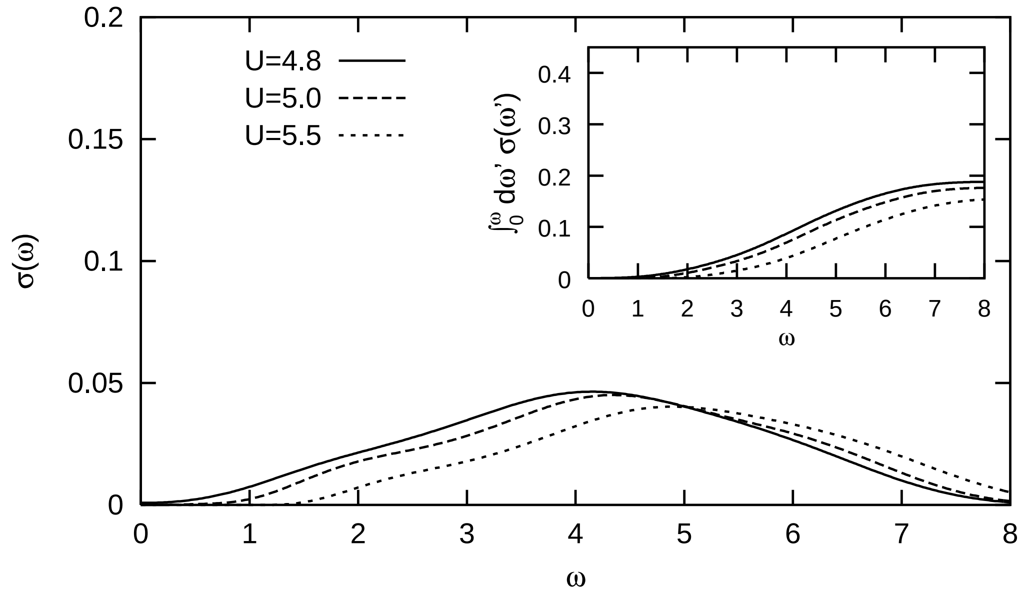



**Figure 4.42:** Optical conductivity $\sigma(\omega)$ for a fully disordered model or Stumpf's layout of the Bethe lattice at $T = 0.05$ in the insulating phase analogous to Fig. 4.36.

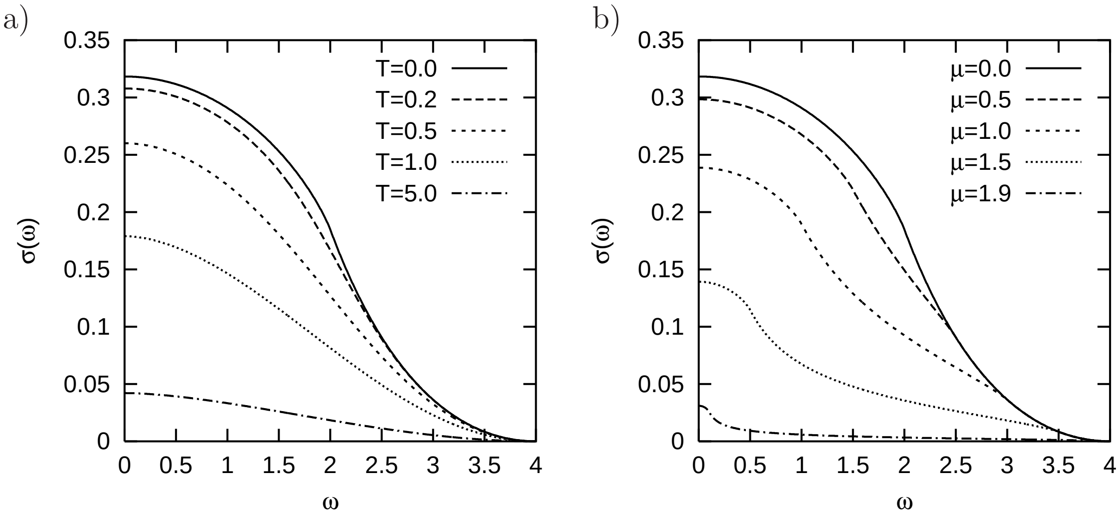



**Figure 4.43:** Optical conductivity $\sigma(\omega)$ for a noninteracting ($U = 0$) model with semi-elliptic DOS and full disorder or a single-chain layout. **a)** Half-filled case for different temperatures. **b)** Doped case at $T = 0$.

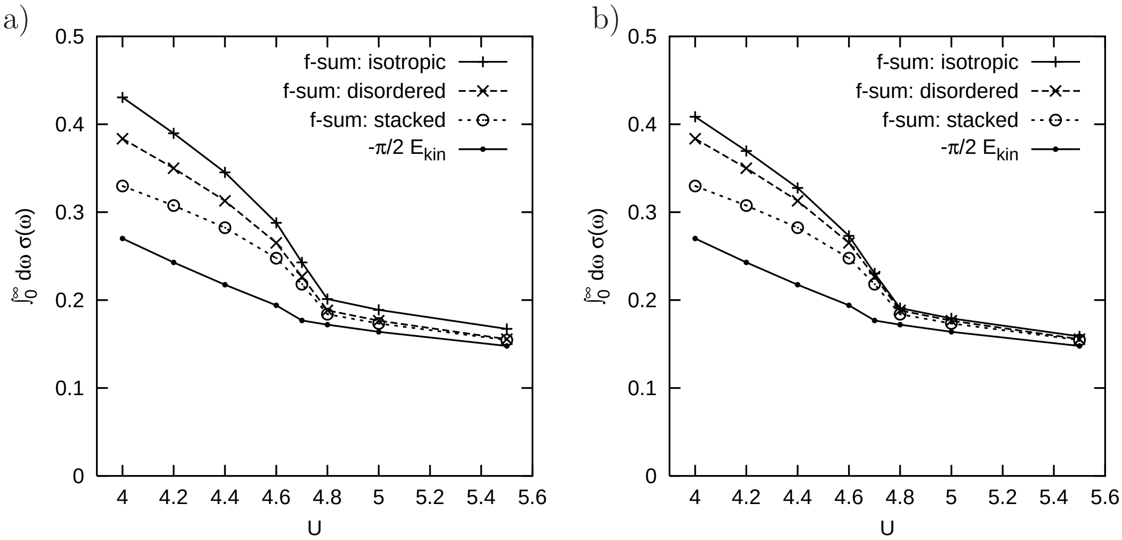



**Figure 4.44:** Optical $f$-sum for models with semi-elliptic DOS, computed by numerical integration of the results for $\sigma(\omega)$ shown in Fig. 4.34 – Fig. 4.42: **a)** scaled to unit $a$ and unit variance of the DOS, **b)** rescaled to unit average hopping distance (division of the $f$-sum of the isotropic model by 1.05406).

limit $U \to 0$ and $T \to 0$ where only states near the Fermi surface can contribute to transport. Averaged over the whole Brillouin zone, the enhancement is only (in general) $\int_{-\infty}^{\infty} d\epsilon\, \tilde{\rho}(\epsilon)$ which equals the average hopping distance $\sum_{D=1}^{\infty} D{t_D^*}^2$; using the values for the hopping amplitudes given in Table 2.1, it evaluates here to 1.05406. A corresponding increase in the $f$-sum in the limits $T \to \infty$ and/or $U \to \infty$ can be read off from the form (4.86) of the $f$-sum rule since in these cases the derivative of the momentum distribution function becomes asymptotically constant (as a function of $\epsilon$). This asymptotic behavior is already well approached for $U = 5.5$ as seen in Fig. 4.45 b. If the asymptotic enhancement is removed by scaling the average hopping distance to unity (Fig. 4.44 b), the $f$-sums of the isotropic and of the stacked model (as well as of the disordered model) are nearly identical.

In Fig. 4.46, partial contributions to the $f$-sum rule are displayed as a function of momentum $\epsilon$ so that the $f$-sums shown above would be obtained by integration over $\epsilon$. The picture differs considerably depending on whether the sum rule is used in terms of the momentum distribution function $n_\epsilon$ as in Fig. 4.46 a or in terms of its derivative as in Fig. 4.46 b . In both cases, however, the isotropic model (thick lines) implies contributions to $\sigma(\omega)$ which arise from smaller values of $|\epsilon|$ than the stacked model (thin lines). This is true in particular for the metallic phase at $U = 4.0$; transport is here dominated by states near the Fermi surface. Such behavior is not only consistent with Fermi-liquid theory, but also most compatible with the DMFT approach: nongeneric momenta with extremal $|\epsilon|$ should not influence results of DMFT calculations.

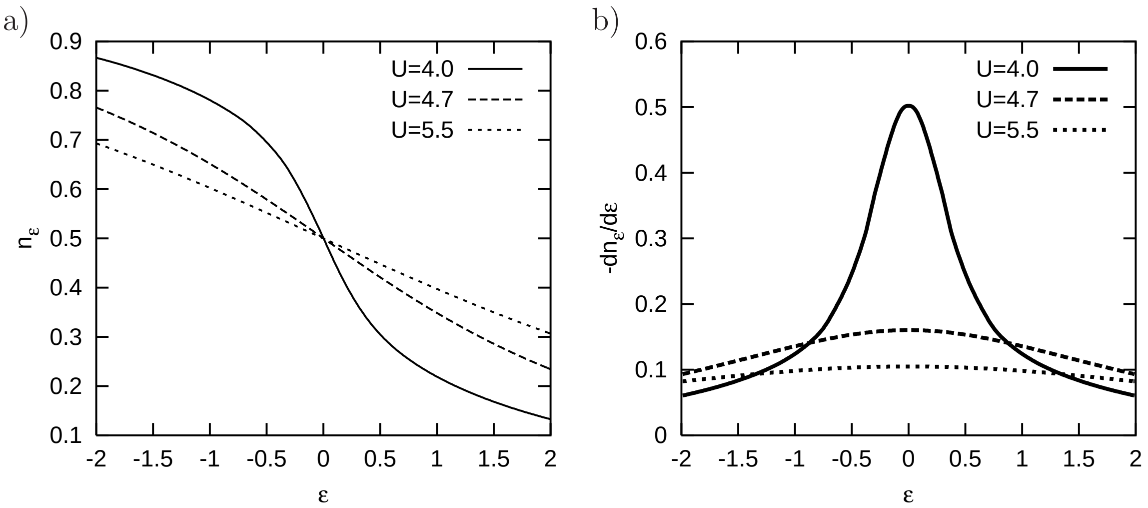



**Figure 4.45: a)** Momentum distribution function $n_\epsilon$ as defined in (4.79) and **b)** (minus) its derivative.

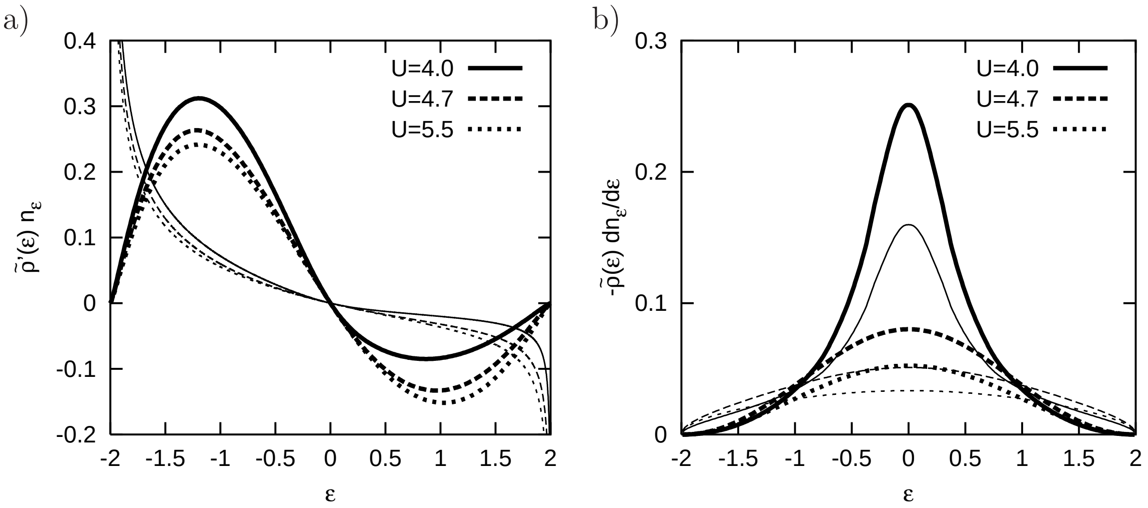



**Figure 4.46:** Contributions to the $f$-sum as a function of "momentum" $\epsilon$ for the redefined Bethe lattice (thick lines) and the stacked Bethe lattice (thin lines) corresponding to the equivalent formulations **a)** (4.86) and **b)** (4.87) of the $f$-sum rule.

## 4.7 Conclusion

In this chapter, we have presented studies of the optical conductivity $\sigma(\omega)$ and its $f$-sum rule in high dimensions with particular focus on "the Bethe lattice", i.e., a class of lattices with semi-elliptic density of states (DOS).

Initially, the project of computing numerical estimates of $\sigma(\omega)$ on the basis of quantum Monte Carlo (QMC) solutions of the half-filled Hubbard model with a semi-elliptic DOS seemed straightforward: The formalism had already been worked out for the hypercubic (hc) lattice and had been applied for general filling based on NCA (Pruschke et al., 1993) and QMC (Jarrell, Freericks, and Pruschke, 1995) calculations. Rozenberg et al. (1995) had even applied the same formalism to the Bethe lattice case (on the basis of ED and IPT data) and extensively discussed the $f$-sum rule in this context. Indeed, our implementation of the algorithm was soon accomplished; using an early version of our set of programs, results were published for a manganite model (Held and Vollhardt, 2000). However, we realized that the conventional treatment and the resulting $f$-sum rule were specific to the hypercubic lattice and not directly applicable to the Bethe lattice. In particular, we found numerical violations of the (incorrect) sum rule stated by Rozenberg. In addition, Janiš (1999) claimed the foundation of the DMFT treatment of conductivities to be incorrect.[53] Thus, we found it necessary to thoroughly check the formalism and to derive extensions for the general case in high dimensions.

We have carefully reviewed the Kubo formalism for the optical conductivity and the simplifications in the limit of infinite dimensionality (or, more generally, infinite coordination number). We have presented the first evaluation of the theory in the case of $t{-}t'$ hopping and have shown that Merino and McKenzie's (2000) (unjustified) use of the conventional hc formalism leads to significant errors in the computation of the transverse conductivity and the thermal conductivity. We have derived a new expression of the optical $f$-sum rule which is universally valid within the DMFT and additionally a form specific to the general dispersion formalism developed in Sec. 2.3.

We have presented a comprehensive discussion of the optical conductivity for lattices with semi-elliptic "Bethe" DOS in infinite dimensions. In particular, we have discovered a flaw in Freericks's (2000) derivation of $\sigma(\omega)$ for a tree-like layout of the Bethe lattice and have shown that the assumption under which Millis (2002) obtained essentially the same result is equally unfounded. We have pointed out that an application of the hc formalism generally implies anisotropic transport and have established the connection between the $f$-sum and a static local property of such models. We have further demonstrated and explained the incoherence of transport resulting from Stumpf's (1999) single-chain layout of the Bethe lattice or from strong offdiagonal disorder and have developed a necessary criterion for coherent transport in the noninteracting limit. We have also presented the first consistent derivation for $\sigma(\omega)$ compatible with a semi-elliptic DOS that implies isotropic transport which is fully coherent in the noninteracting limit. This redefinition of the Bethe lattice (in the DMFT sense) as an isotropic, regular and clean lattice by the general dispersion method and the demonstration that the associated transport properties are robust (with respect to finite dimensionality or hopping range) removes, finally, the pathologies previously associated with the DMFT treatment of transport in connection with non-Gaussian DOSs.

We have written programs for the DMFT computation of conductivities (on the basis of local spectra) for arbitrary lattice types and have developed a scheme that

[53] It has become clear later that Janiš's (1999) statement "*the current-current correlation function and consequently the electrical conductivity do have nontrivial vertex corrections in the mean-field limit contrary to what has been up to now assumed from symmetry considerations within the conserving scheme*" does not apply to the optical conductivity (i.e., in the long-wavelength limit).

is particularly robust in combination with the QMC code and the variant of the maximum entropy method used within this work. The high accuracy of our method has been successfully tested by comparison with a newly established sum rule for the (real-frequency) self-energy. We have presented numerical results for $\sigma(\omega)$ of the half-filled Hubbard model with semi-elliptic DOS near the metal-insulator transition for all valid definitions, i.e., for the incoherent models, the anisotropic stacked model, and for the new isotropic "redefined Bethe lattice". Comparing the qualitative features, the Drude weight, and the $f$-sums, we have shown that the exact definition does matter in general, in particular for not too large interaction.

We have pointed out in the beginning of this chapter that $f$-sum rules as discussed above may be regarded as artifacts of the single-band assumption. While they may still be useful as approximate partial sum rules, a different, universal $f$-sum rule is obtained in a full model. We have only briefly touched upon additional features in $\sigma(\omega)$ which are specific to multi-band models or to non-Bravais lattices. Furthermore we had to limit the discussion of finite dimensionality to a fully local theory where the bubble formula for $\sigma(\omega)$ remains valid; vertex corrections not included in this approximation are expected to increase the static conductivity since the importance of umklapp scattering is reduced (Uhrig and Vollhardt, 1995). On the other hand, a reduction of the static conductivity and a broadening of the Drude peak is expected to result from impurities and from electron-phonon interaction, which were also neglected in this chapter.

# Chapter 5

# Realistic Modeling of Strongly Correlated Materials

Well-controlled studies of abstract models like the one-band Hubbard model as presented in the preceding chapters shed light on important physical phenomena, even when the models themselves are neither fully realistic nor complete (full frustration, single-band assumption, etc.). Still, one ultimate goal of condensed matter theory is a controlled description of materials which is as realistic as possible and allows for a material-specific calculation of a wide range of properties. One step into this direction, although not fully controlled, is the recently developed LDA+DMFT method which we will apply in this chapter for modeling the doped transition metal oxide $La_{1\text{-}x}Sr_xTiO_3$. We will introduce the material under investigation and demonstrate the failure of LDA to reproduce its photoemission spectrum in Sec. 5.1, then discuss conventional density functional theory in Sec. 5.2 and introduce the LDA+DMFT method in Sec. 5.3. Numerical results will be presented and compared with experiments in Sec. 5.4.

The work presented in this chapter was started in a collaboration with the Optics of Metals Laboratory at the Institute of Metal Physics of the Ural Division of the Russian Academy of Sciences in Yekaterinburg lead by Prof. V. I. Anisimov. In particular, all LDA calculations were performed by Igor Nekrasov from Anisimov's group. The initial QMC and MEM results presented in subsection 5.4.1 were mainly calculated by Nekrasov and Held using Held's multi-band version of the QMC code while the photoemission spectra were obtained using an analysis program written as part of this thesis. The QMC computations at higher accuracy and the study of the impact of different MEM procedures on the resulting spectra presented in subsection 5.4.2 as well as the calculation of transport properties in subsection 5.4.3 were also performed by the author of this thesis.

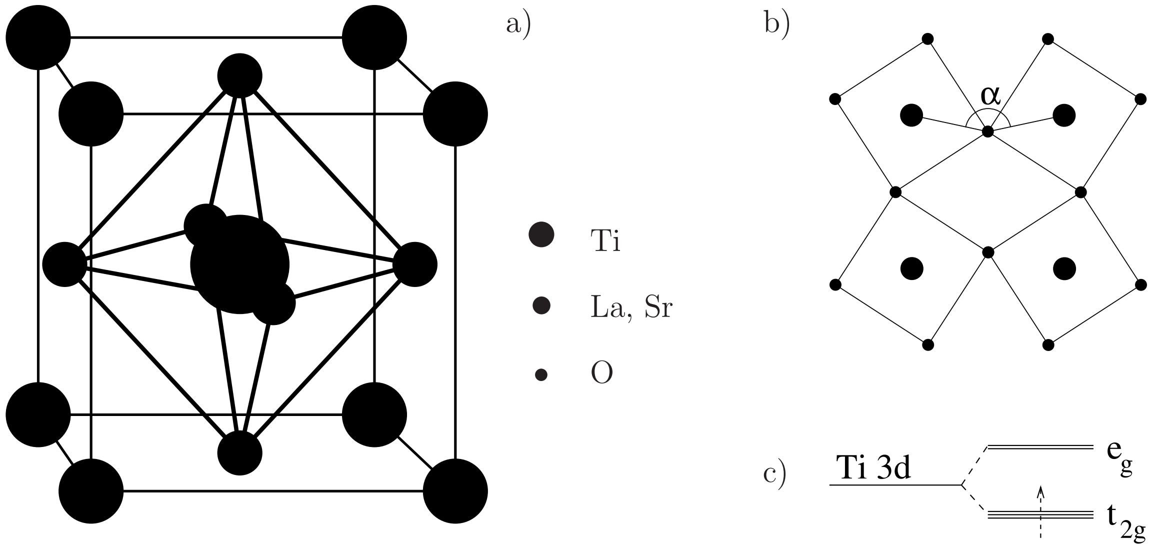



**Figure 5.1:** Structure of $La_{1\text{-}x}Sr_xTiO_3$ : **a)** Idealized (cubic) perovskite structure: Ti ions are octahedrally coordinated by oxygen in a cubic matrix of sites occupied by either La or Sr (presumably disordered). The size of spheres in the figure is not related to ionic radii. **b)** Orthorhombic distortion, i.e., alternating tilting of oxygen octahedra which reduces the angle Ti-O-Ti to $\alpha \approx 155°$. **c)** Scheme of band degeneracy; the $t_{2g}$ band contains one electron per unit cell for $LaTiO_3$ and is empty for $SrTiO_3$.

## 5.1 $La_{1\text{-}x}Sr_xTiO_3$, a Strongly Correlated Transition Metal Oxide

$La_{1\text{-}x}Sr_xTiO_3$ belongs to a large class of $3d$ transition metal oxides of composition $ABO_3$, where A is a rare earth or alkalide (or a nonstoichiometric combination thereof) and B a transition metal and which crystallize in the orthorhombically distorted perovskite ($GdFeO_3$-type) structure. The idealized cubic perovskite structure is shown in Fig. 5.1 a. In the distorted perovskite, the almost perfectly rigid oxygen octahedra tilt alternatingly in all space directions so that the angle $\alpha$ (Ti-O-Ti) is reduced from its cubic value of 180° to, e.g., $\alpha \approx 155°$ for $La_{1\text{-}x}Sr_xTiO_3$ (MacLean, Ng, and Greedan, 1979), see Fig. 5.1 b.[1] Formally, pure $SrTiO_3$ is an ionic insulator where all ions reach noble gas configurations.[2] In fact, this material is a band insulator which is seen both in experiment and band structure calculations and makes this material well-suited as substrate or spacer material for high-$T_c$ superconducting films.[3] Both experiment and band-structure calculations imply that La donates 3

[1]The degree of distortion can in general be estimated from ionic radii in terms of the tolerance factor $f = (r_A + r_O)/(\sqrt{2}(r_B + r_O))$, where $f \approx 0.8$ corresponds to $\alpha \approx 140°$ and $f \approx 0.95$ corresponds to $\alpha \approx 157°$, respectively (Imada et al., 1998).

[2]O needs 2 electrons to obtain the [Ne] configuration; the atomic configuration is [Kr] $5s^2$ for Sr, [Ar] $3d^1 4s^2$ for Ti, and [Xe] $5d^1 6s^2$ for La.

[3]The chemically reduced component, $SrTiO_{3-\delta}$, is a superconductor with $T_c \lesssim 0.5$ K (Schooley, Hosler, Ambler, Becker, Cohen, and Koonce, 1965).

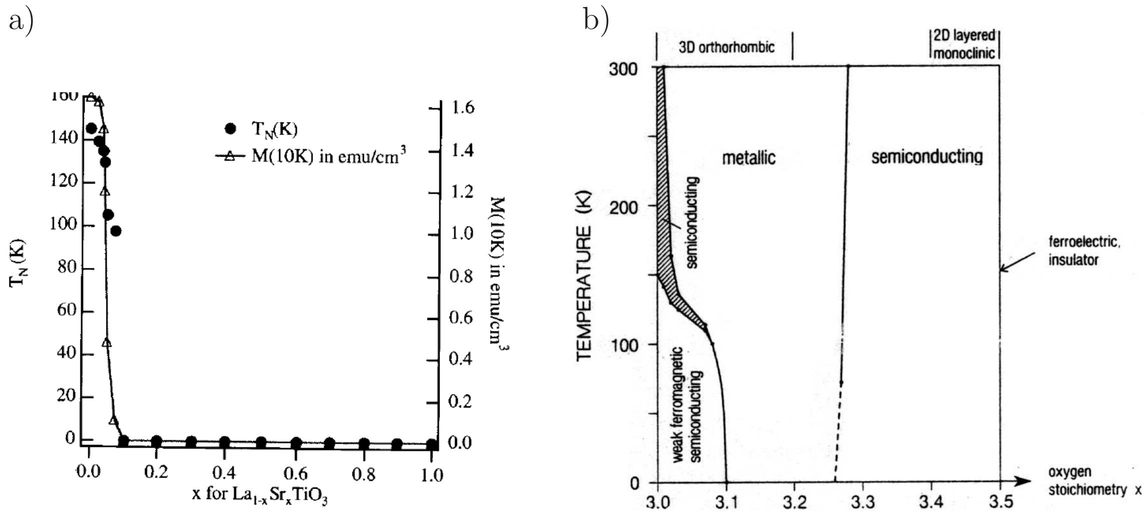



**Figure 5.2: a)** Néel temperature $T_{\rm N}$ and remanent magnetization at $T = 10$ K for $La_{1\text{-}x}Sr_xTiO_3$ as a function of doping (Hays et al., 1999). **b)** Phase diagram for $LaTiO_{3+x}$ (Lichtenberg et al., 1991).

electrons, i.e., formally reaches noble gas configuration in $La_{1\text{-}x}Sr_xTiO_3$ as well, so that Ti obtains the configuration $3d^1$ in pure $LaTiO_3$. As discussed in App. A.1, cubic symmetry splits the $3d$ band into a twofold degenerate $e_g$ and a threefold degenerate $t_{2g}$ band as depicted in Fig. 5.1 c. While in principle the reduced symmetry due to distortion should further split up the $t_{2g}$ band, this effect is negligible in practice since the local ligand field of the Ti ion is dominated by the local octahedral oxygen coordination which retains cubic symmetry. The single electron per unit cell therefore occupies predominantly the threefold degenerate $t_{2g}$ bands hybridizing via oxygen $2p$ states. The effective overlap between Ti states of neighboring unit cells and, consequently, the effective band width (in a single-particle picture) is reduced by orthorhombic distortion compared to the idealized cubic structure for $La_{1\text{-}x}Sr_xTiO_3$ (and even more for $YTiO_3$ where $\alpha \approx 142°$).

Experimentally, pure $LaTiO_3$ is an antiferromagnetic insulator with a Néel temperature estimated between $T_{\rm N} \approx 125$K (Eitel and Greedan, 1986) and $T_{\rm N} \approx 145$ K (Hays, Zhou, Markert, and Goodenough, 1999); see Fig. 5.2. The Ti magnetic moment saturates at 0.45 $\mu_{\rm B}$ and the energy gap is about 0.2 eV (Goral, Greedan, and MacLean, 1982; Crandles, Timusk, Garrett, and Greedan, 1994). At doping $x \approx 0.05$, $La_{1\text{-}x}Sr_xTiO_3$ undergoes an insulator-to-metal transition and becomes a correlated paramagnetic metal with a strongly enhanced susceptibility and electronic specific heat coefficient (Pickett, Erwin, and Ethridge, 1998). Apparently, the fabrication of samples of controlled stoichiometry and even the a posteriori analysis of their composition are still nontrivial tasks for experimentalists. Often, only the formal deviation $\delta = x + y/2$ from integer $t_{2g}$ filling of a composition $La_{1-x}Sr_xTiO_{3+y}$ is stated in results (Lichtenberg et al., 1991) as seen in Fig. 5.2 b. For $y > 0$ such a formula does no longer correspond to a unit cell since there is no place in the structure

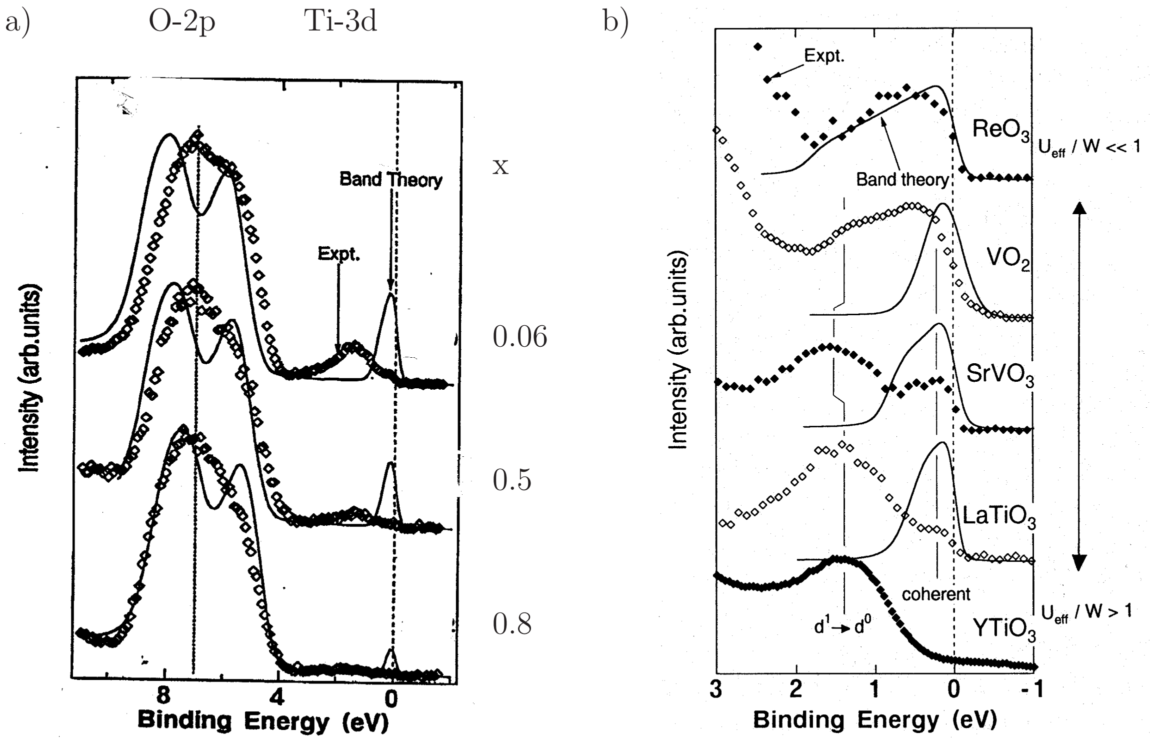



**Figure 5.3: a)** Photoemission spectra (squares) in comparison with LDA calculations (solid lines) for $La_{1-x}Sr_xTiO_3$ (Fujimori, Hase, Nakamura, Namatame, Fujishima, Tokura, Abbate, de Groot, Czyzyk, Fuggle, Strebel, Lopez, Domke and Kaindl, 1992). **b)** PES over reduced energy range of various transition metal oxides with $d^1$ configuration with increasing deviations from LDA calculations from top to bottom. The curve "$LaTiO_3$" gives a magnified view of the PES of $La_{1-x}Sr_xTiO_3$ for $x = 0.06$ (Fujimori, Hase, Namatame, Fujishima, Tokura, Eisaki, Uchida, Takegahara and de Groot, 1992).

for interstitial oxygen; instead vacancies form at La and Ti sites in this case (Hays et al., 1999). As seen in Fig. 5.2 a, recent measurements (with 1% relative accuracy of the oxygen content) show a continuous decay of the remanent magnetization with increased Sr doping while the Néel temperature jumps discontinuously with an intermediate plateau. This behavior implies the existence of a two-phase region including AF order for $La_{1-x}Sr_xTiO_3$ with $0.045 \leq x \leq 0.08$ (Hays et al., 1999).

The importance of correlations is directly seen in Fig. 5.3 a, a comparison between measured photoemission spectra (PES) and theoretical estimates based on LDA band structure calculations. Evidently, the oxygen $2p$ bands lying far below the Fermi energy are well reproduced by LDA; in contrast, spectral weight corresponding to $t_{2g}$ bands is predicted by LDA as a peak at the Fermi energy while the experiment shows contributions at around -1.5 eV. The disagreement is strongest for small doping $x$ (i.e., large occupancy $1 - x$ of the Ti $t_{2g}$ band) which is the case we will study in the following. Both a magnified view of the $t_{2g}$ portion of the PES for $La_{1-x}Sr_xTiO_3$ for $x = 0.06$, i.e., just outside the antiferromagnetic phase region, and a comparison to other transition metal oxides with $d^1$ configuration are shown in Fig. 5.3 b.

## 5.2 Density Functional Theory and Local Spin Density Approximation

Density functional theory (DFT) is, originally, an approach for calculating ground state properties via a mapping of the full interacting many-body problem onto an effective single-particle problem. It is based on the observation that the ground state energy of any $N$-electron system can be written as a functional of the electronic density $\rho(\boldsymbol{r})$ alone (Hohenberg and Kohn, 1964). Following Levy's (1979) proof, we split up the full electronic Hamiltonian (1.2) into internal contributions $\hat{T} + \hat{V}_{ee}$ and the external contribution $\hat{V}_{\text{ext}}$ with $\langle \hat{V}_{\text{ext}} \rangle = \int \mathrm{d}\boldsymbol{r} V_{\text{ext}}(\boldsymbol{r})\rho(\boldsymbol{r}) =: E_{\text{ext}}[\rho]$ . Since both the kinetic energy $\hat{T}$ and the Coulomb electron-electron interaction $\hat{V}_{ee}$ have a universal form, the functional

$$F[\rho] := \min_{\psi} \left\{ \langle\psi|\hat{T} + \hat{V}_{ee}|\psi\rangle \;\middle|\; \langle\psi|\hat{\rho}(\boldsymbol{r})|\psi\rangle = \rho(\boldsymbol{r}) \right\}, \tag{5.1}$$

where $\hat{\rho}(\boldsymbol{r})$ measures the density at site $\boldsymbol{r}$ and $\psi$ denotes a fermion many-body wave function, is indeed universal,[4] i.e., a functional of the density alone. In terms of this functional, the basic theorems of DFT read

$$E[\rho] := F[\rho] + \int \mathrm{d}\boldsymbol{r}\, \hat{V}_{\text{ext}}\, \rho(\boldsymbol{r}) \quad \geq \quad E_{\text{GS}}\,, \tag{5.2}$$

$$F[\rho_{\text{GS}}] + \int \mathrm{d}\boldsymbol{r}\, \hat{V}_{\text{ext}}\, \rho_{\text{GS}}(\boldsymbol{r}) \quad = \quad E_{\text{GS}}\,, \tag{5.3}$$

i.e., the energy functional $E[\rho]$ defined in (5.2) is both an upper bound to the true ground state energy $E_{\text{GS}}$ for arbitrary density distribution and coincides with $E_{\text{GS}}$ for the true ground state density distribution $\rho_{\text{GS}}$.[5] The first theorem (5.2) follows from an application of the usual variational principle $\langle\psi|\hat{T} + \hat{V}_{ee} + \hat{V}_{\text{ext}}|\psi\rangle \geq E_{\text{GS}}$ to the wave function $\psi_{\min}$ which minimizes the r.h.s. expression in (5.1). On the other hand, each ground state wave function $\psi_{\text{GS}}$ also minimizes $\hat{T} + \hat{V}_{ee}$ in (5.1) for fixed ground state density which proves (5.3). If the universal functional $F[\rho]$ was known explicitly and could be evaluated efficiently the ground state energy could be obtained in an arbitrary external potential by minimizing $E[\rho]$ defined in (5.2). Other ground state properties could then be obtained using corresponding (possibly multi-valued) functionals. In practice, however, a controlled computation of $F[\rho]$ along the lines of (5.1) (which would require the solution of jellium models with constrained density) is not possible.

In this situation, it is useful to split up the energy functional differently (Kohn and Sham, 1965),

$$E[\rho] = T_0[\rho] + E_{\text{ext}}[\rho] + E_{\text{Hartree}}[\rho] + E_{\text{xc}}[\rho], \tag{5.4}$$

[4] Note that the proof reproduced here implies universality of the functional $F[\rho]$ only for fixed particle number $N$. Implicitly, one usually assumes that the thermodynamic limit of the formalism is well defined.

[5] One can show that even the external potential is a functional of the ground state density (Hohenberg and Kohn, 1964). This finding is, however, not used in actual calculations and does not extend to spin-resolved DFT with inclusion of magnetic fields (Capelle and Vignale, 2001).

where $T_0[\rho]$ is not the true kinetic energy (Jones and Gunnarsson, 1989), but the kinetic energy of a noninteracting $N$-particle system with density distribution $\{\rho(\boldsymbol{r})\}$; $E_{\text{ext}}$ was defined above and the Hartree energy is given by

$$E_{\text{Hartree}} = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,V_{ee}(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r})\rho(\boldsymbol{r}')\,. \tag{5.5}$$

The first three terms in (5.4) can be treated exactly, so that the methodological problem is reduced to finding a good approximation for the exchange-correlation energy for which (5.4) is a defining equation in connection with (5.1), (5.2), and (5.5). This energy functional can be assumed to be much less dependent on details of the density distribution than the full functional which includes very large kinetic energy terms. Given an (approximate) explicit expression for $E_{\text{xc}}[\rho]$, the ground state properties $E_{\text{GS}}$ and $\{\rho_{\text{GS}}\}$ of some specified system are then found numerically by a self-consistent solution of the Kohn-Sham equations,

$$\Big[-\frac{\hbar^2}{2m_e}\nabla + V_{ext}(\boldsymbol{r}) + \int \mathrm{d}\boldsymbol{r}'\rho(\boldsymbol{r}')\,V_{\text{ee}}(\boldsymbol{r}-\boldsymbol{r}') + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\boldsymbol{r})}\Big]\,\varphi_i(\boldsymbol{r}) = \epsilon_i\,\varphi_i(\boldsymbol{r})\,, \tag{5.6}$$

i.e., a Schrödinger-like equation for a set of single-particle wave functions $\{\varphi_i\}_{i=1}^N$ which reproduce the density $\rho(\boldsymbol{r}) = \sum_{i=1}^N |\varphi_i(\boldsymbol{r})|^2$. Note that these wave functions and the associated eigenvalues $\epsilon_i$ have no direct meaning in the context of the original many-body problem.

In the DFT calculations presented in this work we follow the usual approach of treating the functional $E_{\text{xc}}$ in local spin density approximation (LDA or LSDA),[6]

$$E_{\text{xc}}[\rho] \approx \int \mathrm{d}\boldsymbol{r}\ \rho(\boldsymbol{r})\ v_{\text{xc}}^{\text{LDA}}\big(\rho_\uparrow(\boldsymbol{r}),\rho_\downarrow(\boldsymbol{r})\big), \tag{5.7}$$

where the functional dependence of $v_{\text{xc}}^{\text{LDA}}$ is usually taken from parameterizations of solutions of simplified jellium models (von Barth and Hedin, 1972; Gunnarsson and Lundqvist, 1976) or from the QMC solution by Ceperley and Alder (1980) of the jellium model (Vosko and Wilk, 1980; Perdew and Zunger, 1981). Different parameterizations lead to slightly different results, but are all free from adjustable parameters. For exact properties of $E_{\text{xc}}[\rho]$ and possible general approximations beyond LDA see, e.g., the review by Jones and Gunnarsson (1989).

Even though the LDA problem (5.6) with (5.7) is conceptually much simpler than the original many-body problem, its numerically stable solution in the presence of singular ion potentials and for a relatively large number of electrons is still a nontrivial task. For condensed matter problems[7] one may distinguish full-potential from

[6]It is easy to check that the above derivations directly generalize to functionals $F[\rho_\uparrow,\rho_\downarrow]$ and $E[\rho_\uparrow,\rho_\downarrow]$ of the spin density. While in principle the approaches with and without explicit spin resolution are equivalent (cf. footnote 5 on page 249) since $\{\rho_\uparrow\}$ and $\{\rho_\downarrow\}$ are functionals of $\{\rho\}$, it is found that better approximations to the exchange-correlation energy are available in terms of the spin resolved densities.

[7]Specialized methods and, in particular, basis sets for the single-particle wave functions $\varphi_i$ exist for quantum chemistry problems.

pseudopotential approaches where the latter use weak effective potentials for valence electrons in order to get rid of core oscillations which would require very large basis sets in plane-wave expansions. While newer codes allow for efficient full-potential calculations in a linear augmented plane wave (LAPW) basis [as, e.g., Blaha, Schwarz, and Luitz's (1999) package VIENNA], the problem can be avoided by using atomic-like basis sets at least near the ionic cores. In the muffin-tin orbital (MTO) approach, a minimal basis characterized by angular momentum quantum numbers is constructed where the external potential only acts inside a spherical "muffin tin". The energy dependence of the MTOs is linearized and an energy-independent set obtained in the LMTO method which is then reliable only for valence states (Andersen, 1975). Lifting the constraint of minimalism, more localized[8] "tight binding" basis functions were introduced in the TB-LMTO approach (Andersen, Pawlowska, and Jepsen, 1986) which is usually used in connection with the atomic sphere approximation (ASA) of a locally spherical resulting potential.

It should be stressed that DFT as a ground state method cannot directly be used for the determination of excitation spectra. In principle, apart from true ground state properties, only the sizes of energy gaps can be computed as differences of ground state energies for systems with $N$ and $N+1$ electrons, respectively. In practice, however, the eigenvalues $\epsilon_i$ of the single particle problem (5.6) often compare favorably with experiment and are regularly regarded as band structures of the interacting system.

The various LDA methods have been used successfully for the computation of bulk properties of materials, in particular metals, as well as surface or impurity effects. While their superiority over preceding DFTs such as the Thomas-Fermi approximation can be explained from the vastly improved treatment of the kinetic energy, the accuracy of LDA, an effective single-particle theory without adjustable parameters ("ab initio") came as a surprise not only to its inventors. Still, LDA fails for important classes of materials with strong electronic correlations like $La_{1\text{-}x}Sr_xTiO_3$, which limits its general predictive power. In the next section we will discuss extensions of the LDA approach which address this problem.

## 5.3 LDA+DMFT

Since the LDA uses input from the jellium model, it is not surprising that the resulting errors are larger for localized than for extended states. One may suspect that both local correlations and the self-interaction contribute significantly to these errors for localized orbitals. In fact, self-interaction corrected (SIC-LDA) calculations yield a much better estimate for the total energy of isolated atoms than plain LDA which can be attributed to an improved treatment of core states.[9] Using a basis with localized

[8]Conventional MTOs fall off as $r^{-(l+1)}$ for angular momentum $l$.

[9]For the total energy of, e.g., Ne atoms, LDA is even inferior to the Hartree-Fock theory (HF). For ionization energies, however, the agreement with experiment is significantly better for LSDA than either for LDA or HF with no general improvement from inclusion of SIC (Gunnarsson and Jones, 1981).

orbitals, self-interaction corrections can also be applied for transition metal oxides yielding improved predictions for band gaps and moments (Svane and Gunnarsson, 1990). The important advantage of SIC-LDA over LDA (and the reason we mention the method here) is that SIC provides a mechanism which allows the solutions to localize and that it effectively generates a Hubbard-type interaction which may split occupied and unoccupied states.

A similar effect arises in the LDA+U method which explicitly introduces mean-field type local interactions between electrons in correlated orbitals of degeneracy $M$, e.g., $d$ orbitals (Anisimov, Zaanen, and Andersen, 1991),

$$E_{\mathrm{LDA+U}} = E_{\mathrm{LDA}} + \frac{1}{2} \sum_{m\sigma \neq m'\sigma'} (U - J\delta_{\sigma\sigma'})\,(n_{m\sigma} - \bar{n})(n_{m'\sigma'} - \bar{n}) \tag{5.8}$$

The LDA energy is unchanged in the unpolarized case, when the expectation value $n_{m\sigma}$ of the occupancy of each state equals $\bar{n} := \frac{1}{2M}\sum_{m=1}^{M}\sum_{\sigma} n_{m\sigma}$. Orbital and/or spin order, however, generically lead to negative contributions of the interaction term in (5.8) which may outweigh an concurrent increase of the LDA energy and thus stabilize ordered insulating phases for integer filling. Both interaction parameters $U$ and $J$ can be estimated from constrained supercell LDA calculations (Gunnarsson, Andersen, Jepsen, and Zaanen, 1989). Thus, in contrast to plain LDA, the LDA+U method allows for an “ab initio” modeling of insulating, long-range ordered phases (e.g., antiferromagnetic or orbital order) within the DFT framework, but still cannot describe paramagnetic Mott insulators or strongly renormalized quasiparticle states. Such problems are clearly beyond static effective single-particle theories.

More general approaches which also allow for a microscopic treatment of on-site correlations have to be formulated in terms of a true many-body Hamiltonian. Ideally, one could treat interactions between electrons in some localized orbitals of an open shell explicitly (without resorting to static mean-field approximations) and use LDA for all remaining electron-electron interactions as well as for electron-ion interactions and kinetic energy. In practice, we can at least avoid static mean-field approximations for the explicitly treated interactions and approximately extract a tight-binding model from LDA calculations which we then have to correct approximately for double counting. For this purpose, we interpret the Lagrange parameters $\epsilon_i$ of the LDA solution of the Kohn-Sham equations (5.6) for a tight-binding LMTO basis (Andersen et al., 1986; Jepsen and Andersen, 2000) as defining a band structure which may be represented in terms of hopping matrix elements (vanishing on the diagonal: $t_{ilm,ilm} = 0$) and on-site energies,

$$\hat{H}^{\mathrm{LDA}} = \sum_{ilm,jl'm',\sigma} \left( t_{ilm,jl'm'} \hat{c}^{\dagger}_{ilm,\sigma} \hat{c}_{jl'm',\sigma} + \varepsilon_{il} \hat{n}_{ilm,\sigma} \delta_{ilm,jl'm'} \right) . \tag{5.9}$$

Here, the indices $i$ and $j$ label atomic sites, $\sigma$ labels spin, and $l$, $m$, $l'$, $m'$ are orbital indices.[10] Up to double counting corrections, we will regard (5.9) as the

[10]Neither the orbital basis nor the distinction between the index components $l$ and $m$ is unique. Here, we have adopted the convention that degenerate orbitals share a common value of $l$, but can be distinguished by the index $m$.

noninteracting part of our model. It is neither possible nor necessary to include all possible on-site interactions shown in (1.9). Instead, one has to concentrate on terms which can be treated more accurately in a many-body tight-binding model approach than in LDA and which dominate the correlation physics missed by plain LDA. In a minimal extension of the LDA, one may restrict the treatment to contributions of the density-density type[11]

$$H_{\mathrm{corr}} = \frac{1}{2} \sum_{\substack{ilm,l'm',\sigma\sigma' \\ (lm\sigma)\neq(l'm'\sigma')}} U^{i}_{lm,l'm'}\, \hat{n}_{ilm,\sigma}\, \hat{n}_{il'm',\sigma'}\,. \tag{5.10}$$

For the application to $La_{1\text{-}x}Sr_xTiO_3$, we will only include on-site interactions within the threefold degenerate Ti $t_{2g}$ orbitals, i.e., restrict the sum in (5.10) to $i = i_{\mathrm{Ti}}$ and $l = l' = l_{t_{2g}}$. Furthermore, we crudely approximate intra-orbital and interorbital electron-electron interaction by a common mean parameter, i.e.,

$$U^{i}_{lm,l'm'} = U\delta_{ii_{\mathrm{Ti}}}\delta_{ll'}\delta_{ll_{t_{2g}}}\,. \tag{5.11}$$

The parameter $U$ may be approximately calculated in constrained supercell LDA calculations (Gunnarsson et al., 1989) or estimated from experiment, e.g., high-energy spectroscopy. Note that the LDA estimate depends both on the particular basis set used in LDA and on the selection of interacting orbitals: $U$ decreases if more orbitals (e.g., the Ti $3d$ $e_g$ orbitals) are considered as noninteracting and are thus allowed to participate in screening. In order to avoid double counting, one should in general take into account that the Coulomb interaction (5.11) is already contained in LDA in some averaged way. A guess for the corresponding contribution has been derived from the atomic limit (Anisimov, Poteryaev, Korotin, Anokhin, and Kotliar, 1997),

$$E^{\mathrm{LDA}}_{\mathrm{corr}} = \frac{1}{2}Un_{t_{2g}}(n_{t_{2g}} - 1)\,, \tag{5.12}$$

where $n_{t_{2g}}$ denotes the total occupation of the correlated Ti $3d$ $t_{2g}$ orbitals and which gives rise to a shift of the on-site potential,

$$\Delta\epsilon_{i_{\mathrm{Ti}}l_{t_{2g}}} = U(n_{t_{2g}} - \frac{1}{2})\,. \tag{5.13}$$

We point out that the parabolic interpolation of atomic energies for fractional occupation number $n_{t_{2g}}$ inherent in (5.12) and (5.13) is unphysical, at least for $n_{t_{2g}} < 1$ (where the contribution (5.12) becomes negative). Therefore, the Hartree estimate $E^{\mathrm{Hartree}}_{\mathrm{corr}} = Un_{t_{2g}}(M - 1/2)/(2M)$ (for orbital degeneracy $M$) might be considered as an alternative.[12] We will in the following consider the Hamiltonian

$$\hat{H}_{\mathrm{LDA+corr}} = \hat{H}^{0}_{LMTO} + U \sum_{\substack{i',m\sigma,m'\sigma' \\ (m\sigma\neq m'\sigma')}} \hat{n}^{t_{2g}}_{i'm\sigma}\hat{n}^{t_{2g}}_{i'm'\sigma'}\,, \tag{5.14}$$

[11]It is also possible to include exchange interactions analogous to the $J$ term in (5.8) on equal footing (Zölfl, Pruschke, Keller, Poteryaev, Nekrasov, and Anisimov, 2000; Held, Keller, Eyert, Vollhardt and Anisimov, 2001) which may lead to further corrections for LDA at the cost of destroying SU(2) symmetry and introducing additional parameters.

[12]For the calculations presented in this chapter, the double counting correction $\Delta\epsilon$ will be irrelevant (see below).

where $i'$ only sums over Ti sites and the double-counting corrected LDA Hamiltonian is assumed as

$$\hat{H}^0_{\rm LDA} = \hat{H}_{\rm LMTO} - U(n_{t_{2g}} - \frac{1}{2}) \sum_{i',m\sigma} \hat{n}^{t_{2g}}_{i'm\sigma} \, . \tag{5.15}$$

In lack of exact methods for the solution of the general multi-band Hubbard model, approximative solutions are needed just like for pure model calculations. A treatment within the DMFT seems natural and appropriate since it treats the local dynamics induced by the additional local interactions exactly and was also the first choice historically (Anisimov et al., 1997). Lichtenstein and Katsnelson (1998) used DMFT, the fluctuation-exchange approximation (FLEX), and the Hubbard-I approximation alternatively and named the general approach LDA$^{++}$. Drchal, Janiš, and Kudrnovský (1999) combined FLEX with the coherent potential approximation (CPA) for the inclusion of disorder effects.

The DMFT solution of (5.14) is best formulated in reciprocal space, where $\hat{H}^0_{\rm LDA}$ has the matrix elements

$$\big(\hat{H}^0_{\rm LDA}(\boldsymbol{k})\big)_{qlm,q'l'm'} = \big(\hat{H}_{\rm LDA}(\boldsymbol{k})\big)_{qlm,q'l'm'} - \delta_{qlm,q'l'm'}\delta_{ql,q_{\rm Ti}l_{t_{2g}}} U(n_{t_{2g}} - \frac{1}{2}) \, . \tag{5.16}$$

Here, $q$ is an index of the atoms in the elementary unit cell. Generalizing from the 1-band Bravais lattice case discussed in subsection 1.2.3, the lattice problem can be mapped to a multi-band SIAM.[13] Now, on-site lattice Green function $G$, impurity Green function $\mathcal{G}$, and self-energy $\Sigma$ are matrices where each index denotes an orbital within the unit cell, $\alpha \equiv (q,l,m)$. The $\boldsymbol{k}$ integrated lattice Dyson equation and the impurity Dyson equation read for (complex) frequency $z$:

$$\begin{aligned} G^\sigma_{\alpha\alpha'}(z) &= \int \frac{{\rm d}\boldsymbol{k}}{V_{\rm B}} \Big[ z\delta_{\alpha,\alpha'} - \big(H^0_{\rm LDA}(\boldsymbol{k})\big)_{\alpha\alpha'} - \Sigma^\sigma_{\alpha\alpha'}(z)\Big]^{-1} & (5.17)\\ \big[G^\sigma_{\alpha\alpha'}(z)\big]^{-1} &= \big[\mathcal{G}^\sigma_{\alpha\alpha'}(z)\big]^{-1} - \Sigma^\sigma_{\alpha\alpha'}(z) \, , & (5.18)\end{aligned}$$

where $[\dots]^{-1}$ denotes matrix inversion (in $\alpha,\alpha'$). The impurity (or cluster) problem, expressed in mixed notation using imaginary time $\tau$ and index $n$ for Matsubara frequency $i\omega_n$ takes the form:

$$G^{\sigma n}_{\alpha\alpha'} = -\frac{1}{Z} \int \mathcal{D}[\psi]\mathcal{D}[\psi^*] \; \psi^{\sigma n}_\alpha \psi^{\sigma n *}_{\alpha'} {\rm e}^{\mathcal{A}[\psi,\psi^*,\mathcal{G}]} \, , \tag{5.19}$$

with the action

$$\begin{aligned} \mathcal{A}[\psi,\psi^*,\mathcal{G}] &= \sum_{n,\sigma,\alpha,\alpha'} \psi^{\sigma n *}_\alpha \big[\mathcal{G}^{\sigma n}_{\alpha\alpha'}\big]^{-1} \psi^{\sigma n}_{\alpha'} \\ &\quad - \frac{U}{2} \sum_{\substack{m,\sigma,m',\sigma' \\ (m\sigma)\neq(m'\sigma')}} \int_0^\beta {\rm d}\tau \, \psi^{\sigma n^*}_{q_{\rm Ti}l_{t_{2g}}m} \, \psi^{\sigma n}_{q_{\rm Ti}l_{t_{2g}}m} \, \psi^{\sigma' n *}_{q_{\rm Ti}l_{t_{2g}}m'} \, \psi^{\sigma' n}_{q_{\rm Ti}l_{t_{2g}}m'} \, . \qquad (5.20)\end{aligned}$$

[13] In general, the resulting problem appears rather as a finite cluster with on-site interactions and frequency dependent on-site "energies" and "hopping terms" very similar to the systems studied in DCA (cf. App. A.3). Only in absence of hybridization between different orbitals the DMFT mapping leads to a conventional multi-band impurity without mixing terms. Since we will neglect small hybridization terms found in the LDA calculations for $La_{1-x}Sr_xTiO_3$ , however, the cluster will indeed reduce to a single impurity in our calculations.

Only in the last term in (5.20) we have given up generality and thus had to split up the compound orbital index $\alpha$ to single out the strongly correlated Ti $3d$ $t_{2g}$ orbitals.

Both experiment and our LDA calculations (see Sec. 5.4) indicate that the low-temperature properties of $La_{1\text{-}x}Sr_xTiO_3$ are dominated by Ti $3d$ $t_{2g}$ states which hybridize only weakly with other orbitals. Since these states are degenerate, the LDA Hamiltonian is diagonal in this subspace, i.e.,

$$\left(\hat{H}^0_{\mathrm{LDA}}(\boldsymbol{k})\right)_{q_{\mathrm{Ti}}l_{t_{2g}}m,q'l'm'} = \hat{H}^0_{t_{2g}}(\boldsymbol{k})\,\delta_{q_{\mathrm{Ti}}l_{t_{2g}}m,q'l'm'}\,. \tag{5.21}$$

We can thus restrict the explicit treatment to the $t_{2g}$ subspace, in which $G$, $\mathcal{G}$, and $\Sigma$ become diagonal. Due to the special choice $U = V$, $J = 0$, the DMFT problem is then equivalent to a single-band Ising-spin-5/2 Hubbard model, i.e., the derivations of subsection 1.2.3 apply when the replacement $\sigma \to (\sigma, m)$ is made in (1.30)-(1.35). In particular, the LDA input reduces to the $t_{2g}$ DOS (which is supplemented by the estimate for $U$ from constrained supercell LDA calculations).

In this work, we solve the impurity problem using (multi-band) QMC and refer to the resulting general method as LDA+DMFT(QMC). This distinguishes our calculations of photoemission spectra for $La_{1\text{-}x}Sr_xTiO_3$ from earlier work, where analytic approximations have been used for the solution of the impurity problem, i.e., LDA+DMFT(IPT) (Anisimov et al., 1997) and LDA+DMFT(NCA) (Zölfl et al., 2000). The LDA+DMFT(QMC) method was first used by Katsnelson and Lichtenstein (2000) for ferromagnetic iron and applied for the calculation of PES of $Sr_2RuO_4$ by Liebsch and Lichtenstein (2000).

## 5.4 Results for $La_{1\text{-}x}Sr_xTiO_3$

In this section, we present numerical results of the LDA+DMFT(QMC) method applied to $La_{1\text{-}x}Sr_xTiO_3$ for slight doping $x = 0.06$ which drives the system just outside the antiferromagnetic phase and into a paramagnetic strongly renormalized phase. For notational convenience, we will in the following often omit units. Unless stated otherwise, energies are measured in eV, frequencies in eV/$\hbar$, densities of states in $\mathrm{eV}^{-1}$ (or $\hbar$ $\mathrm{eV}^{-1}$), temperatures in eV/$k_{\mathrm{B}}$, imaginary times in $\hbar$ $\mathrm{eV}^{-1}$, and conductivities in arbitrary units.

### 5.4.1 Density of States and Photoemission Spectra

Since the predominant effect of the substitution of La by Sr in $La_{1\text{-}x}Sr_xTiO_3$ is a reduction of the Ti $3d$ $t_{2g}$ band filling while the shape of this band remains essentially unchanged,[14] the LDA spectra were calculated for pure $LaTiO_3$. Furthermore, the orthorhombically distorted crystal structure of this material was replaced by a cubic unit cell. As discussed also in Sec. 5.1, the primary resulting error is an overestimation of the effective bandwidth. The density functional theory problem was solved using

[14]This working hypothesis was later checked by comparison to LDA spectra obtained for doping $x = 0.5$.

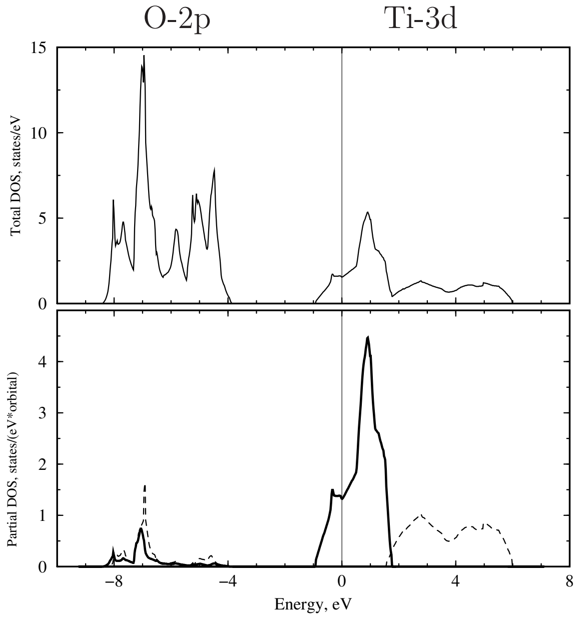



**Figure 5.4:** Upper part: LDA DOS $\rho(\epsilon)$ for $|\epsilon| < 10$ eV. The band is clearly split into contributions with O-2$p$ and Ti-3$d$ character, respectively. Lower part: partial Ti-3$d$ LDA DOS. The $t_{2g}$ contribution (solid line) forms the peak at the Fermi level, the $e_g$ contribution (dashed lines) is primarily located above.

the Stuttgart TB-LMTO-ASA program in version 4.7 (Jepsen and Andersen, 2000) within the basis set Ti$(4s, 4p, 3d)$ La$(6s, 6p, 5d)$ O$(2s, 2p)$ at a Wigner Seitz radius of 2.37 atomic units for Ti. The average on-site interaction for $t_{2g}$ orbitals was calculated as $U \approx 4.2$ eV which contrasts with an estimate of $U \approx 3.2$ eV previously obtained within LMTO-ASA in orthogonal representation (Solovyev, Hamada, and Terakura, 1996). Since LDA is a ground state method, its numerical results apply to temperature $T = 0$.

The low-energy part (measured from the chemical potential) of the resulting full LDA DOS is shown in the upper part of Fig. 5.4. In comparing to the lower part of Fig. 5.4, which displays only the Ti 3$d$ contribution (note the different scales), split up into $t_{2g}$ and $e_g$ parts, we find a situation typical of early[15] transition metal oxides: The band at about -8 to -4 eV corresponds to the completely filled O 2$p$ bands, while the Ti 3$d$ band is located near the Fermi level. Very importantly, the $t_{2g}$ orbitals account for nearly all the weight of the peak at the Fermi energy and have very small overlap with other bands. Therefore, we can disregard all $t_{2g}$ contributions to the DOS below about -1 eV. Then the $t_{2g}$ band has a total width (edge to edge) of about 2.8 eV with a root-mean-square spread in the energy distribution of 0.65 eV. Double counting corrections not yet included in the plot should move the bottom of the $t_{2g}$ band down compared to the rest by at most 2 eV. We note that the O 2$p$ bands also lead to some contribution near the Fermi energy which will also be neglected in the

[15]Here, "early" refers to the position of the transition metal in the periodic table of elements, i.e., denotes a small number of valence $d$ electrons per atom.

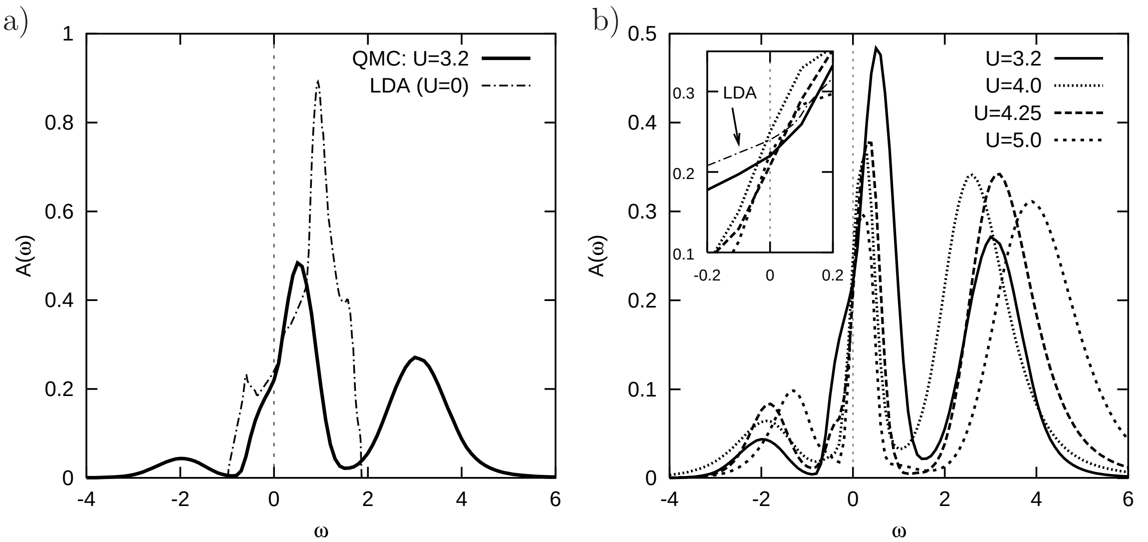



**Figure 5.5: a)** Partial Ti $3d$ $t_{2g}$ LDA+DMFT(QMC) DOS for $T = 0.1$ eV and $U = 3.2$ eV in comparison to LDA. **b)** Partial Ti $3d$ $t_{2g}$ LDA+DMFT(QMC) DOS for $T = 0.1$ for different interaction strengths. Inset: magnified view for small energies (including LDA).

following except when we discuss inverse photoemission (see Fig. 5.14).

For the solution of (5.14) for $La_{1-x}Sr_xTiO_3$ ($x = 0.06$) within the DMFT, the self-consistency equations were iterated to convergence using QMC at a relatively high temperature $k_B T = 0.1$ eV, i.e., $T \approx 1160$ K and with a discretization $\Delta\tau = 0.25\text{eV}^{-1}$. Here, the filling was fixed to $1-x$ by adjusting the chemical potential (set to 0 in the resulting spectra). Then, 40 measurements for the imaginary-time Green function $G(\tau_i)$ on the discretized grid were performed with fixed self-energy. Finally, real-time partial $t_{2g}$ spectra were obtained using Sandvik's MEM program. As seen in Fig. 5.5 a, the explicit inclusion of a finite on-site interaction $U = 3.2$ eV changes the shape of the spectrum near the Fermi energy dramatically: in addition to the quasiparticle peak, a remnant of the noninteracting LDA DOS, lower and upper Hubbard bands appear. The spectral density $A(\omega)$ right at the Fermi energy is almost pinned to its noninteracting value, even at this comparatively high temperature. The result of corresponding calculations is shown in Fig. 5.5 b also for $U = 4$ eV, $U = 4.25$ eV, and $U = 5.0$ eV. The spread of interaction values 3.2 eV $\lesssim U \lesssim$ 5.0 eV reflects the intrinsic uncertainty in the LDA estimate for $U$. Clearly, the quasiparticle weight decreases with increasing $U$. The relatively large difference between the spectra obtained for $U = 4.0$ and $U = 4.25$ indicates significant intrinsic uncertainties which we will discuss in subsection 5.4.2.

We compare our QMC result for $U = 4.0$ with approximate solutions of the DMFT problem in Fig. 5.6 at the same temperature $T = 0.1$ eV, i.e., $T \approx 1160$ K. Most prominently, IPT misses the quasiparticle peak at this temperature since this approximation underestimates the associated Kondo temperature [cf. related observations for a three-band Hubbard model with Bethe DOS by Kajueter and Kotliar (1997)]. Moreover, the tails of the IPT spectral functions appear to have too much

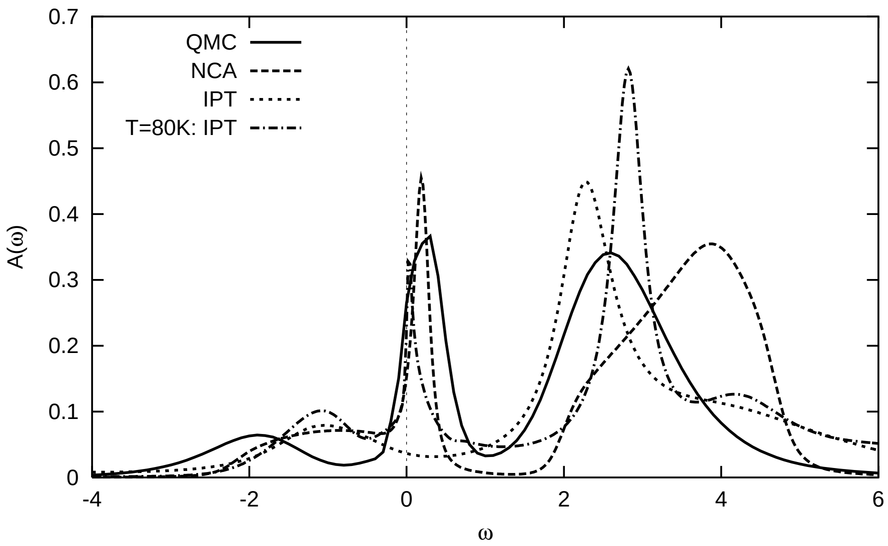



**Figure 5.6:** Partial Ti $3d$ $t_{2g}$ LDA+DMFT(X) DOS for $T = 0.1$ eV and $U = 4$ eV for different approaches $X$ for the impurity problem. Also shown is the LDA+DMFT(IPT) result for $T = 0.008$ eV.

weight. When IPT is applied for lower temperatures $T \approx 80$ K, the quasiparticle peak is retained, but with far too small weight; the overall agreement with QMC remains fair. NCA comes off much better in comparison. However, this approximation underestimates both the weight and the width of the quasiparticle peak, cannot resolve the lower Hubbard band (but produces a shoulder too close to the quasiparticle peak instead), and shows strong deviations in the shape of the upper Hubbard band. The multi-band generalizations of NCA will usually fail at even higher temperatures than conventional NCA due to violations of proper time order in processes involving several orbitals. From the numerical comparison we conclude that the method of solution of the DMFT equations does matter.

In photoemission spectroscopy (PES), the surface of a probe is irradiated with photons of a well-defined energy above a threshold given by its work function. The intensity distribution of emitted photo electrons as a function of electron energy and angle (with energy and angle of incident photons as well as temperature, pressure etc. as parameters) is then characteristic of the material.[16] Conventionally, a distinction is made between X-ray photoemission (XPS, previously also called electron spectroscopy for chemical analysis, ESCA) and ultraviolet photoemission (UPS) depending on whether the photons originate from X-ray tubes or helium discharge lamps; more recently, however, synchrotrons provide for highly luminous sources of photons with extremely narrow energy distribution which is continuously tunable over both the XPS and UPS energy range. Within the "sudden approximation" (no

[16]While Einstein explained the photo effect in 1905 in terms of quantized radiation, i.e., photons (Nobel prize 1921), its application as an analysis tool for chemical and electronic structure of solids is due to pioneering work by Siegbahn and coworkers carried out since the fifties (Nobel prize 1981).

interaction of the outgoing photo electron with the remaining hole) and second order response theory one obtains an expression of the Fermi golden rule form for the steady photo current (Schaich and Ashcroft, 1971; Bansil and Lindroos, 1999),

$$I(E_i, \hbar\omega) = \frac{2\pi e}{\hbar} \sum_{i,f} \left|\langle\psi_f|\hat{V}|\psi_i\rangle\right|^2 \delta(E_f - E_i - \hbar\omega)\,. \tag{5.22}$$

Here, $\psi_i$ ($\psi_f$) are initial (final) states with energy $E_i$ ($E_f$) of the semi-infinite solid, $\hat{V} = -e(\hat{\boldsymbol{p}} \cdot \boldsymbol{A} + \boldsymbol{A} \cdot \hat{\boldsymbol{p}})/(2mc)$ is the interaction with the electromagnetic field, and full initial as well as empty final states have been assumed. For localized electron orbitals, both the dipole selection rules and the transport to the surface determine the matrix elements $\langle\psi_f|\hat{V}|\psi_i\rangle$. Neglecting both its structure and assuming a featureless continuum of final states one can draw a direct connection between PES and the single-particle spectrum $A(E) = -\frac{1}{\pi}\mathrm{Im}\, G(E)$ (at temperature $T$):

$$I(E) \propto \int_{-\infty}^{\infty} \mathrm{d}E'\, A(E') \frac{1}{1 + \mathrm{e}^{-(E'-\mu)/(k_B T)}} R_\Delta(E - E') + B \tag{5.23}$$

Here, $R_\Delta(E - E') = \mathrm{e}^{-(E-E')^2/(2\Delta^2)}/\sqrt{2\pi\Delta^2}$ is a Gaussian broadening function associated with the experimental resolution,[17] and $B$ a background contribution arising from photo electrons which have been inelastically scattered within the solid. We stress that (5.23) is at best a crude approximation for contributions arising from orbitals within a few top layers of the material. Photo electrons generated in the bulk cannot escape since the typical mean free path for photo electrons in the energy range of 10 to 100 eV in crystals is only a few nanometers. Within such a close distance to the surface, however, surface effects may be important, at least quantitatively. Partially, the importance of final-state and transport effects can be assessed by measuring the same spectra using different incident photon energies.

In Fig. 5.7, we compare theoretical PES calculated via (5.23) for $La_{1\text{-}x}Sr_xTiO_3$ with a broadening width $\Delta = 0.3$ eV (and $B = 0$) with two experimental results. In one experiment, polycrystalline samples of $LaTiO_{3.03\pm0.01}$, characterized by thermogravimetry analysis, were cooled down to $T = 80$ K, scraped in situ, and irradiated with photons of the energy $h\nu = 48$ eV. Here, the stated resolution is $0.2-0.3$ eV (Fujimori, Hase, Namatame, Fujishima, Tokura, Eisaki, Uchida, Takegahara and de Groot, 1992). In a more recent experiment, PES for the nominally same material grown by the floating-zone method were measured at 23 K and 150 K using a He discharge lamp ($h\nu = 21.2$ eV) with a nominal resolution of 30 meV (Yoshida et al., 2002).[18] Qualitatively, good agreement is observed for $U = 5$ eV, the strongest interaction value being shown in Fig. 5.7: both the peak position at $-1.5$ eV $\lesssim E \lesssim -1$ eV and the suppression of spectral weight near the Fermi energy are

[17] The spectra are also Lorentzian broadened by the finite lifetime of the final states, i.e., the decay of the photo holes left behind in the bulk.

[18] Note that here the authors claim to see antiferromagnetic order even for the doped system with $T_N \approx 112$ K. Thus, only the measurement at $T = 150$ K corresponds to the paramagnetic phase modeled in our simulations.

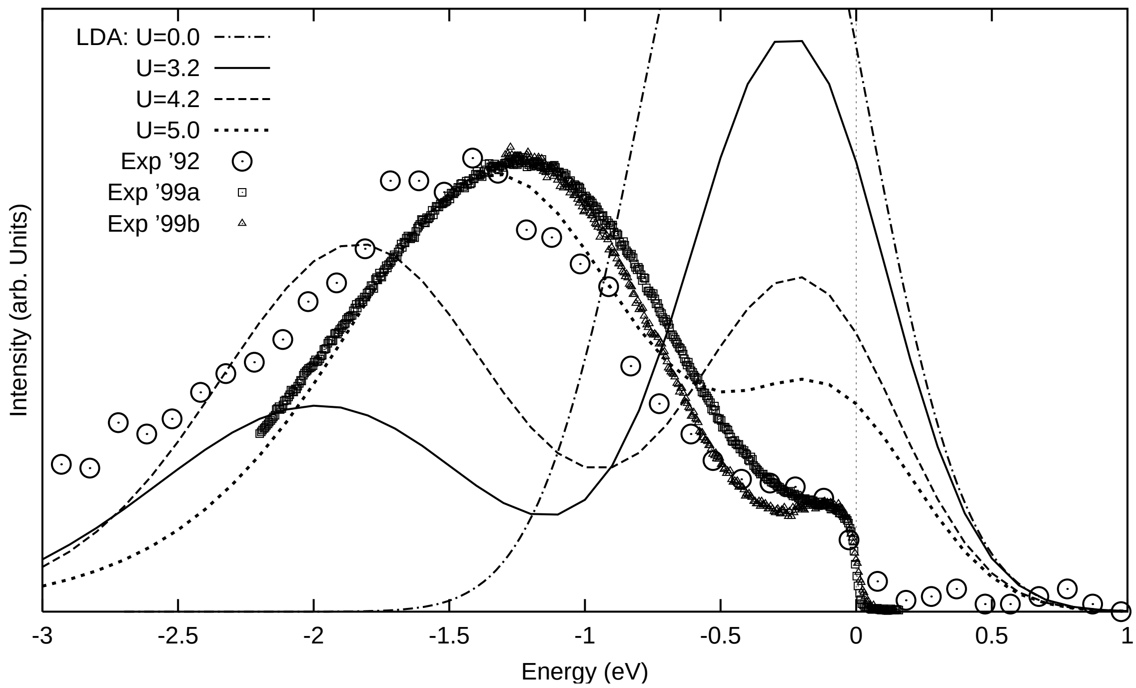



**Figure 5.7:** $La_{1-x}Sr_xTiO_3$ photoemission spectra (x=0.06): LDA+DMFT(QMC) estimates for different $U$ (QMC simulations at $T = 1160$ K, Fermi function for $T \approx 0$ K) in comparison with experiments at $T = 80$ K (Fujimori, Hase, Namatame, Fujishima, Tokura, Eisaki, Uchida, Takegahara and de Groot, 1992), $T = 23$ K (a), and $T = 150$ K (b) (Yoshida et al., 2002). The agreement with experiments improves with increasing $U$ (used in the simulation).

predicted correctly. Although the relative weight originating from the quasiparticle peak is computed to be almost a factor of 5 higher than seen experimentally, the improvement over LDA is impressive. The value of $U \approx 5$ eV which best reproduces the experimental spectra is a bit higher than estimated from LDA, but still within the error range, in particular when the reduction of band width by orthorhombic distortion is considered. The remaining differences may be due to disorder,[19] to a reduced hybridization at the surface, to deficiencies of the theoretical PES treatment, to simplifications inherent in the electronic model, to the too high temperature used in the QMC cimulations, and to numerical errors related to the QMC and MEM treatment of the DMFT problem. This last technical point will be addressed in the following subsections where we also leave the ground of published collaborative work.

We point out that due to the high spectral weight present in $\rho(\epsilon)$ just above the Fermi energy (cf. Fig. 5.5), the spectral weight near the Fermi energy of the resulting PES is very sensitive to the exact position of the Fermi energy. In fact,

[19] Note that, while the experimentally realized form of doping with 3% O should correspond to 6% doping with Sr, the materials are clearly not fully equivalent electronically. In particular, disorder should have much stronger effects for $LaTiO_{3.03\pm0.01}$.

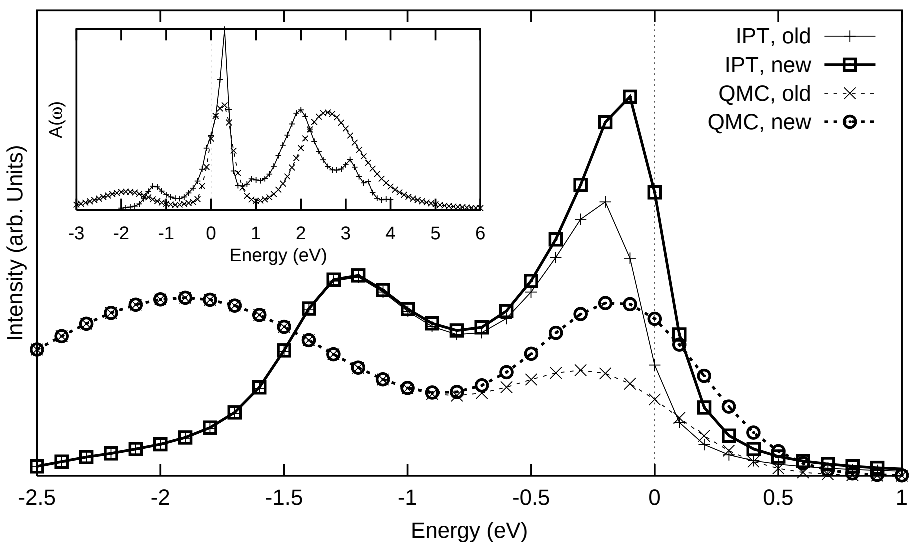



**Figure 5.8:** Correction of erroneous shift of the chemical potential in (5.23) to -0.1 eV. The "old" IPT PES ($U = 2$ eV, $T = 125$ K) published in (Anisimov et al., 1997) is incorrect, i.e., does not correspond to the DOS computed in the paper (see inset). A corrected curve ("new") shows a much more prominent peak at the Fermi level. A recalculation of the QMC PES for $U = 4$ eV and $T = 1160$ K shown in our preprint (cond-mat/0005207v1) leads to similarly significant changes.

---

an error present in the broadening program [i.e., an implementation of (5.23)] used by Anisimov's group, i.e., an erroneous shift of the chemical potential by one grid point, led to strong deviations for the PES when compared to a new (correct) code written as part of this thesis. This observation necessitated new calculations and explains the differences between a first preprint (cond-mat/0005207v1) and the actual publications (Nekrasov, Held, Blümer, Poteryaev, Anisimov, and Vollhardt, 2000; Held, Nekrasov, Blümer, Anisimov, and Vollhardt, 2001) of the work presented in this subsection. While in earlier LDA publications the problem should have been virtually unobservable due to a finer frequency grid, it clearly shows up in the first LDA+DMFT publication (Anisimov et al., 1997). This is seen in Fig. 5.8: both in IPT for $U_{\text{eff}} = 2$ eV, $T = 125$ K and the QMC calculation for $U = 4.0$ eV, $T = 1160$ K the weight of the peak at the Fermi energy was underestimated by about a factor of 2.

### 5.4.2 Influence of Discretization Errors

While the QMC algorithm used for the treatment of the LDA+DMFT problem in this chapter can in principle yield results of arbitrary precision in the limit of infinite computing resources, it is clear that statistical noise, incomplete convergence of the self-consistency cycle and the Trotter discretization error $\Delta\tau$ in conjunction with approximations made for the Fourier transform (see chapter 3) may impact the results of actual calculations. Even worse, for a finite discretization $\Delta\tau$, the incomplete information on the imaginary axis has to be supplemented by a maximum entropy assumption in order to obtain real-frequency spectra. Obviously, the qualitative importance of these errors depends on the final quantity that is calculated. For the PES shown in the previous subsection, the broadening procedure renders details of the initial spectra relatively unimportant. In contrast, the computation of optical conductivity data which we will present in subsection 5.4.3 involves comparisons with the noninteracting DOS and is far less well conditioned: in the computation of $\Sigma(\omega)$, sums can essentially cancel so that the relative error explodes. Thus, reliable results for $\sigma(\omega)$ can only be expected if both the QMC procedure and the MEM algorithms fulfill high standards.

Although the particular algorithm for analytic continuation of imaginary-time QMC data used for obtaining the spectra shown in Fig. 5.5 yields spectra compatible with the measured data, it takes improper account of the associated errors. Consequently, the spectra selected by the algorithm may not be the best maximum entropy spectra. In extensive studies, we implemented more accurate MEM schemes (still using Sandvik's code), which we will characterize below, and also tested a variety of default models.[20] Still, none of the methodological improvements yielded significantly more reliable spectra and derived self-energies. In order to check for convergence, to improve on the statistical accuracy, and to generate enough input data for Jarrell's fully consistent maximum entropy program, we performed additional longer QMC simulations for the same parameter sets. Again, no general improvement was found using the variants of Sandvik's scheme. Jarrell's program failed to yield sensible spectra at all. In our view, the only likely explanation for the fact that an improved treatment of statistical errors did not in general lead to improved spectra is that statistical errors are not the important ones; instead they are dominated by systematic errors. After having established convergence and well thermalized Ising spin ensembles in the QMC simulations, the only remaining candidate as a source for such systematic errors is the finite imaginary-time discretization $\Delta\tau$ in connection with the empirical "smoothing" scheme employed for the discrete Fourier transformation. As discussed in Sec. 3.4, this scheme leads to large errors near a metal-insulator transition in the half-filled single-band case. More generally, one has to expect problems also in the present multi-band case with noninteger filling when correlations become strong. While we cannot test for the influence of the "smoothing" trick alone without rewriting the QMC program, the total systematic error including the unavoidable Trotter error may be estimated from changing the value of $\Delta\tau$, which we will do in the remainder of this subsection.

[20] In the present multi-band case, the choice of a nontrivial appropriate default model is much more difficult than in the symmetric single-band case covered in Sec. 3.8. As an alternative to the previously used flat default model (energy range -20 eV $\leq \omega \leq$ 20 eV with a grid of 0.1 eV) we used Gaussian default models of various widths and with the symmetry axis approximately matched to the band center as well as tighter models derived from a rigid-band assumption which becomes good in the atomic limit. While it is relatively easy to suppress the unphysically long negative-energy tails visible, e.g., in Fig. 5.5 with such default models, a too tight cut-off at high frequencies is likely to lead to artifacts in the main part of the spectrum. For an improved treatment, it seems promising to use QMC estimates for double and triple occupancy for fine-tuning the default model for each particular data set.

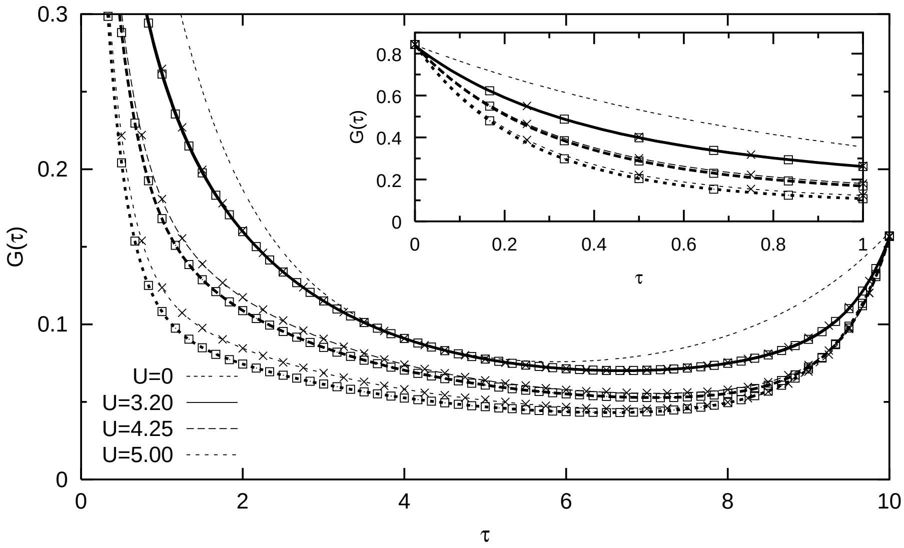



**Figure 5.9:** Imaginary time Green function $G(\tau)$ for $\Delta\tau = 0.25$ (crosses) or $\Delta\tau = 0.167$ (squares). The lines connecting the data points correspond via (1.55) to the spectra shown in Fig. 5.5 and Fig. 5.12, respectively. The curve "$U = 0$" shows the bare LDA estimate for $G(\tau)$ for comparison.

---

In Fig. 5.9 we compare the imaginary-time Green function $G(\tau)$ as previously calculated for $\Delta\tau = 0.25$ with the result of new measurements for $\Delta\tau = 0.167$, i.e., with the number of time slices increased from 40 to 60. This step requires a 3-4 fold increase of computer time and should eliminate roughly between a third (for a linear extrapolation law in $\Delta\tau$) and half (for a quadratic extrapolation law) of the original systematic error. For "weak" interaction, i.e., for $U = 3.2$ of the order of the bandwidth, differences upon reducing the discretization error are hardly detectable on the scale of the figure (while being significant at least for $1 \lesssim \tau \lesssim 2$). In contrast, the improved accuracy clearly changes the results for $U = 4.25$ and, particularly, $U = 5$. It is important to realize that statistical errors are of the order of the linewidths (of the thin lines) and that the full systematic errors are expected to equal 2 or 3 times the observed differences. Thus, systematic errors exceed statistical errors by more than an order of magnitude which fully explains all the problems encountered with maximum entropy analytic continuation described above.[21] The solid lines connecting the $\Delta\tau = 0.25$ data points shown in Fig. 5.9 have been obtained by applying (1.55) to the spectra shown in Fig. 5.5; practically indistinguishable curves were obtained, however, by also using "backward" analytic continuation (1.55) for spectra obtained in variations of the maximum entropy method. Physically, the reduced values for $G(\tau)$ at $\tau \approx \beta/2$ point towards decreasing spectral weight near the Fermi energy, i.e.,

[21]We note that the statistical error of the self-energy effectively gives rise to a systematic error for the measurements of $G(\tau)$ since the self-energy defining the bath green function was kept fixed for all measurements. This error also exceeds the statistical error calculated in the measurements but is (for $U = 4.25$ or $U = 5.0$) well below the $\Delta\tau$ error.

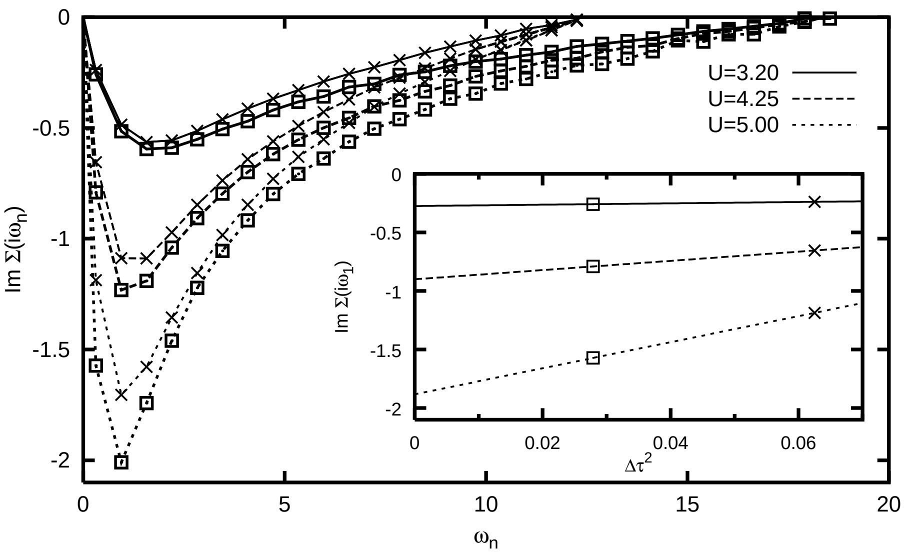



**Figure 5.10:** Imaginary part of self-energy for $\Delta\tau = 0.25$ (crosses) or $\Delta\tau = 0.167$ (squares). Inset: Extrapolation $\Delta\tau \to 0$ of the value at the first Matsubara frequency $\omega_1 = \pi T$ assuming the best case of a quadratic law. Lines are guides to the eye only.

to an increase of correlation effects.

The discretization error leads to even more prominent changes in the imaginary part of the self-energy on the imaginary axis $\operatorname{Im}\Sigma(i\omega_n)$ shown in Fig. 5.10. On the one hand, the Nyquist cutoff frequency where the computed self-energy unphysically vanishes (instead of decaying like $1/\omega$) is pushed to higher frequencies; on the other hand, the absolute value of $\operatorname{Im}\Sigma$ increases significantly at small frequencies upon decreasing $\Delta\tau$. The latter tendency is, again, a sign that correlation effects are stronger than originally estimated. In fact, the quadratic extrapolation suggests that the quasiparticle weight for $U = 5$ is $Z \approx 0.15$ rather than the value $Z \approx 0.21$ measured for $\Delta\tau = 0.25$.

Due to lack of computing resources in the final stages of this work we could not obtain sufficient data for a controlled application of Jarrell's MEM program. It would also seem more efficient to eliminate systematic errors first before taking details of the statistical error distribution too seriously. Instead, we continued to use the previously applied scheme: for a data set of $50 \lesssim N \lesssim 100$ measurements (bins) of $G^\sigma_\alpha(\tau_i)$ for band index $\alpha \in \{1, 2, 3\}$, spin $\sigma \in \{\uparrow, \downarrow\}$, and time slice $0 \le i \le \Lambda = 60$, correlations between measurements at different $\tau_i$ and between different spins at the same time slice are neglected when computing averages $G^{\text{av}}_\alpha(\tau_i)$ and the associated errors $\Delta G_\alpha(\tau_i)$ for each band separately.[22] Each band is analytically

[22]Here, a more proper treatment would at least average over spins before calculating errors since the spin-averaged quantities fluctuate much less. In fact, the cleanest approach at this point would involve averaging over spins and orbitals at each time slice before starting the error calculations since we seek information about the paramagnetic phase. On the other hand, the increased error bars generated by the procedures outlined in the text partially mimic systematic errors and can thus limit the impact of unphysical $\Delta\tau$ errors on the resulting spectra.

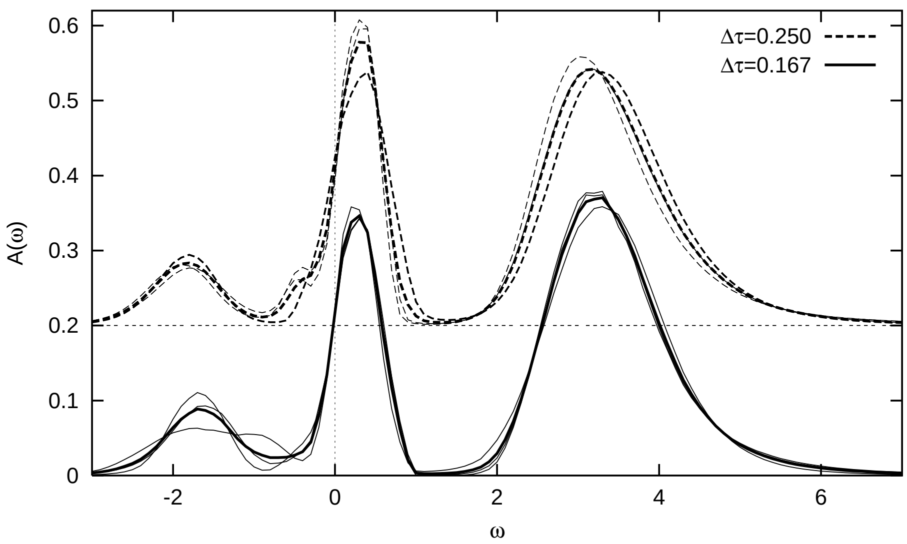



**Figure 5.11:** Illustration of the MEM strategy "b" used throughout this chapter: data sets (here for $U = 4.25$) for each degenerate $t_{2g}$ band are analytically continued independently to the real axis (thin lines); the partial spectra are then averaged (thick lines). For enhanced clarity of the plot, the ordinate has been shifted for $\Delta\tau = 0.25$ by 0.2.

continued independently. An average over all three bands then defines the final result for each parameter set. As illustrated in Fig. 5.11 for $U = 4.25$ (and $\beta = 10$), the partial spectra (thin lines) show substantial deviations from each other and from the average (thick lines) for both $\Delta\tau = 0.25$ and $\Delta\tau = 0.167$. The observed differences may be regarded as a lower bound on the intrinsic errors. While in general these deviations are not significantly reduced for our new data, the accuracy at small frequencies $|\omega| \lesssim 0.5$ seems improved.

A comparison of the old and new spectra for $U = 3.2$, $U = 4.25$, and $U = 5.0$ is presented in Fig. 5.12. In addition to the data for $\Delta\tau = 0.25$ (same as shown in Fig. 5.5) and data for $\Delta\tau = 0.167$ computed according to the scheme detailed in the last paragraph and denoted as "b", a tighter fit was obtained for comparison by averaging the data over all bands first and applying MEM to the resulting single-band problem ("a"). We stress that the new spectra still leave room for improvement (even for fixed $\Delta\tau = 0.167$); still, the observed changes, i.e., the shift in peak positions and the reduced spectral weight at small frequencies for large $U$ represent genuine improvements over the previously obtained estimates. Unless specified otherwise, we will in the following show results derived from the set "b" of spectra.

**Improved Results** Figure 5.13 depicts photoemission spectra derived from the new spectral functions with $\Delta\tau = 0.167$ (thick lines) in comparison with those already shown in Fig. 5.5. For the new estimates we also used the temperature $T = 1160$ K in the Fermi function in order to fulfill the sum rule: when the spectra are properly normalized, the integral over the PES computed according to (1.55) equals the band

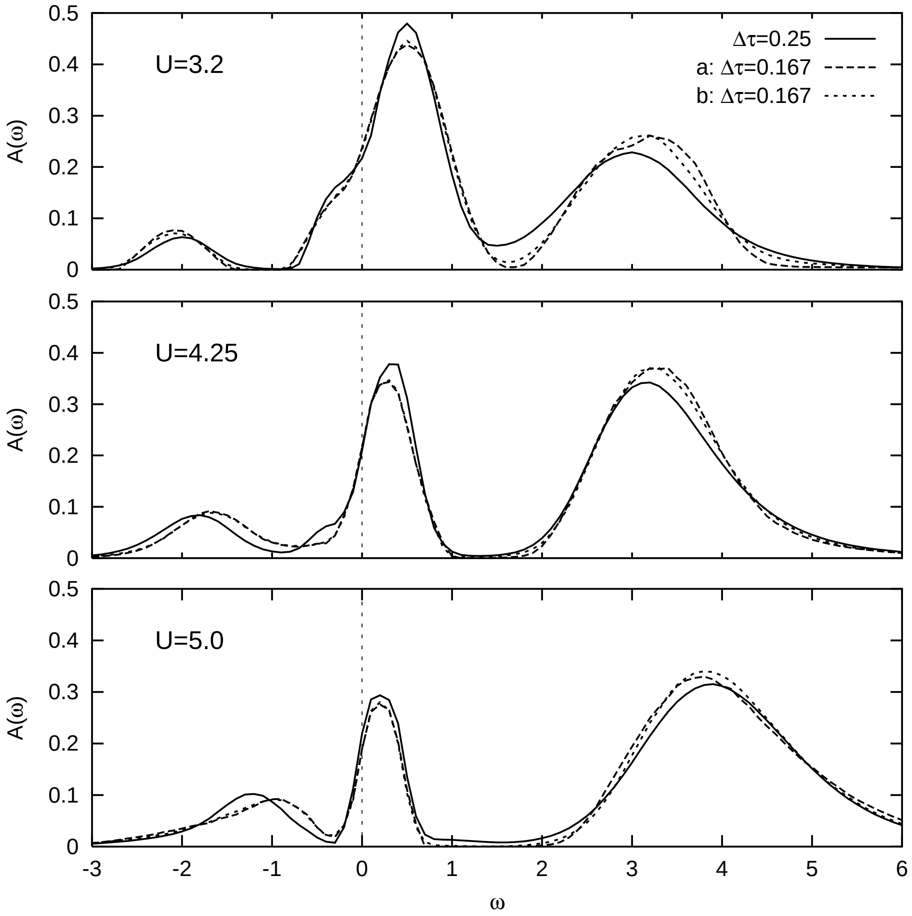



**Figure 5.12:** Spectra (interacting DOS): new results for $\Delta\tau = 0.167$ using variations of the MEM technique (see text) in comparison to the spectra for $\Delta\tau = 0.25$ shown in Fig. 5.5.

filling to very high accuracy.[23] The increased accuracy makes the peaks at the Fermi energy less prominent; this correlation effect would appear even stronger had both sets of curves been calculated using the same Fermi function. Furthermore, the shape and position of the lower Hubbard bands are modified and the peak positions become more systematic as a function of $U$.

The upper part of the interacting DOS is in principle also observable experimentally via inverse photoemission, which we model by formally applying (5.23) for negative temperatures. In this case, the noninteracting $e_g$ bands situated above the LDA Fermi energy become very important as is evident in Fig. 5.14. Here, we have used $U = 5.0$ and ignored a possible shift of the Fermi energy, which would also shift the $t_{2g}$ band relative to the remainder of the spectrum. The total PES includes a contribution associated with the O $2p$ band which then fills the gap in the $t_{2g}$ contribution at about 1 eV. In lack of true inverse PES experiments, we compare to

[23]The estimates shown in Fig. 5.7, using $T \ll \Delta\omega = 0.1$ eV in the Fermi function, corresponded to significantly lower band fillings. Therefore, they had been rescaled in (Nekrasov et al., 2000).

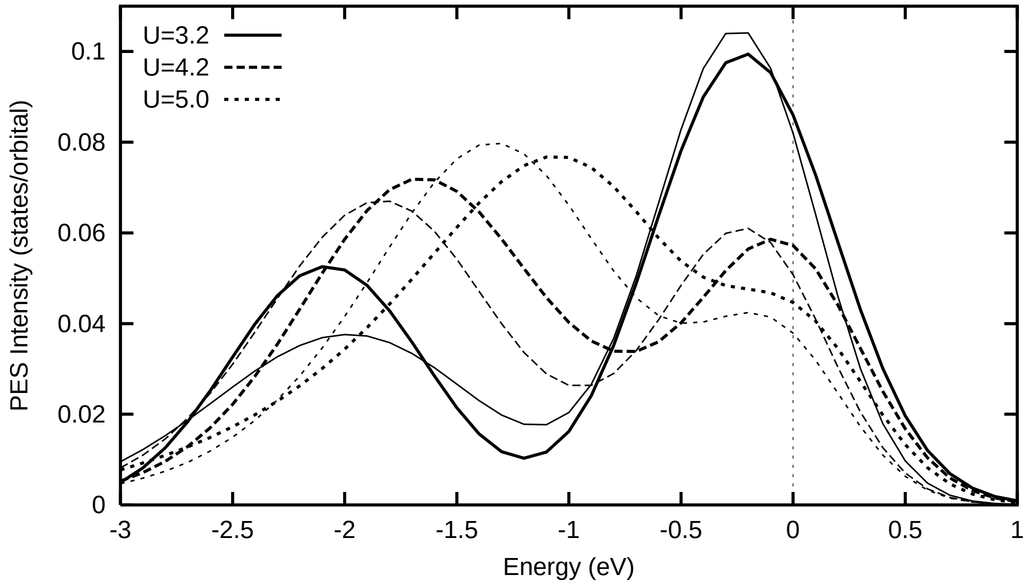



**Figure 5.13:** $La_{1\text{-}x}Sr_xTiO_3$ photoemission spectra (x=0.06): LDA+DMFT(QMC) estimates for $\Delta\tau = 0.167$ and $T = 1160$ K (thick lines) in comparison to the previously obtained results shown in Fig. 5.7 (thin lines). The longer tails at positive energies are due to our use of $T = 1160$ K also in the Fermi function (see text).

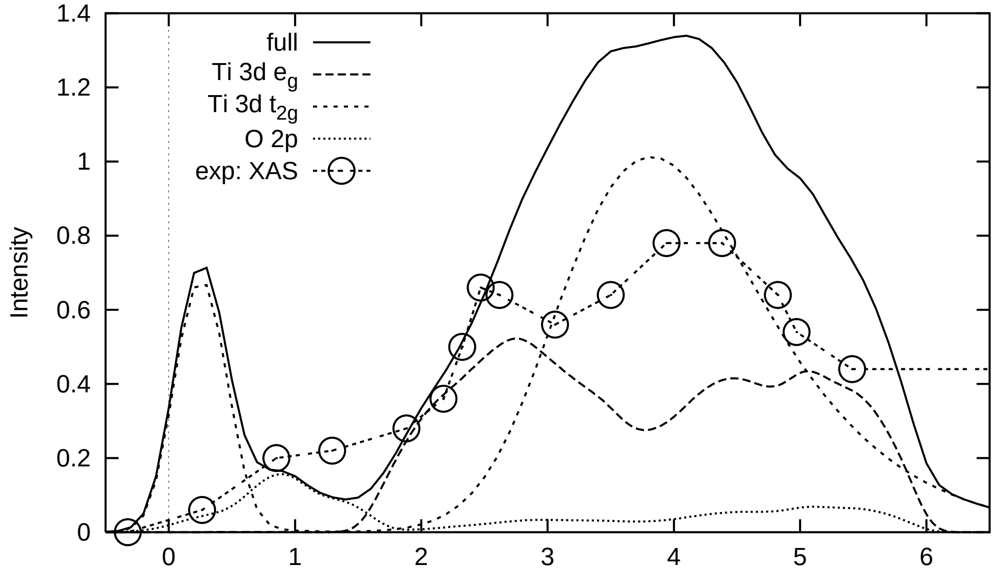



**Figure 5.14:** Inverse photoemission: full theoretical estimate and contributions arising from the correlated $t_{2g}$ band as well as uncorrelated $e_g$ and O $2p$ bands in comparison with experimental XAS data (Abbate et al., 1991). For the theoretical curves, $U = 5$ eV and a fixed Fermi energy (see text) have been assumed.

results of X-ray absorption spectroscopy (XAS) measurements. In this technique, incident X-ray photons excite core electrons which, in a crude approximation, give rise to the inverse photo effect.[24] In Fig. 5.14, we have adjusted the intensity scale of the experimental curves (Abbate et al., 1991) to be of the order of magnitude of the theoretical curves and have matched the (experimental) absorption edge with the (calculated) Fermi energy. While the XAS peak observed at 2 eV $\lesssim \omega \lesssim$ 5 eV is in good correspondence with the combined contributions of the upper Hubbard band and the uncorrelated $e_g$ band, the low intensity at the Fermi energy seems incompatible with the theory. Here, the situation is similar to PES, where the theory also predicts more weight near the Fermi level than is seen experimentally.

### 5.4.3 Optical Conductivity

Another interesting spectral property is the optical conductivity $\sigma(\omega)$ which can be determined experimentally via measurements of the reflectivity and a Kramers-Kronig transformation or, more directly, by transmission experiments (see Sec. 4.1). Within the DMFT, the diagrammatic expression for $\sigma(\omega)$ reduces to a bubble diagram which can be evaluated given the local self-energy and the quantity $\tilde{\rho}(\epsilon)$ defined in (2.2). The latter deviates from the noninteracting DOS $\rho(\epsilon)$ by a factor $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ which takes the nonconstant Fermi velocity within the Brillouin zone into account (see chapter 4). While it should be possible to compute $\tilde{\rho}(\epsilon)$ within LDA, this important ingredient for computations of $\sigma(\omega)$ has not (yet) been obtained within our collaboration. In this situation, the general formalism developed in Sec. 2.3 provides for a reasonable starting point by interpreting the LDA DOS $\rho^{\mathrm{LDA}}(\epsilon)$ as the defining property of an isotropic tight-binding model in $d = \infty$. For this model, one can then compute $\tilde{\rho}(\epsilon)$ and thereby $\sigma(\omega)$ on the basis of estimates for $\Sigma(\omega)$ derived from the QMC/MEM spectra. For the original three-dimensional cubic model, this method introduces an additional approximation which (like the DMFT) also becomes exact in high dimensions. This is illustrated in Fig. 5.15 a for the $t-t'$ lattice with $a^* = -0.3333$ (cf. Fig. 2.5 and subsection 4.5.3). Here, the exact evaluation of $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ in three and four dimensions (thick lines) is seen to agree well with the result of an application of (2.62) to the exact finite-dimensional DOS (thin lines), in particular in the region $\epsilon \lesssim 2$ which comprises about 95% of the DOS. The application of this $d = \infty$ approach is also far more accurate than assuming a constant $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon) \equiv 1$ (which would not be correct for any finite-dimensional lattice).

In Fig. 5.15 b we can see that our formalism, i.e., the application of (2.62) to the LDA DOS indeed yields an estimate for $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ which vanishes at the band edges as required physically.[25] Consequently, $\tilde{\rho}(\epsilon)$ is suppressed at the band edges compared to $\rho^{\mathrm{LDA}}(\epsilon)$. The suppression of peaks is also a realistic effect which, however, might

[24] Clearly, matrix elements and selection rules become even more important for this second order process (in the number of electron-photon vertices) than for the regular PES discussed in subsection 5.4.1.

[25] In order to correct for the bandwidth of 0.65 (square-root of the variance of the LDA energy distribution) instead of unity, the expressions for $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ and $\tilde{\rho}(\epsilon)$ stated in Sec. 2.3 have to be multiplied by the squared inverse bandwidth. Only then is the average hopping distance given by the norm of $\tilde{\rho}(\epsilon)$.

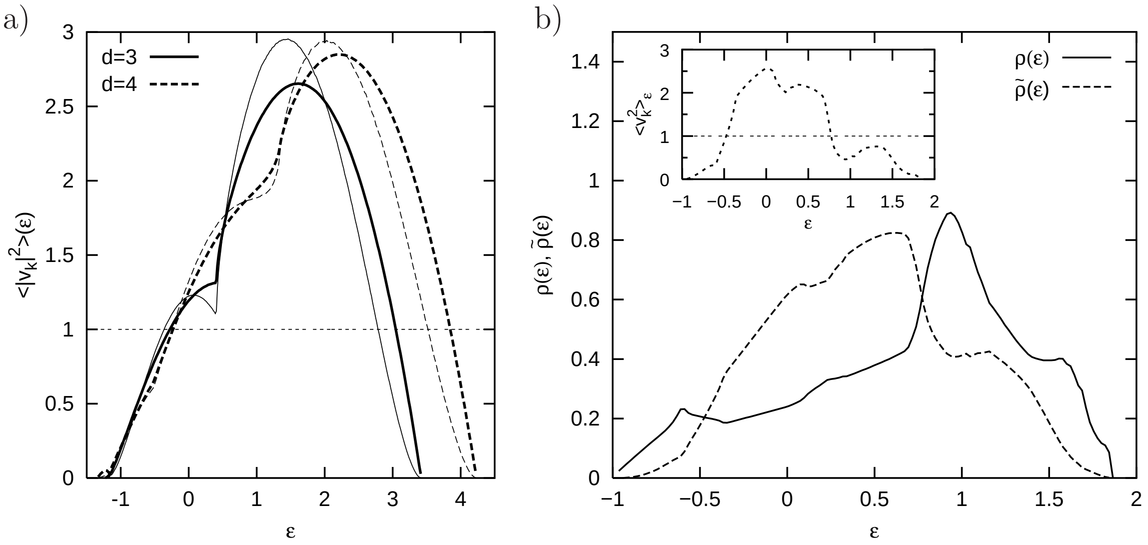



**Figure 5.15: a)** Exact $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ of the $t-t'$ lattice with $a^* = -0.3333$ (thick lines) in comparison with a computation via (2.62) from the DOS (thin lines). **b)** LDA $t_{2g}$ DOS $\rho(\epsilon)$ (solid line) for La$_{1\text{-}x}$Sr$_x$TiO$_3$, derived energy-dependent $d=\infty$ expectation value of the squared Fermi velocity $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ (inset), and $\tilde{\rho}(\epsilon) = \langle|\boldsymbol{v_k}|^2\rangle(\epsilon)\,\rho(\epsilon)$ (dashed line in main figure). The average squared hopping distance is 1.10.

be less pronounced in finite dimensions. Assuming that the LDA spectrum used throughout this chapter is correct, we would therefore expect the true curve for $\tilde{\rho}(\epsilon)$ to lie between both curves shown in the main part of Fig. 5.15 b. Thus, we will for comparison also compute $\sigma(\omega)$ for constant $\langle|\boldsymbol{v_k}|^2\rangle$ and regard the deviations as an upper bound to intrinsic errors.

As a set of intermediate results, we show estimates for the imaginary part of the self-energy on the real axis, $\operatorname{Im}\Sigma(\omega)$ in Fig. 5.16. These were calculated by numerical inversion of the lattice Dyson equation (1.30) using a program written as part of this thesis. Our algorithm (see also subsection 4.6.1) implements a modified Newton search with underrelaxation, i.e., the step width is reduced by a factor of 0.5 compared to the usual Newton method.[26] Note that the Dyson equation is local[27] in $\omega$ and that the Jacobian needed for the Newton method (with 2 dimensions corresponding to the real and imaginary parts, respectively) can easily be computed numerically. When no causal solution exists for $\Sigma(\omega_0)$ given $G(\omega)$ at an isolated frequency point $\omega_0$, the program issues a warning and interpolates the solution using the adjacent grid points: $\Sigma(\omega_0) = \big(\Sigma(\omega_0 - \Delta\omega) + \Sigma(\omega_0 + \Delta\omega)\big)/2$.

[26] Analogous to the procedure used in Sec. 4.4, the MEM spectra (with a frequency resolution of $\Delta\omega = 0.1$) have been interpolated (here to $\Delta\omega = 0.02$) before performing the Kramers-Kronig transformation and inverting the Dyson equation.

[27] By "local" we here mean that $\Sigma(\omega_0)$ is for fixed noninteracting DOS only a function of $\omega_0$ and the complex quantity $G(\omega_0)$. Therefore, the inversion of the Dyson equation can be done independently for different frequencies $\omega_0$. The relation between $\Sigma(\omega)$ and $A(\omega) = -\frac{1}{\pi}\operatorname{Im} G(\omega)$ is not local, however, since the Kramers-Kronig transform relates $\operatorname{Re} G(\omega_0)$ to $\operatorname{Im} G(\omega)$ for all $\omega$.

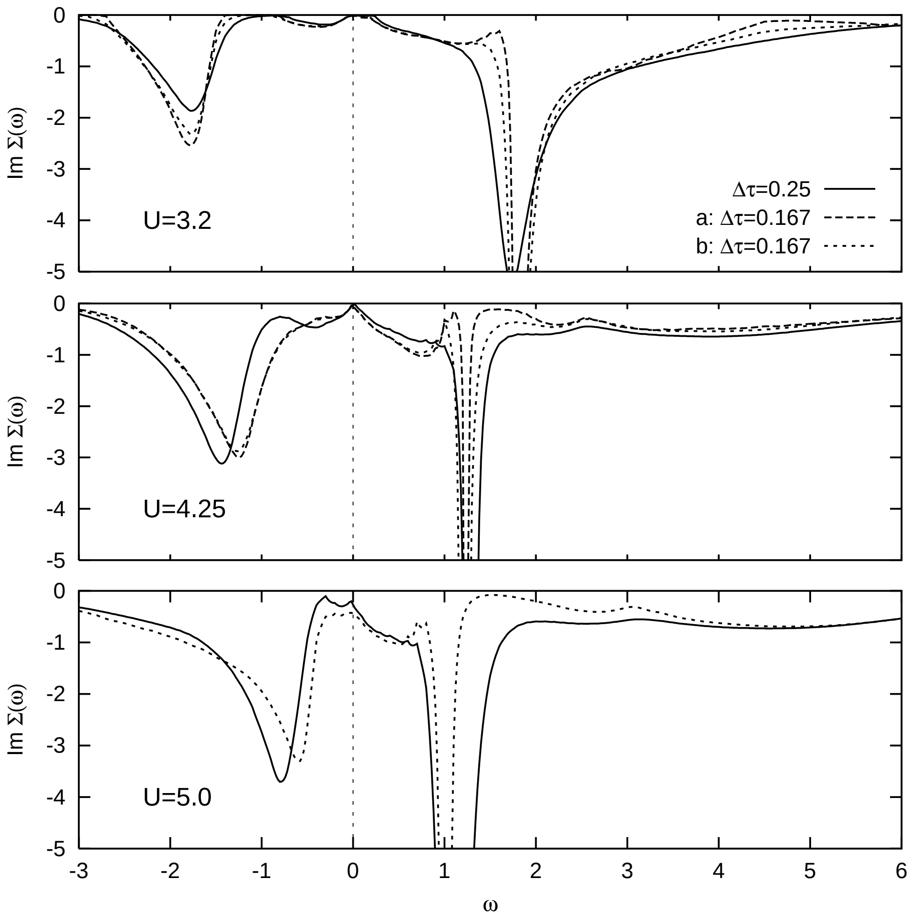



**Figure 5.16:** Imaginary part of the self-energy on the real axis computed from the MEM spectra shown in Fig. 5.12. The variations "a" and "b" of the MEM scheme are explained in the discussion of Fig. 5.12.

---

Significant changes are found in the results for $\Sigma(\omega)$ as shown in Fig. 5.16 when the discretization is reduced from $\Delta\tau = 0.25$ to $\Delta\tau = 0.167$. In particular, the shapes of $\Sigma(\omega)$ near the Fermi energy come closer to the expected quadratic form for the more accurate data corresponding to $\Delta\tau = 0.167$. In comparison, differences observed from variations of the MEM scheme ("a" and "b") are small except near singularities where the precise value of $\Sigma$ is less important.[28]

The corresponding estimates for $\sigma(\omega)$ are presented in Fig. 5.17 on a semilogarithmic scale, using the lattice input $\tilde{\rho}(\epsilon)$ from Fig. 5.15 b. Evidently, the $\Delta\tau$ effects are important quantitatively; in particular, peak positions shift and the Drude peak decreases for $\Delta\tau = 0.167$ (and $U \geq 4.25$). Thus, correlation effects are treated more adequately at lower $\Delta\tau$.

Figure 5.18, the main result of this subsection, shows our best estimate for the

[28]For $U = 5.0$, no reliable causal solution $\Sigma(\omega)$ was found using the spectrum obtained in MEM scheme "a" due to numerical errors.

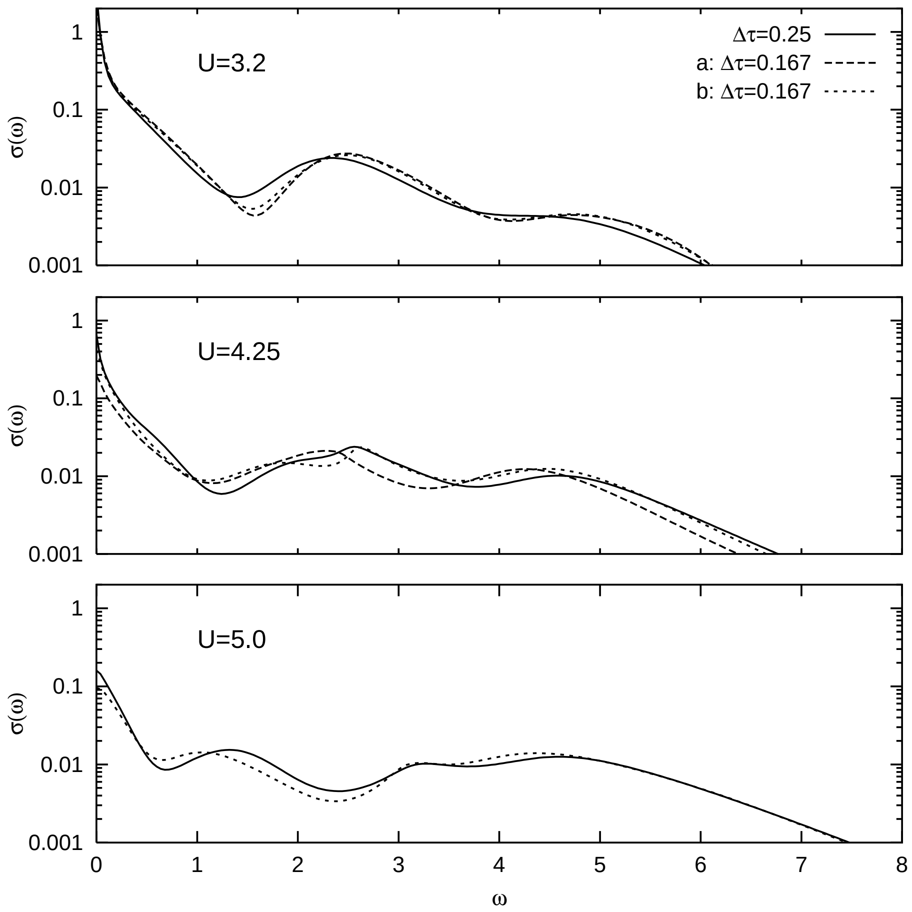



**Figure 5.17:** Optical conductivity $\sigma(\omega)$ for $La_{1-x}Sr_xTiO_3$ using the $d = \infty$ estimate for $\tilde{\rho}(\epsilon)$ depicted in Fig. 5.15 b. The variations "a" and "b" of the MEM scheme are explained in the discussion of Fig. 5.12.

---

real part of the optical conductivity $\sigma(\omega)$ for $U = 3.2$, $U = 4.25$, and $U = 5.0$ (thick lines) at $T = 0.1$ (using set "b" for $\Delta\tau = 0.167$) in comparison with results that the same formalism yields for a constant squared Fermi velocity $\langle|\boldsymbol{v_k}|^2\rangle$. The inclusion of a nonconstant $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ clearly impacts the total spectral weight and its distribution for all choices of the interaction. In general, features become more pronounced. Quantitatively, this effect becomes less important for strong interactions: at $U = 5$, the relative error associated with a neglect of the energy dependence of $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ never exceeds about 20%. From this fact we conclude that also a full finite-dimensional LDA evaluation of $\tilde{\rho}(\epsilon)$ or, equivalently, of $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ would not lead to dramatic changes compared to our estimate. Therefore, we regard our result for $U = 5$ as a reasonable LDA+DMFT(QMC) estimate for the $3d$ $t_{2g}$ contribution to the optical conductivity of $La_{1-x}Sr_xTiO_3$ with doping $x = 0.06$. On the other hand, much more reliable results for all observables discussed in this section could be expected when using a state-of-the-art QMC program that does not suffer from large $\Delta\tau$ errors due to an inefficient implementation of the Fourier-transformation

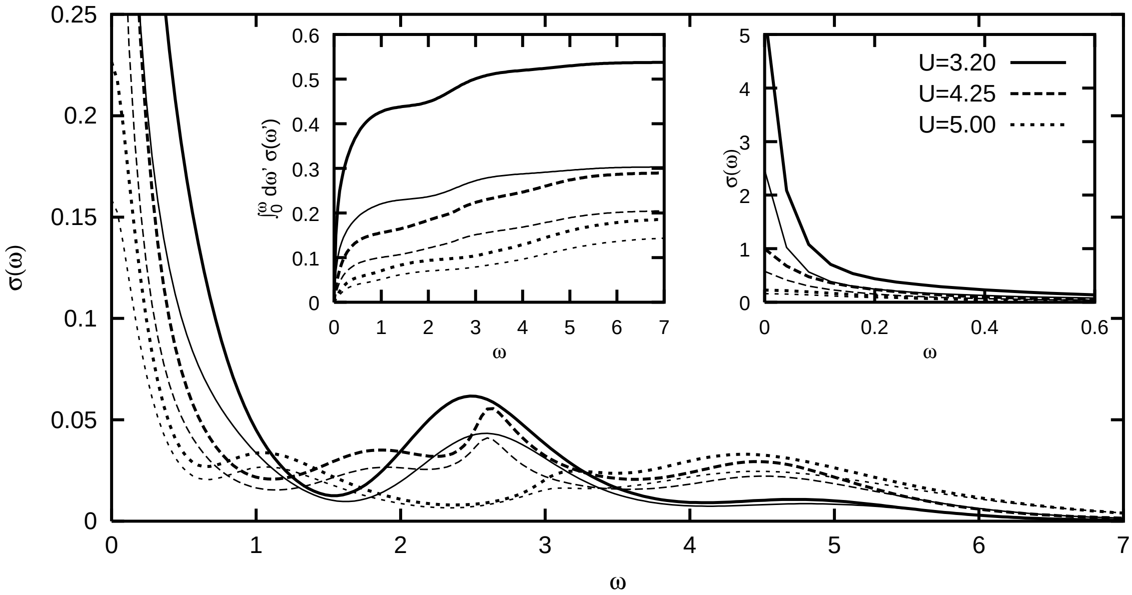



**Figure 5.18:** QMC estimates for the $t_{2g}$ contribution to the optical conductivity $\sigma(\omega)$ of $La_{1\text{-}x}Sr_xTiO_3$ at $T = 0.1$ (thick lines). For the thin lines, $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$ is set to the constant valid for the hc lattice with NN hopping.

(cf. subsection 3.4.1 and App. C).

When comparing the theoretical transport data to the experimental results reproduced in the following, we also have to keep in mind that interband excitations have been completely neglected so far. A full consistent calculation would require knowledge of the dispersions of all relevant bands and would have to take relative positions of atoms in the unit cell into account (see, e.g., Ahn and Millis, 2000). Then all contributions not involving the correlated $t_{2g}$ band could be calculated using the usual LDA bubble formula for noninteracting electrons. The pure $t_{2g}$ contribution could be evaluated as before, but using the LDA estimate for $\langle|\boldsymbol{v_k}|^2\rangle(\epsilon)$. Finally, transitions between the $t_{2g}$ band and other bands (e.g., the $e_g$ or the O $2p$ bands) would involve both quasiparticles with finite and infinite lifetime. In lack of this data we can only estimate that additional contributions to $\sigma(\omega)$ will appear at least for $\omega \gtrsim 1$ eV.

Since we are not aware of experimental data for the optical conductivity of $La_{1\text{-}x}Sr_xTiO_3$ for the exact doping level $x = 0.06$, we reproduce a set of results obtained for doped $LaTiO_3$ with different stated compositions in chronological order. Measurements of $\sigma(\omega)$ (solid lines) on melt-grown samples over the full Sr doping range (Fujishima, Tokura, Arima, and Uchida, 1992) are shown in Fig. 5.19. An interpolation between the results for "$x = 0.9$" (here "x" denotes the La content) and "$x = 1.0$" or "$x = 0.98$" should be compared to our results. We note that this experiment indicates a minimum in $\sigma(\omega)$ at $\omega \approx 1.2$ eV and that the maximum of $\sigma(\omega)$ at small frequencies exceeds this minimum by a factor of $5 - 6$. Since the contribution at about 6 eV remains visible in Fig. 5.19 a in the limit $x \to 0$, i.e., for an empty $t_{2g}$

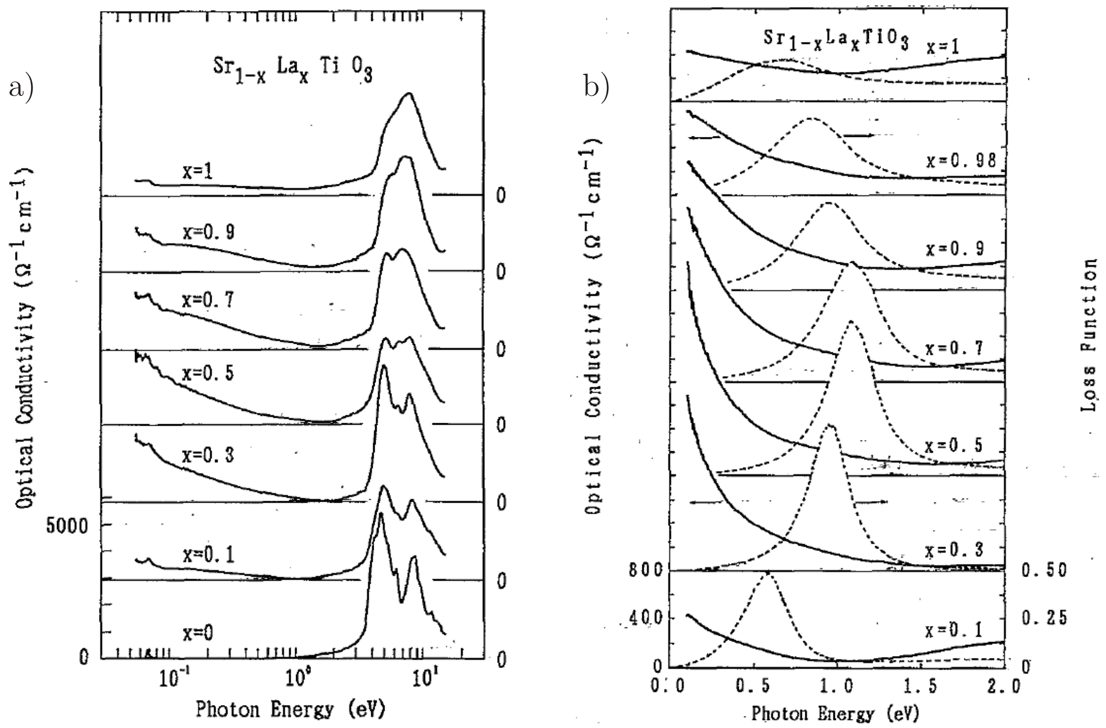



**Figure 5.19: a)+b)** Optical conductivity $\sigma(\omega)$ (solid lines) as measured by Fujishima et al. (1992) at $T \approx 300$ K. Since $x$ here corresponds to the $t_{2g}$ band filling, $x \approx 1$ is the strongly correlated limit. **b)** Low-energy part with a linear scale and electronic energy-loss function $\mathrm{Im}\,(-1/\epsilon)$ (dashed lines).

band, a large part of it must be due to excitations involving other bands. Note that a tail of this contribution extends down to about 1 eV. The width of the remnant of a Drude peak is even much larger than found in our high-temperature calculations which points towards strong disorder. Curve "C" in Fig. 5.20 a shows $\sigma(\omega)$ measured for $La_{1-x}SrTiO_3$ with $x \approx 0.07$ (Crandles et al., 1994). Here, a sharp Drude peak is observed at very low frequency $\omega \lesssim 0.1$ eV while for larger frequencies the curve is almost flat. If this is significant, disorder effects seem to be less important in this case.

In both cases, there is no obvious agreement with any of the results shown in Fig. 5.18, i.e., with the (intraband absorption only) LDA+DMFT theory. While the minimum experimentally seen in $\sigma(\omega)$ slightly above 1 eV would find correspondence in the theoretical results for $U = 4.25$, the computed Drude weight seems still too large even for $U = 5.0$.

The remaining three figures for $\sigma(\omega)$ in doped titanates are all due to Tokura's group. While Fig. 5.20 b relates to $LaTiO_{3+\delta/2}$ and shows noisy peaks at low frequencies (Okimoto et al., 1995), a curve for $La_{1-x}Sr_xTiO_3$ with $x = 0.1$ has been added to a selection of the same curves in Fig. 5.21 a (Katsufuji et al., 1995). In the latter diagram, also the low-frequency parts have been removed. Finally, Fig. 5.21 b apparently shows the same measured data (solid lines) as Fig. 5.21 a, but now ex-

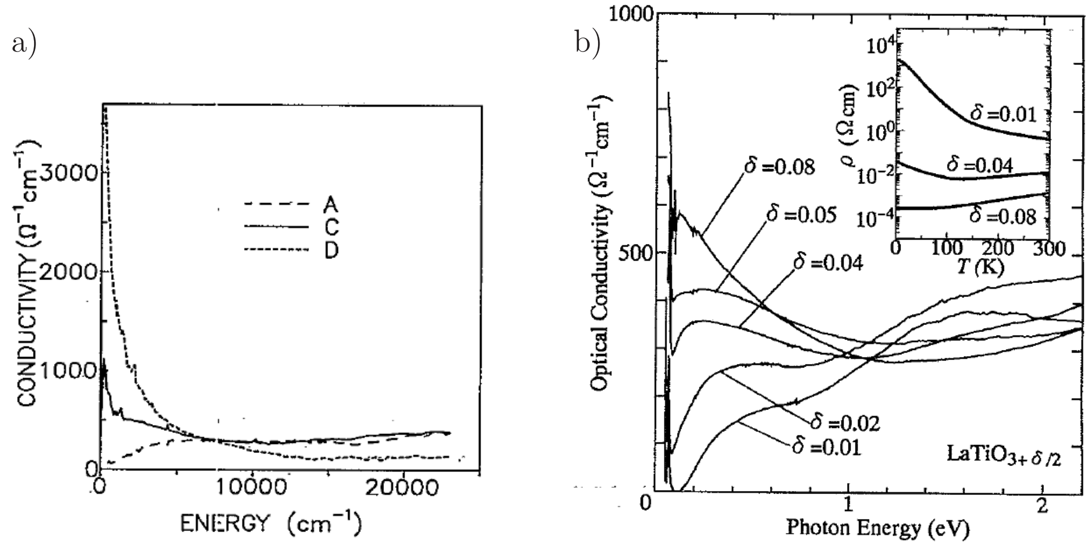



**Figure 5.20: a)** Optical conductivity $\sigma(\omega)$ for $La_{1-x}TiO_3$ with $x \approx 0.07$ for curve C and $x \approx 0.12$ for curve D at $T \approx 300$ K (Crandles et al., 1994). The measurements extend to $\omega \approx 3$ eV (1 eV = ($\hbar c$) 8066 cm$^{-1}$). **b)** Optical conductivity $\sigma(\omega)$ for $LaTiO_{3+\delta/2}$ at $T \approx 300$ K (Okimoto et al., 1995). Inset: resistivity as a function of temperature.

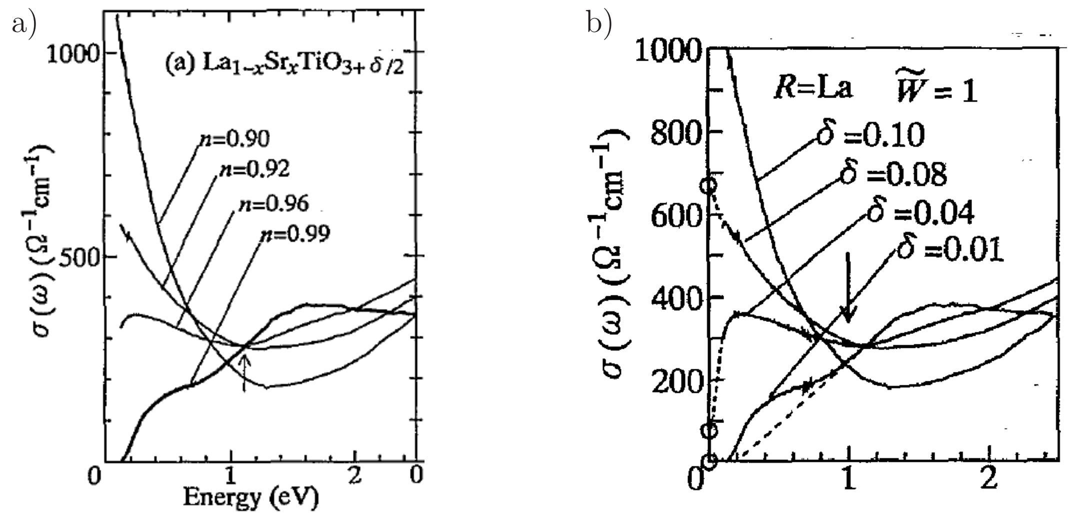



**Figure 5.21: a)** Optical conductivity $\sigma(\omega)$ for $La_{1\text{-}x}Sr_xTiO_{3+\delta/2}$ with $n = 1 - (x + \delta)$ at $T \approx 300$ K: here, the result for $n = 0.1$ is claimed to have a Sr content of $x = 0.1$ (Okimoto et al., 1995). **b)** Optical conductivity $\sigma(\omega)$ for $La_{1-x}Ca_xTiO_3$ at $T \approx 300$ K (Katsufuji et al., 1995). Here, the dashed lines extrapolate to the measured dc conductivity (open circles). Very surprisingly, the measured curves in a) and b) appear to be absolutely identical.

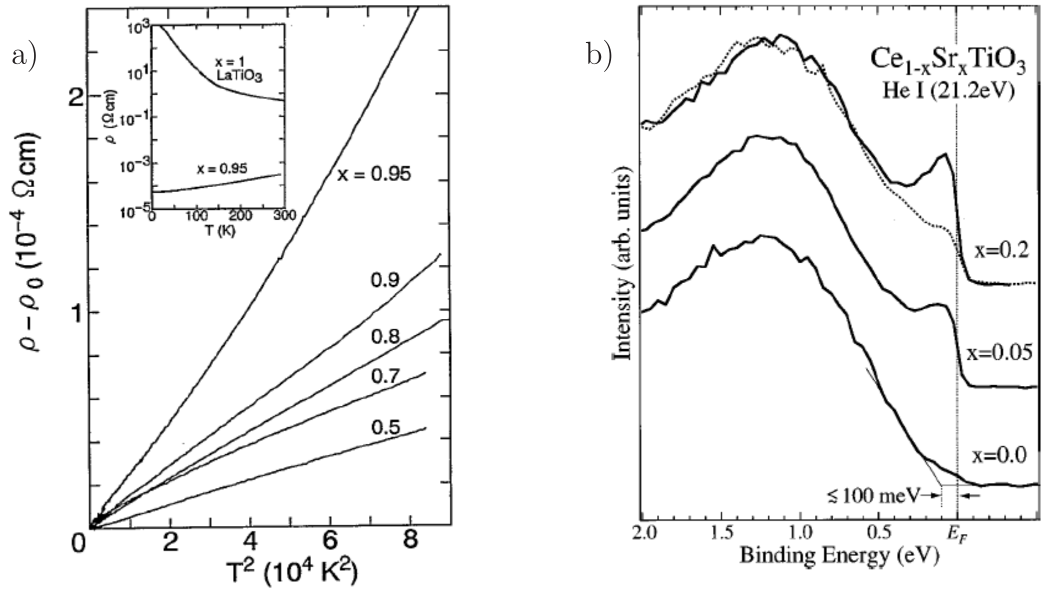



**Figure 5.22: a)** Resistivity of $Sr_{1-x}La_xTiO_3$ (Tokura et al., 1993). **b)** PES for $Ce_{1-x}Sr_xTiO_3$ (Yokoya et al., 1999).

trapolated (dashed lines) towards independently measured data points for the dc conductivity (circles). Katsufuji and Tokura's (1999) claim that their data was taken for $La_{1-x}Ca_xTiO_3$ is very likely erroneous, even though Ca doping might in general have effects similar to doping with Sr. In the relevant doping range $\delta \approx 0.06$, these measurements indicate an almost constant function $\sigma(\omega)$ which is clearly incompatible with our data and points to the presence of strong disorder.

One limitation inherent in our results is the high temperature of $T \approx 1160$ K used in the QMC calculations. This might not be such a bad approximation, e.g., for room temperature systems, as it first appears. On the one hand, due to the restriction to purely electronic degrees of freedom and to the neglect of spatial fluctuations in the DMFT, the effect of high temperatures is likely to be reduced compared to a full calculation. Furthermore, the qualitative electronic properties of $La_{1-x}Sr_xTiO_3$ appear to be unchanged over an extremely wide temperature range as illustrated in Fig. 5.22 a: For this material, the resistivity is nearly linear as a function of $T^2$ both in the correlated ($x \approx 0.95$, i.e., $\delta \approx 0.05$) and in the strongly doped regime ($0.5 \lesssim x \lesssim 0.9$) (Tokura et al., 1993). Similar results have recently been obtained by Hays et al. (1999) for powder samples and by Gariglio, Seo, Fompeyrine, Locquet, and Triscone (2001) for epitaxial films.[29] Thus, at least the coefficient of the low-temperature resistivity can be measured at and above room temperature.

Figure 5.22 b shows PES for $Ce_{1-x}Sr_xTiO_3$ (Yokoya et al., 1999) which are sur-

[29]Gariglio et al. (2001) concentrated on the strongly doped case ($\delta \approx 0.5$) and attributed the behavior of the resistivity to a polaronic mechanism. The relevant lattice distortion is here the dynamic tilting of the oxygen octahedra which gives rise to an optical phonon mode at $\hbar\omega_0 \approx 80$ K.

prisingly similar to the corresponding plot for $La_{1-x}Sr_xTiO_3$ shown in Fig. 5.7. This demonstrates that our results may apply to a relatively large class of materials and that, within this class, $La_{1-x}Sr_xTiO_3$ may not be the system best described by our approach.

## 5.5 Conclusion

In this chapter, we have given an introduction to the ab initio density functional theory, to its local density approximation (LDA), and to the hybrid LDA+DMFT method in enough detail to expose the various inherent approximations. We have presented published results (Nekrasov et al., 2000) for $La_{1-x}Sr_xTiO_3$; the relatively good agreement of the computed photoemission data with experiments clearly demonstrates the significant improvement over conventional LDA calculations. We have also discussed some subtleties of the theoretical treatment of photoemission spectra (PES) and have demonstrated the numerical significance of an error in the PES code which we had discovered and corrected.

New results obtained by Nekrasov for the doping dependence of the quasiparticle weight and of the susceptibility of $La_{1-x}Sr_xTiO_3$ (unpublished) and a planned joint publication (Nekrasov, Blümer, Held, Anisimov, and Vollhardt, 2001) as well as our experiences from treating the metal insulator transition of the half-filled Hubbard model (cf. chapter 3) using the uncorrected QMC code prompted us to question the reliability of the LDA+DMFT(QMC) results in the strongly correlated regime. Therefore, we have repeated all original calculations using an increased number of imaginary time slices in QMC ($\Delta\tau = 0.167$ instead of $\Delta\tau = 0.25$) and testing variations of the maximum entropy method (MEM). As expected, we found significant corrections to the quasiparticle weights and also to the MEM spectra. Our more accurate treatment also leads to a slightly reduced relative weight in the PES near the Fermi energy which is, however, still too high in comparison with experiment. Furthermore, we have computed x-ray absorption (XAS) spectra. While the overall agreement with experiment was seen to be reasonable when including the contributions of Ti $e_g$ and O $2p$ bands (in addition to the Ti $t_{2g}$ bands), the weight at the Fermi energy was again overestimated by theory.

Finally, we have computed the first LDA+DMFT(QMC) estimate of the Ti $t_{2g}$ contribution to the optical conductivity $\sigma(\omega)$ of $La_{1-x}Sr_xTiO_3$. One prerequisite for this calculation was the development of a program for the inversion of the general $\boldsymbol{k}$-integrated Dyson equation. In lack of momentum resolved LDA data, we also needed to determine the effect of a nonconstant Fermi velocity by application of the general dispersion formalism developed in Sec. 2.3 to the LDA density of states. A detailed comparison of the numerical results with experiment is difficult. On the one hand, experimental results for $\sigma(\omega)$ differ greatly with respect to the presence or width of a Drude peak. On the other hand, contributions of orbitals not included in our theory can be expected to be significant. Taking these uncertainties into account, the overall agreement seems reasonable; again, the most prominent discrepancy is that the theory predicts too much weight at small frequencies.

One unrealistic aspect of all LDA+DMFT(QMC) calculations presented in this chapter was the high temperature $T \approx 1160$ K. A reduction of the simulation temperature to room temperature would be desirable; such a project is clearly realistic, in particular when using an improved QMC code (cf. App. C) for which the discretization $\Delta\tau = 0.25$ is sufficiently accurate. One might also want to investigate the influence of non-uniform on-site coupling terms, the importance of hybridization terms in a full-Hamiltonian calculation, or the impact of the orthorhombic distortion.

We conclude that the LDA+DMFT(QMC) treatment is vastly superior over the sole application of LDA already at the present stage. Still, the material $La_{1-x}Sr_xTiO_3$ is less metallic at $x = 0.06$ than the theory predicts. If this finding holds up in improvements of the theoretical treatment and in experiments which are more bulk-sensitive, the remaining discrepancies may be due to disorder (which is clearly present, but should be washed out since Ti has 8 La/Sr nearest neighbors) or to (precursors of) antiferromagnetic order.

# Summary

The characterization of a material as metallic or insulating, the study of transitions between such states and the development of suitable models is clearly of great fundamental and practical interest. The one-band Hubbard model is a conceptionally relatively simple model which is potentially relevant in this context since its primary parameter, the on-site interaction $U$ triggers a transition or a crossover from metallic behavior at small $U$ to insulating behavior at large $U$ and half filling. A reduction of complexity of this model is achieved by the dynamical mean-field theory (DMFT); due to its nonperturbative character, this method is reliable in the range of interest, i.e., for intermediate to strong coupling. The resulting model certainly misses important properties of specific materials such as $V_2O_3$; nevertheless, it is a good candidate for controlled numerical studies of the more abstract metal-insulator transition (MIT) problem.

We have presented quantum Monte Carlo (QMC) results for a fully frustrated version of the Hubbard model and constructed its phase diagram. We have resolved a controversy regarding the existence and extent of the MIT coexistence region and have succeeded in the precise determination of the first-order transition line; a task that was previously considered too difficult for numerically exact methods. The new methods employed for this purpose and a more direct (and yet untested) approach developed in this work should prove fruitful for broader application in the future. Further research will also profit from our discovery and correction of problems in some implementations of the QMC method.

The MIT phase diagram Fig. 3.50, a main result of this work, clearly does not represent a complete scenario for any material featuring a MIT. This follows already from the finite (and large) entropy predicted in the insulating ground state which violates Nernst's law. This contradiction may to some degree be regarded as an artifact of the DMFT limit: For any fixed disorder distribution and finite hopping range, long-range order is expected to set in below some critical temperature in any finite dimension $d$. A more realistic and consistent DMFT description is possible when the low-temperature phase is ordered already in $d = \infty$, i.e., for partial frustration. We have discussed the qualitative implications of variable frustration in the (nongeneric) case that the properties of the paramagnetic phase remain unchanged. The question whether the topology of the phase diagram of $V_2O_3$ (with a first-order paramagnetic metal-insulator transition and an antiferromagnetic low-temperature phase) can be reproduced for a lattice with realistic frustration is still open.

The theory of transport in infinite dimensions was previously restricted to the hypercubic lattice (or anisotropic stacked lattices) with nearest-neighbor hopping.

We have generalized the formalism and derived explicit general expressions for the optical $f$-sum rule. We have constructed the first regular lattice with finite (semi-elliptic) band edges in $d = \infty$ and have thus obtained the first DMFT results for coherent and isotropic transport compatible with a semi-elliptic density of states.

We have reported on progress in the realistic modeling of strongly correlated electron systems using the hybrid LDA+DMFT method. This recently developed method starts with the ab initio density functional theory and treats the missing correlations within DMFT. Evaluating the resulting multi-band problem using QMC, we have computed photoemission and x-ray absorption spectra for the transition metal oxide $La_{1\text{-}x}Sr_xTiO_3$. The results represent a significant improvement over conventional LDA calculations. We have discussed the impact of the QMC discretization error and computed a first estimate of the $t_{2g}$ contribution to the optical conductivity.

The unifying project of this thesis has been the modeling of strongly correlated electron systems in the vicinity of a metal-insulator transition by application of the dynamical mean-field theory which becomes exact in the limit of infinite dimensionality. Within this project, we have studied single-band as well as multi-band models, pursued both pure model Hamiltonian approaches and the ab initio LDA+DMFT approach and computed both local and transport properties. A large amount of numerical work, i.e., the planning, organization, and analysis of several thousand QMC simulations using tens of thousands of hours of supercomputer time was clearly essential for obtaining the accurate results which we have presented in this thesis. Furthermore, we have written a large number of computer programs and have modified and extended some others. Many of our results, however, could not have been obtained without the development or improvement of methods and formalisms as well as analytical calculations. It is this more theoretical than numerical aspect of our work which we found particularly rewarding.

# Appendix A

# Additions to "Models and Methods"

## A.1 Extensions of the Hubbard Model

The one-band Hubbard model represents a highly idealized view of strongly correlated materials. Some of its shortcomings are best characterized by writing down minimal model extensions which could be used for adding the missing physics. While most of the extensions collected in this section will not be used for actual calculations in this thesis, we will discuss the implications of extended hopping in chapter 2 and study a 3-band model based on LDA input data in chapter 5.

The applicability of the one-band Hubbard model to $d$ or $f$ electron systems is a priori questionable since the partially filled bands correspond to atomic orbitals which are 5-fold and 7-fold degenerate (for each spin direction), respectively. While bands in a lattice are more complicated than orbitals of isolated atoms, the remaining degeneracy can be inferred from symmetry considerations alone. Often it is useful to consider a cubic representation of the angular part of atomic $d$ orbitals,

$$\begin{aligned}&|d_{xy}\rangle \propto (|2,2\rangle - |2,-2\rangle)\,, \quad |d_{yz}\rangle \propto (|2,1\rangle + |2,-1\rangle)\,, \quad |d_{zx}\rangle \propto (|2,1\rangle - |2,-1\rangle)\\ &|d_{x^2-y^2}\rangle \propto (|2,2\rangle + |2,-2\rangle)\,, \quad |d_{3z^2-r^2}\rangle \propto |2,0\rangle, \end{aligned} \tag{A.1}$$

expressed in terms of eigenfunctions of the angular momentum operator,

$$\boldsymbol{l}^2|l,m\rangle = \hbar^2 l(l+1)|l,m\rangle, \quad l_z|l,m\rangle = \hbar m|l,m\rangle\,. \tag{A.2}$$

In lattices with cubic symmetry the five $d$ orbitals are energetically split into the $t_{2g}$ orbitals ($|d_{xy}\rangle$, $|d_{yz}\rangle$, $|d_{zx}\rangle$) and the $e_g$ orbitals ($|d_{x^2-y^2}\rangle$, $|d_{3z^2-r^2}\rangle$), which give rise to one threefold degenerate and one twofold degenerate band, respectively. Lower symmetry can lift the remaining degeneracies; e.g., in the trigonal case the $t_{2g}$ orbitals are further split into one nondegenerate $a_{1g}$ and one twofold degenerate $e_g^\pi$ band. Thus, it is possible that in some $d$ systems only one band crosses or touches the Fermi surface which then justifies the one-band assumption made in subsection 1.1.4 and used in chapter 3. In general, however, the inclusion of several orbitals per site is important. An SU(2)-invariant generalization of the Hubbard model where the interaction is still local but the valence band is degenerate then contains additional coupling terms[1] (Anderson, 1961; Dworin and Narath, 1970)

$$\begin{aligned}\hat{H}_{m\text{-band}} &= -t \sum_{\langle ij\rangle,\nu\sigma} \hat{c}^{\dagger}_{i\nu\sigma}\hat{c}_{j\nu\sigma} + U\sum_{i\nu}\hat{n}_{i\nu\uparrow}\hat{n}_{i\nu\downarrow} \\ &\quad + V_0 \sum_{i;\nu<\nu';\sigma\sigma'} \hat{n}_{i\nu\sigma}\hat{n}_{i\nu'\sigma'} + F_0 \sum_{i;\nu<\nu';\sigma\sigma'} \hat{c}^{\dagger}_{i\nu\sigma}\hat{c}^{\dagger}_{i\nu'\sigma'}\hat{c}_{i\nu\sigma'}\hat{c}_{i\nu'\sigma}, \qquad (A.3)\end{aligned}$$

where $\nu$,$\nu'$ (with $1 \le \nu \le m$, $1 \le \nu' \le m$) are band indices. The exchange term parameterized by the Hund's rule coupling $F_0$ can be rewritten as

$$\hat{H}_{F_0} = -2F_0 \sum_{i,\nu<\nu'} \left( \hat{\boldsymbol{S}}_{i\nu}\cdot\hat{\boldsymbol{S}}_{i\nu'} + \frac{1}{4}\hat{n}_{i\nu}\hat{n}_{i\nu'} \right) \qquad (A.4)$$

with $\hat{\boldsymbol{S}}_{i\nu} = \frac{1}{2}\sum_{\sigma\sigma'} \hat{c}^{\dagger}_{i\nu\sigma}\boldsymbol{\tau}_{\sigma\sigma'}\hat{c}_{i\nu\sigma'}$ being the spin operator for orbital $\nu$ at site $i$ and the Pauli matrices $\boldsymbol{\tau}_{\sigma\sigma'}$. In quantum Monte Carlo (QMC) simulations, the spin-flipping terms implicit in (A.4) lead to a numeric (minus-sign) problem. Therefore, one here usually replaces the Heisenberg interaction part of (A.4) by an Ising-type interaction at the cost of breaking the SU(2) symmetry (Motome and Imada, 1997; Held and Vollhardt, 1998). Since $\hat{s}^z_{i\nu}\hat{s}^z_{i\nu'} = \sigma\sigma'\hat{n}_{i\nu\sigma}\hat{n}_{i\nu'\sigma'}/4 = (2\delta_{\sigma\sigma'}-1)\,\hat{n}_{i\nu\sigma}\hat{n}_{i\nu'\sigma'}/4$, one can write this modified multi-band Hubbard Hamiltonian as

$$\hat{H}^z_{m\text{-band}} = \sum_{\langle ij\rangle,\nu\sigma} \epsilon_{\boldsymbol{k}\nu}\hat{c}^{\dagger}_{\boldsymbol{k}\nu}\hat{c}_{\boldsymbol{k}\nu} + U\sum_{i\nu}\hat{n}_{i\nu\uparrow}\hat{n}_{i\nu\downarrow} + \sum_{i;\nu<\nu';\sigma\sigma'} (V_0 - \delta_{\sigma\sigma'}F_0)\,\hat{n}_{i\nu\sigma}\hat{n}_{i\nu'\sigma'}. \qquad (A.5)$$

The interaction $U$ between electrons within each orbital is always larger than the interorbital density-density interaction $V_0$. The smaller exchange coupling $F_0$ can trigger ferromagnetic and (possibly coexisting) orbital order.[2]

While in general the inclusion of different nondegenerate bands involves an additional large number of parameters, one can model (and simulate using QMC, e.g., within the DMFT) the influence of a half-filled, locally polarized and relatively inert subband as observed in manganites by adding a Kondo-type interaction between the itinerant electrons and some spin degree of freedom, $-2\sum_\nu\sum_i \hat{\boldsymbol{s}}_{i\nu}\cdot\hat{\boldsymbol{S}}_i$. Here, $\hat{\boldsymbol{s}}_{i\nu}$ are the spin operators for the itinerant twofold degenerate $e_g$ band while $\hat{\boldsymbol{S}}_i$ models a localized spin-3/2 originating from the $t_{2g}$ band (Held and Vollhardt, 2000).

An additional complication in compounds is the presence of several inequivalent ions per unit cell. Such systems can be modeled microscopically by introducing different on-site energies, Coulomb interactions, and hopping matrix elements within a unit cell or more phenomenologically by using renormalized hopping matrix elements between just a subset of orbitals (while others are integrated out). One example of the

[1] Here, an on-site pair hopping term which only contributes when one orbital is doubly and another singly occupied is neglected (Held and Vollhardt, 1998).

[2] In the application of the LDA+DMFT method in chapter 5 we will set $U = V_0$ and $F_0 = 0$; in the absence of long range order and for small filling $n \lesssim 1$ the resulting error is not expected to exceed other sources of errors. In general, local exchange terms only become negligible for extended orbitals (an extreme case being valence molecular orbitals in $C_{60}$ Fullerides).

former approach is the "Emery model" or "Varma-Schmitt-Rink-Abrahams model" for electron holes in the copper oxide planes in high-$T_c$ superconductors (Varma, Schmitt-Rink, and Abrahams, 1987; Emery, 1987)[3]

$$\begin{aligned}\hat{H}_{\text{sc}} &= -t_{pd}\sum_{\langle i,j\rangle,\nu,\sigma}(\hat{p}^{\dagger}_{j\nu\sigma}\hat{d}_{i\sigma}+\text{h.c.}) - t_{pp}\sum_{\langle j,j'\rangle,\nu,\sigma}(\hat{p}^{\dagger}_{j\nu\sigma}\hat{p}_{j'\nu\sigma}+\text{h.c.}) + \epsilon_d\sum_i \hat{n}^d_i + \epsilon_p\sum_{j\nu}\hat{n}^p_{j\nu} \\ &\quad + U_d\sum_i \hat{n}^d_{i\uparrow}\hat{n}^d_{i\downarrow} + U_p\sum_{j\nu}\hat{n}^p_{j\nu\uparrow}\hat{n}^p_{j\nu\downarrow} + V\sum_{\langle ij\rangle,\nu}\hat{n}^d_i\hat{n}^p_{j\nu}\,. \end{aligned} \tag{A.6}$$

The latter, more phenomenological approach is used for modeling the compound $La_{1\text{-}x}Sr_xTiO_3$ using an explicit treatment of the Ti $3d$ $t_{2g}$ orbitals only in chapter 5.

Since the restriction to local interactions in the Hubbard model is a crude approximation to the Coulomb interaction, one may expect significant improvements by including nearest-neighbor terms of the Coulomb interaction. For a one-band model, these may be written in the form (Hirsch, 1989; Strack and Vollhardt, 1994)

$$\begin{aligned}\hat{V}^{\text{NN}}_{\text{1-band}} &= \sum_{\langle i,j\rangle}\Big[V\hat{n}_i\hat{n}_j - 2F(\hat{\boldsymbol{S}}_i\hat{\boldsymbol{S}}_j + \frac{1}{4}\hat{n}_i\hat{n}_j) \\ &\quad + X\sum_\sigma(\hat{c}^{\dagger}_{i\sigma}\hat{c}_{j\sigma}+\text{h.c.})(\hat{n}_{i\,-\sigma}+\hat{n}_{j\,-\sigma}) + F'(\hat{c}^{\dagger}_{i\uparrow}\hat{c}^{\dagger}_{i\downarrow}\hat{c}_{j\downarrow}\hat{c}_{j\uparrow}+\text{h.c.})\Big]. \end{aligned} \tag{A.7}$$

Here, the $V$ term and the $F$ term are completely analogous to the onsite interactions between orbitals in (A.3), while the additional terms involve density-dependent hopping ($X$) and pair hopping ($F'$). These nearest-neighbor contributions of the Coulomb interactions can, e.g., drive ferromagnetism (Hirsch, 1989; Kollar, Strack, and Vollhardt, 1996; Blümer, 1996; Wahle et al., 1998), charge density waves, phase separation, or superconductivity (Hirsch and Marsiglio, 2000).

Disorder is always present to some degree in real materials, in particular in doped systems (as considered in chapter 5), and is potentially very important, e.g., in the context of Anderson localization. A straightforward generalization of the one-band Hubbard model leads to

$$\hat{H}_{\text{disorder}} = \sum_{i\sigma}(\epsilon_i-\mu)\,\hat{n}_{i\sigma} + \sum_{\langle i,j\rangle,\sigma} t_{ij}\left(\hat{c}^{\dagger}_{i\sigma}\hat{c}_{j\sigma}+\text{h.c.}\right) + \sum_i U_i(\hat{n}_{i\uparrow}-\frac{1}{2})(\hat{n}_{i\downarrow}-\frac{1}{2}), \tag{A.8}$$

where site-diagonal terms, i.e., the lattice potentials $\epsilon_i$ and the Coulomb interaction $U_i$ and/or the (intersite) hopping terms $t_{ij}$ may be allowed to vary [for a treatment within DMFT, see, e.g., Dobrosavljević and Kotliar (1993) and Ulmke (1998)].

The inclusion of phonons is possible in terms of the Holstein-Hubbard model, in which the Hubbard model is supplemented by the electron-phonon interaction plus a phonon energy (Holstein, 1959; Zhong and Schüttler, 1992; Freericks and Jarrell, 1995*b*)

$$\hat{H}_{\text{H-H}} = \hat{H}_{\text{Hub}} + g\sum_i R_i(\hat{n}_{i\uparrow}+\hat{n}_{i\downarrow}-1) + \frac{1}{2}\sum_i\left(\frac{P_i^2}{M}+M\Omega^2R_i^2\right). \tag{A.9}$$

[3]Here, $i$ and $j$ denote the unit cell for Cu $3d_{x^2-y^2}$ orbitals and O $2p_x$ or $2p_y$ orbitals, respectively. The orbital index $\nu = x$ or $\nu = y$ distinguishes the oxygen orbitals (living on different sites); unrestricted summation over omitted spin indices is implied.

Here, $R_i$ ($P_i$) label the phonon position (momentum) at site $i$, $g$ is a coupling strength, $M$ is the phonon mass and $\Omega$ its frequency.[4]

Finally, the relativistic spin-orbit interaction can be approximately reintroduced in the form

$$\hat{H}_{SO} = \xi \sum_{\nu\nu'\sigma\sigma'} \langle \nu\sigma | \boldsymbol{L} \cdot \boldsymbol{S} | \nu'\sigma' \rangle c^{\dagger}_{\nu\sigma} c_{\nu'\sigma'}, \tag{A.10}$$

where $L$ is the angular momentum operator, $\boldsymbol{S}$ the vector of Pauli matrices, and the parameter $\xi$ depends on the system (see, e.g., Bruno, 1993; Harle, 2000). In the presence of magnetic order, this interaction may lead to an easy or hard axis. Within DMFT, the behavior of antiferromagnetic systems with a strong easy axis can also be modeled phenomenologically by restricting local spin moments to point along this axis which then leads to metamagnetic phase transitions of both first and second order upon variation of an external magnetic field (Held, Ulmke, and Vollhardt, 1996; Held, Ulmke, Blümer, and Vollhardt, 1997).

## A.2 Characterization of Generic Momenta

In the following, we will first extend the definition of nongeneric momenta given in subsection 1.2.2 to finite dimensions, then show that in all dimensions nongeneric momenta can be characterized as being close to the origin $\mathbf{0}$ or to the antiferromagnetic wave vector $\boldsymbol{Q}$ and, finally, draw some conclusions for the $d \to \infty$ limit of the hypercubic lattice and the DMFT in general.

In the following, we will exploit that the expression $\eta_{\boldsymbol{q}} = \lim_{d\to\infty} \frac{1}{d} \sum_{\alpha=1}^{d} \cos(q_\alpha)$ introduced in the context of (1.23) for the $d = \infty$ hc lattice naturally extends to finite dimensions. Since $\eta_{\boldsymbol{q}}$ always varies in the range $[-1, 1]$ and $\eta_{\boldsymbol{q}} \neq 0$ defines nongeneric vectors $\boldsymbol{q}$ in $d = \infty$, one can identify vectors $\boldsymbol{q}$ with $|\eta_{\boldsymbol{q}}| > \eta_0$ (for any $\eta_0 > 0$) with strong momentum correlation in any finite dimension $d$ and loosely call them nongeneric, too. Furthermore, we can estimate $\eta_{\boldsymbol{q}}$ from the vector norm $|\boldsymbol{q}|$ which for momenta in the first Brillouin zone, unit lattice spacing $a = 1$, and dimension $d$ varies between $|\mathbf{0}| = 0$ and $|\boldsymbol{Q}| = \pi\sqrt{d}$:

$$\eta_{\boldsymbol{q}} = -1 + \frac{1}{d} \sum_{\alpha=1}^{d} (1 - \cos(\boldsymbol{q}_\alpha)) \leq \frac{|\boldsymbol{q}|^2}{2d} - 1 \tag{A.11}$$

Therefore, $\eta$ is strictly negative for all momenta inside a hypersphere of radius $\sqrt{2d}$ while $\eta < -0.5$ for momenta $|\boldsymbol{q}| < \sqrt{d}$. Analogously, $\eta > 0$ ($\eta > 0.5$) for the corresponding hyperspheres around $\boldsymbol{Q}$. Thus, we have identified convex sets of "nongeneric" momenta illustrated in Fig. A.1 a which are associated with strong energetic correlations via (1.23) and which acquire $\epsilon_{\boldsymbol{q}} = -\infty$ ($\epsilon_{\boldsymbol{q}} = \infty$) in the limit $d \to \infty$ on the hc lattice. In Fig. A.1 b, we show the probability distribution of $\eta_q$ for $d = 1, 2, 3$ as

[4] In principle, "phonon coordinate" can be translated as "ion coordinate" (which is $d$-dimensional). Note, however, that the single Einstein phonon mode considered here only has a single coordinate, i.e., is one-dimensional. A generalization to (anisotropic) $d$-dimensional coupling would require the inclusion of several phonon frequencies.

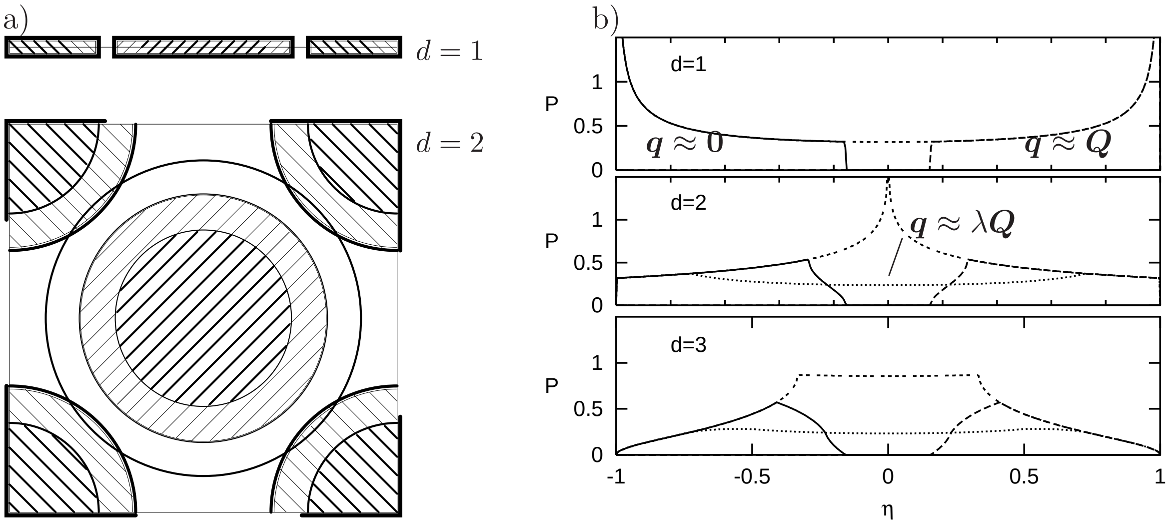



**Figure A.1: a)** Distribution of "nongeneric" momenta $\boldsymbol{q}$ for the hypercubic lattice in dimensions 1 and 2. Light (heavy) shading indicates states with $|\boldsymbol{q}| < \sqrt{2d}$ ($|\boldsymbol{q}| < \sqrt{d}$) which implies $|\eta_{\boldsymbol{q}}| > 0$ ($|\eta_{\boldsymbol{q}}| > 0.5$), when estimated from (A.11). The thick circle indicates the rescaled average length $\pi\sqrt{d/3}$ of vectors in $d \to \infty$. **b)** Full probability distribution $P(\eta_{\boldsymbol{q}})$ (rescaled hc DOS) for $d = 1, 2, 3$ and partial distributions corresponding to hyperspheres with $|\boldsymbol{q}| < \sqrt{2d}$ (solid lines) or $|\boldsymbol{q} - \boldsymbol{Q}| < \sqrt{2d}$ (dashed lines) (cf. shaded areas in a) or to momenta close to the diagonal (dotted lines).

well as partial distributions corresponding to the restricted set of momenta $|\boldsymbol{q}| < \sqrt{2d}$ and $|\boldsymbol{q}-\boldsymbol{Q}| < \sqrt{2d}$. While the criterion (A.11) is obviously not sharp, it clearly selects states with extremal values of $\eta_{\boldsymbol{q}}$ (and, for a hc lattice, also with extremal energy $\epsilon_{\boldsymbol{q}}$). For comparison, the partial probability distributions are also shown for momenta which are close to the diagonal extending from $\boldsymbol{0}$ to $\boldsymbol{Q}$. While this set of vectors with $\min_\lambda\{|\boldsymbol{q} - \lambda\boldsymbol{Q}|\} < \pi^2 d/12 - \pi^2\sqrt{d/180}$ (vectors one standard deviation closer to the diagonal than the average in the reduced Brillouin zone) has overlap with the other two sets for fixed (finite) dimension, it does not pick out extremal $\eta_{\boldsymbol{q}}$ and can therefore not be associated with nongeneric momenta.[5] We stress that the Euclidean radius of the selected hyperspheres is of the order of the longest vector within the Brillouin zone in all dimensions, i.e., is not small in high dimensions.

Therefore, transitions with small momentum transfer (measured by Euclidean length) imply small energy transfer in any dimension, also in the limit $d = \infty$. Consequently, the concept of a Fermi surface remains meaningful in this limit. The Fermi body of states occupied at low $T$ even remains simply connected within the Brillouin zone in the noninteracting limit. Thus, energies are not random in momentum space.

[5]This finding contradicts statements made in Gebhard's book which we cite here for comparison: *"Furthermore, on the hypercubic lattice, some correlations remain in the nesting direction,* $\boldsymbol{q} = \lambda\boldsymbol{Q}$ *(*$|\lambda| \leq 1$*). For all other* $\boldsymbol{q}$ *with* $|\boldsymbol{q} - \lambda\boldsymbol{Q}| > \mathcal{O}(\sqrt{1/d})$*, however, the two energies are completely uncorrelated,..." (Gebhard, 1997).* While it is clear that strictly on the diagonal $\boldsymbol{q} = \lambda\boldsymbol{Q}$ (with $0 \leq \lambda \leq 1$) all values $\lambda \neq 0.5$ correspond to finite $\eta_{\boldsymbol{q}} = \cos(\lambda\pi)$, this is evidently not true for its surroundings.

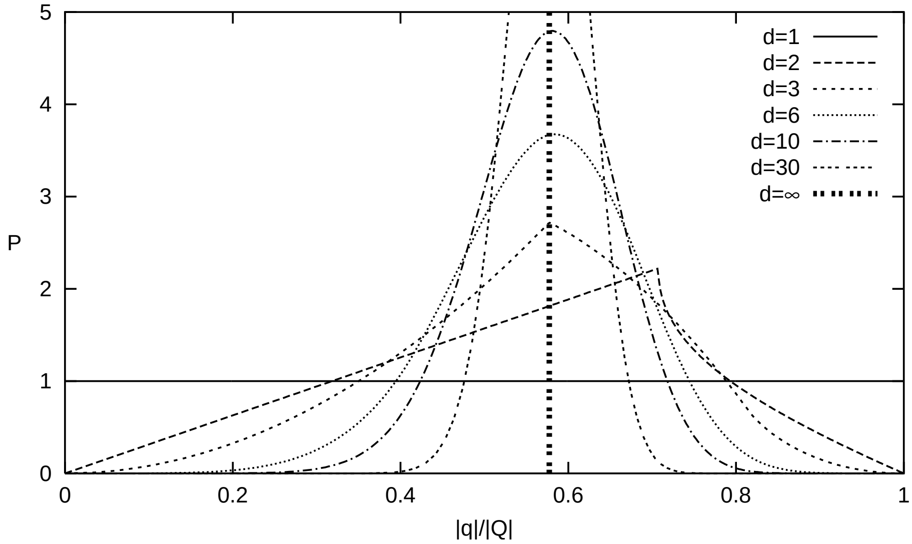



**Figure A.2:** Probability distribution $P(|\boldsymbol{q}|/|\boldsymbol{Q}|)$ of the Euclidean vector length of momenta $|\boldsymbol{q}|$ within the Brillouin zone for hypercubic lattices, normalized by $|\boldsymbol{Q}| = \pi\sqrt{d}$. In the limit $d \to \infty$, almost all $\boldsymbol{q}$ vectors (i.e., a unit fraction) are of length $\pi\sqrt{d/3}$.

On the other hand, small momenta and small momentum transfers have little phase space in high dimensions which leads to the irrelevance of momentum conservation in integrations over the full Brillouin zone. The Fermi surface cuts the Brillouin zone boundary infinitely often which is not surprising: each of the $2d$ hypersurfaces defined by $k_\alpha = \pm\pi$ is cut once for the hc lattice. The somewhat unphysical consequence is strongly enhanced umklapp scattering. As illustrated in Fig. A.2, the distribution of $|\boldsymbol{q}|$ is peaked around $\pi\sqrt{d/3}$ with vanishing width for $d \to \infty$. Thus, a unit fraction of all vectors has exactly the same lengths in this limit. In this sense, the generic $\boldsymbol{q}$ vectors are characterized by an intersection of the surface of a hypersphere of radius $\pi\sqrt{d/3}$ with a hypercube of side length $2\pi$.

## A.3 Dynamical Cluster Approximation, CDMFT, and Random Dispersion Approximation

We will here shortly characterize two approaches, the dynamical cluster approximation (DCA) and the CDMFT ("C" for "cellular"), which generalize the DMFT by introducing some aspects of finite dimensionality without being systematic in $1/Z$. We will also discuss the random dispersion approximation which is not an extension but an alternative approach to the DMFT in the limit of $Z \to \infty$.

The DCA is a generalization of the DMFT which introduces momentum dependence for the self-energy (Hettler, Tahvildar-Zadeh, Jarrell, Pruschke, and Krishnamurthy, 1998; 2000). This isaccomplished by dividing the Brillouin zone into regular patches characterized by some average momentum. The self-energy is then

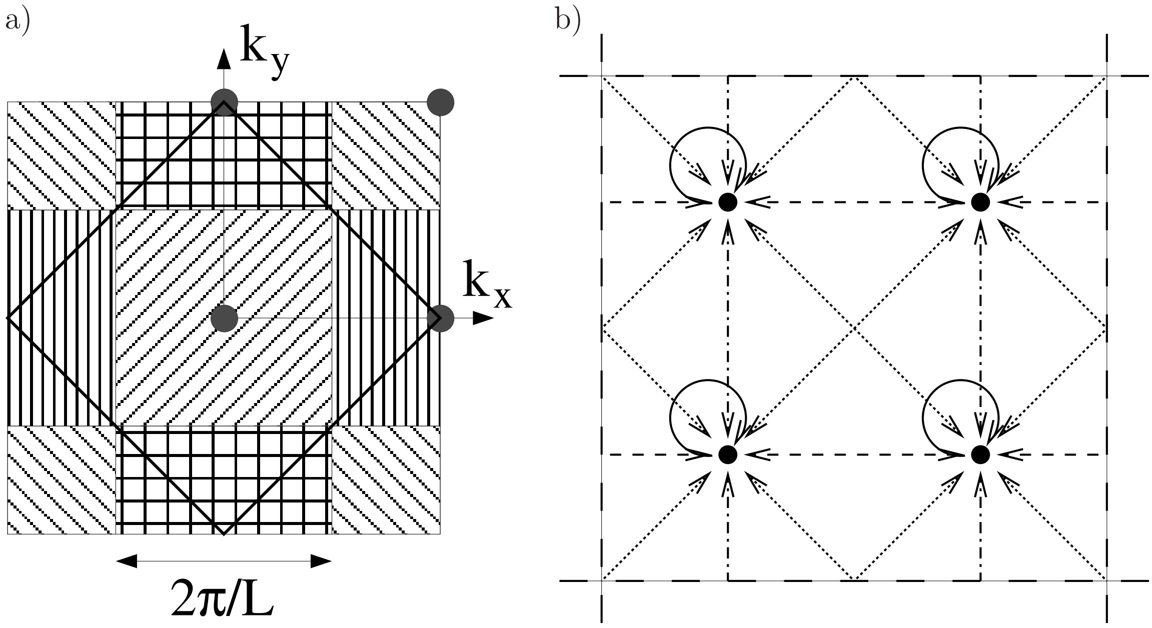



**Figure A.3:** DCA. **a)** Illustration of course-graining of the Brillouin zone for $d = 2$ and $L = 2$ [taken from Hettler et al. (2000)]. **b)** Real-space scheme of the resulting impurity model. Each impurity is not only influenced by a local bath (a frequency-dependent on-site energy), but also coupled by frequency-dependent hopping matrix elements to other impurities (arrows). The couplings also extend across the periodic cluster boundaries.

---

assumed to depend only on the coarse-grained momentum which may be chosen as $K_{\alpha l} = 2\pi l/L$, where $0 \leq l < L$ and $1 \leq \alpha \leq d$ labels the dimensions. This is illustrated in Fig. A.3 a for $d = 2$ and $L = 2$ where the Brillouin zone is mapped onto $L^2 = 4$ coarse-grained momenta. Given a self-energy with constant functional form (with respect to $\omega$) on each of the patches one may apply the obvious generalization of (1.30) for computing a coarse-grained lattice Green function which in turn defines a generalized bath Green function by using (1.35) for each $K$. The corresponding generalized real-space impurity problem, illustrated in Fig. A.3 b now involves $L^d$ impurities which do not only couple to frequency-dependent baths (which here appear as generalized on-site energies), but are also coupled by generalized hopping processes with frequency-dependent amplitudes. This problem has to be solved in presence of on-site interactions on each impurity in order to calculate a new Green function for the periodic cluster which is then interpreted as a course-grained lattice Green function so that the iteration scheme can be closed. The method evidently reduces to the DMFT for $L = 1$, but involves the solution of a problem even more complicated than a periodic finite-size Hubbard model in the general case. It can be shown to be causal at any $L$ and to approach the exact finite-dimensional model for $L \to \infty$; therefore it constitutes a very attractive interpolating scheme between finite and infinite dimensions.

The distinctive feature of CDMFT methods (Lichtenstein and Katsnelson, 2000; Kotliar, Savrasov, and Pálsson, 2001) (where “C” stands for “cluster” or “cellular”)

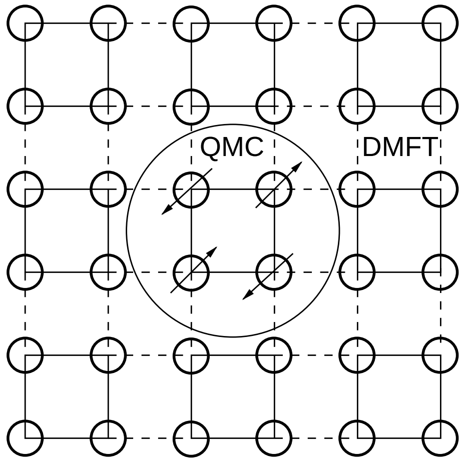



**Figure A.4:** CDMFT: the lattice is broken up in clusters (here with 2x2 sites); one of these is treated explicitly (e.g., using QMC) while the influence of the others is captured in a cavity picture analogous to the DMFT (scheme adapted from Lichtenstein and Katsnelson (2000)).

is the direct formulation in real (not momentum) space. Here, the lattice is divided up into free finite-size clusters, one of which is treated explicitly and embedded into a medium which gives a mean-field estimate of the effect of the surrounding clusters as illustrated in Fig. A.4. In contrast to the DCA, the cluster is here free, i.e., does not have periodic boundary conditions. Consequently, the analogue of (1.35) cannot be formulated in momentum space, but involves explicit matrix inversions in the cluster indices. Again, one can show causality and the interpolating character towards finite dimensions.

At least conceptionally, one should expect the CDMFT methods to yield a better description for short-range correlations within the first few nearest-neighbor shells than the DCA since here the cluster Hamiltonian is in direct correspondence to the lattice Hamiltonian. On the other hand, CDMFT breaks the lattice translational symmetry[6] which is preserved in DCA so that the DCA should give a better overall momentum resolution for the same number of cluster sites (and $L > 2$). First results for the $d = 2$ Hubbard model show that the nonlocal extensions of the DMFT indeed introduce $d$-wave type superconducting correlations (Lichtenstein and Katsnelson, 2000; Maier, Jarrell, Pruschke, and Keller, 2000). Very recently, finite-size effects were found to be smaller for the DCA treatment in a comparative study (Maier and Jarrell, 2002). For a detailed pedagogical review of the DCA method, see Jarrell, Maier, Hettler, and Tahvildarzadeh (2001).

The random-dispersion approximation (RDA) is not an extension, but an alternative to the DMFT approach. It avoids the self-consistency cycle inherent in all

[6] Since edge sites of the CDMFT cluster have a reduced number of direct hopping bonds, they are relatively stronger coupled to the effective medium and can, thus, be expected to be *more* mean-field like than inner sites. This is opposite to a direct calculation on a finite cluster with open boundary conditions where edge sites have *less* mean-field character.

solutions of the DMFT (and its extensions mentioned above). Instead, it is based on the factorization of the noninteracting 2-particle density of states (1.24), which is assumed to apply to all momenta except for $\boldsymbol{q} = \boldsymbol{0}$ and, where appropriate, for $\boldsymbol{q} = \boldsymbol{Q}$. Such a system can in principle be obtained for a given DOS as follows (Gebhard, 1997): for a finite system, one randomly assigns the kinetic energy $\epsilon_{\boldsymbol{k}}$ to each $\boldsymbol{k}$ point in the Brillouin zone where the probability $P(\epsilon)$ of finding the energy $\epsilon_{\boldsymbol{k}}$ in the semi-open interval $[\epsilon - \Delta\epsilon/2, \epsilon + \Delta\epsilon/2[$ is given in terms of the DOS by $P(\epsilon) = \rho(\epsilon)\Delta\epsilon$. In practice, one usually starts with a smooth dispersion and applies a random permutation to the momenta, so that the dispersion is again random, but the DOS of each realization exactly equals the (discretized) target DOS. Since the randomization destroys the topology of the lattice anyway, the method can be formally applied in any dimension; for large enough systems it is claimed to be self-averaging. Due to the neglect of correlations in the noninteracting 2-particle DOS, the RDA makes a large error for processes with small momentum transfer, even in large dimensions (cf. the discussion in subsection 1.2.2) which might affect thermodynamic properties although the associated phase space becomes vanishingly small. Still, it is not apparent how this should be a source of the large discrepancies between RDA estimates for the paramagnetic metal-insulator transition (Noack and Gebhard, 1999) in comparison to all recent DMFT estimates (see chapter 3). Up to now, numerical results for the RDA have suffered from considerable finite-size effects, since the number of $\boldsymbol{k}$ points was restricted to a value of 14 or smaller so far in an exact diagonalization scheme. Reliable evaluation methods for bigger systems are needed for the RDA problem, e.g., an efficient formulation of the density matrix renormalization group (DMRG) method (White, 1992; 1993) in $\boldsymbol{k}$-space, in order to clarify if the different transition scenario is a genuine property of the RDA theory.

# Appendix B

# Hyperdiamond Lattice

The hyperdiamond (hd) lattice is a bipartite non-Bravais lattice which generalizes the $d = 2$ honeycomb lattice and the $d = 3$ diamond lattice to arbitrary dimensions. In addition to the usual definition and recursive construction (Santoro, Airoldi, Sorella, and Tosatti, 1993) we will here use a representation in $d + 1$ dimensions (van Dongen, 1997). As we will show, only the latter clearly exposes the isotropy of the hyperdiamond lattice in the long-wavelength limit and allows for a computation of the optical conductivity not only in $d = \infty$ but, within the local approximation for self-energy $\Sigma$ and particle-hole irreducible vertex $\Gamma$, also for arbitrary finite dimensions (see chapter 4). Although an extension to longer range hopping is straightforward we will in this section only consider NN hopping.

A general characterization of the hyperdiamond lattice in dimension $d \geq 1$ is (Santoro et al., 1993)

1. each lattice site has $d + 1$ nearest neighbors (NN),
2. the angle between each pair of $d+1$ unit[1] primitive lattice vectors $\boldsymbol{\tau}_l$ connecting a site to its NN is constant,
$$\boldsymbol{\tau}_l \cdot \boldsymbol{\tau}_m = \gamma_d, \tag{B.1}$$
3. for $d > 1$, there is no site-centered inversion symmetry, but only bond-centered inversion symmetry,

where the last statement follows from the first two. The lack of inversion symmetry implies that the unit vectors $\{\boldsymbol{\tau}_l\}$ cannot be the same on $A$ and $B$ sublattices. Instead they are related by an inversion operation, $\boldsymbol{\tau}_l^B = -\boldsymbol{\tau}_l^A$. Thus it is useful to rewrite the noninteracting Hamiltonian by explicitly separating $A$ and $B$ lattice sites,

$$\begin{aligned}
\hat{H}_0 &= -t \sum_{\langle i,j\rangle,\sigma} \left( \hat{c}^\dagger_{\boldsymbol{r}_i\sigma} \hat{c}_{\boldsymbol{r}_j\sigma} + \hat{c}^\dagger_{\boldsymbol{r}_i\sigma} \hat{c}_{\boldsymbol{r}_j\sigma} \right) \\
&= -t \sum_{i\in A,\sigma} \sum_l \hat{c}^\dagger_{\boldsymbol{r}_i+\boldsymbol{\tau}_l^A,\sigma} \hat{c}_{\boldsymbol{r}_i\sigma} - t \sum_{j\in B,\sigma} \sum_l \hat{c}^\dagger_{\boldsymbol{r}_j+\boldsymbol{\tau}_l^B,\sigma} \hat{c}_{\boldsymbol{r}_j\sigma} \\
&= -t \Big[ \sum_{i\in A,\sigma} \sum_l \hat{c}^\dagger_{\boldsymbol{r}_i+\boldsymbol{\tau}_l,\sigma} \hat{c}_{\boldsymbol{r}_i\sigma} + \sum_{j\in B,\sigma} \sum_l \hat{c}^\dagger_{\boldsymbol{r}_j-\boldsymbol{\tau}_l,\sigma} \hat{c}_{\boldsymbol{r}_j\sigma} \Big],
\end{aligned} \tag{B.2}$$

[1] As usual, we define the lattice spacing such that NNs have unit distance.

where $\langle i,j\rangle$ denotes NN hopping bonds and $\boldsymbol{\tau}_l \equiv \boldsymbol{\tau}_l^A$.

For the hyperdiamond lattice (hd), these hopping vectors can be constructed recursively:

1. start with a regular chain,

$$\boldsymbol{\tau}_1 = (1), \quad \boldsymbol{\tau}_2 = (-1) \quad \text{for } d = 1 \tag{B.3}$$

2. given a set $\{\boldsymbol{\tau}_l'\}$ of primitive lattice vectors in dimension $d-1$, define

$$\boldsymbol{\tau}_l = \begin{cases} (\sqrt{1-\gamma_d^2}\,\boldsymbol{\tau}_l', \gamma_d), & l = 1,2,\ldots,d \\ \underbrace{(0,0,\ldots,0,1)}_{d \text{ components}}, & l = d+1 \end{cases} \tag{B.4}$$

$$\gamma_d = \frac{\gamma_{d-1}}{1-\gamma_{d-1}} = -\frac{1}{d}\,. \tag{B.5}$$

Now, the noninteracting tight-binding Hamiltonian (with NN hopping) in Fourier transformed form remains offdiagonal in the sublattice index $\alpha \in \{A, B\}$,

$$H_0 = \begin{pmatrix} 0 & H_1 \\ H_1^* & 0 \end{pmatrix}, \quad H_1 = -t \sum_{m=1}^{d+1} \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{\tau}_m}, \tag{B.6}$$

where the eigenvalues form two energy bands $\epsilon_{n\boldsymbol{k}} = \pm|H_1|$. This expression can be simplified by a transformation in momentum space ($\alpha = 1,\ldots,d$),

$$k_\alpha' = \sum_\beta A_{\alpha\beta} k_\beta, \quad A_{\alpha\beta} = (\boldsymbol{a}_\alpha)_\beta, \quad \boldsymbol{a}_\alpha = \boldsymbol{\tau}_{d+1} - \boldsymbol{\tau}_\alpha\,. \tag{B.7}$$

Eliminating an overall phase factor $\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{\tau}_{d+1}}$ we have

$$\epsilon_{n\boldsymbol{k}} = \pm t\Big|1 + \sum_{\alpha=1}^{d} \mathrm{e}^{ik_\alpha'}\Big|, \quad -\pi \leq k_\alpha' \leq \pi, \tag{B.8}$$

which provides a convenient way to compute the DOS and shows that perfect nesting is not possible on the bipartite hd lattice. In the limit $d \to \infty$, where the hopping amplitude has to be scaled as $t = t^*/(d+1)$, the DOS is readily evaluated analytically,

$$\rho^{\mathrm{hd}}(\epsilon) = \frac{|\epsilon|}{t^{*2}} \mathrm{e}^{-\epsilon^2/t^{*2}}\,. \tag{B.9}$$

Calculations of properties which directly depend on $\boldsymbol{k}$ (i.e., not only via $\epsilon_{n\boldsymbol{k}}$), in particular transport properties, however, require knowledge of the transformation matrix $(A_{\alpha\beta})$. As a general example for this problem let us consider the Fermi velocity $\boldsymbol{v}_{\boldsymbol{k}} = -\nabla\epsilon_{\boldsymbol{k}}$ for an energy function defined in terms of transformed $k$ vectors, $\epsilon_{\boldsymbol{k}} = f(\boldsymbol{C}\boldsymbol{k})$ with arbitrary function $f$ and matrix $\boldsymbol{C}$. In the present case, where the functional form of $f$ is explicitly known, one can first compute the derivative $\boldsymbol{v}'_{\boldsymbol{k}'} = -\nabla_{\boldsymbol{k}'} f(\boldsymbol{k}')$ and then apply the transformation

$$\begin{aligned} \frac{\partial f(\boldsymbol{C}\boldsymbol{k})}{\partial k_\alpha} &= \sum_\beta \frac{\partial f}{\partial k'_\beta} \frac{\partial (\boldsymbol{C}\boldsymbol{k})_\beta}{\partial k_\alpha} = \sum_\beta A_{\beta\alpha} \frac{\partial f}{\partial k'_\beta} & \text{(B.10)} \\ \boldsymbol{v}_{\boldsymbol{k}} &= \boldsymbol{C}^T \boldsymbol{v}'_{\boldsymbol{k}'}\big|_{\boldsymbol{k}'=\boldsymbol{C}\boldsymbol{k}} \,. & \text{(B.11)} \end{aligned}$$

While the transformation matrix defined via (B.3)-(B.5) and (B.7) can be evaluated in any finite dimension, this is in general a complicated task and it is unclear how to take the limit $d \to \infty$ of (B.10). Furthermore, symmetries are hidden so that it is not even apparent if observables which transform like a tensor such as, e.g., the optical conductivity will be isotropic in all dimensions.

Thus, we switch to an alternative representation of the hyperdiamond lattice which embeds the $d$-dimensional hd lattice into $d+1$ host dimensions in the following way (van Dongen, 1997): starting with a $d+1$-dimensional hypercube with NN hopping we remove every second bond such that the unit vectors connecting to NNs are for sublattices $A$ and $B$

$$\tilde{\boldsymbol{\tau}}_l^A = \boldsymbol{e}_l, \quad \tilde{\boldsymbol{\tau}}_l^B = -\boldsymbol{e}_l \,. \tag{B.12}$$

Here, $\{\boldsymbol{e}_l\}$ are the usual cartesic unit vectors in $d+1$ dimensions; the tilde ($\sim$) indicates properties in the expanded $d+1$-dimensional vector space. For sake of definiteness we define the origin to belong to the $A$ sublattice. Figure B.1 illustrates this embedding for $d = 2$, i.e., the honeycomb lattice.

It is clear from (B.2) that a tight-binding Hamiltonian defined on a periodic lattice with fixed uniform NN hopping is fully defined by the topology of the hopping bonds and invariant under distortions. Since the lattice defined by (B.12) has $d+1$ NN bonds with a constant angle (here: $\pi/2$) between each pair of them, it (more precisely: each sheet of sites connected by hopping bonds) is indeed topologically equivalent to the usual hyperdiamond lattice and has both the same noninteracting DOS and the same single-particle properties for arbitrary additional local interaction. The noninteracting dispersion is given by

$$\epsilon_{\tilde{\boldsymbol{k}}} = \pm\frac{1}{2}\sqrt{(\epsilon^{hc}_{\tilde{\boldsymbol{k}}})^2 + \tilde{\epsilon}^2_{\tilde{\boldsymbol{k}}}}, \tag{B.13}$$

where $\epsilon^{hc}_{\tilde{\boldsymbol{k}}} = -2t\sum_{l=1}^{d+1}\cos(\tilde{k}_l)$ and $\tilde{\epsilon}_{\tilde{\boldsymbol{k}}} = 2t\sum_{l=1}^{d+1}\sin(\tilde{k}_l)$. Again, one of the $\tilde{k}_l$ could be singled out and be fixed for calculations of the DOS, exposing the equivalence of the dispersions (B.8) and (B.13). Since $\langle(\epsilon^{hc}_{\tilde{\boldsymbol{k}}})^2\rangle = \langle\tilde{\epsilon}^2_{\tilde{\boldsymbol{k}}}\rangle = 2(d+1)t^2$, we can read off from (B.13) that the variance of the DOS of the hyperdiamond lattice is exactly $(d+1)t^2 = t^{*2}$ in any dimension. Note, however, that in the present derivation $\tilde{\boldsymbol{k}}$ remains untransformed within the (halved) Brillouin zone which allows for a direct calculation of the Fermi velocity (in $d+1$ dimensions),

$$(\tilde{\boldsymbol{v}}_{\tilde{\boldsymbol{k}}})_l = \frac{t}{\epsilon_{\tilde{\boldsymbol{k}}}}\left(\epsilon^{hc}_{\tilde{\boldsymbol{k}}}\sin(\tilde{k}_l) + \tilde{\epsilon}_{\tilde{\boldsymbol{k}}}\cos(\tilde{k}_l)\right) \,. \tag{B.14}$$

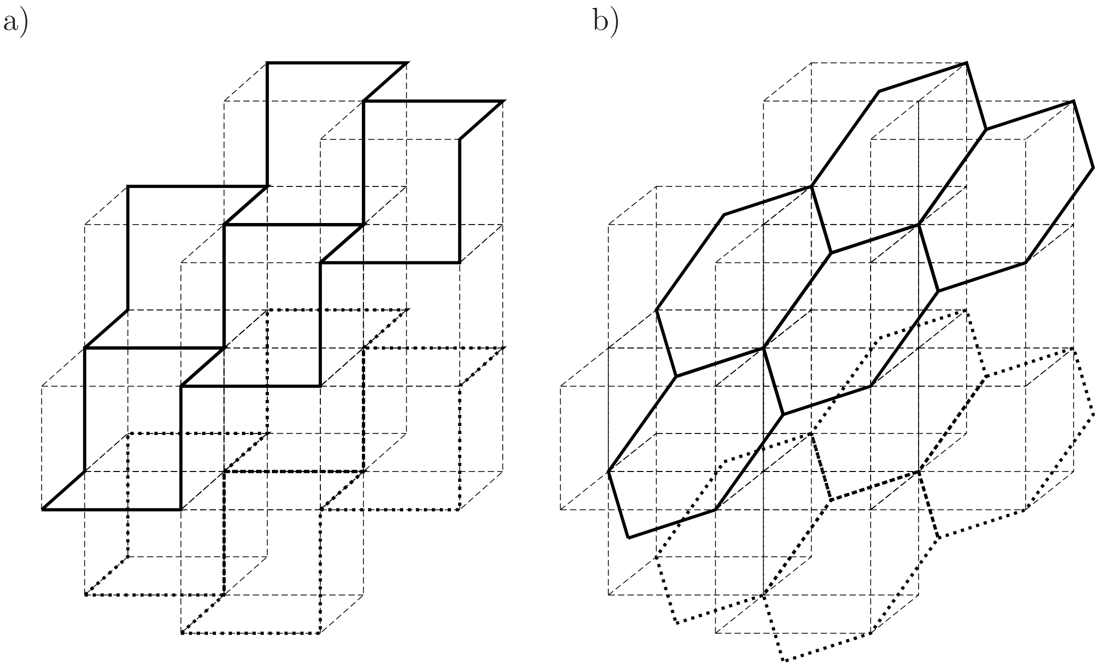



**Figure B.1:** Honeycomb lattice ($d = 2$) embedded in $d + 1 = 3$ dimensions. **a)** Hopping bonds, i.e., distribution of nonzero matrix elements according to (B.12). **b)** Projection of $d$-dimensional sublattices onto planes.

---

Since the spectrum of a Hamiltonian is unchanged under any basis transformation we can perform the $\boldsymbol{k}$ sum in (2.1) directly in the $d+1$-dimensional representation. The numerical result, Fig. B.2, shows that the hd DOS vanishes at $\epsilon = 0$ for $d \geq 2$. More precisely, for $|\epsilon| \to 0$, $\rho(\epsilon) \propto |\epsilon \ln(\epsilon)|$ for the diamond lattice ($d = 3$) and $\rho(\epsilon) \propto |\epsilon|$ otherwise. Thus, the hyperdiamond lattice is a semimetal at half filling in the noninteracting limit.

In order to extract transport properties one has to eliminate the extra dimension by projecting back onto the physical $d$-dimensional space, a hyperplane in $d + 1$ dimensions. In a first step we classify the sites belonging to each planar sheet by considering the quantity $s(\boldsymbol{r}) := \sum_{\alpha=1}^{d} r_\alpha$. We define the origin with $s(\boldsymbol{r} = \boldsymbol{0}) = 0$ to belong to the $A$ sublattice. From (B.12) we can read off that the neighboring $B$ sites connected by hopping bonds have $s = +1$. More generally, hopping from $A$ to $B$ sublattice increases $s$ by one, while hopping from $B$ to $A$ always decreases $s$ by one. Thus, all sites $\boldsymbol{r}_i^n$ belonging to the planar sheet $n$ are identified by $\lfloor s(\boldsymbol{r}_i^n)/2 \rfloor = n$, where $n \in \mathbb{Z}$ and $\lfloor x \rfloor$ is the largest integer $n \leq x$. Thus, the extent of each sheet in the $d+1$-dimensional space-diagonal direction

$$\mathbf{1} = \frac{1}{\sqrt{d+1}}(1, 1, \ldots, 1) \tag{B.15}$$

is finite with Euclidean thickness $1/\sqrt{d+1}$ so that any transport must vanish in this direction. Projecting out direction $\mathbf{1}$ and using $\tilde{\boldsymbol{\tau}}_l \cdot \mathbf{1} = 1/\sqrt{d+1}$ we obtain for the

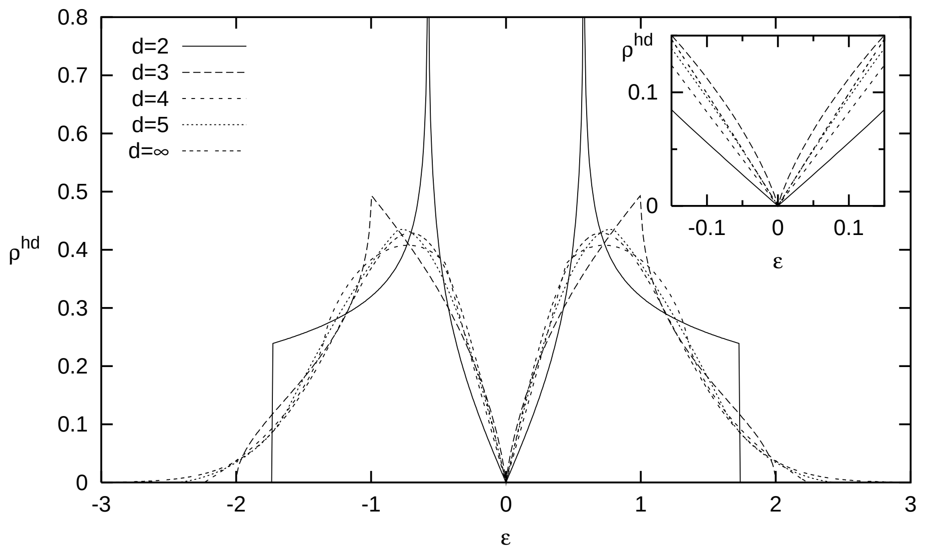



**Figure B.2:** DOS of the hd lattice.

unit hopping vectors,

$$\boldsymbol{\tau}_l = \sqrt{\frac{d+1}{d}}\big(\tilde{\boldsymbol{\tau}}_l - \frac{1}{\sqrt{d+1}}\mathbf{1}\big)\ . \tag{B.16}$$

After the projection, these are no longer orthogonal,

$$\boldsymbol{\tau}_l \cdot \boldsymbol{\tau}_m = \frac{d+1}{d}(\tilde{\boldsymbol{\tau}}_l - \frac{1}{\sqrt{d+1}}\mathbf{1})(\tilde{\boldsymbol{\tau}}_l - \frac{1}{\sqrt{d+1}}\mathbf{1}) \tag{B.17}$$

$$= \frac{d+1}{d}(\delta_{lm} - \frac{2}{d+1} + \frac{1}{d+1}) = \begin{cases} 1 & \text{for } l = m \\ -\frac{1}{d} & \text{for } l \neq m \end{cases}\ , \tag{B.18}$$

but fulfill the original definitions of the hyperdiamond lattice and are therefore identical to those defined in (B.1) up to a rotation.

Computations of tensor-like properties can be easily performed in the new formalism. First we note that in $d+1$ dimensions all cartesic directions are equivalent. Thus, any tensor $\tilde{\boldsymbol{C}}$ in the $d+1$-dimensional representation which is not dependent on coordinates must obey this symmetry and have the form

$$\tilde{\boldsymbol{C}} = \begin{pmatrix} c_1 & c_2 & \cdots & c_2 \\ c_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_2 \\ c_2 & \cdots & c_2 & c_1 \end{pmatrix}\ . \tag{B.19}$$

This matrix has the $d$ fold eigenvalue $c_\parallel = c_1 - c_2$ plus the nondegenerate eigenvalue $c_\perp = c_1 + dc_2$. As expected the eigenvector corresponding to $c_\perp$ is $\mathbf{1}$, the direction orthogonal to the planes. For transport properties, $c_\perp$ must vanish, $c_2 = -c_1/d$, and we obtain

$$\boldsymbol{C} = c_1 \frac{d+1}{d} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} . \tag{B.20}$$

Thus, we have not only proven that transport properties are isotropic on the hyperdiamond lattice in the long-wavelength limit[2] for any dimension $d$, but also constructed a simple algorithm for the computation provided that the corresponding quantity can be computed for the model in $d+1$ dimensions as defined by (B.12). Numerical results for transport related quantities of the hd lattice can be found in subsection 4.5.2.

[2]Note that for the honeycomb lattice the isotropy also follows from the simultaneous presence of mirror symmetries $x \to -x$, $y \to -y$ through the center of a bond in $x$ or $y$ direction and the rotational symmetry $(x, y) \to (x\cos(2\pi/n) + y\sin(2\pi/n))$ with $n > 2$ (here $n = 3$).

# Appendix C

# Using Splines for Fourier-Transforming Imaginary-Time Green Functions

As discussed in subsection 3.4.1, the Fourier transformations of Green functions from imaginary time to Matsubara frequencies and vice versa that arise in the QMC solution of the DMFT problem are not trivial. All problems can be traced back to the discreteness of the QMC estimates $\bar{G}_l$ for time slices $\tau_l = l\Delta\tau$. However, the true imaginary-time Green function $G(\tau)$ is known to be a smooth curve with continuous derivatives (of all orders) on the interval $[0, \beta]$. Thus, reasonable results can be expected by performing the Fourier transform not for the discrete raw data, but for a fitted smooth curve. In this appendix, we will give a short summary of work carried out recently in collaboration with Knecht. More details can be found in Knecht's (2002) diploma thesis and in upcoming publications.

The simplest practical approach in this context consists of a direct interpolation of the discrete QMC data by a cubic spline (with a continuous second derivative) as implemented by Krauth. While this step suffices for closing the self-consistency equations without any further adjustments (like Ulmke's smoothing trick), it leads to nonanalytic behavior of the self-energy near and beyond the Nyquist frequency as illustrated in Fig. 3.11. In our view, this problem can be traced back to the fact that the natural spline chosen by Krauth is inadequate for this problem: By definition, the second derivative of a natural spline vanishes at its boundaries. However, all even derivatives of the true Green function are maximal at the edges of the interval $[0, \beta]$. The resulting misfit leads to unphysical ringing as illustrated in Fig. C.1. Here, the noninteracting Green function for a semi-elliptic DOS is chosen as an example since it can be computed with arbitrary precision. Furthermore, moderate interactions do not lead to qualitative changes in the Green function so that the example is representative. It is clearly seen that a fit of a discrete set of data points (here for $\Lambda = 40$ and $\beta = 100$, i.e., a large discretization $\Delta\tau = 2.5$) using a natural cubic spline with continuous second derivative leads to a large error oscillating with the Nyquist frequency. While the initial misfit at $\tau = 0$ is similarly large for a natural Akima spline, the oscillation decays significantly faster due to jumps in the second derivative

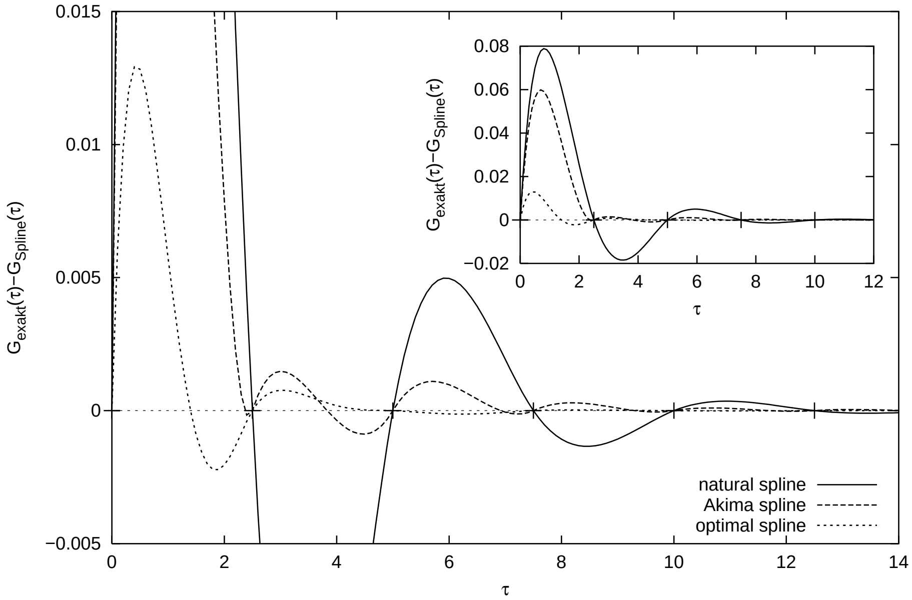



**Figure C.1:** Difference between exact noninteracting Green function (for semi-elliptic DOS and $\beta = 100$) and cubic spline approximations to the discretized function for $\Lambda = 40$. The strong oscillations observed for the natural spline and for the Akima spline are due to the unphysically vanishing second derivatives (at $\tau = 0$) of these splines. An optimal choice of the boundary condition reduces the discrepancies by an order of magnitude and doubles the oscillation frequency. The inset shows the full curves for small $\tau$.

for this fit. The discrepancies can be reduced by about an order of magnitude and a smooth curve can be obtained by choosing an appropriate boundary condition (short-dashed line) which has here been obtained by minimizing the function

$$F(G^{(2)}(0)) = \sum_{j=1}^{L-1} \left( \frac{G^{(2)}((j+1)\Delta\tau)}{G((j+1)\Delta\tau)^2} - \frac{G^{(2)}(j\Delta\tau)}{G(j\Delta\tau)^2} \right)^2 . \tag{C.1}$$

It is possible to avoid the minimization by computing the correct second derivative $G^{(2)}(0)$ analytically (for arbitrary interaction); in any case, however, a cubic spline interpolation of the full Green function suffers from the problem that derivatives of fourth and higher order vanish on segments of the splines while they are large for the true Green function, in particular at $\tau = 0$.

Consequently, a further significant reduction of errors can be expected when the spline approximation is applied to the difference between the measured (discretized) Green function $\{\bar{G}_l\}$ and a good estimate for $G$ when the Fourier transform is known exactly for the latter. Such a scheme can be based on the following exact high-frequency expansion for the self-energy (Potthoff et al., 1997):

$$\Sigma_\sigma(\omega) = U\,(\langle \hat{n}_{-\sigma}\rangle - \frac{1}{2}) + U^2\, \frac{\langle \hat{n}_{-\sigma}\rangle(1 - \langle \hat{n}_{-\sigma}\rangle)}{\omega} + \mathcal{O}(\omega^{-2}). \tag{C.2}$$

One possible choice of a model self-energy with this asymptotic behavior (including the $\omega^{-1}$ term) which is nonsingular (and purely imaginary) on the imaginary axis is given by

$$\Sigma_{\text{model},\sigma}(\omega) = U\left(\langle \hat{n}_{-\sigma}\rangle - \frac{1}{2}\right) + \frac{1}{2}U^2 \langle \hat{n}_{-\sigma}\rangle\big(1 - \langle \hat{n}_{-\sigma}\rangle\big)\Big(\frac{1}{\omega+\omega_0} + \frac{1}{\omega-\omega_0}\Big). \qquad \text{(C.3)}$$

While the quality of the low-frequency part of this fit could be tuned by adjusting the parameter $\omega_0$, the value of this parameter is irrelevant for the following as long as it is not much larger than the bandwidth; we will choose $\omega_0 = 1$. By evaluating $\Sigma_{\text{model}}$ on the imaginary axis, the corresponding Green function $G_{\text{model}}$ can be computed for an arbitrary number of Matsubara frequencies. Consequently, the Fourier transformation to imaginary time is unproblematic (when the "free" term $1/(i\omega_n)$ is taken care of analytically). Thus, the difference of the Green functions can be accurately evaluated at all time slices. Since the second derivative of the transformed model Green function at $\tau = 0$ exactly reproduces that of the true Green function, the difference is well represented by a natural spline. The Green function is obtained as a function of Matsubara frequencies by Fourier transforming the oversampled spline and adding the Matsubara-frequency model Green function to the result.

Predictions for the imaginary-time self-energy of the half-filled Hubbard model with semi-elliptic DOS at $T = 0.1$ resulting from the scheme described above for a range of discretizations $0.1 \leq \Delta\tau \leq 0.4$ are shown in Fig. C.2 for the correlated metallic phase at $U = 4.0$ and in Fig. C.3 for the insulating phase at $U = 5.0$. All results show the correct asymptotic behavior at large frequencies. Oscillatory behavior which is initially observed at intermediate frequencies (see insets) becomes undetectable for $\Delta\tau \lesssim 0.2$ in both figures. The convergence at small frequencies is good; it will be further discussed below.

One fundamental difference of the QMC code previously used in our group (based on Ulmke's smoothing trick, cf. subsection 3.4.1) is that it does not make predictions for the self-energy beyond the Nyquist frequency as seen in Fig. C.4 and Fig. C.5. Even worse, the $\Delta\tau$ error is also large at intermediate and small frequencies. In fact, the insulating solutions are even qualitatively wrong at small frequencies for large values of $\Delta\tau$.

In contrast, the results of the improved QMC method developed within this thesis and used throughout chapter 3 depicted in Fig. C.6 and Fig. C.7 show good convergence in the most important small-frequency region. In fact, the $\Delta\tau$ error is asymptotically almost identical to that of the new optimal scheme at small frequencies, at least in the range $\Delta\tau \leq 0.25$ used within this thesis.

This is also seen in Fig. C.8, where the Eliashberg estimate of the quasiparticle weight $Z$ (cf. subsection 3.2.3) is shown as a function of $\Delta\tau^2$. Both in the metallic and in the insulating phase, the new approach yields the smallest errors which are nearly perfectly quadratic in $\Delta\tau$. The systematic errors of the QMC program used in chapter 3 are identical (within statistical errors) at small $\Delta\tau$ and only slightly larger for $\Delta\tau \gtrsim 0.25$. In contrast, Ulmke's method leads to $\Delta\tau$ errors which are larger by a full order of magnitude in the insulating case.

While these results of Knecht's (2002) thesis reproduced in this appendix confirm that the QMC code used in chapter 3 is indeed adequate for studies of the Mott

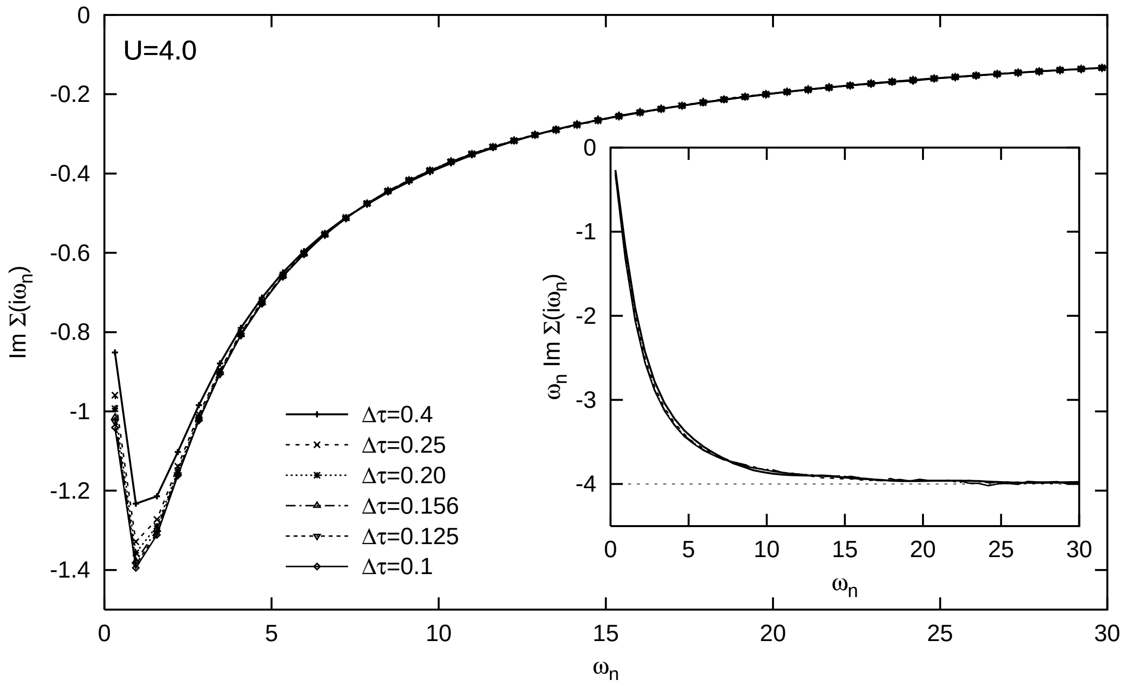



**Figure C.2:** Imaginary part of the self-energy $\Sigma(i\omega_n)$ for $T = 0.1$ and $U = 4.0$ computed using the new Fourier transformation scheme described in this appendix for various values of $\Delta\tau$. Inset: Computed $\omega_n \operatorname{Im} \Sigma(i\omega_n)$ in comparison with the exact asymptotic value -4.

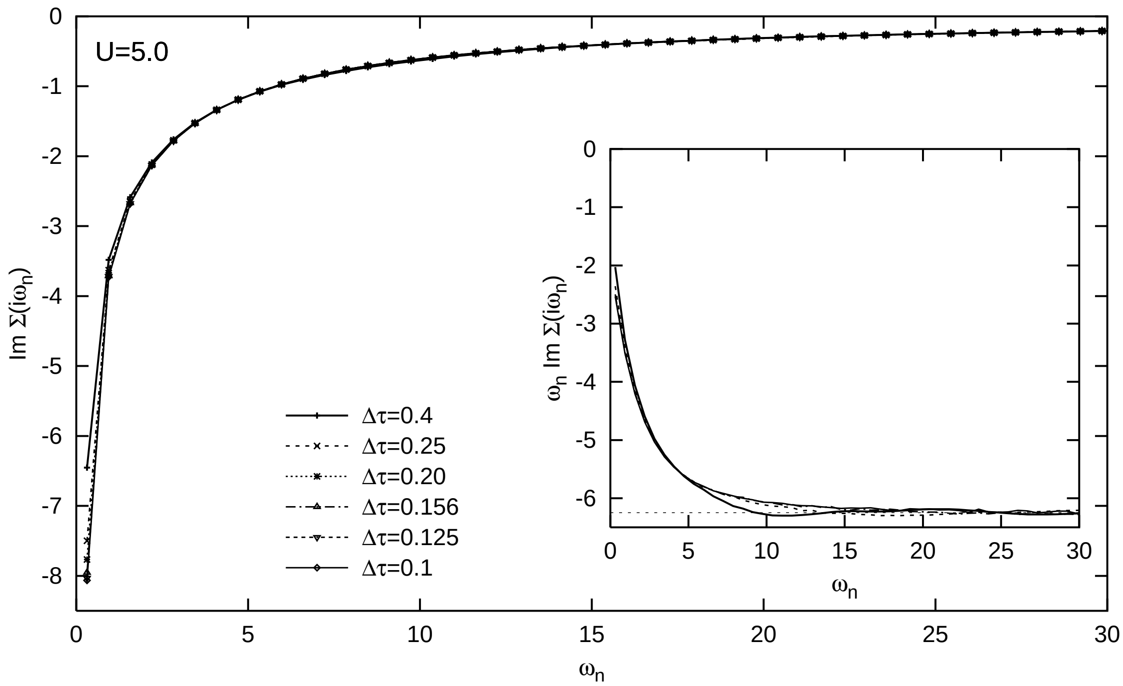



**Figure C.3:** Imaginary part of the self-energy $\Sigma(i\omega_n)$ for $T = 0.1$ and $U = 5.0$ computed using the new Fourier transformation scheme described in this appendix for various values of $\Delta\tau$. Inset: Computed $\omega_n \operatorname{Im} \Sigma(i\omega_n)$ in comparison with the exact asymptotic value -6.25.

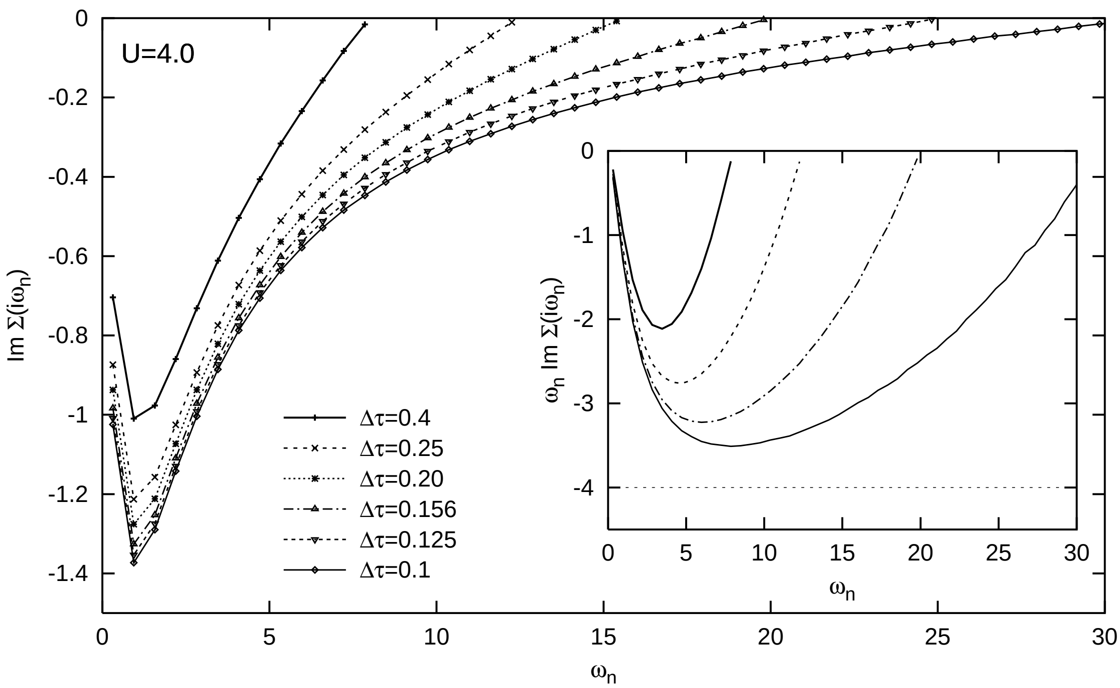



**Figure C.4:** Imaginary part of the self-energy $\Sigma(i\omega_n)$ for $T = 0.1$ and $U = 4.0$ computed using Ulmke's smoothing scheme for various values of $\Delta\tau$. Inset: Computed $\omega_n \operatorname{Im}\Sigma(i\omega_n)$ in comparison with the exact asymptotic value -4.

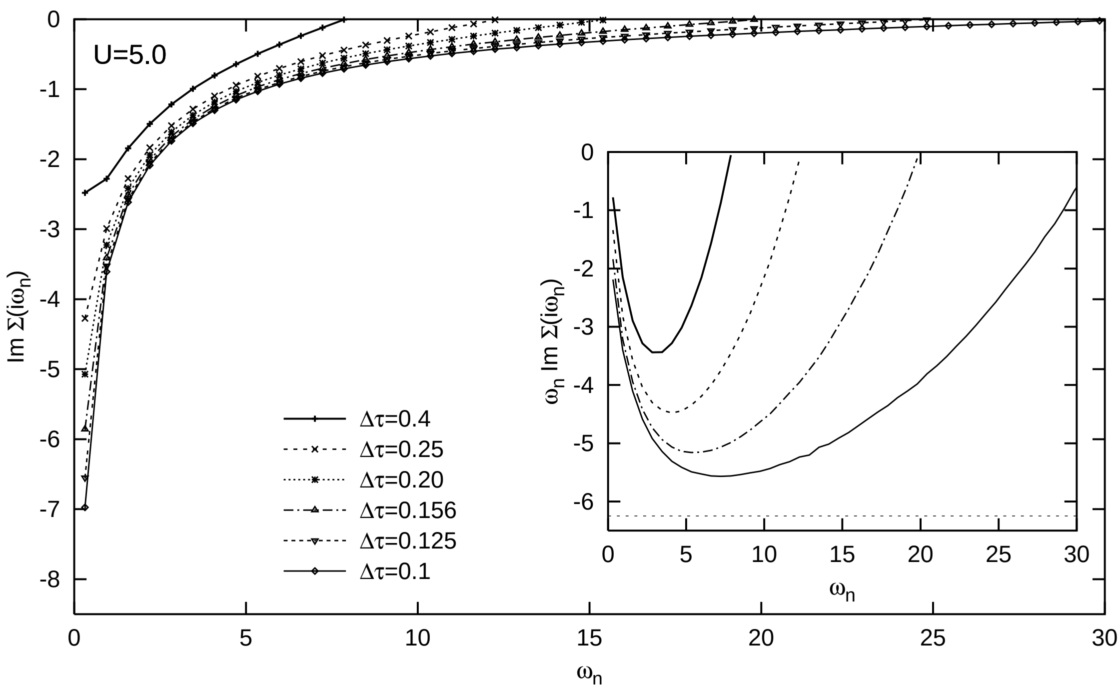



**Figure C.5:** Imaginary part of the self-energy $\Sigma(i\omega_n)$ for $T = 0.1$ and $U = 5.0$ computed using Ulmke's smoothing scheme for various values of $\Delta\tau$. Inset: Computed $\omega_n \operatorname{Im}\Sigma(i\omega_n)$ in comparison with the exact asymptotic value -6.25.

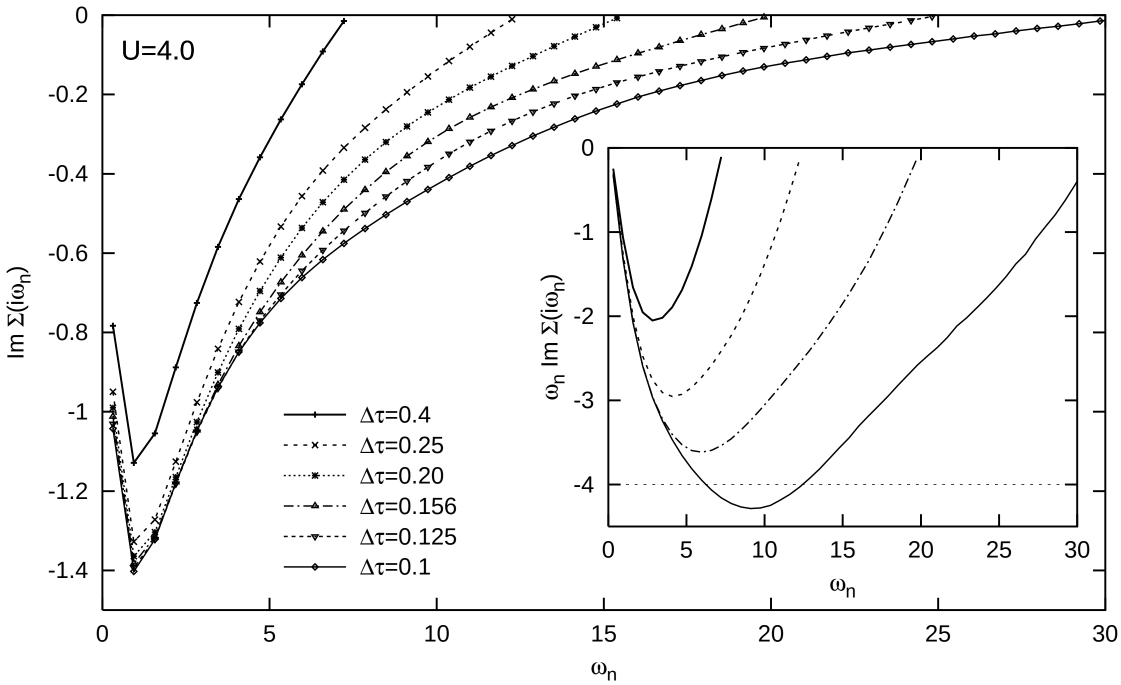



**Figure C.6:** Imaginary part of the self-energy $\Sigma(i\omega_n)$ for $T = 0.1$ and $U = 4.0$ resulting from the QMC code used throughout chapter 3 for various values of $\Delta\tau$. Inset: Computed $\omega_n \operatorname{Im}\Sigma(i\omega_n)$ in comparison with the exact asymptotic value -4.

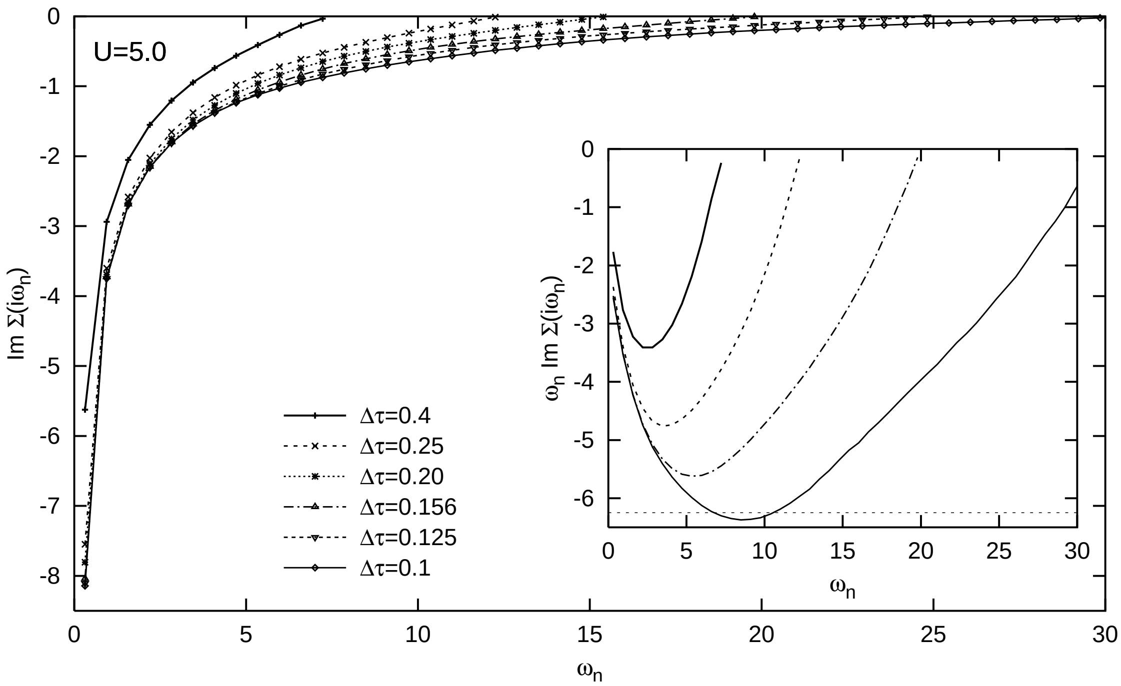



**Figure C.7:** Imaginary part of the self-energy $\Sigma(i\omega_n)$ for $T = 0.1$ and $U = 5.0$ resulting from the QMC code used throughout chapter 3 for various values of $\Delta\tau$. Inset: Computed $\omega_n \operatorname{Im}\Sigma(i\omega_n)$ in comparison with the exact asymptotic value -6.25.

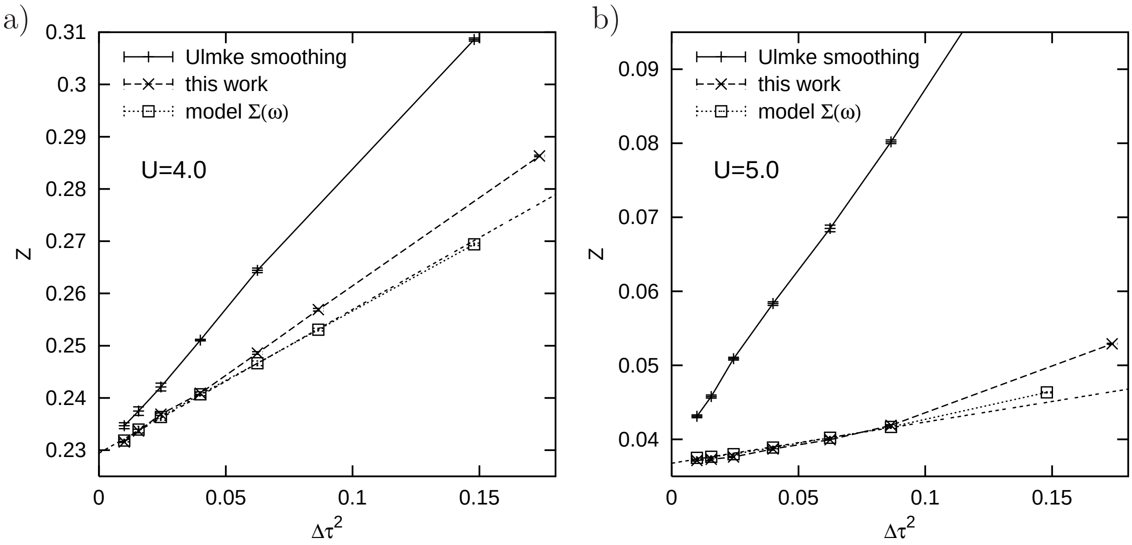



**Figure C.8:** QMC estimates of the quasiparticle weight $Z$ for semi-elliptic DOS at $T = 0.1$ as a function of $\Delta\tau^2$.

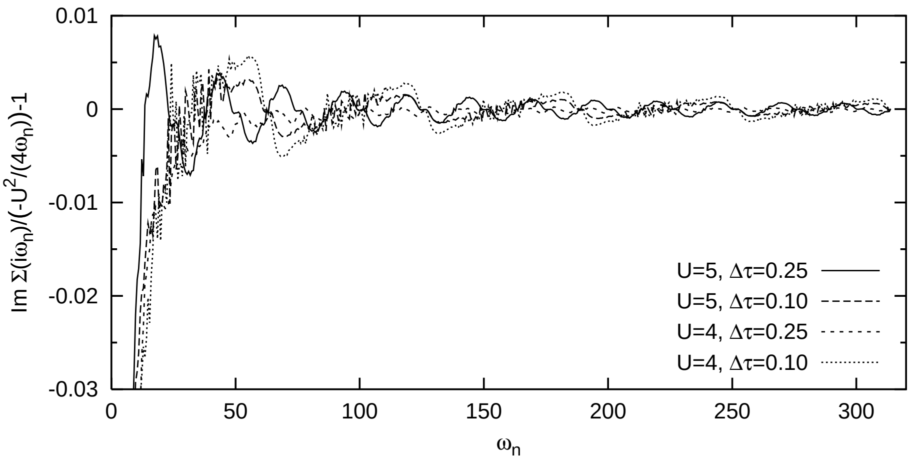



**Figure C.9:** Relative deviation of $\operatorname{Im}\Sigma(i\omega_n)$ (as computed using the new Fourier-transformation scheme for $T = 0.1$) from the asymptotic large-frequency form $-U^2/(4\omega_n)$; for the low-frequency parts of $\operatorname{Im}\Sigma(i\omega_n)$, see Fig. C.2 and Fig. C.3.

transition they also show that even more accurate results can be calculated at lower computational cost using the new code. We note that the concept of the method described here is similar to Jarrell's QMC implementation. The latter is, however, less stable since its Fourier transformation relies on a numerical IPT calculation. At least for a symmetric DOS at half filling, our method is unconditionally stable for arbitrarily large frequencies without the need for bandpass filters. This is illustrated in Fig. C.9. Evidently, even *relative* deviations from the asymptotic behavior never exceed 0.01 at large frequencies ($|\omega_n| \gtrsim 20$) and decay fast in the limit $|\omega| \to \infty$

which is here cut off by the (constant) number of 1000 Matsubara frequencies used in the QMC program. Note that the large deviations from the asymptotic behavior at relatively low frequencies are physical; indeed, these deviations are the nontrivial result of the QMC calculations.

For further practical applications of the new QMC code and for the generalization of the asymptotic expansion of the self-energy to the multi-band case, see Knecht (2002). Interestingly, the multi-band case is more complicated: here, the expansion of the self energy does not only involve expectation values of densities, but also of pairwise double occupancies.

# Appendix D

# Linear Response to Electromagnetic Fields

In this appendix, we review the formalism necessary for identifying the quantized Hamiltonian for particles in an electromagnetic field and give a short summary of linear-response theory.

## D.1 Electromagnetic Interaction Hamiltonian and Choice of Gauge

Classically, the equation of motion for a nonrelativistic particle with charge $q$ and momentum $m\boldsymbol{v}$ at position $\boldsymbol{r}$ in an external electromagnetic field is[1]

$$m\ddot{\boldsymbol{r}} = q\boldsymbol{E}(\boldsymbol{r},t) + \frac{q}{c}\dot{\boldsymbol{r}} \times \boldsymbol{B}(\boldsymbol{r},t), \tag{D.1}$$

where the right hand side is the Lorentz force. The electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ can be expressed in terms of a vector potential $\boldsymbol{A}$ and a scalar potential $\phi$,

$$\boldsymbol{B}(\boldsymbol{r},t) = \nabla \times \boldsymbol{A}(\boldsymbol{r},t); \quad \boldsymbol{E}(\boldsymbol{r},t) = -\nabla\phi(\boldsymbol{r},t) - \frac{1}{c}\frac{\partial \boldsymbol{A}(\boldsymbol{r},t)}{\partial t}. \tag{D.2}$$

A generalization for quantum systems is best derived via the manifestly Lorentz invariant action which for a world line beginning at $a = (ct_1, \boldsymbol{r}_2)$ and ending at $b = (ct_2, \boldsymbol{r}_2)$ reads [see, e.g., Landau and Lifschitz (1992)]

$$S = \int_a^b (-mc^2\,\mathrm{d}\tau - \sum_\alpha \frac{q}{c} A_\alpha\,\mathrm{d}x^\alpha), \tag{D.3}$$

Here, $d\tau = dt/\gamma$, where $\gamma^{-1} = \sqrt{1-(v/c)^2}$; $x^\alpha = (ct, \boldsymbol{r})$ and $A^\alpha = (\phi, \boldsymbol{A})$ are 4-vectors. Rewriting (D.3) in terms of a time integral we can identify the Lagrangian

[1] Here, we use the cgs system. All equations may be transformed into the SI system by multiplying $\boldsymbol{B}$, $\boldsymbol{A}$, and $A_\alpha$ by the speed of light, $c$.

($\boldsymbol{v} = \dot{\boldsymbol{r}}$),

$$\mathcal{L} = -mc^2\sqrt{1-\frac{v^2}{c^2}} + \frac{q}{c}\boldsymbol{A}\cdot\boldsymbol{v} - q\phi. \tag{D.4}$$

Expressing the Hamiltonian,

$$\mathcal{H} = \boldsymbol{v}\cdot\frac{\partial\mathcal{L}}{\partial\boldsymbol{v}} - \mathcal{L} = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} + q\phi, \tag{D.5}$$

in terms of the canonical momentum $\boldsymbol{p}$ (component-wise derivative),

$$\boldsymbol{p} = \frac{\partial\mathcal{L}}{\partial\boldsymbol{v}} = \frac{m\boldsymbol{v}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{q}{c}\boldsymbol{A} \tag{D.6}$$

one arrives at the final relativistic expression,

$$\begin{aligned} \mathcal{H}(\boldsymbol{p},\boldsymbol{r},t) &= \sqrt{m^2c^4 + c^2\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2} + q\phi(\boldsymbol{r},t) && \text{(D.7)} \\ &\approx mc^2\left(1 + \frac{1}{2}\Big(\frac{\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)}{mc}\Big)^2 - \frac{1}{4}\Big(\frac{\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)}{mc}\Big)^4\right) + q\phi(\boldsymbol{r},t)\,. && \text{(D.8)} \end{aligned}$$

In the nonrelativistic limit, only the leading nonconstant term needs to be kept,

$$\mathcal{H}(\boldsymbol{p},\boldsymbol{r},t) \approx \mathcal{H} + \frac{1}{2m}\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2 + q\phi(\boldsymbol{r},t). \tag{D.9}$$

The Lorentz force given in (D.1) is easily reproduced from the corresponding Lagrangian.

The potentials $\boldsymbol{A}$, $\phi$ are not uniquely defined by the physical fields $\boldsymbol{E}$, $\boldsymbol{B}$; the latter remain unchanged under a gauge transformation

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{r},t) &\longrightarrow \boldsymbol{A}'(\boldsymbol{r},t) = \boldsymbol{A}(\boldsymbol{r},t) + \nabla\chi(\boldsymbol{r},t) && \text{(D.10)} \\ \phi(\boldsymbol{r},t) &\longrightarrow \phi'(\boldsymbol{r},t) = \phi(\boldsymbol{r},t) - \frac{1}{c}\frac{\partial\chi(\boldsymbol{r},t)}{\partial t}. && \text{(D.11)} \end{aligned}$$

The Coulomb gauge $\nabla\cdot\boldsymbol{A} = 0$ (also called transverse gauge) is particularly suitable for the study of interacting electron systems, because $\phi$ acts instantaneously in this case. An oscillatory electric field is then represented by the field $\boldsymbol{A}$ only. Note that this treatment implies an infinitely growing vector potential in the limit $\omega \to 0$ which at least a priori gives reason for caution in connection with linear response theory. The final expressions for the dc conductivity can, however, also be computed directly with the same result (Luttinger, 1964; Mahan, 1990).

## D.2 Linear Response Theory

Let a system be subjected to a time-dependent perturbation $F(t)$ which couples through an operator $\hat{A}$,

$$\hat{H}(t) = \hat{H}_0 + \hat{H}_1(t) = \hat{H}_0 - \hat{A}F(t)\,. \tag{D.12}$$

We assume that the function is switched on adiabatically,

$$F(t) \stackrel{t\to-\infty}{\longrightarrow} 0; \quad \Big|\frac{\mathrm{d}}{\mathrm{d}t}F(t)\Big| < \infty\,. \tag{D.13}$$

Then, the evolution of the density operator for an ensemble of perturbed systems at finite temperature from the unperturbed density operator,

$$\hat{\rho}_0 = \frac{1}{Z_0}\mathrm{e}^{-\beta\hat{H}_0}, \quad Z_0 = \mathrm{Tr}\left\{\mathrm{e}^{-\beta\hat{H}_0}\right\} = \sum_n \mathrm{e}^{-\beta E_n}\,, \tag{D.14}$$

is determined by the von Neumann equation,

$$\hat{\rho}(t=-\infty) = \hat{\rho}_0; \quad i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = [\hat{H}(t), \hat{\rho}(t)]\,. \tag{D.15}$$

Due to the invariance of the trace with respect to cyclical permutations, expectation values of operators may be equivalently expressed in the Schrödinger or in the interaction picture,

$$\langle\hat{B}\rangle_F := \mathrm{Tr}\left\{\hat{B}\,\hat{\rho}(t)\right\} = \mathrm{Tr}\left\{\hat{B}_I(t)\,\hat{\rho}_I(t)\right\}. \tag{D.16}$$

Here, we denote expectation values in the perturbed ensemble as $\langle\ldots\rangle_F$ and operators in the interaction picture as $\hat{X}_I(t) := \mathrm{e}^{i\hat{H}_0 t/\hbar}\hat{X}\mathrm{e}^{-i\hat{H}_0 t/\hbar}$. The corresponding equation of motion for the density operator plus the initial condition are equivalent to the integral equation

$$\hat{\rho}_I(t) = \hat{\rho}_0 + \frac{i}{\hbar}\int_{-\infty}^{t}\mathrm{d}t'\,[\hat{A}_I(t'), \hat{\rho}_I(t')]F(t')\,, \tag{D.17}$$

which may be solved iteratively. Inserting the unperturbed density operator in the right hand side of (D.17) yields the lowest order (i.e., linear) correction. Using $\mathrm{Tr}\left\{\hat{B}_I(t)[\hat{A}_I(t), \hat{\rho}_0]\right\} = \mathrm{Tr}\left\{\hat{\rho}_0[\hat{B}_I(t), \hat{A}_I(t)]\right\}$ one, thus, obtains to linear order

$$\langle\hat{B}\rangle_F(t) = \langle\hat{B}\rangle_0 + \int_{-\infty}^{\infty}\mathrm{d}t'\,F(t')\,\chi_{\hat{B},\hat{A}}(t,t')\,, \tag{D.18}$$

with the retarded susceptibility

$$\chi_{\hat{B},\hat{A}}(t,t') = \frac{i}{\hbar}\langle[\hat{B}_I(t), \hat{A}_I(t')]\rangle_0\Theta(t-t')\,. \tag{D.19}$$

Specializing on harmonic time dependence of the perturbation,

$$F(t) = F_0\mathrm{e}^{-i(\omega+i\delta)t}, \tag{D.20}$$

one can also derive the formulation in frequency space,

$$\langle\hat{B}\rangle_F(\omega) = \langle\hat{B}\rangle_0 + \langle\langle\hat{B}, \hat{A}\rangle\rangle\,(\omega+i\delta)\,F(\omega) \tag{D.21}$$

using the definition

$$\langle\langle \hat{B}, \hat{A} \rangle\rangle\,(\omega + i\delta) \equiv \chi_{\hat{B},\hat{A}}(\omega + i\delta) = \frac{i}{\hbar} \int\limits_0^\infty \mathrm{d}t \langle [\hat{B}_I(t), \hat{A}] \rangle_0 \, \mathrm{e}^{i(\omega+i\delta)t} \,. \tag{D.22}$$

Performing the trace for an eigenbasis of the unperturbed Hamiltonian, inserting the identity $1 = \sum_m |m\rangle\langle m|$ between the operators, and reordering terms one obtains the (thermal) spectral representation,

$$\begin{aligned} \langle\langle \hat{B}, \hat{A} \rangle\rangle\,(\omega + i\delta) \;&=\; \frac{i}{\hbar} \frac{1}{Z_0} \sum_{n,m} \mathrm{e}^{-\beta E_n} \int\limits_0^\infty \mathrm{d}t \, \Big( \mathrm{e}^{it(\omega + i\delta + \frac{E_n - E_m}{\hbar})} \langle n|\hat{B}|m\rangle\langle m|\hat{A}|n\rangle \\ &\qquad - \mathrm{e}^{it(\omega + i\delta - \frac{E_n - E_m}{\hbar})} \langle n|\hat{A}|m\rangle\langle m|\hat{B}|n\rangle \Big) &\text{(D.23)} \\ &=\; -\frac{1}{Z_0} \sum_{n,m} \frac{\langle n|\hat{B}|m\rangle\langle m|\hat{A}|n\rangle}{\hbar(\omega + i\delta) + E_n - E_m} \left( \mathrm{e}^{-\beta E_n} - \mathrm{e}^{-\beta E_m} \right) . &\text{(D.24)} \end{aligned}$$

If we further assume that the spectrum of $\hat{H}_0$ is bounded, we find from (D.24) for the limiting behavior,

$$\langle\langle \hat{B}, \hat{A} \rangle\rangle\,(\omega + i\delta) \overset{\omega\to\infty}{\longrightarrow} -\frac{1}{\hbar\omega} \langle [\hat{B}, \hat{A}] \rangle_0 + \mathcal{O}\big(\frac{1}{\omega^2}\big) \,. \tag{D.25}$$

For the usual case $\hat{A} = \hat{B}$ the leading term in (D.25) vanishes so that the susceptibility $\chi_{\hat{B},\hat{B}}$ then falls off at least as $1/\omega^2$. Finally, the imaginary part of the susceptibility can be simplified when $\langle n|\hat{B}|m\rangle\langle m|\hat{A}|n\rangle$ is real. In particular, one obtains for $\hat{B} = \hat{A}^\dagger$ (i.e., for $\hat{B} = \hat{A}$ when $\hat{A}$ is hermitian),

$$\begin{aligned} \mathrm{Im}\,\langle\langle \hat{B}^\dagger, \hat{B} \rangle\rangle\,(\omega) \;&=\; \frac{\pi}{\hbar Z_0} \sum_{n,m} \big|\langle m|\hat{B}|n\rangle\big|^2 \delta(\omega + \frac{E_n - E_m}{\hbar}) \big( \mathrm{e}^{-\beta E_n} - \mathrm{e}^{-\beta E_m} \big) &\text{(D.26)} \\ &\overset{T\to 0}{\longrightarrow} \frac{\pi}{\hbar} \sum_m \big|\langle m|\hat{B}|0\rangle\big|^2 \big( \delta(\omega - \frac{E_m - E_0}{\hbar}) - \delta(\omega + \frac{E_m - E_0}{\hbar}) \big) \,. &\text{(D.27)} \end{aligned}$$

# Bibliography


Abbate, M., F. M. F. de Groot, J. C. Fuggle, A. Fujimori, Y. Tokura, Y. Fujishima, O. Strebel, M. Domke, G. Kaindl, J. van Elp, B. T. Thole, G. A. Sawatzky, M. Sacchi, and N. Tsuda, 1991, Phys. Rev. B **44**, 5419.

Abramowitz, M., and I. A. Stegun, 1972, *Handbook of Mathematical Functions* (Dover, New York).

Abrikosov, A. A., and I. M. Khalatnikov, 1958, Sov. Phys. JETP **6**, 888.

Ahn, K. H., and A. J. Millis, 2000, Phys. Rev. B **61**, 13545.

Altarelli, M., D. L. Dexter, H. M. Nussenzveig, and D. Y. Smith, 1972, Phys. Rev. B **6**, 4502.

Andersen, O. K., 1975, Phys. Rev. B **12**, 3060.

Andersen, O. K., Z. Pawlowska, and O. Jepsen, 1986, Phys. Rev. B **34**, 5253.

Anderson, P. W., 1961, Phys. Rev. **124**, 41.

Anisimov, V. I., A. I. Poteryaev, M. A. Korotin, A. O. Anokhin, and G. Kotliar, 1997, J. Phys. Cond. Matter **9**, 7359.

Anisimov, V. I., J. Zaanen, and O. K. Andersen, 1991, Phys. Rev. B **44**, 943.

Ashcroft, N. W., and N. D. Mermin, 1976, *Solid State Physics* (Holt-Saunders, Philadelphia).

Baeriswyl, D., D. K. Campbell, J. M. P. Carmelo, F. Guinea, and E. Louis, 1995, *The Hubbard Model: Its Physics and Mathematical Physics* (Plenum Press, New York).

Bansil, A., and M. Lindroos, 1999, Phys. Rev. Lett. **83**, 5154.

Bao, W., C. Broholm, G. Aeppli, P. Dai, J. M. Honig, and P. Metcalf, 1997, Phys. Rev. Lett. **78**, 507.

Bao, W., C. Broholm, G. Aeppli, S. A. Carter, P. Dai, T. F. Rosenbaum, J. M. Honig, P. Metcalf, and S. F. Trevino, 1998, Phys. Rev. B **58**, 12727.

Bao, W., C. Broholm, S. A. Carter, T. F. Rosenbaum, G. Aeppli, S. F. Trevino, P. Metcalf, J. M. Honig, and J. Spalek, 1993, Phys. Rev. Lett. **71**, 766.

Batson, P. E., and J. Silcox, 1983, Phys. Rev. B **27**, 5224.

Berlin, T. H., and M. Kac, 1952, Phys. Rev. **86**, 821.

Bethe, H. A., 1930, Ann. Physik **5**, 325.

Bethe, H. A., 1931, Z. Phys. **71**, 205.

Bethe, H. A., 1935, Proc. Roy. Soc. London A **150**, 552.

Bickers, N. E., and D. J. Scalapino, 1989, Ann. Physik **193**, 206.

Bickers, N. E., and S. R. White, 1991, Phys. Rev. B **43**, 8044.

Bickers, N. E., D. J. Scalapino, and S. R. White, 1989, Phys. Rev. Lett. **62**, 961.

Bickers, N. E., D. L. Cox, and J. W. Wilkins, 1987, Phys. Rev. B **36**, 2036.

Blaha, P., K. Schwarz, and J. Luitz, 1999, *WIEN97, A Full Potential Linearized Augmented Plane Wave Package for Calculating Crystal Properties* (Techn. Univ. Wien, Vienna).

Blankenbecler, R., D. J. Scalapino, and R. L. Sugar, 1981, Phys. Rev. D **24**, 2278.

Bloch, F., 1928, Z. Phys. **52**, 555.

Blümer, N., 1996, Diploma thesis, RWTH Aachen.

Born, M., and R. Oppenheimer, 1927, Ann. Physik **84**, 457.

Brinkman, W. F., and T. M. Rice, 1970, Phys. Rev. B **2**, 4302.

Brout, R., 1960, Phys. Rev. **118**, 1009.

Brout, R., 1961, Phys. Rev. **122**, 469.

Bruno, P., 1993, in *Magnetismus von Festkörpern und Grenzflächen, 24. IFF Ferienkurs* (Forschungszentrum Jülich, ISBN 3-89336-110-3), p. 24.1.

Bryan, R. K., 1990, Eur. Biophys. J **18**, 165.

Bryan, R. K., and J. Skilling, 1980, Mon. Not. R. Astr. Soc. **191**, 69.

Bulla, R., 1999, Phys. Rev. Lett. **83**, 136.

Bulla, R., 2000, Advances In Solid State Physics **40**, 169.

Bulla, R., T. A. Costi, and D. Vollhardt, 2001, Phys. Rev. B **64**, 045103.

Caffarel, M., and W. Krauth, 1994, Phys. Rev. Lett. **72**, 1545.

Capelle, K., and G. Vignale, 2001, Phys. Rev. Lett. **86**, 5546.

Ceperley, D. M., and B. J. Alder, 1980, Phys. Rev. Lett. **45**, 566.

Ceperley, D., M. Mascagni, L. Mitas, F. Saied, and A. Srinivasan, 1998, SPRNG, Version 1.0; http://sprng.cs.fsu.edu.

Chattopadhyay, A., A. J. Millis, and S. Das Sarma, 2000, Phys. Rev. B **61**, 10738.

Chitra, R., and G. Kotliar, 1999, Phys. Rev. Lett. **83**, 2386.

Chung, W., and J. K. Freericks, 1998*a*, Phys. Rev. B **57**, 11955.

Chung, W., and J. K. Freericks, 1998*b*, Preprint cond-mat/9802120.

Costi, T. A., A. C. Hewson, and V. Zlatić, 1994, J. Phys. Cond. Matter **6**, 2519.

Crandles, D. A., T. Timusk, J. D. Garrett, and J. E. Greedan, 1994, Phys. Rev. B **49**, 16207.

Czycholl, G., 2000, *Theoretische Festkörperphysik* (Vieweg, Braunschweig).

Dagotto, E., 1994, Rev. Mod. Phys. **66**, 763.

Deisz, J. J., D. W. Hess, and J. W. Serene, 1995, in *Recent Progress In Many Body Theories*, Vol. 4 (Plenum, New York).

Deisz, J., M. Jarrell, and D. L. Cox, 1990, Phys. Rev. B **42**, 4869.

Dobrosavljević, V., and G. Kotliar, 1993, Phys. Rev. Lett. **71**, 3218.

Dobrosavljević, V., and G. Kotliar, 1994, Phys. Rev. B **50**, 1430.

Drchal, V., V. Janiš, and J. Kudrnovský, 1999, Phys. Rev. B **60**, 15664.

Dworin, L., and A. Narath, 1970, Phys. Rev. Lett. **25**, 1287.

Economou, E. N., 1979, *Green's Functions in Quantum Physics* (Springer Verlag, Berlin).

Ehrenreich, H., H. R. Philipp, and B. Segall, 1963, Phys. Rev. **132**, 1918.

Eitel, M., and J. E. Greedan, 1986, Journal of the Less-Common Metals **116**, 95.

Emery, V. J., 1987, Phys. Rev. Lett. **58**, 2794.

Engelbrecht, J. R., M. Randeria, and L. Zhang, 1992, Phys. Rev. B **45**, 10135.

Fetter, A. L., and J. D. Walecka, 1971, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York).

Freericks, J. K., 2000, private communication.

Freericks, J. K., and M. Jarrell, 1995*a*, Phys. Rev. Lett. **74**, 186.

Freericks, J. K., and M. Jarrell, 1995*b*, Phys. Rev. Lett. **75**, 2570.

Fujimori, A., I. Hase, H. Namatame, Y. Fujishima, Y. Tokura, H. Eisaki, S. Uchida, K. Takegahara, and F. de Groot, 1992, Phys. Rev. Lett. **69**, 1796.

Fujimori, A., I. Hase, M. Nakamura, H. Namatame, Y. Fujishima, Y. Tokura, M. Abbate, F. M. F. de Groot, M. T. Czyzyk, J. C. Fuggle, O. Strebel, F. Lopez, M. Domke, and G. Kaindl, 1992, Phys. Rev. B **46**, 9841.

Fujishima, Y., Y. Tokura, T. Arima, and S. Uchida, 1992, Phys. Rev. B **46**, 11167.

Galitskii, V. M., 1958, Sov. Phys. JETP **34**, 104.

Gariglio, S., J. W. Seo, J. Fompeyrine, J.-P. Locquet, and J.-M. Triscone, 2001, Phys. Rev. B **63**, 161103.

Gebhard, F., 1990, Phys. Rev. B **41**, 9452.

Gebhard, F., 1997, *The Mott Metal-Insulator Transition* (Springer-Verlag, Berlin).

Georges, A., and G. Kotliar, 1992, Phys. Rev. B **45**, 6479.

Georges, A., and J. S. Yedidia, 1991, Phys. Rev. B **43**, 3475.

Georges, A., and W. Krauth, 1992, Phys. Rev. Lett. **69**, 1240.

Georges, A., and W. Krauth, 1993, Phys. Rev. B **48**, 7167.

Georges, A., G. Kotliar, and Q. Si, 1992, Int. J. Mod. Phys. B **6**, 705.

Georges, A., G. Kotliar, and W. Krauth, 1993, Z. Phys. B **92**, 313.

Georges, A., G. Kotliar, W. Krauth, and M. Rozenberg, 1996, Rev. Mod. Phys. **68**, 13.

Goral, J. P., J. E. Greedan, and D. A. MacLean, 1982, J. Solid State Chem. **43**, 244.

Gubernatis, J. E., M. Jarrell, R. N. Silver, and D. S. Sivia, 1991, Phys. Rev. B **44**, 6011.

Gunnarsson, O., and B. I. Lundqvist, 1976, Phys. Rev. B **13**, 4274.

Gunnarsson, O., and R. O. Jones, 1981, Solid State Commun. **37**, 249.

Gunnarsson, O., O. K. Andersen, O. Jepsen, and J. Zaanen, 1989, Phys. Rev. B **39**, 1708.

Gutzwiller, M. C., 1963, Phys. Rev. Lett. **10**, 159.

Halboth, C. J., and W. Metzner, 2000, Phys. Rev. B **61**, 7364.

Han, J. E., M. Jarrell, and D. L. Cox, 1998, Phys. Rev. B **58**, R4199.

Harle, M., 2000, Diploma Thesis, Universität Augsburg.

Harris, A. B., and R. V. Lange, 1967, Phys. Rev. **157**, 295.

Hays, C. C., J.-S. Zhou, J. T. Markert, and J. B. Goodenough, 1999, Phys. Rev. B **60**, 10367.

Hecht, E., 1997, *Optics* (Addison-Wesley, New York).

Heine, V., and L. F. Mattheiss, 1971, J. Phys. C **4**, L191.

Held, K., 1999, Ph.D. Thesis, Universität Augsburg.

Held, K., and D. Vollhardt, 1998, Eur. Phys. J. B **5**, 473.

Held, K., and D. Vollhardt, 2000, Phys. Rev. Lett. **84**, 5168.

Held, K., G. Keller, V. Eyert, D. Vollhardt, and V. Anisimov, 2001, Phys. Rev. Lett. **86**, 5345.

Held, K., I. A. Nekrasov, N. Blümer, V. I. Anisimov, and D. Vollhardt, 2001, Int. J. Mod. Phys. B **15**, 2611.

Held, K., M. Ulmke, and D. Vollhardt, 1996, Mod. Phys. Lett. B **10**, 203.

Held, K., M. Ulmke, N. Blümer, and D. Vollhardt, 1997, Phys. Rev. B **56**, 14469.

Hettler, M., A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, 1998, Phys. Rev. B **58**, R7475.

Hettler, M., A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, 2000, Phys. Rev. B **61**, 12739.

Hirsch, J. E., 1983, Phys. Rev. B **28**, 4059.

Hirsch, J. E., 1989, Phys. Rev. B **40**, 2354.

Hirsch, J. E., and F. Marsiglio, 2000, Phys. Rev. B **62**, 15131.

Hirsch, J. E., and R. M. Fye, 1986, Phys. Rev. Lett. **56**, 2521.

Hofstetter, W., 1998, private communication.

Hohenberg, P., and W. Kohn, 1964, Phys. Rev. **136**, B864.

Holstein, T., 1959, Ann. Phys. (N. Y.) **8**, 325.

Honerkamp, C., and M. Salmhofer, 2001, Phys. Rev. Lett. **87**, 187004.

Hubbard, J., 1963, Proc. Roy. Soc. London A **276**, 238.

Hubbard, J., 1964*a*, Proc. Roy. Soc. London A **277**, 237.

Hubbard, J., 1964*b*, Proc. Roy. Soc. London A **281**, 401.

Imada, M., A. Fujimori, and Y. Tokura, 1998, Rev. Mod. Phys. **70**, 1039.

Janiš, V., 1999, Phys. Rev. Lett. **83**, 2781.

Janiš, V., and D. Vollhardt, 1992, Int. J. Mod. Phys. B **6**, 731.

Jarrell, M., 1992, Phys. Rev. Lett. **69**, 168.

Jarrell, M., 1997, in *Numerical Methods for Lattice Quantum Many-Body Problems*, edited by D. Scalapino (Addison Wesley, Reading).

Jarrell, M., D. S. Sivia, and B. Patton, 1990, Phys. Rev. B **42**, 4804.

Jarrell, M., H. Akhlaghpour, and T. Pruschke, 1993, in *Quantum Monte Carlo Methods in Condensed Matter Physics*, edited by M. Suzuki (World Scientific, Singapore).

Jarrell, M., J. K. Freericks, and T. Pruschke, 1995, Phys. Rev. B **51**, 11704.

Jarrell, M., T. Maier, M. H. Hettler, and A. N. Tahvildarzadeh, 2001, Europhys. Lett. **56**, 563.

Jepsen, O., and O. K. Andersen, 2000, The STUTTGART TB-LMTO-ASA program, Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany.

Jones, R. O., and O. Gunnarsson, 1989, Rev. Mod. Phys. **61**, 689.

Joo, J., and V. Oudovenko, 2001, Phys. Rev. B **64**, 193102.

Kajueter, H., and G. Kotliar, 1997, Int. J. Mod. Phys. B **11**, 729.

Kalinowski, E., and F. Gebhard, 2002, J. Low Temp. Phys. **126**, 979.

Kanamori, J., 1963, Prog. Theor. Phys. **30**, 275.

Katsnelson, M. I., and A. I. Lichtenstein, 2000, Phys. Rev. B **61**, 8906.

Katsufuji, T., and Y. Tokura, 1999, Phys. Rev. B **60**, 7673.

Katsufuji, T., Y. Okimoto, and Y. Tokura, 1995, Phys. Rev. Lett. **75**, 3497.

Kawarabayashi, T., 1993, Technical report of ISSP A (Tokyo).

Keiter, H., and J. C. Kimball, 1970, Phys. Rev. Lett. **25**, 672.

Khurana, A., 1990, Phys. Rev. Lett. **64**, 1990.

Kittel, C., 1996, *Introduction to Solid State Physics* (Wiley, New York).

Knecht, C., 2002, Diploma Thesis, Universität Mainz.

Kohn, W., 1964, Phys. Rev. **133**, A 171.

Kohn, W., and L. J. Sham, 1965, Phys. Rev. **140**, A1133.

Kollar, M., 1998, Ph.D. Thesis, Universität Augsburg.

Kollar, M., 2002, Int. J. Mod. Phys. B **6**, 3491.

Kollar, M., R. Strack, and D. Vollhardt, 1996, Phys. Rev. B **53**, 9225.

Kotliar, G., E. Lange, and M. J. Rozenberg, 2000, Phys. Rev. Lett. **84**, 5180.

Kotliar, G., S. Y. Savrasov, and G. Pálsson, 2001, Phys. Rev. Lett. **87**, 186401.

Krishna-murthy, H. R., J. W. Wilkins, and K. G. Wilson, 1980, Phys. Rev. B **21**, 1003.

Kuhn, W., 1925, Z. Phys. **33**, 408.

Kurata, M., R. Kikuchi, and T. Watari, 1953, J. Chem. Phys. **21**, 434.

Kuwamoto, H., J. M. Honig, and J. Appel, 1980, Phys. Rev. B **22**, 2626.

Landau, L. D., 1957*a*, Sov. Phys. JETP **3**, 920.

Landau, L. D., 1957*b*, Sov. Phys. JETP **5**, 101.

Landau, L. D., 1959, Sov. Phys. JETP **35**, 70.

Landau, L. D., and E. M. Lifschitz, 1990, *Lehrbuch der Theoretischen Physik, Band 8; 5. Auflage* (Akademie-Verlag, Berlin).

Landau, L. D., and E. M. Lifschitz, 1992, *Lehrbuch der Theoretischen Physik, Band 2; 12. Auflage* (Akademie-Verlag, Berlin).

Levy, M., 1979, Proc. Natl. Acad. Sci. (USA) **76**, 6062.

Lichtenberg, F., D. Widmer, J. G. Bednorz, T. Williams, and A. Reller, 1991, Z. Phys. B **82**, 211.

Lichtenstein, A. I., and M. I. Katsnelson, 1998, Phys. Rev. B **57**, 6884.

Lichtenstein, A. I., and M. I. Katsnelson, 2000, Phys. Rev. B **62**, R9283.

Lieb, E. H., and F. Y. Wu, 1968, Phys. Rev. Lett. **20**, 1445.

Liebsch, A., and A. Lichtenstein, 2000, Phys. Rev. Lett. **84**, 1591.

Luttinger, J. M., 1964, Phys. Rev. **135**, A1505.

MacLean, D. A., H.-N. Ng, and J. E. Greedan, 1979, J. Solid State Chem. **30**, 35.

Mahan, G. D., 1990, *Many-Particle Physics*, 2nd edn (Plenum Press, New York).

Maier, T., and M. Jarrell, 2002, Phys. Rev. B **65**, 041104.

Maier, T., M. Jarrell, T. Pruschke, and J. Keller, 2000, Phys. Rev. Lett. **85**, 1524.

Mattheiss, L. F., 1972, Phys. Rev. B **5**, 290.

McWhan, D. B., J. P. Remeika, J. P. Maita, H. Okinaka, K. Kosuge, and S. Kachi, 1973, Phys. Rev. B **7**, 326.

McWhan, D. B., J. P. Remeika, T. M. Rice, W. F. Brinkman, J. P. Maita, and A. Menth, 1971, Phys. Rev. Lett. **27**, 941.

Merino, J., and R. H. McKenzie, 2000, Phys. Rev. B **61**, 7996.

Mermin, N. D., and H. Wagner, 1966, Phys. Rev. Lett. **17**, 1133.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, 1953, J. Chem. Phys. **21**, 1087.

Metzner, W., and D. Vollhardt, 1989, Phys. Rev. Lett. **62**, 324.

Metzner, W., P. Schmit, and D. Vollhardt, 1992, Phys. Rev. B **45**, 2237.

Mielke, A., and H. Tasaki, 1993, Commun. Math. Phys. **158**, 341.

Millis, A. J., 2002, private communication.

Moeller, G., Q. Si, G. Kotliar, M. J. Rozenberg, and D. S. Fisher, 1995, Phys. Rev. Lett. **74**, 2082.

Motome, Y., and M. Imada, 1997, J. Phys. Soc. Jap. **66**, 1872.

Mott, N. F., 1968, Rev. Mod. Phys. **40**, 677.

Mott, N. F., 1990, *Metal-Insulator Transitions*, 2nd edn (Taylor & Francis, London).

Müller-Hartmann, E., 1989*a*, Z. Phys. B **74**, 507.

Müller-Hartmann, E., 1989*b*, Z. Phys. B **76**, 211.

Negele, W., and H. Orland, 1987, *Quantum Many-Particle Systems* (Addison-Wesley, New York).

Nekrasov, I. A., K. Held, N. Blümer, A. I. Poteryaev, V. I. Anisimov, and D. Vollhardt, 2000, Eur. Phys. J. B **18**, 55.

Nekrasov, I. A., N. Blümer, K. Held, V. I. Anisimov, and D. Vollhardt, 2001, unpublished.

Noack, R. M., and F. Gebhard, 1999, Phys. Rev. Lett. **82**, 1915.

Nozières, P., and D. Pines, 1958, Phys. Rev. **109**, 741.

Nozières, P., and D. Pines, 1959, Phys. Rev. **113**, 1254.

Nozières, P., and D. Pines, 1989, *Theory of quantum liquids* (Addison-Wesley, New York).

Okimoto, Y., T. Katsufuji, Y. Okada, T. Arima, and Y. Tokura, 1995, Phys. Rev. B **51**, 9581.

Peierls, R., 1936, Proc. Roy. Soc. London A **154**, 207.

Penn, R. G., 1966, Phys. Rev. **142**, 350.

Perdew, J. P., and A. Zunger, 1981, Phys. Rev. B **23**, 5048.

Pickett, W. E., S. C. Erwin, and E. C. Ethridge, 1998, Phys. Rev. B **58**, 1201.

Potthoff, M., T. Wegner, and W. Nolting, 1997, Phys. Rev. B **55**, 16132.

Powell, C. J., and J. B. Swan, 1959, Phys. Rev. **115**, 869.

Pruschke, T., and N. Grewe, 1989, Z. Phys. B **74**, 439.

Pruschke, T., D. L. Cox, and M. Jarrell, 1993, Phys. Rev. B **47**, 3553.

Pruschke, T., M. Jarrell, and J. K. Freericks, 1995, Adv. Phys. **44**, 187.

Pruschke, T., W. Metzner, and D. Vollhardt, 2001, J. Phys. Cond. Matter **13**, 9455.

Reiche, F., and W. Thomas, 1925, Z. Phys. **34**, 510.

Rice, T. M., and D. B. McWhan, 1970, IBM J. Res. Develop. p. 251.

Rozenberg, M. J., 1997, Phys. Rev. B **55**, 4855.

Rozenberg, M. J., G. Kotliar, and X. Y. Zhang, 1994, Phys. Rev. B **49**, 10181.

Rozenberg, M. J., G. Kotliar, H. Kajüter, G. A. Thomas, D. H. Rapkine, J. M. Honig, and P. Metcalf, 1995, Phys. Rev. Lett. **75**, 105.

Rozenberg, M. J., I. H. Inoue, H. Makino, F. Iga, and Y. Nishihara, 1996, Phys. Rev. Lett. **76**, 4781.

Rozenberg, M. J., R. Chitra, and G. Kotliar, 1999, Phys. Rev. Lett. **83**, 3498.

Rozenberg, M. J., X. Y. Zhang, and G. Kotliar, 1992, Phys. Rev. Lett. **69**, 1236.

Sandvik, A. W., and D. J. Scalapino, 1995, Phys. Rev. B **51**, 9403.

Santoro, G., M. Airoldi, S. Sorella, and E. Tosatti, 1993, Phys. Rev. B **47**, 16216.

Scalapino, D. J., S. R. White, and S. C. Zhang, 1992, Phys. Rev. Lett. **68**, 2830.

Schaich, W. L., and N. W. Ashcroft, 1971, Phys. Rev. B **8**, 2452.

Schiller, A., and K. Ingersent, 1995, Phys. Rev. Lett. **75**, 113.

Schlipf, J., 1998, Ph.D. Thesis, Universität Augsburg.

Schlipf, J., M. Jarrell, P. G. J. van Dongen, N. Blümer, S. Kehrein, T. Pruschke, and D. Vollhardt, 1999, Phys. Rev. Lett. **82**, 4890.

Schooley, J. F., W. R. Hosler, E. Ambler, J. H. Becker, M. L. Cohen, and C. S. Koonce, 1965, Phys. Rev. Lett. **14**, 305.

Schweitzer, H., and G. Czycholl, 1991, Phys. Rev. Lett. **67**, 3724.

Shastry, B. S., and B. Sutherland, 1990, Phys. Rev. Lett. **65**, 243.

Shastry, B. S., B. I. Shraiman, and R. R. P. Singh, 1993, Phys. Rev. Lett. **70**, 2004.

Silver, R. N., D. S. Sivia, and J. E. Gubernatis, 1990, Phys. Rev. B **41**, 2380.

Silver, R. N., J. E. Gubernatis, D. S. Sivia, and M. Jarrell, 1990, Phys. Rev. Lett. **65**, 496.

Skilling, J., and R. K. Bryan, 1984, Mon. Not. R. astr. Soc. **211**, 111.

Smith, D. Y., and E. Shiles, 1978, Phys. Rev. B **17**, 4689.

Solovyev, I., N. Hamada, and K. Terakura, 1996, Phys. Rev. B **53**, 7158.

Staudt, R., M. Dzierzawa, and A. Muramatsu, 2000, Eur. Phys. J. B **17**, 411.

Strack, R., and D. Vollhardt, 1994, Phys. Rev. Lett. **72**, 3425.

Stratton, J. A., 1941, *Electromagnetic Theory* (McGraw-Hill, New York).

Stumpf, M. P. H., 1999, Ph.D. Thesis, University of Oxford.

Su, G., and M. Suzuki, 1998, Phys. Rev. B **58**, 117.

Suzuki, M., 1976, Prog. Theor. Phys. **56**, 1454.

Svane, A., and O. Gunnarsson, 1990, Phys. Rev. Lett. **65**, 1148.

Takahashi, M., 1977, J. Phys. C **10**, 1289.

Thomas, W., 1925, Naturwissenschaften **13**, 627.

Thorpe, M. F., 1981, in *Excitations in Disordered Solids*, edited by M. F. Thorpe (Plenum, New York).

Tokura, Y., Y. Taguchi, Y. Okada, Y. Fujishima, T. Arima, K. Kumagai, and Y. Iye, 1993, Phys. Rev. Lett. **70**, 2126.

Tong, N., S. Shen, and F. Pu, 2001, Phys. Rev. B **64**, 235109.

Trotter, H. F., 1959, Proc. Am. Math. Soc. **10**, 545.

Uhrig, G., and D. Vollhardt, 1995, Phys. Rev. B **52**, 5617.

Uhrig, G. S., and R. Vlaming, 1993, J. Phys. Cond. Matter **5**, 2561.

Ulmke, M., 1995, Ph.D. Thesis, KFA Jülich.

Ulmke, M., 1998, Eur. Phys. J. B **1**, 301.

Ulmke, M., V. Janiš, and D. Vollhardt, 1995, Phys. Rev. B **51**, 10411.

v. Baltz, R., 1997, in *Spectroscopy and Dynamics of Collective Excitations in Solids*, edited by D. Bartolo (Plenum Press, New York), p. 303.

van Dongen, P. G. J., 1994*a*, Phys. Rev. B **50**, 14016.

van Dongen, P. G. J., 1994*b*, Phys. Rev. B **49**, 7904.

van Dongen, P. G. J., 1994*c*, Phys. Rev. Lett. **72**, 3258.

van Dongen, P. G. J., 1997, private communication.

van Dongen, P. G. J., 2000, private communication.

van Dongen, P. G. J., 2001, private communication.

van Dongen, P. G. J., and D. Vollhardt, 1989, Phys. Rev. B **40**, 7252.

van Dongen, P. G. J., F. Gebhard, and D. Vollhardt, 1989, Z. Phys. B **76**, 199.

Varma, C. M., S. Schmitt-Rink, and E. Abrahams, 1987, Solid State Commun. **62**, 681.

Vlaming, R., and D. Vollhardt, 1992, Phys. Rev. B **45**, 4637.

von Barth, U., and L. Hedin, 1972, J. Phys. C **5**, 1629.

Vosko, S. H., and L. H. Wilk, 1980, Phys. Rev. B **22**, 3812.

Wahle, J., N. Blümer, J. Schlipf, K. Held, and D. Vollhardt, 1998, Phys. Rev. B **58**, 12749.

Wegner, F., 1976, Z. Phys. B **25**, 327.

Weiss, P., 1907, J. de Physique **6 (4)**, 661.

Weisstein, E. W., 1998, *The CRC Concise Encyclopedia of Mathematics* (CRC Press, London).

White, S. R., 1992, Phys. Rev. Lett. **69**, 2863.

White, S. R., 1993, Phys. Rev. B **48**, 10345.

Wilson, K. G., 1975, Rev. Mod. Phys. **47**, 773.

Yokoya, T., T. Sato, H. Fujisawa, T. Tkahashi, A. Chainani, and M. Onoda, 1999, Phys. Rev. B **59**, 1815.

Yoshida, T., A. Ino, T. Mizokawa, A. Fujimori, Y. Taguchi, T. Katsufuji, and Y. Tokura, 2002, Europhys. Lett. **59**, 258.

Zaránd, G., D. L. Cox, and A. Schiller, 2000, Phys. Rev. B **62**, R16227.

Zhang, X. Y., M. J. Rozenberg, and G. Kotliar, 1993, Phys. Rev. Lett. **70**, 1666.

Zhong, J., and H.-B. Schüttler, 1992, Phys. Rev. Lett. **69**, 1600.

Zölfl, M. B., T. Pruschke, J. Keller, A. I. Poteryaev, I. A. Nekrasov, and V. I. Anisimov, 2000, Phys. Rev. B **61**, 12810.

# Index

# List of Publications

Parts of this thesis were published in the following articles:

J. Schlipf, M. Jarrell, P. G. J. van Dongen, N. Blümer, S. Kehrein, Th. Pruschke, and D. Vollhardt, "*Absence of Hysteresis at the Mott-Hubbard Metal-Insulator Transition in Infinite Dimensions*", Phys. Rev. Lett. **82**, 4890 (1999).

I. A. Nekrasov, K. Held, N. Blümer, A. I. Poteryaev, V. I. Anisimov, and D. Vollhardt, "*Calculation of photoemission spectra of the doped Mott insulator* $La_{1\text{-}x}Sr_xTiO_3$ *using LDA+DMFT(QMC)*", Eur. Phys. J. B **18**, 55 (2000).

K. Held, I. A. Nekrasov, N. Blümer, V. I. Anisimov, and D. Vollhardt, "*Realistic Modeling of Strongly Correlated Electron Systems: An Introduction to the LDA+DMFT Approach*", Int. J. Mod. Phys. B **15**, 2611 (2001).

N. Blümer, K. Held, G. Keller, and D. Vollhardt, "*Metal-Insulator Transitions and Realistic Modelling of Strongly Correlated Electron Systems*", in NIC-Symposium 2001, eds. H. Rollnik and D. Wolf, NIC Series Vol. **9**, p. 347 (Forschungszentrum Jülich, 2002).

K. Held, I. A. Nekrasov, G. Keller, V. Eyert, N. Blümer, A. K. McMahan, R. T. Scalettar, T. Pruschke, V. I. Anisimov, and D. Vollhardt, "*The LDA+DMFT Approach to Materials with Strong Electronic Correlations*" in Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms, eds. J. Grotendorst, D. Marx and A. Muramatsu, NIC Series Vol. **10**, p. 175 (Forschungszentrum Jülich, 2002) and preprint cond-mat/0112079.

Calculations of the optical conductivity based on the programs developed as part of this thesis appeared in:

K. Held and D. Vollhardt, "*Electronic Correlations in Manganites*", Phys. Rev. Lett. **84**, 1568 (2000).

# Curriculum Vitae

| | |
|---|---|
| Name | Nils Michael Blümer |
| Date of birth | 31 August 1970 |
| Place of birth | München |
| Nationality | German |

**Education**

| | |
|---|---|
| 09/77 – 07/81 | Primary School in München |
| 08/81 – 05/90 | Grammar School: Schillergymnasium Münster |
| 05/90 | Abitur |

**Studies**

| | |
|---|---|
| 10/91 – 07/96 | Study of Physics and Mathematics at the RWTH Aachen |
| 05/95 – 05/96 | Diploma thesis under supervision of Prof. Dr. D. Vollhardt: "Quanten-Monte-Carlo-Untersuchungen von magnetischen Phasen im Hubbard-Modell mit Nächst-Nachbar-Wechselwirkungen" |
| 07/96 | Diploma in Physics ("sehr gut") |
| 08/96 – 08/97 | Graduate studies at the University of Illinois at Urbana-Champaign under Fulbright scholarship, supervised by Prof. Dr. D. Ceperley |
| 09/96 | PhD qualifying exam passed (on free shot) at the UIUC |
| 08/00 | Degree "Master of Science" awarded by the UIUC |
| since 09/97 | PhD studies supervised by Prof. Dr. D. Vollhardt, Universität Augsburg |
| 04/99 – 05/99 | Research stay with Prof. Dr. Mark Jarrell, University of Cincinnati |

**Employments**

| | |
|---|---|
| 06/90 – 08/91 | Community service: Zivildienst at the Center for Heart Surgery, Universität Münster (IT department) |
| 10/92 – 07/96 | Teaching Assistant at the RWTH Aachen |
| 01/97 – 07/97 | Teaching Assistant and Research Assistant at the UIUC |
| 09/97 – 12/00 | Research Assistant at the Chair of Theoretical Physics III, Universität Augsburg |
| 08/98 – 04/00 | System-Administrator at the Chair of Theoretical Physics III, Universität Augsburg |
| since 01/01 | Research Assistent in the group of Prof. Dr. P. G. J. van Dongen, Universität Mainz |

# Acknowledgements

My sincere gratitude goes to Prof. Dr. D. Vollhardt for proposing challenging and rewarding subjects for my thesis, for numerous fruitful discussions, continuing support, and for creating a stimulating environment at the forefront of science.

A am also indebted to Prof. Dr. Peter van Dongen for making himself available whenever I needed a competent partner for discussions. Our increasingly close collaboration, his advice, and his encouragement are deeply appreciated.

I thank Prof. Dr. A. Kampf for his interest in my work and for helpful suggestions regarding the optical conductivity. I am grateful to PD Dr. Ralf Bulla for his insisting criticism against not fully thought-out ideas and to PD Dr. Stefan Kehrein for his openness to serious discussions of fundamental problems. I thank Dr. Karsten Held, Dr. Jan Schlipf, Dr. Marcus Kollar, and Dr. Martin Ulmke for their friendship and for numerous instructive and inspiring discussions. My time in Augsburg would have been less enjoyable and less productive without Dr. Philipp Brune, Dr. Walter Hofstetter, Dr. Matthias Vojta, Dr. Theo Costi, Dr. Bruce Normand, Dr. Corine Roth, and Georg Keller. I have profited from scientific exchange with Dr. Krzysztof Byczuk and Prof. Dr. Václav Janiš and from the collaboration with Dr. Igor Nekrasov. Thanks are also due to Ralf Utermann, Dr. Markus Zahn, Dr. Robert Schulz, and Dr. Ulrich Schneider, for a common good time as system administrators. Last, but not least, Matthias Hense, Martina Herold, and U. Öttl contributed to the good atmosphere in Augsburg.

I thank PD Dr. Reinhard Noack for his enjoyable company, for countless valuable scientific, political, and social discussions, and for his contributions in building up the IT infrastructure in Mainz. I am grateful to Carsten Knecht for our fruitful collaboration and for his total commitment to the research carried out under my guidance. I also enjoyed the interaction and common lunches with Dr. Thomas Schauerte, Markus Himmerich, Markus Lippert, and E. Helf.

The research presented in this thesis has significantly profited from earlier experience won at the UIUC, in particular from my interaction with Prof. Dr. David Ceperley, Prof. Dr. Richard Martin, Dr. John Shumway, and Dr. Burkhard Militzer.

My research stay with Prof. Dr. Mark Jarrell in Cincinnati has been a unique experience. I am deeply grateful for his hospitality, for the training in advanced aspects of quantum Monte Carlo and maximum entropy methods, and for our continuing collaboration. I also thank Dr. Carey Huscroft for illuminating discussions.

Useful discussions on the optical conductivity are acknowledged with Prof. Dr. Jim Freericks, Prof. Dr. David Logan, and Prof. Dr. A. J. Millis. I have profited from numerous scientific exchanges with Prof. Dr. Gabriel Kotliar, Prof. Dr. Florian Gebhard, PD Dr. Thomas Pruschke, and PD Dr. Michael Potthoff. I thank Dr. N. Tong for providing details on his method for treating the Mott transition and Dr. Marcelo Rozenberg for making some QMC solutions available.

Extra thanks go to Georg Keller, Carsten Knecht, and Dr. Marcus Kollar and, in particular, to Dr. Karsten Held, PD Dr. Reinhard Noack, and Prof. Dr. Peter van Dongen for carefully reading draft versions of parts of this thesis and for providing valuable comments.

Finally, very special thanks are due to my parents and to Birgit and Stefanie.