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 $1 \leq j \leq N$),

$$\big(X_{\mathrm{MC}}\big)^2 \;=\; \frac{1}{N}\sum_{j=1}^{N} x_{l_j} \tag{1.45}$$

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

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_{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

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

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.