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, σ_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, α 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 α which may be chosen so that $\chi^2 = \Lambda$ (historic MEM). Alternatively, one may use the value of α 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 ω_j of frequencies for which A is going to be computed), and a starting guess for α, a simple MEM program thus both searches for the spectrum $\{A(\omega_j)\}$ with maximum probability $P(A|\bar{G},m,\alpha)$ for given α using, e.g., the Newton-Raphson method and, in an outer loop, searches for the best value of α.

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.