7 Continuous-time QMC Solvers for Electronic Systems in Fermionic and Bosonic Baths

Fakher F. Assaad
Institut für Theoretische Physik und Astrophysik
Universität Würzburg

Contents

E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein
DMFT at 25: Infinite Dimensions
Modeling and Simulation Vol. 4
Forschungszentrum Jülich, 2014, ISBN 978-3-89336-953-9
http://www.cond-mat.de/events/correl14

Continuous-time quantum Monte Carlo (CT-QMC) methods are a tool of choice to solve correlated electron problems embedded in bosonic or fermonic baths [1]. This is precisely the problem that is encountered in dynamical mean-field theories (DMFT), where in the limit of infinite coordination number the environment can be replaced by a fermionic bath [2]. In DMFT, the physics of the Hubbard model in the limit of infinite coordination number maps onto that of the single-impurity Anderson model (SIAM). This mapping provides invaluable insight into the important problem of the Mott transition [3]. The beauty of the CT-QMC algorithms lie in their flexibility. Rather than being Hamiltonian-based – like the auxiliary-field QMC method (see Ref. [4] for a review), they are action-based and allow the simulation of effective low-energy models after having integrated out high-energy degrees of freedom. This aspect of the method has spurred many applications. In the domain of model-building, screening effects by high energy bands, which can be taken into account within the constrained random phase approximation (cRPA) [5], naturally lead to a low-energy effective model with retarded interactions which only has an action-based formulation. Retarded interactions are also obtained in the context of electron-phonon interactions. Here, one can integrate out the bosonic phonon modes at the expense of a retarded interaction [6–9]. Other applications of the algorithm have been introduced in the realm of topological insulators [10, 11]. In this context, helical liquids can only be realized as the edge theory of a quantum spin Hall insulator [12]. In many cases, correlation effects can be neglected in the bulk but are dominant on the edge [13]. Thereby, one can retain interactions along the edge of the system and view the bulk as a bath, which one can readily integrate out [14–17]. Further applications of the CT-QMC include for example formulations along the Keldysh contour (see Ref. [18]) or applications within the realm of extensions to DMFT methods to include spatial fluctuations. Here one can mention cluster generalizations such as the dynamical cluster or cellular DMFT approximations [19], the dual fermion approach [20], the dynamical vertex approximation [21], or extended DMFT [22].

There is a price to the flexibility of the CT-QMC algorithms. In the best-case scenario – absence of a sign problem – the computational time scales as the third power of the Euclidean volume; to be more precise $(N_{imp}\beta)^3$, where β corresponds to the inverse temperature and N_{imp} to the number of impurities. This scaling has to be contrasted with the auxiliary-field methods [4], which have linear scaling in β. Such algorithms have recently been used in the context of DMFT [23]. In the worst case, all stochastic methods are prone to the so-called negative sign problem, which effectively leads to a signal-to-noise ratio that grows exponentially in the Euclidian volume. There is to date no solution to this problem. Different algorithms or different formulations of the same algorithm can lead to very different sign problems. Clever tricks such as the fermion-bag approach can sometimes solve the problem in special situations [24].

Another issue when opting for stochastic algorithms – as opposed to NCA or tensor-network based approaches – is the fact that QMC operates on the imaginary time axis. To produce spectral functions on the real-frequency axis so as to compare with experiments, analytical continuation is necessary. This is a numerically ill-conditioned problem which limits the precision for the calculation of spectral functions. This issue may be especially severe when considering multi-orbital problems with complicated spectral line-shapes. The only solution to this prob-

lem is to work directly on the real time axis, or to analytically continue spectral functions with *simple* line shapes.

The CT-QMC approach is the method of choice for thermodynamics and equal time correlation functions. It is unbiased. Given an adequate error analysis, exact results are reproducible within the cited error-bars. The CT-QMC methods have different formulations. The interaction expansion (CT-INT) [25] and auxiliary-field (CT-AUX) [26, 27] approaches turn out to be identical, with CT-INT being the more general. The CT-INT is a *weak coupling* method and to access the strong coupling it is more convenient to use the hybridization expansion CT-HYB [28].

The organization of this lecture is the following. In section 2, we concentrate on the CT-INT [25, 26] approach in the context of the SIAM. We will review this algorithm in detail, since it has the potential for tackling lattice problems. In section 3, we will cover the basic ideas of the CT-HYB algorithm [28]. The CT-HYB is certainly the method of choice for multi-orbital models in the strong coupling limit. In section 4, we provide a generalization of the CT-INT method to tackle problems with bosonic baths. Here, the bosonic bath corresponds to a phonon mode that after integration leads to a retarded interaction.

For completeness, we have included an appendix that reviews the basic ideas of the Monte Carlo method as well as a discussion of error analysis. This is an important aspect of the implementation of CT-INT algorithms.

1 The single impurity Anderson model (SIAM)

The Anderson impurity model [29] describes the formation and screening of local magnetic moments in a metallic host. The metallic host is described by a conduction electron band with dispersion relation $\epsilon(\mathbf{k})$. The impurity state is described by a localized Kramers doublet orbital. The localized nature of the impurity orbital obliges one to include the Coulomb repulsion in terms of a Hubbard U. Finally, a hybridization matrix element $V_{\mathbf{k}}$ allows for charge transfer between the localized orbital and extended Bloch states. In second quantization, the model is given by

$$\hat{H}_{\mathrm{SIAM}} = \sum_{\mathbf{k},\sigma} \epsilon(\mathbf{k})\, \hat{c}^{\dagger}_{\mathbf{k},\sigma}\hat{c}_{\mathbf{k},\sigma} + \sum_{\mathbf{k},\sigma} \left(V_{\mathbf{k}}\, \hat{c}^{\dagger}_{\mathbf{k},\sigma}\hat{f}_{\sigma} + \bar{V}_{\mathbf{k}}\, \hat{f}^{\dagger}_{\sigma}\hat{c}_{\mathbf{k},\sigma} \right) + \epsilon_f \sum_{\sigma} \hat{n}_{\sigma} + U\, \hat{n}_{\uparrow}\hat{n}_{\downarrow}. \qquad (1)$$

Here, $\hat{c}^{\dagger}_{\mathbf{k},\sigma}$ creates a Bloch electron with z-component of spin σ, $\hat{f}^{\dagger}_{\sigma}$ an electron on the Kramers doublet localized orbital and $\hat{n}_{\sigma} = \hat{f}^{\dagger}_{\sigma}\hat{f}_{\sigma}$. A discussion of the physics described by the single impurity Anderson model can be found in Ref. [30]. In the spirit of action based CT-QMC algorithms, it is convenient to integrate out the conduction electrons. To carry out this step, we introduce fermion coherent states $|c, f\rangle$ that satisfy

$$\hat{c}_{\mathbf{k},\sigma}|c, f\rangle = c_{\mathbf{k},\sigma}|c, f\rangle, \quad \hat{f}_{\sigma}|c, f\rangle = f_{\sigma}|c, f\rangle, \qquad (2)$$

with $c_{\mathbf{k},\sigma}$ and f_{σ} being Grassmann variables. Using standard many body formalism, reviewed for example in [31], the partition function of the SIAM is given by a path integral over Grass-

mann variables

$$Z_{\mathrm{SIAM}} = \mathrm{Tr}\left[e^{-\beta \hat{H}_{\mathrm{SIAM}}}\right] = \int D\left\{c^\dagger c f^\dagger f\right\} e^{-\int_0^\beta d\tau \left[\sum_{\mathbf{k},\sigma} c^\dagger_{\mathbf{k},\sigma}(\tau)\frac{\partial}{\partial\tau} c_{\mathbf{k},\sigma}(\tau) + \sum_\sigma f^\dagger_\sigma(\tau)\frac{\partial}{\partial\tau} f_\sigma(\tau) + H_{\mathrm{SIAM}}\left(c^\dagger, c, f^\dagger, f\right)\right]} . \tag{3}$$

Since the Grassmann variables satisfy anti-periodic boundary conditions in β, we can define the Fourier transform

$$f_\sigma(i\omega_m) = \frac{1}{\sqrt{\beta}} \int_0^\beta d\tau e^{i\omega_m \tau} f_\sigma(\tau) \tag{4}$$

with $\omega_m = (2m+1)\,\pi/\beta$ a fermionic Matsubra frequency. An equivalent equation holds for the conduction electrons.
Owing to the fact that the action is bilinear in the conduction electrons one can integrate them out with a Gaussian integration to obtain our final result

$$Z_{\mathrm{SIAM}} = \int D\left\{f^\dagger f\right\} e^{-S(f^\dagger, f)} \quad \text{with}$$
$$S(f^\dagger, f) = -\int_0^\beta d\tau d\tau' \sum_\sigma f^\dagger_\sigma(\tau)\, G_0^{-1}(\tau - \tau')\, f_\sigma(\tau') + U \int_0^\beta d\tau\, f^\dagger_\uparrow(\tau) f_\uparrow(\tau) f^\dagger_\downarrow(\tau) f_\downarrow(\tau). \tag{5}$$

Here $G_0(\tau - \tau') = -\langle T \hat{f}_\sigma(\tau) \hat{f}^\dagger_\sigma(\tau')\rangle_0$ corresponds to the non-interacting f-Green function. The Gaussian integration yields

$$G_0^{-1}(\tau - \tau') = -\delta(\tau - \tau')\left[\frac{\partial}{\partial \tau'} + \epsilon_f\right] + \Delta(\tau - \tau') \;\text{ with }\; \Delta(i\omega_m) = \sum_{\mathbf{k}} \frac{|V_{\mathbf{k}}|^2}{i\omega_m - \epsilon(\mathbf{k})}. \tag{6}$$

The above equation is the starting point for both the CT-INT and CT-HYB algorithms. The CT-INT follows the idea of expanding in the interaction term, whereas the CT-HYB expands in the hybridization. In this lecture, we will assume that the fermonic bath possesses $U(1)$ gauge symmetry. Generalizations of the CT-INT to account for superconducting leads can be found in Refs. [32, 33].

2 CT-INT

In this section, we will describe in some detail the implementation of the CT-INT and show that it is equivalent to the CT-AUX

2.1 The partition function

Anticipating the elimination of the sign problem in some cases and to establish the equivalence between the CT-INT and CT-AUX algorithms, we will rewrite the Hubbard interaction as

$$\frac{U}{2} \sum_{s=\pm 1} \left(\hat{n}_\uparrow - \rho/2 - s\delta\right)\left(\hat{n}_\downarrow - \rho/2 + s\delta\right). \tag{7}$$

Here we have introduced a additional Ising variable s, and ρ corresponds to the average electronic density. We will discuss the additional parameter δ later on in this section. Starting from the action in Eq. (5), we can Taylor expand in the Hubbard interaction to obtain

$$\frac{Z_{\mathrm{SIAM}}}{Z_0} = \sum_{n=0}^{\infty} \left(\frac{-U}{2}\right)^n \frac{1}{n!} \int_0^\beta \mathrm{d}\tau_1 \sum_{s_1} \cdots \int_0^\beta \mathrm{d}\tau_n \sum_{s_n} \prod_\sigma \langle [n_\sigma(\tau_1) - \alpha_\sigma(s_1)] \cdots [n_\sigma(\tau_n) - \alpha_\sigma(s_n)] \rangle_0 \tag{8}$$

The expectation value $\langle \bullet \rangle_0$ is taken with respect to the non-interacting SIAM with partition function Z_0 and we have used the short cut notation

$$\alpha_\sigma(s) = \rho/2 + \sigma s \delta \; . \tag{9}$$

The thermal expectation value is the sum over all diagrams, connected and disconnected, of a given order n. Using the general formulation of Wick's theorem, this sum can be expressed as a determinant where the entries are the Green functions of the non-interacting system

$$\langle [n_\sigma(\tau_1) - \alpha_\sigma(s_1)] \cdots [n_\sigma(\tau_n) - \alpha_\sigma(s_n)] \rangle_0 =$$

$$\det \underbrace{\begin{pmatrix} g_0(\tau_1,\tau_1) - \alpha_\sigma(s_1) & g_0(\tau_1,\tau_2) & \cdots & g_0(\tau_1,\tau_n) \\ g_0(\tau_2,\tau_1) & g_0(\tau_2,\tau_2) - \alpha_\sigma(s_2) & \cdots & g_0(\tau_2,\tau_n) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ g_0(\tau_n,\tau_1) & g_0(\tau_n,\tau_2) & \cdots & g_0(\tau_n,\tau_n) - \alpha_\sigma(s_n) \end{pmatrix}}_{\equiv \mathbf{M}_\sigma(C_n)} , \tag{10}$$

where we have defined the Green function:

$$g_0(\tau_1,\tau_2) = \langle T \hat{f}_\sigma^\dagger(\tau_1) \hat{f}_\sigma(\tau_2) \rangle_0 \; , \tag{11}$$

which we have assumed to be spin-independent. In the above, T corresponds to the time ordering. The product of the two determinants is nothing but the sum over connected and disconnected Feynman diagrams. The summation over individual Feynman diagrams reduces the negative sign problem and, as we will see later, eliminates it altogether for a class of problems. A configuration C_n is defined by the n Hubbard vertices and Ising spins introduced in Eq. (7)

$$C_n = \{ [\tau_1, s_1] \cdots [\tau_n, s_n] \} \, . \tag{12}$$

With the short-hand notation

$$\sum_{C_n} = \sum_{n=0}^{\infty} \int_0^\beta \mathrm{d}\tau_1 \sum_{s_1} \cdots \int_0^\beta \mathrm{d}\tau_n \sum_{s_n} \, , \tag{13}$$

the partition function can conveniently be written as

$$\frac{Z_{\mathrm{SIAM}}}{Z_0} = \sum_{C_n} W(C_n), \text{ with } W(C_n) = \left(-\frac{U}{2}\right)^n \frac{1}{n!} \prod_\sigma \det \mathbf{M}_\sigma(C_n) \; . \tag{14}$$

Here $\mathbf{M}_\sigma(C_n)$ is the $n \times n$ matrix of Eq. (10).

2.2 Observables and Wick's theorem

Observables $\hat{O}(\tau)$ can now be computed with

$$\langle \hat{O}(\tau) \rangle = \frac{\sum_{C_n} W(C_n) \langle\langle \hat{O}(\tau) \rangle\rangle_{C_n}}{\sum_{C_n} W(C_n)} \; , \tag{15}$$

where for $\hat{O}(\tau) = \prod_\sigma \hat{O}_\sigma(\tau)$ we have

$$\langle\langle \hat{O}(\tau) \rangle\rangle_{C_n} = \frac{\prod_\sigma \langle T \left[\hat{n}_\sigma(\tau_1) - \alpha_\sigma(s_1)\right] \cdots \left[\hat{n}_\sigma(\tau_n) - \alpha_\sigma(s_n)\right] \hat{O}_\sigma(\tau) \rangle_0}{\prod_\sigma \langle T \left[\hat{n}_\sigma(\tau_1) - \alpha_\sigma(s_1)\right] \cdots \left[\hat{n}_\sigma(\tau_n) - \alpha_\sigma(s_n)\right] \rangle_0} \; . \tag{16}$$

We will compute the single-particle Green function and then show that any many-particle Green function can be expressed in terms of this quantity. This statement corresponds to Wick's theorem, which holds when expanding around a Gaussian theory.
Using the determinant identity given by Eq. (10), one will readily see that the single-particle Green function is given by the ratio of two determinants:

$$\langle\langle T \hat{f}^\dagger_\sigma(\tau) \hat{f}_\sigma(\tau') \rangle\rangle_{C_n} = \frac{\det B_\sigma(C_n)}{\det M_\sigma(C_n)} \tag{17}$$

where

$$B_\sigma(C_n) = \begin{pmatrix} & & g_0(\tau_1, \tau') \\ \mathbf{M}_\sigma(C_n) & & \vdots \\ & & g_0(\tau_n, \tau') \\ g_0(\tau, \tau_1) & \ldots g_0(\tau, \tau_n) & g_0(\tau, \tau') \end{pmatrix} . \tag{18}$$

To compute the ratio of the two determinants, we use the determinant identity

$$\det(\mathbf{A} + \mathbf{u} \otimes \mathbf{v}) = \det(\mathbf{A}) \left(1 + \mathbf{v} \cdot \mathbf{A}^{-1} \mathbf{u}\right) \tag{19}$$

as well as the Sherman-Morrison formula

$$(\mathbf{A} + \mathbf{u} \otimes \mathbf{v})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u} \otimes \mathbf{v}\mathbf{A}^{-1}}{1 + \mathbf{v} \cdot \mathbf{A}^{-1}\mathbf{u}} \tag{20}$$

to obtain

$$\begin{aligned} &\det\left(\mathbf{A} + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2\right) = \\ &\quad \det(\mathbf{A}) \left[\left(1 + \mathbf{v}_1 \cdot \mathbf{A}^{-1}\mathbf{u}_1\right)\left(1 + \mathbf{v}_2 \cdot \mathbf{A}^{-1}\mathbf{u}_2\right) - \left(\mathbf{v}_2 \cdot \mathbf{A}^{-1}\mathbf{u}_1\right)\left(\mathbf{v}_1 \cdot \mathbf{A}^{-1}\mathbf{u}_2\right)\right] \; , \end{aligned} \tag{21}$$

where the outer product is given by $(\mathbf{u} \otimes \mathbf{v})_{i,j} = \mathbf{u}_i \mathbf{v}_j$ and the scalar product by $\mathbf{u} \cdot \mathbf{v} = \sum_i \mathbf{u}_i \mathbf{v}_i$. Eq. (20) can be formally derived by Taylor-expanding $(1 + \mathbf{A}^{-1}\mathbf{u} \otimes \mathbf{v})^{-1}$. Eq. (19) can equally be formally demonstrated by using the fact that $\det(\mathbf{A}) = \exp \mathrm{Tr} \log(\mathbf{A})$.
Decomposing the $\mathbf{B}_\sigma(C_n)$ matrix as

$$\mathbf{B}_\sigma(C_n) = \begin{pmatrix} & & 0 \\ \mathbf{M}_\sigma(C_n) & & \vdots \\ & & 0 \\ 0 & \ldots 0 & 1 \end{pmatrix} + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2 \tag{22}$$

with

$$\mathbf{u}_1 = (g_0(\tau_1,\tau'),\ldots,g_0(\tau_n,\tau'),g_0(\tau,\tau')-1)\,,\quad (\mathbf{v}_1)_i = \delta_{i,n+1}$$

and

$$(\mathbf{u}_2)_i = \delta_{i,n+1},\quad \mathbf{v}_2 = (g_0(\tau,\tau_1),\ldots,g_0(\tau,\tau_n),0)$$

yields the following expression for the single-particle Green function

$$g(\tau,\tau')_{C_n} \equiv \langle\langle T\hat{f}_\sigma^\dagger(\tau)\hat{f}_\sigma(\tau')\rangle\rangle_{C_n} = g_0(\tau,\tau') - \sum_{r,s=1}^{n} g_0(\tau,\tau_r)\left(\mathbf{M}_\sigma(C_n)^{-1}\right)_{r,s} g_0(\tau_s,\tau')\,. \quad (23)$$

An important consequence of a continuous-time formulation is that one can compute the Green function directly in Matsubara frequencies. With the Fourier transformation of Eq. (4) one obtains:

$$g(i\omega_m,i\omega_m')_{C_n} = \delta_{\omega_m,\omega_m'}g_0(i\omega_m) - g_0(i\omega_m)\left(\frac{1}{\beta}\sum_{r,s=1}^{n} e^{-i\omega_m\tau_r}\left(\mathbf{M}_\sigma(C_n)^{-1}\right)_{r,s} e^{i\omega_m'\tau_s}\right) g_0(i\omega_m'). \quad (24)$$

For a given configuration of vertices C_n, translation symmetry in imaginary time is broken such that $g(i\omega_m,i\omega_m')_{C_n}$ has to be a function of two Matsubara frequencies. Clearly, translation symmetry has to be restored after summation over the configurations C_n has been carried out. One can use this fact to define improved estimators for the Green function.

Higher-order Green functions may be computed by using the matrix identity demonstrated in App. B [32]. Here we consider Green functions of the form $\langle T\hat{f}_\sigma^\dagger(1)\hat{f}_\sigma(1')\ldots\hat{f}_\sigma^\dagger(m)\hat{f}_\sigma(m')\rangle$. For every configuration C_n, a relation similar to Wick's theorem can be found, which greatly simplifies the calculation of higher-order Green functions. The application of the ordinary Wick theorem to the denominator and the numerator of Eq. (16) yields

$$\langle\langle T\hat{f}_\sigma^\dagger(1)\hat{f}_\sigma(1')\ldots\hat{f}_\sigma^\dagger(m)\hat{f}_\sigma(m')\rangle\rangle_{C_n} = \frac{\det\mathbf{B}_\sigma(C_n)}{\det\mathbf{M}_\sigma(C_n)}, \quad (25)$$

where now $\mathbf{B}_\sigma(C_n)$ is a $\mathbb{C}^{(n+m)\times(n+m)}$ matrix given by

$$\mathbf{B}_\sigma(C_n) = \begin{pmatrix} & & & g_0(\tau_1,1') & \ldots & g_0(\tau_1,m') \\ & \mathbf{M}_\sigma(C_n) & & \vdots & \ddots & \vdots \\ & & & g_0(\tau_n,1') & \ldots & g_0(\tau_n,m') \\ g_0(1,\tau_1) & \ldots & g_0(1,\tau_n) & g_0(1,1') & \ldots & g_0(1,m') \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_0(m,\tau_1) & \ldots & g_0(m,\tau_n) & g_0(m,1') & \ldots & g_0(m,m') \end{pmatrix}. \quad (26)$$

Defining the matrices $\mathbf{B}_\sigma^{\mathbf{ij}}(C_n) \in \mathbb{C}^{(n+1)\times(n+1)}$,

$$\mathbf{B}_\sigma^{\mathbf{ij}}(C_n) = \begin{pmatrix} & & g_0(\tau_1,i') \\ & \mathbf{M}_\sigma(C_n) & \vdots \\ & & g_0(\tau_n,i') \\ g_0(j,\tau_1,) & \ldots & g_0(j,\tau_n) \quad g_0(j,i') \end{pmatrix}, \quad (27)$$

we can make use of the determinant identity (145) yielding

$$\frac{\det \mathbf{B}_{C_n}}{\det \mathbf{M}_{C_n}} = \frac{1}{(\det \mathbf{M}_{C_n})^n} \det \begin{pmatrix} \det \mathbf{B}^{\mathbf{11}}_{C_n} & \dots & \det \mathbf{B}^{\mathbf{1m}}_{C_n} \\ \vdots & \ddots & \vdots \\ \det \mathbf{B}^{\mathbf{m1}}_{C_n} & \dots & \det \mathbf{B}^{\mathbf{mm}}_{C_n} \end{pmatrix}. \tag{28}$$

From Eq. (17) and (18), it is obvious that $\det \mathbf{B}^{\mathbf{ij}}_{\sigma}(C_n)/\det \mathbf{M}_{\sigma}(C_n)$ is identical to the contribution of the configuration C_n to the one particle Green's function $\langle T \hat{f}^{\dagger}_{\sigma}(j) \hat{f}_{\sigma}(i') \rangle$. Hence, Wick's theorem holds for every configuration C_n and is given by

$$\begin{aligned} \langle\langle T \hat{f}^{\dagger}_{\sigma}(1) \hat{f}_{\sigma}(1') \dots \hat{f}^{\dagger}_{\sigma}(m) \hat{f}_{\sigma}(m') \rangle\rangle_{C_n} = \\ \det \begin{pmatrix} \langle\langle T \hat{f}^{\dagger}_{\sigma}(1) \hat{f}_{\sigma}(1') \rangle\rangle_{C_n} & \dots & \langle\langle T \hat{f}^{\dagger}_{\sigma}(1) \hat{f}_{\sigma}(m') \rangle\rangle_{C_n} \\ \vdots & \ddots & \vdots \\ \langle\langle T \hat{f}^{\dagger}_{\sigma}(m) \hat{f}_{\sigma}(1') \rangle\rangle_{C_n} & \dots & \langle\langle T \hat{f}^{\dagger}_{\sigma}(m) \hat{f}_{\sigma}(m') \rangle\rangle_{C_n} \end{pmatrix}. \end{aligned} \tag{29}$$

This demonstrates that knowing the single particle Green function $\langle\langle T \hat{f}^{\dagger}_{\sigma}(\tau) \hat{f}_{\sigma}(\tau') \rangle\rangle_{C_n}$ suffices to compute any observable.

2.3 The negative sign problem

For each configuration of vertices C_n, we are able to compute arbitrary correlation functions. Due to the dimension of the configuration space, it is prohibitively expensive to carry out the summation over C_n exactly. One will thus opt for a stochastic Monte Carlo approach which, for completeness sake, is reviewed in Appendix A.

A prerequisite for applying the Monte Carlo approach is that the *weight*, $W(C_n)$, of Eq. (14) is positive. For quantum systems, a positive formulation is not always possible, and one will decide to sample $|W(C_n)|$. Thereby, we will carry out the Monte Carlo evaluation with

$$\langle \hat{O}(\tau) \rangle = \frac{\sum_{C_n} W(C_n) \langle\langle \hat{O}(\tau) \rangle\rangle_{C_n}}{\sum_{C_n} W(C_n)} = \frac{\frac{\sum_{C_n} |W(C_n)| \mathrm{sign}(W(C_n)) \langle\langle \hat{O}(\tau) \rangle\rangle_{C_n}}{\sum_{C_n} |W(C_n)|}}{\frac{\sum_{C_n} |W(C_n)| \mathrm{sign}(W(C_n))}{\sum_{C_n} |W(C_n)|}} \tag{30}$$

and separately compute the numerator and denominator. The denominator corresponds to the average sign. On general grounds, one can argue that it is given by the ratio of two partition functions and thereby decays exponentially with the inverse temperature and the number of impurities N_{imp}

$$\langle \mathrm{sign} \rangle = \frac{\sum_{C_n} W(C_n)}{\sum_{C_n} |W(C_n)|} \simeq e^{-\Delta \beta N_{\mathrm{imp}}} . \tag{31}$$

In other words, at low temperatures there is a next to perfect cancellation of positive and negative weights. The sign problem is a consequence of the law of large numbers which states that the error on the average sign scales as

$$\sigma_{\mathrm{sign}} \sim \frac{1}{\sqrt{T_{\mathrm{CPU}}}} , \tag{32}$$

where $T_{\rm CPU}$ corresponds to the computational time. Obviously, to obtain sensible results, we will require that

$$\frac{\sigma_{\rm Sign}}{\langle {\rm sign} \rangle} \ll 1 \tag{33}$$

such that

$$T_{\rm CPU} \gg e^{2\Delta\beta N_{\rm imp}} \, . \tag{34}$$

Hence the required CPU time will scale exponentially with inverse temperature and number of impurities. Note that one can counter the sign problem if one can define an improved estimator for the average sign such that the fluctuations are greatly suppressed! Let us furthermore note that the pre-factor Δ is formulation dependent. One can for instance mention recent work of Huffman *et al.* [24], who have found a CT-INT formulation for spin-polarized electron problems at half-band filling that is free of the sign problem. Hence in this case $\Delta = 0$. There are other problems which are free of the negative sign problem. To show this, for the special case of attractive and repulsive Hubbard interactions, we will consider the mapping of the CT-INT to the CT-AUX [26, 27, 34]. This mapping allows the use of results derived in the framework of the Hirsch-Fye [35] and auxiliary-field QMC (for a review see Ref. [4]) algorithms to argue for the absence of the sign problem.

Let us start with the repulsive Hubbard interaction, which we will write as

$$H_U = \frac{U}{2} \sum_{s=\pm 1} \left[(\hat{n}_\uparrow - 1/2) + s\delta \right] \left[(\hat{n}_\uparrow - 1/2) - s\delta \right] = - \frac{U\left(\delta^2 - 1/4\right)}{2} \sum_{s=\pm 1} e^{\alpha s \hat{m}} \tag{35}$$

with

$$\cosh(\alpha) - 1 = \frac{1}{2} \, \frac{1}{\delta^2 - 1/4}$$

and magnetization $\hat{m} = \hat{n}_\uparrow - \hat{n}_\downarrow$. This identity relies on the fact that $\hat{m}^4 = \hat{m}^2$ and requires $\delta > 1/2$. Using this identity, the weight of a vertex configuration is given by

$$W(C_n) = \left(\frac{U\left(\delta^2 - 1/4\right)}{2} \right)^n \frac{1}{n!} \, \langle T e^{\alpha s_1 \hat{m}(\tau_1)} \ldots e^{\alpha s_n \hat{m}(\tau_n)} \rangle_0 \tag{36}$$

where $\langle \bullet \rangle_0$ corresponds to the thermal expectation value with respect to the non-interacting model. Several comments are in order. (i) The notation $K = U\beta(\delta^2 - 1/4)$ makes the mapping to the CT-AUX explicit [27]. (ii) For $\delta > 1/2$, $U\left(\delta^2 - 1/4\right)$ is positive for the repulsive case. (iii) For a given set of Ising fields, the thermal expectation value has precisely the same structure as in the Hirsch-Fye and auxiliary-field QMC algorithms [4]. It hence follows that the CT-INT, CT-AUX, Hirsch-Fye and auxiliary-field QMC algorithms have the same sign problem for repulsive Hubbard interactions. Thus, as shown in Ref. [36], the SIAM is free of the negative sign problem. By the same token, one can argue that a class of one-dimensional problems [6] and problems with particle-hole symmetry such as the Kane-Mele-Hubbard model [37, 13] are sign problem free if formulated within the CT-INT. Studies of correlation effects in one-dimensional helical liquids [15] hinge on this observation.

For attractive interactions, one can use a similar identity as above. In this case, we have to adopt a different convention for the δ-shift

$$H_U = \frac{U}{2} \sum_{s=\pm 1} \left[(\hat{n}_\uparrow - 1/2) + s\delta \right] \left[(\hat{n}_\uparrow - 1/2) + s\delta \right] = \frac{U\left(\delta^2 - 1/4\right)}{2} \sum_{s=\pm 1} e^{\alpha s (\hat{n}-1)} \tag{37}$$

with $\hat{n} = \hat{n}_\uparrow + \hat{n}_\downarrow$. Again, the above equation relies on the fact that $(\hat{n}-1)^4 = (\hat{n}-1)^2$ and the same equation as above holds for α. Thus, for the attractive case, the weight reads

$$W(C_n) = \left(-\frac{U\left(\delta^2 - 1/4\right)}{2} \right)^n \frac{1}{n!} \langle T e^{\alpha s_1 (\hat{n}(\tau_1)-1)} \dots e^{\alpha s_n (\hat{n}(\tau_n)-1)} \rangle_0 \,. \tag{38}$$

Since $U < 0$ and $\delta > 1/2$, $-U\left(\delta^2 - 1/4\right)$ is positive. Furthermore, provided that the non-interacting model factorizes into identical spin-up an spin-down real representable Hamiltonians, the thermal expectation value reads

$$\langle T e^{\alpha s_1 (\hat{n}(\tau_1)-1)} \dots e^{\alpha s_n (\hat{n}(\tau_n)-1)} \rangle_0 = \left[\langle T e^{\alpha s_1 \left(\hat{n}_\uparrow(\tau_1) - 1/2\right)} \dots e^{\alpha s_n \left(\hat{n}_\uparrow(\tau_n) - 1/2\right)} \rangle_{0,\uparrow} \right]^2 , \tag{39}$$

which is manifestly positive. Here $\langle \bullet \rangle_{0,\uparrow}$ corresponds to the thermal expectation value in the spin-up sector. Factorization is not necessarily required for the absence of the sign problem in the presence of attractive interactions. In general, time-reversal-symmetric fermionic problems where time reversal symmetry is present for *every configuration* C_n, are free of the minus sign problem. This follows essentially from Kramers theorem and is proven in Ref. [38].

2.4 The Monte Carlo sampling

At this point, we will assume that the weight is positive such that we can carry out Monte Carlo importance sampling. Only two moves are required to guarantee ergodicity: addition and subtraction of vertices. Vertex addition corresponds to the proposal

$$C_n = \{ [\tau_1, s_1] \cdots [\tau_n, s_n] \} \to C_{n+1} = \{ [\tau_1, s_1] \cdots [\tau_i, s_i], [\tau', s'], [\tau_{i+1}, s_{i+1}] \cdots [\tau_n, s_n] \} \tag{40}$$

where we add the vertex τ', s' at position i in the string. The proposal probability reads

$$T^0_{C_n \to C_{n+1}} = \underbrace{\frac{1}{n+1}}_{\text{Position in string}} \underbrace{\frac{1}{\beta}}_{\text{Value of } \tau'} \underbrace{\frac{1}{2}}_{\text{Value of } s'} . \tag{41}$$

Vertex removal corresponds to

$$C_n = \{ [\tau_1, s_1] \cdots [\tau_n, s_n] \} \to C_{n-1} = \{ [\tau_1, s_1] \cdots [\tau_i, s_i] \, , [\tau_{i+2}, s_{i+2}] \, , \cdots [\tau_n, s_n] \} \, , \tag{42}$$

where vertex i has been removed. The probability to propose this move reads

$$T^0_{C_n \to C_{n-1}} = 1/n \, , \tag{43}$$

which corresponds to the probability of choosing vertex i under the assumption that each vertex is equally probable. As shown in Appendix A (see Eq. (132)), the Metropolis acceptance reads

$$P_{C\to C'} = \min\left(\frac{T^0_{C'\to C}W(C')}{T^0_{C\to C'}W(C)}, 1\right). \tag{44}$$

Thus

$$\begin{aligned} P_{C_n\to C_{n+1}} &= \min\left(-\frac{U\beta}{(n+1)}\frac{\prod_\sigma \det \mathbf{M}_\sigma(C_{n+1})}{\prod_\sigma \det \mathbf{M}_\sigma(C_n)}, 1\right) \\ P_{C_{n+1}\to C_n} &= \min\left(-\frac{(n+1)}{U\beta}\frac{\prod_\sigma \det \mathbf{M}_\sigma(C_n)}{\prod_\sigma \det \mathbf{M}_\sigma(C_{n+1})}, 1\right) . \end{aligned}$$

Note that in our formulation the ordering of the vertices is important since we have defined the integration without time-ordering $\int_0^\beta \mathrm{d}\tau_1 \cdots \int_0^\beta \mathrm{d}\tau_n$ as opposed to a time-ordered formulation $\int_0^\beta \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \cdots \int_0^{\tau_{n-1}} \mathrm{d}\tau_n$. The reader is encouraged to show that both formulations lead to the same acceptance/rejection ratios. In practical implementations one will also include a move that keeps the vertex number constant but flips the value of the Ising spin. Strictly speaking, this move is not necessary but has the potential of improving the autocorrelation time. For repulsive interactions, this statement follows from the notion that summing over the Ising fields will restore the broken $SU(2)$ spin-symmetry.

2.5 Fast updates

The Monte Carlo dynamics relies on the calculation of ratios of determinants. Such ratios can be computed using the determinant identities of Eq. (21). For instance, under vertex addition we will have to compute for each spin sector

$$\begin{aligned} \frac{\det \mathbf{M}_\sigma(C_{n+1})}{\det \mathbf{M}_\sigma(C_n)} &= \frac{\det \begin{pmatrix} & & g_0(\tau_1,\tau') \\ \mathbf{M}_\sigma(C_n) & & \vdots \\ & & g_0(\tau_n,\tau') \\ g_0(\tau',\tau_1) & \dots g_0(\tau',\tau_n) & g_0(\tau',\tau') - \alpha_\sigma(s') \end{pmatrix}}{\det \mathbf{M}_\sigma(C_n)} \\ &= \frac{\det \left(\begin{pmatrix} & & 0 \\ \mathbf{M}_\sigma(C_n) & & \vdots \\ & & 0 \\ 0 & \dots 0 & 1 \end{pmatrix} + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2 \right)}{\det \mathbf{M}_\sigma(C_n)} \end{aligned}$$

$$= \left(1 + \mathbf{v}_1 \cdot \mathbf{M}_\sigma^{-1}(C_n)\mathbf{u}_1\right)\left(1 + \mathbf{v}_2 \cdot \mathbf{M}_\sigma^{-1}(C_n)\mathbf{u}_2\right) - \left(\mathbf{v}_2 \cdot \mathbf{M}_\sigma^{-1}(C_n)\mathbf{u}_1\right)\left(\mathbf{v}_1 \cdot \mathbf{M}_\sigma^{-1}(C_n)\mathbf{u}_2\right).$$

with

$$\mathbf{u}_1 = (g_0(\tau_1,\tau'), \dots, g_0(\tau_n,\tau'), g_0(\tau',\tau') - \alpha_\sigma(s') - 1), \quad (\mathbf{v}_1)_i = \delta_{i,n+1}$$

and

$$(\mathbf{u}_2)_i = \delta_{i,n+1}, \quad \mathbf{v}_2 = (g_0(\tau', \tau_1), \ldots, g_0(\tau', \tau_n), 0).$$

Carrying out the calculation yields

$$\frac{\det \mathbf{M}_\sigma(C_{n+1})}{\det \mathbf{M}_\sigma(C_n)} = g_0(\tau', \tau') - \alpha_\sigma(s') - \sum_{i,j=1}^{n} g_0(\tau', \tau_i) \left[M_\sigma^{-1}(C_n)\right]_{i,j} g_0(\tau_j, \tau') \,. \tag{45}$$

Hence, provided that the matrix $\mathbf{M}_\sigma^{-1}(C_n)$ is known, computing the ratio involves n^2 operations. The vertex removal takes a very simple form. Assume that we remove the n^{th} vertex of the configuration C_n. Then for a given spin sector, we will have to compute

$$\begin{aligned} \frac{\det \mathbf{M}_\sigma(C_{n-1})}{\det \mathbf{M}_\sigma(C_n)} &= \frac{\det \begin{pmatrix} g_0(\tau_1,\tau_1) - \alpha_\sigma(s_1) & \ldots & g_0(\tau_1,\tau_{n-1}) & 0 \\ \vdots & & \vdots & \vdots \\ g_0(\tau_{n-1},\tau_1) & \ldots & g_0(\tau_{n-1},\tau_{n-1}) - \alpha_\sigma(s_{n-1}) & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}}{\det \mathbf{M}_\sigma(C_n)} \\ &= \frac{\det \left[\mathbf{M}_\sigma(C_n) + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2\right]}{\det \mathbf{M}_\sigma(C_n)} \end{aligned} \tag{46}$$

with

$$\mathbf{u}_1 = -\left(\left[M_\sigma(C_n)\right]_{1,n}, \ldots, \left[M_\sigma(C_n)\right]_{n,n} - 1 \right), \quad (\mathbf{v}_1)_i = \delta_{i,n}$$

and

$$(\mathbf{u}_2)_i = \delta_{i,n}, \quad \mathbf{v}_2 = -\left(\left[M_\sigma(C_n)\right]_{n,1}, \ldots, \left[M_\sigma(C_n)\right]_{n,n-1}, 0 \right).$$

Evaluating the above gives

$$\frac{\det \mathbf{M}_\sigma(C_{n-1})}{\det \mathbf{M}_\sigma(C_n)} = \left[M_\sigma^{-1}(C_n)\right]_{n,n} . \tag{47}$$

Again, provided that we have the matrix $\mathbf{M}_\sigma^{-1}(C_n)$ at hand, the computational cost for computing the ratio for vertex removal is negligible.

Having computed the ratio of determinants, we can compute the acceptance probability, draw a pseudo-random number, and accept or reject the move. If accepted, we will have to upgrade the matrix $\mathbf{M}_\sigma^{-1}(C_n)$. This is readily done with the use of the Sherman-Morrison formula of Eq. (20) and involves n^2 operations.

In some cases, it is desirable to add more than one vertex at a time. For this purpose, it is more useful to use the Woodbury formula

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U\left(C^{-1} + VA^{-1}U\right)^{-1} VA^{-1} \tag{48}$$

with $A \in \mathbb{C}^{n\times n}$, $U \in \mathbb{C}^{n\times k}$, $C \in \mathbb{C}^{k\times k}$ and $V \in \mathbb{C}^{k\times n}$. The Woodbury identity reduces to the Sherman-Morrison formula of Eq. (20) at $k = 1$ and $C = 1$. A discussion of block updates as well as a demonstration of various matrix identities can be found in Ref. [39].

2.6 Average expansion parameter

A crucial issue concerns the average expansion parameter $\langle n \rangle$ since it will determine the average size of the matrix $\mathbf{M}_\sigma(C_n)$. The computational effort to visit each vertex – a sweep – will then scale as $\langle n \rangle^3$. For a general interaction term $\hat{H}_1$, the average expansion parameter is [6]

$$\begin{aligned}\langle n \rangle = & \frac{1}{Z} \sum_n \frac{(-1)^n n}{n!} \int_0^\beta \mathrm{d}\tau_1 \quad \cdots \quad \int_0^\beta \mathrm{d}\tau_n \, \langle T \hat{H}_1(\tau_1) \quad \cdots \quad \hat{H}_1(\tau_n) \rangle_0 \\ = & -\frac{1}{Z} \sum_m \frac{(-1)^m}{m!} \int_0^\beta \mathrm{d}\tau_1 \cdots \int_0^\beta \mathrm{d}\tau_m \int_0^\beta \mathrm{d}\tau \, \langle T \hat{H}_1(\tau_1) \cdots \hat{H}_1(\tau_m) \hat{H}_1(\tau) \rangle_0 \\ = & -\int_0^\beta \mathrm{d}\tau \langle \hat{H}_1(\tau) \rangle \, . \end{aligned} \tag{49}$$

For the Hubbard model, replacing $\hat{H}_1$ by the form of Eq. (7), we obtain

$$\langle n \rangle = -\beta U \left[\langle (\hat{n}_\uparrow - 1/2)(\hat{n}_\downarrow - 1/2) \rangle - \delta^2 \right] , \tag{50}$$

where we have set $\rho = 1/2$. Thus, the computational time for a sweep scales as in the Hirsch-Fye approach, namely as $(\beta U)^3$. The algorithm can be used for lattice models with N correlated sites. In this case, the computational time for a sweep scales as $(N\beta U)^3$, which is more expensive than the auxiliary-field approach, which scales as $\beta U N^3$. As mentioned in the introduction, the advantage of the CT-INT method lies in the fact that it is action-based such that fermionic and bosonic baths can be easily implemented.

3 CT-HYB

In this section we will provide a very succinct overview of the basic formulation of the CT-HYB. For a detailed discussion of the algorithm, the reader is referred to the review article [1] and references therein.

3.1 The partition function

In contrast to the CT-INT, the CT-HYB carries out the expansion in the hybridization matrix $\Delta(\tau - \tau')$. The action of the SIAM is decomposed into local

$$S_{\mathrm{loc}}(f^\dagger, f) = \int_0^\beta d\tau \sum_\sigma f_\sigma^\dagger(\tau) \left[\frac{\partial}{\partial \tau} + \epsilon_f \right] f_\sigma(\tau) + U \int_0^\beta d\tau \, f_\uparrow^\dagger(\tau) f_\uparrow(\tau) f_\downarrow^\dagger(\tau) f_\downarrow(\tau) \tag{51}$$

and hybridization

$$S_{\mathrm{hyb}}(f^\dagger, f) = -\int_0^\beta d\tau d\tau' \sum_\sigma f_\sigma^\dagger(\tau) \, \Delta(\tau - \tau') \, f_\sigma(\tau') \tag{52}$$

parts. Taylor expanding in $S_{\rm hyb}$ gives

$$Z_{\rm SIAM} = \int D\{f^\dagger f\}\ e^{-S_{\rm loc}(f^\dagger,f)} \sum_n \frac{1}{n!} \sum_{\sigma_1 \ldots \sigma_n} \int_0^\beta d\tau_1 d\tau_1' \ldots d\tau_n d\tau_n' \\ \times f^\dagger_{\sigma_1}(\tau_1) f_{\sigma_1}(\tau_1') \ldots f^\dagger_{\sigma_n}(\tau_n) f_{\sigma_n}(\tau_n') \Delta(\tau_1 - \tau_1') \ldots \Delta(\tau_n - \tau_n'). \tag{53}$$

The insight of Ref. [28] is to sum up a set of configurations corresponding to the permutations of the Grassmann variables f. As we will see, the weight of this sum of configurations is given by an $n \times n$ determinant of the hybridization function. This is the crucial step in the basic formulation of the algorithm that leads to the absence of a sign problem for the SIAM. To achieve this, it is convenient to introduce the notation

$$x = (\tau, \sigma), \quad \int dx = \sum_\sigma \int_0^\beta d\tau, \quad \Delta_{x,x'} = \Delta(\tau - \tau')\delta_{\sigma,\sigma'} \tag{54}$$

such that

$$\sum_{\sigma_1 \ldots \sigma_n} \int_0^\beta d\tau_1 d\tau_1' \ldots d\tau_n d\tau_n' f^\dagger_{\sigma_1}(\tau_1) f_{\sigma_1}(\tau_1') \ldots f^\dagger_{\sigma_n}(\tau_n) f_{\sigma_n}(\tau_n') \Delta(\tau_1 - \tau_1') \ldots \Delta(\tau_n - \tau_n') \\ = \int dx_1 dx_1' \ldots dx_n dx_n'\ f^\dagger_{x_1} f_{x_1'} \ldots f^\dagger_{x_n} f_{x_n'}\ \Delta_{x_1,x_1'} \ldots \Delta_{x_n,x_n'} \\ = \frac{1}{n!} \sum_{P \in \mathcal{S}_n} \int dx_1 dx_{P(1)}' \ldots dx_n dx_{P(n)}'\ f^\dagger_{x_1} f_{x_{P(1)}'} \ldots f^\dagger_{x_n} f_{x_{P(n)}'}\ \Delta_{x_1,x_{P(1)}'} \ldots \Delta_{x_n,x_{P(n)}'}. \tag{55}$$

In the above P is a permutation of n objects, and we have merely replicated the result $n!$ times. Using the anti-commuting property of the Grassmann algebra, one can show that

$$f^\dagger_{x_1} f_{x_{P(1)}'} \ldots f^\dagger_{x_n} f_{x_{P(n)}'} = (-1)^P f^\dagger_{x_1} f_{x_1'} \ldots f^\dagger_{x_n} f_{x_n'}$$

where $(-1)^P$ is the sign of the permutation. Since $dx_1 dx_{P(1)}' \ldots dx_n dx_{P(n)}' = dx_1 dx_1' \ldots dx_n dx_n'$ Eq. (55) transforms as

$$\frac{1}{n!} \int dx_1 dx_1' \ldots dx_n dx_n'\ \ f^\dagger_{x_1} f_{x_1'} \ldots f^\dagger_{x_n} f_{x_n'}\ \ \sum_{P \in \mathcal{S}_n} (-1)^P \Delta_{x_1,x_{P(1)}'} \ldots \Delta_{x_n,x_{P(n)}'} \\ = \quad \frac{1}{n!} \int dx_1 dx_1' \ldots dx_n dx_n'\ \ f^\dagger_{x_1} f_{x_1'} \ldots f^\dagger_{x_n} f_{x_n'} \det \begin{pmatrix} \Delta_{x_1,x_1'} & \ldots & \Delta_{x_1,x_n'} \\ \vdots & & \vdots \\ \Delta_{x_n,x_1'} & \ldots & \Delta_{x_n,x_n'} \end{pmatrix} \tag{56}$$

Using the above, the partition function is written as

$$\frac{Z_{\rm SIAM}}{Z_{\rm loc}} = \sum_n \frac{1}{n!^2} \int dx_1 dx_1' \ldots dx_n dx_n'\ \langle f^\dagger_{x_1} f_{x_1'} \ldots f^\dagger_{x_n} f_{x_n'} \rangle_{\rm loc} \det \begin{pmatrix} \Delta_{x_1,x_1'} & \ldots & \Delta_{x_1,x_n'} \\ \vdots & & \vdots \\ \Delta_{x_n,x_1'} & \ldots & \Delta_{x_n,x_n'} \end{pmatrix} \tag{57}$$

where

$$Z_{\rm loc} = \int D\{f^\dagger f\}\, e^{-S_{\rm loc}(f^\dagger, f)}$$

and

$$\langle \bullet \rangle_{\rm loc} = \frac{1}{Z_{\rm loc}} \int D\left\{f^\dagger f\right\} e^{-S_{\rm loc}(f^\dagger, f)} \quad \bullet \, .$$

For the special case where the hybridization function is spin diagonal, one can simplify the above equation. In this case, the spin variables $\{\sigma_1' \ldots \sigma_n'\}$ have to be a permutation of $\{\sigma_1 \ldots \sigma_n\}$ such that

$$\sigma_i' = \sigma_{P(i)} \text{ with } P \in \mathcal{S}_n. \tag{58}$$

Hence,

$$\begin{aligned}
&\int dx_1 dx_1' \ldots dx_n dx_n' \;\; \langle f^\dagger_{x_1} f_{x_1'} \ldots f^\dagger_{x_n} f_{x_n'} \rangle_{\rm loc} \det \begin{pmatrix} \Delta_{x_1,x_1'} & \cdots & \Delta_{x_1,x_n'} \\ \vdots & & \vdots \\ \Delta_{x_n,x_1'} & \cdots & \Delta_{x_n,x_n'} \end{pmatrix} \\
&= \sum_{P\in\mathcal{S}_n} \sum_{\sigma_1 \ldots \sigma_n} \int_0^\beta d\tau_1 d\tau_1' \ldots d\tau_n d\tau_n' \;\; \langle f^\dagger_{\tau_1,\sigma_1} f_{\tau_1',\sigma_{P(1)}} \ldots f^\dagger_{\tau_n,\sigma_n} f_{\tau_n',\sigma_{P(n)}} \rangle_{\rm loc} \\
&\qquad \times \det \begin{pmatrix} \Delta_{(\tau_1,\sigma_1),(\tau_1',\sigma_{P(1)})} & \cdots & \Delta_{(\tau_1,\sigma_1),(\tau_n',\sigma_{P(n)})} \\ \vdots & & \vdots \\ \Delta_{(\tau_n,\sigma_n),(\tau_1',\sigma_{P(1)})} & \cdots & \Delta_{(\tau_n,\sigma_n),(\tau_n',\sigma_{P(n)})} \end{pmatrix}.
\end{aligned}$$

For a given permutation P one can carry out the substitution $\tau_i' = \tau''_{P(i)}$. This substitution leaves the integration measure invariant such that the above reads

$$\begin{aligned}
&\sum_{P\in\mathcal{S}_n} \sum_{\sigma_1 \ldots \sigma_n} \int_0^\beta d\tau_1 d\tau_1'' \ldots d\tau_n d\tau_n'' \;\; \langle f^\dagger_{\tau_1,\sigma_1} f_{\tau''_{P(1)},\sigma_{P(1)}} \ldots f^\dagger_{\tau_n,\sigma_n} f_{\tau''_{P(n)},\sigma_{P(n)}} \rangle_{\rm loc} \\
&\qquad \times \det \begin{pmatrix} \Delta_{(\tau_1,\sigma_1),(\tau''_{P(1)},\sigma_{P(1)})} & \cdots & \Delta_{(\tau_1,\sigma_1),(\tau''_{P(n)},\sigma_{P(n)})} \\ \vdots & & \vdots \\ \Delta_{(\tau_n,\sigma_n),(\tau''_{P(1)},\sigma_{P(1)})} & \cdots & \Delta_{(\tau_n,\sigma_n),(\tau''_{P(n)},\sigma_{P(n)})} \end{pmatrix}.
\end{aligned}$$

One can get rid of the permutation under the integral by reordering the Grassmann variables $f_{\tau''_{P(i)},\sigma_{P(i)}}$ as well as the columns of the matrix. In this process, the minus signs cancel and the sum over the permutations gives a factor $n!$. Hence, the partition function reads

$$\begin{aligned}
\frac{Z_{\rm SIAM}}{Z_{\rm loc}} \;\; = \;\; &\sum_n \frac{1}{n!} \sum_{\sigma_1 \ldots \sigma_n} \int_0^\beta d\tau_1 d\tau_1'' \ldots d\tau_n d\tau_n'' \; \langle f^\dagger_{\tau_1,\sigma_1} f_{\tau_1'',\sigma_1} \ldots f^\dagger_{\tau_n,\sigma_n} f_{\tau_n'',\sigma_n} \rangle_{\rm loc} \\
&\qquad \times \det \begin{pmatrix} \Delta_{(\tau_1,\sigma_1),(\tau_1'',\sigma_1)} & \cdots & \Delta_{(\tau_1,\sigma_1),(\tau_n'',\sigma_n)} \\ \vdots & & \vdots \\ \Delta_{(\tau_n,\sigma_n),(\tau_1'',\sigma_1)} & \cdots & \Delta_{(\tau_n,\sigma_n),(\tau_n'',\sigma_n)} \end{pmatrix}.
\end{aligned} \tag{59}$$

3.2 The Monte Carlo sampling and evaluation of the trace

For the SIAM where the hybridization matrix is spin-diagonal, we can define a configuration as

$$C_n = \{[\tau_1, \tau_1', \sigma_1] \cdots [\tau_n, \tau_n', \sigma_n]\} \tag{60}$$

and the weight of a configuration as

$$W(C_n) = \frac{1}{n!} \langle f^\dagger_{\tau_1,\sigma_1} f_{\tau_1'',\sigma_1} \cdots f^\dagger_{\tau_n,\sigma_n} f_{\tau_n'',\sigma_n} \rangle_{\mathrm{loc}} \det \begin{pmatrix} \Delta_{(\tau_1,\sigma_1),(\tau_1'',\sigma_1)} & \cdots & \Delta_{(\tau_1,\sigma_1),(\tau_n'',\sigma_n)} \\ \vdots & & \vdots \\ \Delta_{(\tau_n,\sigma_n),(\tau_1'',\sigma_1)} & \cdots & \Delta_{(\tau_n,\sigma_n),(\tau_n'',\sigma_n)} \end{pmatrix} \tag{61}$$

such that $\frac{Z_{\mathrm{SIAM}}}{Z_{\mathrm{loc}}} = \sum_{C_n} W(C_n)$. The simplest possible proposal matrix to add a vertex reads,

$$T^0_{C_n \to C_{n+1}} = \underbrace{\frac{1}{n+1}}_{\text{Position in string}} \quad \underbrace{\frac{1}{\beta}}_{\text{Value of } \tau} \quad \underbrace{\frac{1}{\beta}}_{\text{Value of } \tau'} \quad \underbrace{\frac{1}{2}}_{\text{Spin up or spin down}} . \tag{62}$$

To remove a vertex

$$T^0_{C_n \to C_{n-1}} = \frac{1}{n} . \tag{63}$$

With the above we are now in a position to compute the Metropolis acceptance ratio given in Eq. (132). The update of the hybridization matrix follows the same ideas as in the CT-INT.

The computationally expensive part of the CT-HYB algorithm is the evaluation of the trace $\langle \cdot \rangle_{\mathrm{loc}}$. In general, as the expectation value is taken with respect to the local Hamiltonian, the operators can be represented by matrices in a basis of the local Hilbert space.

One particular basis that is often used is the eigenbasis of the local Hamiltonian. This makes the time evolution of the operators trivial but leads to dense operator matrices, which have to be multiplied. Here the main drawback of the CT-HYB becomes apparent: the size of the local Hilbert space D, and thus of the operator matrices, grows exponentially with the number of local degrees of freedom. Thus, simulation of larger systems become unfeasible quite rapidly. Nevertheless, there are strategies to simulate multi-orbital systems as well as small clusters for Cluster-DMFT applications.

To reduce the size of the matrices that have to be multiplied, one can exploit certain symmetries of the Hamiltonian. This allows splitting the matrices into blocks of size $d_i \ll D$ [40], such that $\sum_i d_i = D$. Each block contains all the states associated with a certain value of a conserved quantity. The most obvious conserved quantities are the total f-particle number N_f and the z-component of the total spin S_z. There are well-defined rules governing how an operator connects the different blocks corresponding to different quantum numbers. As an example, the creation operator $\hat{f}^\dagger_\uparrow$ connects the blocks (N_f, S_z) and $(N_f + 1, S_z + 1/2)$.

In Ref. [41], another set of conserved quantities was identified, which leads to even smaller matrix blocks. For each orbital or site, the projection onto single occupation

$$PS_a = (n_{a,\uparrow} - n_{a,\downarrow})^2, \quad a = 1 \ldots M \tag{64}$$

Considered symmetries	None	N_f	N_f, S_z	N_f, S_z, PS
$N = 3$	$64 = 4^3$	20	9	3
$N = 5$	$1024 = 4^4$	252	100	10

Table 1: *Largest matrix block size, with different sets of conserved quantities taken into account.*

commutes with the local Hamiltonian. This makes the whole sequence of $PS = (PS_1, \ldots, PS_M)$ a good quantum number, so that the Hamiltonian can be written in a block-diagonal form with respect to PS. The resulting reduction of the matrix block sizes is shown in table 1.

Besides the matrix code described above, there are other possibilities for calculating the local trace. An algorithm based on a Krylov-subspace method was brought forward in Ref. [42]. Here the particle-number basis is used, so that applying operators to states is trivial. However, now the time evolution becomes more involved. Another recent proposal uses matrix product states for the propagation in time [43].

3.3 Selected applications

The CT-HYB is clearly the method of choice to tackle complex impurity problems at strong coupling and with the full Coulomb repulsion. In the context of the DMFT, extracting the local Green function is essential. In this context it is important to point out the work of Hafermann *et al.* [44], which describes an improved estimator for computing the self-energy.

In a two-orbital model, the inclusion of the Hund's coupling J greatly influences the critical Hubbard interaction U_c for the Mott transition [45]. A three-orbital model studied in Ref. [46] exhibits not only the Mott phase at integer fillings, but also a non-Fermi-liquid frozen-moment phase. In Ref. [47], models with only d-orbitals and with d- and additional oxygen p-orbitals where compared. With the p-orbitals included, the filling of the d-orbitals changes significantly, which thereby leads to a very different low-energy behavior.

Another recent application of the CT-HYB method in the context of DMFT is the study of models for topological Kondo insulators conducted in Ref. [48]. It was found that, starting from the non-interacting case, switching on the Hubbard interaction can drive the system through a series of transitions. In particular, a transition was observed between different topological states that are distinct due to the point-group symmetry of the lattice considered. Meanwhile, in Ref. [49] it was found that the edge spectrum of topological Kondo insulators is governed by the same scale as the bulk heavy-fermion state, namely the coherence scale T_{coh}. This makes it possible to infer information about the bulk coherence from the topological properties of the system.

4 Application of CT-INT to the Hubbard-Holstein model

The CT-INT allows for a very simple and efficient inclusion of phonon degrees of freedom. The path we follow here is to integrate out the phonons in favor of a retarded interaction and then solve the purely electronic model with the CT-INT approach. Starting from the Hubbard-Holstein model with Einstein phonons we show how to integrate out the phonons, describe some details of the algorithm, and then present results for the crossover from adiabatic to anti-adiabatic phonons in the one-dimensional Holstein model.

4.1 Integrating-out the phonons

The Hubbard-Holstein Hamiltonian we consider reads

$$\hat{H} = -\sum_{i,j,\sigma} t_{i,j}\, \hat{c}^{\dagger}_{i,\sigma}\hat{c}_{j,\sigma} + U\sum_{i} \left(\hat{n}_{i,\uparrow} - 1/2\right)\left(\hat{n}_{i,\downarrow} - 1/2\right) + g\sum_{i} \hat{Q}_i \left(\hat{n}_i - 1\right) + \sum_{i}\left(\frac{\hat{P}_i^2}{2M} + \frac{k}{2}\hat{Q}_i^2\right) \tag{65}$$

Here, $\hat{n}_i = \sum_\sigma \hat{n}_{i,\sigma}$ and the last two terms correspond respectively to the electron-phonon coupling, g, and the phonon energy. The Hamiltonian is written such that for a particle-hole symmetric band, half-filling corresponds to chemical potential $\mu = 0$. Opting for fermion coherent states

$$\hat{c}_{i,\sigma}|c\rangle = c_{i,\sigma}|c\rangle \ , \tag{66}$$

$c_{i,\sigma}$ being a Grassmann variable, and a real space representation for the phonon coordinates

$$\hat{Q}_i|q\rangle = q_i|q\rangle \ , \tag{67}$$

the path integral formulation of the partition function reads

$$Z = \int D\{\mathrm{q}\}\, D\left\{c^{\dagger}c\right\} e^{-(S_U + S_{\mathrm{ep}})} \ , \tag{68}$$

with

$$S_U = \int_0^\beta \mathrm{d}\tau \sum_{i,j,\sigma} c^{\dagger}_{i,\sigma}(\tau)\left(\delta_{i,j}\frac{\partial}{\partial\tau} - t_{i,j}\right) c_{j,\sigma}(\tau) + \ U\sum_i (n_{i,\uparrow}(\tau) - 1/2)(n_{i,\downarrow}(\tau) - 1/2)$$

$$S_{\mathrm{ep}} = \int_0^\beta \mathrm{d}\tau \sum_i \left(\frac{M\dot{q_i}^2(\tau)}{2} + \frac{k}{2}q_i^2(\tau) + \ g\, q_i(\tau)(n_i(\tau) - 1)\right). \tag{69}$$

In Fourier space,

$$q_j(\tau) = \frac{1}{\sqrt{\beta N}} \sum_{k,\Omega_m} e^{-i(\Omega_m\tau - kj)} q_{k,m} \ , \tag{70}$$

where Ω_m is a bosonic Matsubara frequency, the electron-phonon part of the action reads

$$S_{\mathrm{ep}} = \sum_{\Omega_m,k} \frac{M}{2}\left(\Omega_m^2 + \omega_0^2\right) q^{\dagger}_{k,m} q_{k,m} + g q_{k,m}\rho^{\dagger}_{k,m} \ ,$$

$$\rho^{\dagger}_{k,m} = \frac{1}{\sqrt{\beta N}} \int \mathrm{d}\tau \sum_j e^{-i(\Omega_m\tau - kj)}(n_j(\tau) - 1) \, . \tag{71}$$

Gaussian integration over the phonon degrees of freedom leads to a retarded density-density interaction

$$\int D\{\mathrm{q}\}\, e^{-S_{\mathrm{ep}}} = e^{\int_0^\beta \mathrm{d}\tau \int_0^\beta \mathrm{d}\tau' \sum_{i,j} [n_i(\tau)-1] D^0(i-j,\tau-\tau') [n_j(\tau')-1]} \,. \tag{72}$$

For Einstein phonons the phonon propagator is diagonal in real space,

$$\begin{aligned} D^0(i-j,\tau-\tau') &= \delta_{i,j} \frac{g^2}{2k} P(\tau-\tau') \;\; \text{with} \\ P(\tau) &= \frac{\omega_0}{2\left(1-e^{-\beta\omega_0}\right)} \left(e^{-|\tau|\omega_0} + e^{-(\beta-|\tau|)\omega_0}\right). \end{aligned} \tag{73}$$

Hence the partition function of the Hubbard-Holstein model takes the form

$$Z = \int D\left\{c^\dagger c\right\} e^{-(S_U - \int_0^\beta \mathrm{d}\tau \int_0^\beta \mathrm{d}\tau' \sum_{i,j} [n_i(\tau)-1] D^0(i-j,\tau-\tau') [n_j(\tau')-1])} \,.$$

In the anti-adiabatic limit, $\lim_{\omega_0\to\infty} P(\tau) = \delta(\tau)$ such that the phonon interaction maps onto an attractive Hubbard interaction of magnitude g^2/k. We are now in a position to apply the CT-INT algorithm by expanding in both the retarded and the Hubbard interactions.

4.2 Formulation of CT-INT for the Hubbard-Holstein model

To avoid the minus-sign problem at least for one-dimensional chains with nearest neighbor hopping matrix element t, we rewrite the phonon retarded interaction as

$$H_P(\tau) = -\frac{g^2}{4k} \int_0^\beta \mathrm{d}\tau' \sum_{i,\sigma,\sigma'} \sum_{s=\pm 1} P(\tau-\tau') \left[n_{i,\sigma}(\tau) - \alpha_+(s)\right] \left[n_{i,\sigma'}(\tau') - \alpha_+(s)\right]. \tag{74}$$

For each phonon vertex, we have introduced an Ising variable s. Summation over this Ising field reproduces, up to a constant, the original interaction. Since the phonon term is attractive, the adequate choice of signs is $\alpha_+(s) \equiv 1/2 + s\delta$, irrespective of the spin σ and σ'. Following Eq. (7), we rewrite the Hubbard term as

$$H_U(\tau) = \frac{U}{2} \sum_{i,s} \prod_\sigma \left(n_{i,\sigma}(\tau) - \alpha_\sigma(s)\right). \tag{75}$$

To proceed with a description of the implementation of the algorithm, it is useful to define a general vertex

$$V(\tau) = \{i, \tau, \sigma, \tau', \sigma', s, b\} \,, \tag{76}$$

where b defines the type of vertex at hand, Hubbard ($b = 0$) or phonon ($b = 1$). For this vertex, we define a sum over the available phase space

$$\sum_{V(\tau)} = \sum_{i,\sigma,\sigma',s,b} \int_0^\beta \mathrm{d}\tau', \tag{77}$$

a weight

$$w\left[V(\tau)\right] = \delta_{b,0}\frac{U}{2} - \delta_{b,1}P(\tau-\tau')\frac{g^2}{4k}\,, \tag{78}$$

as well as

$$\begin{aligned} H\left[V(\tau)\right] = \quad & \delta_{b,0}\delta_{\sigma,\uparrow}\delta_{\sigma',\downarrow}\delta(\tau-\tau')\left[n_{i,\uparrow}(\tau)-\alpha_+(s)\right]\left[n_{i,\downarrow}(\tau)-\alpha_-(s)\right] \\ & + \delta_{b,1}\left[n_{i,\sigma}(\tau)-\alpha_+(s)\right]\left[n_{i,\sigma'}(\tau')-\alpha_+(s)\right]. \end{aligned} \tag{79}$$

With the above definitions, the partition function can now be written as

$$\frac{Z}{Z_0} = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_0^\beta \mathrm{d}\tau_1 \sum_{V_1(\tau_1)} w[V_1(\tau_1)]\cdots\int_0^\beta \mathrm{d}\tau_n \sum_{V_n(\tau_n)} w\left[V_n(\tau_n)\right]\ \langle T\hat{H}\left[V_1(\tau_1)\right]\cdots\hat{H}\left[V_n(\tau_n)\right]\rangle_0. \tag{80}$$

As for the Hubbard model, a configuration consists of a set of vertices $C_n = \{V_1(\tau_1),\ldots,V_n(\tau_n)\}$. With the short-hand notation

$$\sum_{C_n} = \sum_n \int_0^\beta d\tau_1 \sum_{V(\tau_1)} \cdots \int_0^\beta d\tau_n \sum_{V(\tau_n)} \tag{81}$$

and

$$W(C_n) = \frac{(-1)^n}{n!}\, w[V_1(\tau_1)]\cdots w\left[V_n(\tau_n)\right]\ \langle T\hat{H}\left[V_1(\tau_1)\right]\cdots\hat{H}\left[V_n(\tau_n)\right]\rangle_0 \tag{82}$$

the partition function takes a form amenable to Monte Carlo sampling

$$\frac{Z}{Z_0} = \sum_{C_n} W(C_n). \tag{83}$$

The Monte Carlo sampling follows precisely the scheme presented in Sec. 2.4, namely the addition and removal of vertices. To be more specific, we consider the following form for vertex addition

$$T^0_{C_n\to C_{n+1}} = P_U\ \frac{1}{n+1}\ \frac{1}{L}\ \frac{1}{\beta}\ \frac{1}{2} + (1-P_U)\ \frac{1}{n+1}\ \frac{1}{L}\ \frac{1}{\beta}\ P(\Delta\tau)\ \left(\frac{1}{2}\right)^3. \tag{84}$$

The first term corresponds to the addition of a Hubbard vertex and has the same form as in Eq. (41). Note the additional factor $1/L$, which corresponds to the choice of the lattice site. This move is carried out with probability P_U. The second term corresponds to the addition of the phonon vertex and is carried out with probability $1-P_U$. The factor $P(\Delta\tau)$ allows a direct sampling of the phonon propagator. In particular, for a randomly chosen value of τ, $\tau' = \tau + \Delta\tau$. The factor $(1/2)^3$ accounts for the choice of the Ising field, as well as for the choice of the spin variables σ,σ' entering into the density-density two-body term. The vertex removal retains the same form as for the SIAM

$$T^0_{C_n\to C_{n-1}} = \frac{1}{n}\,. \tag{85}$$

4.3 The quarter-filled Holstein model from adiabatic to anti-adiabatic phonons

The concept of pre-formed fermion pairs or bosonic degrees of freedom that condense to form a superfluid can be found in many domains of correlated quantum many-body systems. Examples include the resonating valence bond theory of high-temperature superconductivity [50], Mott metal-insulator transitions in cold atoms [51], or transitions between charge-density-wave and superconducting states in the family of dichalcogenides [52]. As an interesting application of the CT-INT to the Holstein model, we show that these concepts can be carried over to one dimension where, in the absence of continuous symmetry breaking, the phase transition is replaced by a crossover.

In the anti-adiabatic limit $\omega_0 \to \infty$, $P(\tau)$ in equation (73) reduces to a Dirac δ-function, facilitating the above-mentioned mapping of the Holstein model onto the attractive Hubbard model, with $U = g^2/k$. The ratio of this binding energy and the bandwidth $W = 4t$ gives the dimensionless electron-phonon coupling

$$\lambda = \frac{g^2}{kW} \,. \tag{86}$$

In this section, we set $\lambda = 0.35$ and concentrate on the quarter-filled band with $k_F = \pi/4$.

Figure 1 shows equal-time correlation functions

$$C_O(q) = \sum_r \mathrm{e}^{\mathrm{i}qr} \left(\langle \hat{O}_r^\dagger \hat{O}_0 \rangle - \langle \hat{O}_r^\dagger \rangle \langle \hat{O}_0 \rangle \right) , \tag{87}$$

for charge, $\hat{O}_r = \hat{n}_{r,\uparrow} + \hat{n}_{r,\downarrow}$; spin, $\hat{O}_r = \hat{n}_{r,\uparrow} - \hat{n}_{r,\downarrow}$; pairing, $\hat{O}_r = \hat{c}^\dagger_{r,\uparrow}\hat{c}^\dagger_{r,\downarrow}$; and single-particle, $\hat{O}_r = \hat{c}_{r,\uparrow} + \hat{c}_{r,\downarrow}$, degrees of freedom at various phonon frequencies. In the adiabatic limit $\omega_0/t = 0$, any $\lambda > 0$ leads to an insulating state with $2k_F$ long-range charge order corresponding to the Peierls instability. We choose the coupling strength $\lambda = 0.35$ such that we have a metallic Luther-Emery liquid with dominant $2k_F$ charge correlations for low phonon frequencies and then study the evolution as a function of increasing ω_0/t. In particular, we have verified that for $\omega_0/t = 0.1$, the lowest nonzero frequency considered in the following, there is no long-range order; this can be seen from the finite-size dependence of the charge susceptibility [53].

The density (or charge) structure factor is plotted in Fig. 1(a). For classical phonons ($\omega_0 = 0$) the Peierls instability leads to long-range $2k_F$ charge order at zero temperature. As discussed above, quantum lattice fluctuations (occurring for $\omega_0 > 0$) can melt this order and lead to a state with dominant but power-law $2k_F$ charge correlations [8], as confirmed by the cusp at $2k_F = \pi/2$ in Fig. 1(a). The magnitude of the peak at $q = 2k_F$ initially decreases and then saturates upon increasing the phonon frequency, signalling competing ordering mechanisms as well as enhanced lattice fluctuations. The linear form of the charge structure factor at long wavelengths [see figure 1(a)] indicates a $1/r^2$ power-law decay of the real-space charge correlations and hence a metallic state.

In Fig. 1(b), we present the pair correlation function in the onsite s-wave channel. In contrast to the density correlator, which picks up diagonal order, $P(r)$ detects off-diagonal order charac-

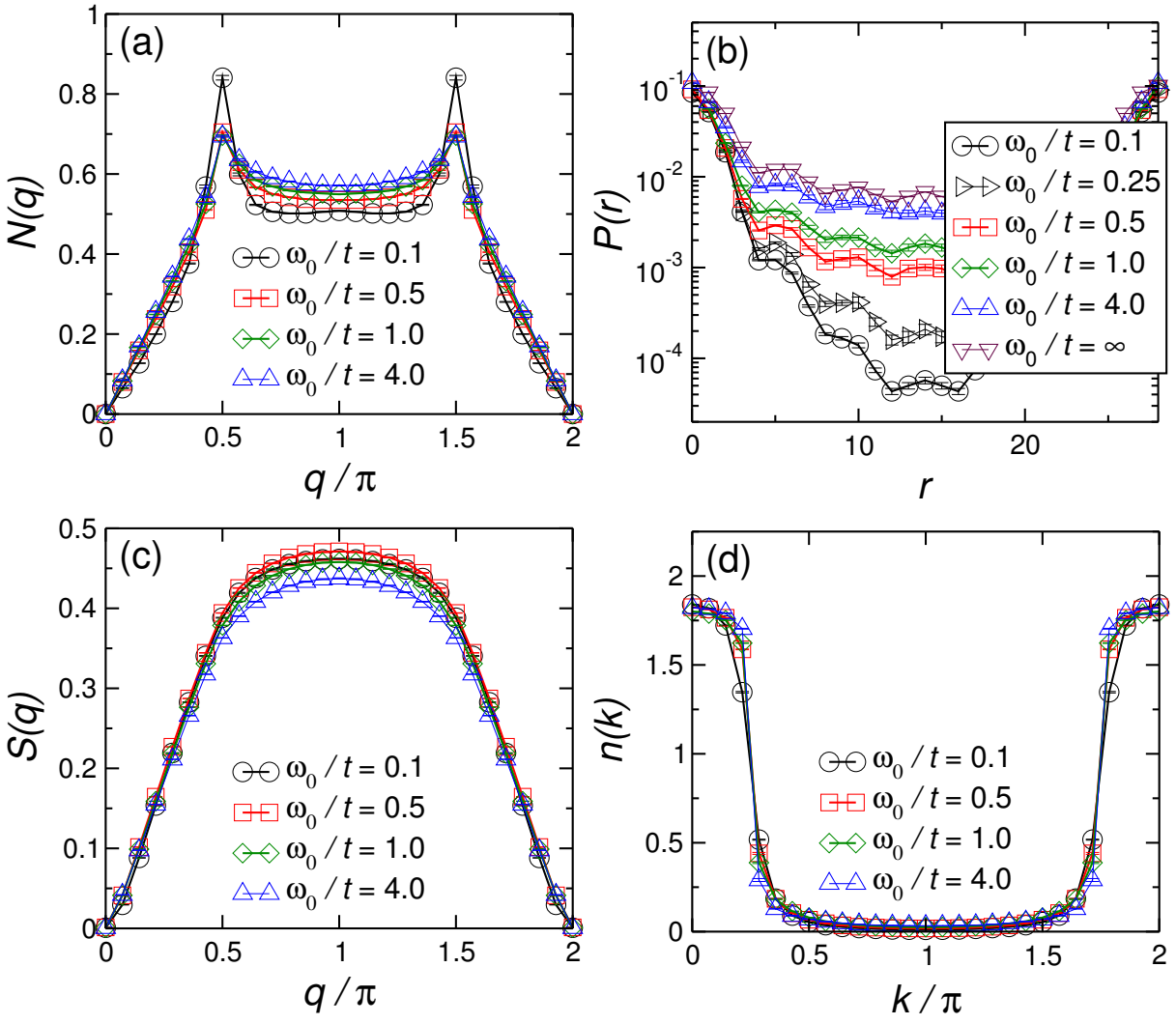

Fig. 1: *Static correlation functions for different values of the phonon frequency* ω_0/t *at* $\lambda = 0.35$ *for a quarter-filled band* ($n = 0.5$). *The panels show (a) the charge structure factor, (b) the pairing correlator, (c) the spin structure factor, and (d) the momentum distribution function. Here* $L = 28$ *and* $\beta t = 40$.

teristic of a superconducting state. In the Peierls state obtained for classical phonons, diagonal long-range charge order leads to an exponential decay of pairing correlations at long distances. The fluctuations resulting from a finite phonon frequency close the charge gap and render the pairing correlations critical. Comparing figures 1(a) and 1(b), we see that the suppression of the $2k_F$ charge correlations is accompanied by an increase of the pairing correlations, especially at large distances. A possible interpretation is that with increasing phonon frequency, the trapping of bipolarons in the $2k_F$ lattice modulation gives way to a "*condensation*" (in the usual sense of superfluidity in one dimension) of those preformed pairs. A detailed study of the dynamics of the observed crossover can be found in Ref. [54].

5 Concluding remarks

The aim of this lecture was to give a detailed theoretical overview of the CT-INT and to comment on the general formulation of the CT-HYB. For the SIAM, the CT-QMC methods are easy to implement and the interested reader is encouraged to try. It is essential to point out once again that the CT-QMC methods are action based such that the pool of potential applications is extremely large.

Acknowledgments

I would like to thank T. Lang, D. Luitz, M. Aulbach, F. Goth, M. Hohenadler, J. Werner, M. Weber and J. Hofmann who have implemented and generalized CT-QMC methods thus triggering many questions and discussions. Support of the Deutsche Forschungsgemeinschaft through FOR 1346 is gratefully acknowledged.

A Basic principles of Monte Carlo methods in statistical mechanics

In this section, we show how to use the Monte Carlo method to compute integrals of the form

$$\langle O \rangle_P = \int_\Omega \mathrm{d}x \; P(x)\, O(x)\,. \tag{88}$$

Ω is a discrete or continuous configuration space with elements x, $P(x)$ is a probability distribution on this space,

$$\int_\Omega \mathrm{d}x P(x) = 1 \;\text{ and }\; P(x) \geq 0 \;\; \forall\, x \in \Omega, \tag{89}$$

and O is an observable or random variable.
To illustrate why stochastic methods are useful, let us assume Ω to be a subspace of $\mathbb{R}^d$ with volume $V = L^d$. One can break up Ω into hypercubes of linear dimension h and approximate the integral by a Riemann sum. The required number of function evaluations N then scales as $V/h^d = (L/h)^d = e^{d \ln(L/h)}$. Hence, the numerical effort – which is nothing but the number of function evaluations – grows exponentially with the dimension d. In contrast, stochasitc methods provide an estimate of the integral with statistical uncertainty scaling as the inverse square root of the number of function evaluations, irrespective of the dimensionality d. Hence, in the large-d limit, stochastic methods become attractive. This result stems from the central limit theorem.

A.1 The central limit theorem

Before proceeding in illustrating and proving the central limit theorem, let us introduce some notation. We will denote by $P_O(\mathcal{O})$ the probability that the observable O takes the value $\mathcal{O}$. Hence in terms of integrals over the configuration space,

$$P_O(\mathcal{O}) = \int_\Omega \mathrm{d}x \; P(x)\, \delta(O(x) - \mathcal{O}) \tag{90}$$

such that

$$\langle O \rangle_p = \int \mathrm{d}\mathcal{O} \; P_O(\mathcal{O})\, \mathcal{O}, \;\text{ and }\; \langle O^2 \rangle_p = \int \mathrm{d}\mathcal{O} \; P_O(\mathcal{O})\, \mathcal{O}^2. \tag{91}$$

Suppose that we have a set of configurations $\{x_i \,|\, i \in 1 \ldots N\}$ distributed according to the probability distribution $P(x)$, then we can approximate $\langle O \rangle_P$ by

$$\langle O \rangle_P \;\sim\; \frac{1}{N} \sum_{i=1}^{N} \underbrace{O(x_i)}_{=\mathcal{O}_i} = X. \tag{92}$$

Clearly, X will depend upon the chosen set $\{x_i \,|\, i \in 1 \ldots N\}$. Hence the relevant question is the distribution of X, $\mathcal{P}(X)$, for a given value of N. The central limit theorem tells us that

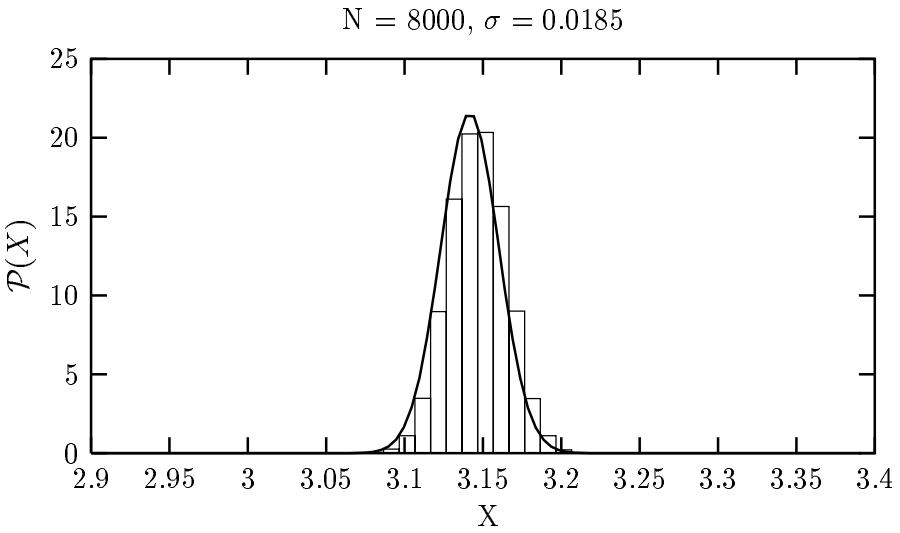

Fig. 2: *Boxes correspond to the distribution results obtained after 10000 simulations. For each simulation we draw $N = 8000$ points. For a single simulation, we obtain $\sigma = 0.0185$. The heavy line corresponds to the result of central limit theorem with above value of σ.*

provided that the $\mathcal{O}_1 \cdots \mathcal{O}_N$ are statistically independent and that N is large enough $\mathcal{P}(X)$ reads

$$\begin{aligned} \mathcal{P}(X) &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left[-\frac{(X - \langle O \rangle_P)^2}{2\sigma^2} \right] \\ &\text{with } \sigma^2 = \frac{1}{N} \left(\langle O^2 \rangle_P - \langle O \rangle_P^2 \right). \end{aligned} \tag{93}$$

Thus irrespective of the dimension d, the convergence to the exact result scales as $1/\sqrt{N}$, and the width of the above normal distribution, σ, corresponds to the statistical error. For practical purposes, one estimates σ by

$$\sigma^2 \approx \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^{N} O(x_i)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} O(x_i) \right)^2 \right). \tag{94}$$

More general methods for estimating the error are discussed in section A.2.
Before demonstrating the central limit theorem we give a simple example, the evaluation of the number π obtained with

$$\pi = 4 \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \, \Theta(1 - x^2 - y^2). \tag{95}$$

Here Θ is the Heaviside function, $\Theta(x) = 1$ for $x > 0$ and vanishes otherwise. In this example we have $P(x, y) \equiv 1$. To generate a sequence of N points $(x, y)_i$ from this probability distribution, we draw random numbers, x, y, in the interval $[0, 1]$. For $N = 8000$ we obtain an error $\sigma = 0.0185$ with the use of Eq. (94). To check the central limit theorem, we repeat the simulation 10000 times with different random numbers. Fig. (2) shows the thus obtained distribution which compares well to the result of the central limit theorem.

We now demonstrate the central limit theorem.
Proof: The probability of obtaining the result X reads

$$\mathcal{P}(X) = \int \mathrm{d}\mathcal{O}_1 \cdots \mathrm{d}\mathcal{O}_N \; P_O(\mathcal{O}_1, \mathcal{O}_2, \cdots, \mathcal{O}_N) \; \delta\left(X - \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}_i \right). \tag{96}$$

The assumption that the $\mathcal{O}_i$ are statistically independent means that the combined probability factorizes

$$P_O(\mathcal{O}_1, \mathcal{O}_2, \cdots, \mathcal{O}_N) = P_O(\mathcal{O}_1) P_O(\mathcal{O}_2) \cdots P_O(\mathcal{O}_N).$$

Furthermore, using the representation $\delta(x) = \frac{1}{2\pi}\int \mathrm{d}\lambda \, e^{i\lambda x}$ of the Dirac-δ function, we can reduce the above expression of $\mathcal{P}(X)$ to

$$\begin{aligned}
\mathcal{P}(X) &= \frac{1}{2\pi}\int \mathrm{d}\lambda \; \mathrm{d}\mathcal{O}_1 \cdots \mathrm{d}\mathcal{O}_N \; P_O(\mathcal{O}_1) \cdots P_O(\mathcal{O}_N) \, e^{i\lambda\left(X - \frac{1}{N}\sum_i \mathcal{O}_i\right)} \\
&= \frac{1}{2\pi}\int \mathrm{d}\lambda \; e^{i\lambda X} \left(\int \mathrm{d}\mathcal{O} \, P_O(\mathcal{O}) \, e^{-i\frac{\lambda}{N}\mathcal{O}} \right)^N \\
&= \frac{N}{2\pi}\int \mathrm{d}\lambda \; e^{-NS(\lambda, X)} \;\; \text{with} \;\; S(\lambda, X) = -i\lambda X - \ln \int \mathrm{d}\mathcal{O} \, P_O(\mathcal{O}) \, e^{-i\lambda\mathcal{O}}
\end{aligned} \tag{97}$$

As $N \to \infty$ the saddle point approximation becomes exact

$$\begin{aligned}
\mathcal{P}(X) &\simeq \frac{N}{2\pi}\int \mathrm{d}\lambda \; e^{-N\left(S(\lambda^*(X), X) + \frac{(\lambda^*(X)-\lambda)^2}{2} \frac{\partial^2}{\partial\lambda^2} S(\lambda, X)\big|_{\lambda=\lambda^*(X)} \right)} \\
&= \sqrt{\frac{N}{2\pi \, \frac{\partial^2}{\partial\lambda^2} S(\lambda, X)\big|_{\lambda=\lambda^*(X)}}} \; e^{-NS(\lambda^*(X), X)} \;\; \text{with} \;\; \left.\frac{\partial}{\partial\lambda} S(\lambda, X)\right|_{\lambda=\lambda^*(X)} = 0
\end{aligned} \tag{98}$$

such that $\lambda^*(X)$ is determined by:

$$X = \frac{\int \mathrm{d}\mathcal{O} \; P_O(\mathcal{O}) \, \mathcal{O} \, e^{-i\lambda^*(X)\mathcal{O}}}{\int \mathrm{d}\mathcal{O} \; P_O(\mathcal{O}) \, e^{-i\lambda^*(X)\mathcal{O}}}. \tag{99}$$

To proceed, we again use that for large values of N we can expand around the saddle point

$$\begin{aligned}
\left.\frac{d}{dX} S(\lambda^*(X), X)\right|_{X=X^*} &= \left[\underbrace{\left.\frac{\partial}{\partial\lambda} S(\lambda, X)\right|_{\lambda=\lambda^*(X)}}_{=0} \frac{d}{dX}\lambda^*(X) + \frac{\partial}{\partial X} S(\lambda^*(X), X) \right]_{X=X^*} \\
&= -i\lambda^*(X^*) = 0.
\end{aligned} \tag{100}$$

Hence with Eq. (99) we have $X^* = \langle O \rangle_P$. Expanding around X^* (note that $S(\lambda^*(X^*), X^*) = 0$) yields

$$\mathcal{P}(X) = \sqrt{\frac{N}{2\pi \, \frac{\partial^2}{\partial\lambda^2} S(\lambda, X^*)\big|_{\lambda=\lambda^*(X^*)}}} \; e^{-N\frac{(X - \langle O\rangle_P)^2}{2} \frac{d^2}{dX^2} S(\lambda^*(X), X)\big|_{X=X^*}}. \tag{101}$$

Using Eqn. (100) and (99) we obtain

$$\left.\frac{d^2}{dX^2}S(\lambda^*(X),X)\right|_{X=X^*} = -i\left.\frac{d\lambda^*(X)}{dX}\right|_{X=X^*} = \frac{1}{\langle O^2\rangle_P - \langle O\rangle_P^2}$$

$$\left.\frac{\partial^2}{\partial\lambda^2}S(\lambda,X^*)\right|_{\lambda=\lambda^*(X^*)} = \langle O^2\rangle_P - \langle O\rangle_P^2 \tag{102}$$

such that

$$\mathcal{P}(X) \simeq \sqrt{\frac{N}{2\pi\left(\langle O^2\rangle_P - \langle O\rangle_P^2\right)}}\, e^{-N\frac{(X-\langle O\rangle_P)^2}{2}\frac{1}{\langle O^2\rangle_P - \langle O\rangle_P^2}}. \tag{103}$$

This completes the demonstration of the central limit theorem. The two important conditions for the validity of the theorem are that (i) the $\mathcal{O}_1\cdots\mathcal{O}_N$ are statistically independent and that (ii) N is large.

A.2 Jackknife and bootstrap methods for error evaluation

The jackknife and bootstrap methods [55] provide alternative ways of estimating the error (94). These methods become particularly useful, if not essential, when one wishes to estimate the error on $f(\langle O_1\rangle,\cdots,\langle O_n\rangle)$, where f is an arbitrary function of n variables. For a given sequence of configurations $\{x_1\cdots x_N\}$ drawn from the probability distribution $P(x)$, the jackknife focuses on the samples that leave out one configuration at a time

$$f_i^J = f\left(\frac{1}{N-1}\sum_{j\neq i}O_1(x_j),\cdots,\frac{1}{N-1}\sum_{j\neq i}O_n(x_j)\right). \tag{104}$$

The error estimate on f is then given by

$$\left(\sigma_f^J\right)^2 \approx N\left(\frac{1}{N}\sum_{i=1}^{N}(f_i^J)^2 - \left(\frac{1}{N}\sum_{i=1}^{N}f_i^J\right)^2\right). \tag{105}$$

One may verify explicitly that for $n=1$ and $f(x)=x$ Eq. (105) reduces to Eq. (94) up to a factor $(N/(N-1))^2$, which tends to unity in the large N limit.

An alternative method for determining errors of f is the bootstrap algorithm. For a given sample of N configurations $\{x_1\cdots x_N\}$ drawn from the probability distribution $P(x)$, we can construct N^N sets of N configurations, $\{x_{i_1}\cdots x_{i_N}\}$ with $i_1\in 1\cdots N$, $i_2\in 1\cdots N$, $\cdots$, $i_N\in 1\cdots N$, which correspond to the ideal bootstrap samples. For a given bootstrap sample, defined by the vector $\mathbf{i}=(i_1,\cdots,i_N)$,

$$f_{\mathbf{i}}^B = f\left(\frac{1}{N}\sum_{k=1}^{N}O_1(x_{i_k}),\cdots,\frac{1}{N}\sum_{k=1}^{N}O_n(x_{i_k})\right). \tag{106}$$

The bootstrap estimate of the error is given by

$$\left(\sigma_f^B\right)^2 \approx \frac{1}{N^N}\sum_{\mathbf{i}}^{N^N}(f_{\mathbf{i}}^B)^2 - \left(\frac{1}{N^N}\sum_{\mathbf{i}}^{N^N}f_{\mathbf{i}}^B\right)^2. \tag{107}$$

Again, one may check that for the special case $n = 1$ and $f(x) = x$, Eq. (107) reduces to Eq. (94). Clearly, when N is large, it is numerically out of reach to generate all of the N^N bootstrap samples. Typically, to estimate the right-hand side of Eq. (107), 200 or more bootstrap samples are generated stochastically. Since each bootstrap sample is equally probable we can generate them with: $i_k = \texttt{trunc}(N * \xi_k + 1)$ where ξ_k is a random number in the interval $[0, 1)$ and the function `trunc` returns an integer by truncating the numbers after the decimal point.

A.3 Markov chains

Our task is now to generate a set of states distributed according to a given probability distribution. Here we will consider a discrete space Ω with N_s states; x runs over all the states $1 \cdots N_s$, and $P(x)$ denotes the probability of the occurrence of the state x. We introduce a Monte-Carlo time t and a time dependent probability distribution $P_t(x)$ that evolves according to a Markov process: the future $(t + 1)$ depends only on the present (t). To define the Markov process, we introduce a matrix $T_{y,x}$, which corresponds to the transition probability from state x to state y. The time evolution of $P_t(x)$ is given by

$$P_{t+1}(y) = \sum_x T_{y,x} P_t(x). \tag{108}$$

T has to satisfy the properties

$$\sum_y T_{y,x} = 1 \;\text{ and }\; T_{y,x} \geq 0\,. \tag{109}$$

Hence, if $P_t(x)$ is a probability distribution then $P_{t+1}(x)$ is also a probability distribution.
T has to be ergodic

$$\forall x,\, y \in \Omega \;\; \exists\, s \,|\; \left(T^s\right)_{y,x} > 0\,. \tag{110}$$

Thus, we are assured to sample the whole phase space provided the above is satisfied. Last, we have the requirement of stationarity

$$\sum_x T_{y,x}\, P(x) = P(y)\,. \tag{111}$$

Once we have reached the desired distribution $P(x)$ we wish to stay there. Stationarity is automatically satisfied if

$$T_{y,x}\, P(x) = T_{x,y}\, P(y) \tag{112}$$

as may be seen by summing on both sides over x. This relation is referred to as *detailed balance* or microreversibility. However, one has to keep in mind that only stationarity and not detailed balance is essential.
Given the above, in the Monte Carlo simulation we will generate the Markov Chain

$$x_1, x_2, \ldots, x_n\ ,$$

where the conditional probability of sampling the state x_{t+1} given the state x_t reads

$$P(x_{t+1}|\, x_t) = T_{x_{t+1},x_t}\,. \tag{113}$$

Our aim is now to show that when $n \to \infty$ the fraction of the time one can expect the Markov process to be in state x is $P(x)$, independent of the initial state x_1. In other words as $n \to \infty$ the set of states $x_1 \cdots x_n$ are distributed according to $P(x)$.
At step t in the Markov chain, the state x will occur on average with probability, $[T^t]_{x,x_1}$. Hence, we have to show that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \left[T^t\right]_{x,x_1} = P(x)\,. \tag{114}$$

We will first show the above under the assumption that T is regular. That is, there is an integer N such that T^N has only positive, non-zero, entries. If T is regular then T is ergodic. However the inverse is not true so that the condition of regularity is more stringent than that of ergodicity. After the demonstration, we will argue on the basis of a simple example that Eq. (114) is equally valid for ergodic but not regular transition matrices.
We now demonstrate Eq. (114) for regular transition matrices.
Proof: We introduce the set of vectors of real numbers, $\mathbf{a}_t$

$$\mathbf{a}_{t+1} = \mathbf{a}_t T \;\text{ and }\; d_t = \max(\mathbf{a}_t) - \min(\mathbf{a}_t) \tag{115}$$

where $\max(\mathbf{a}_t)$ corresponds to the largest element of the vector $\mathbf{a}$. Since

$$\begin{aligned} [\mathbf{a}_{t+1}]_x &= \sum_y [\mathbf{a}_t]_y\, T_{y,x} \le \max(\mathbf{a}_t) \sum_y T_{y,x} = \max(\mathbf{a}_t) \\ \text{and} \qquad [\mathbf{a}_{t+1}]_x &= \sum_y [\mathbf{a}_t]_y\, T_{y,x} \ge \min(\mathbf{a}_t) \sum_y T_{y,x} = \min(\mathbf{a}_t) \end{aligned} \tag{116}$$

the sequence d_t satisfies

$$d_{t+1} \le d_t. \tag{117}$$

We now show that there is a subsequence that is strictly decreasing. Let us consider vector $\mathbf{a}_t$ with $\max(\mathbf{a}_t) = [\mathbf{a}_t]_M = a_M$ and $\min(\mathbf{a}_t) = [\mathbf{a}_t]_m = a_m$. We define the vector $\mathbf{a}^M$ ($\mathbf{a}^m$) by replacing all the elements of $\mathbf{a}_t$ apart from the minimal (maximal) one by a_M (a_m). Hence $\mathbf{a}^m \le \mathbf{a}_t \le \mathbf{a}^M$ where the inequalities hold element-wise. Furthermore let ϵ be the minimal entry in T^N. In light of the assumption that T is regular, $\epsilon > 0$. With those definitions, we have

$$\begin{aligned} [\mathbf{a}_{t+N}]_x &= \left[\mathbf{a}_t T^N\right]_x \le \left[\mathbf{a}^M T^N\right]_x = a_M \sum_{y\neq m} T^N_{y,x} + a_m T^N_{m,x} \\ &= a_M \left(1 - T^N_{m,x}\right) + a_m T^N_{m,x} = a_M - (a_M - a_m) T^N_{m,x} \\ &\le a_M - (a_M - a_m)\epsilon \\ \text{and} \quad [\mathbf{a}_{t+N}]_x &\ge \left[\mathbf{a}^m_t T^N\right]_x = a_m + (a_M - a_m) T^N_{M,x} \;\ge\; a_m + (a_M - a_m)\epsilon. \end{aligned} \tag{118}$$

With the above inequalities,

$$d_{t+N} = \max(\mathbf{a}_{t+N}) - \min(\mathbf{a}_{t+N}) \leq a_M - (a_M - a_m)\epsilon - a_m - (a_M - a_m)\epsilon = d_t(1-2\epsilon) \quad (119)$$

Hence, $d_{t+mN} \leq d_t(1-2\epsilon)^m$ such that the series d_{t+mN} decreases at least exponentially with increasing m. Recalling that $d_{t+1} \leq d_t$, we can find positive numbers τ and b such that

$$d_t \leq be^{-t/\tau}. \quad (120)$$

In particular, we can set $\mathbf{a}_1 = \mathbf{e}_i$ such that $\mathbf{e}_i T$ corresponds to the i^{th} row of the matrix T. Hence, since d_t tends towards zero in the limit $t \to \infty$, the i^{th} row of matrix T^t tends towards a constant when $t \to \infty$. Hence, we have

$$\lim_{t\to\infty} T^t = \begin{pmatrix} \alpha_1 & \cdots & \alpha_1 \\ \alpha_2 & \cdots & \alpha_2 \\ . & . & . \\ . & . & . \\ . & . & . \\ \alpha_{N_s} & \cdots & \alpha_{N_s} \end{pmatrix} \quad (121)$$

where N_s denotes the total number of states. Many comments are now in order.

(i) Since T satisfies Eq. (109), so does T^t for all values of t. Thus $\sum_{x=1}^{N_s} \alpha_x = 1$ and $\alpha_x > 0$. In other words $\boldsymbol{\alpha}$ is a probability vector.

(ii) Due to Eq. (120), $T^t_{x,y} = \alpha_x + \Delta^{(t)}_{x,y}$ with $|\Delta^{(t)}_{x,y}| \leq be^{-t/\tau}$.

(iii) For any probability vector $\mathbf{v}$, $\lim_{t\to\infty} T^t\mathbf{v} = \boldsymbol{\alpha}$.
Hence, there is a unique asymptotic distribution: $\boldsymbol{\alpha}$.

(iv) Due to the stationarity condition, we have $T^t\mathbf{P} = \mathbf{P}$ for all values of t and hence also for $t \to \infty$. Since there is a unique asymptotic distribution, $\boldsymbol{\alpha} = \mathbf{P}$.

The validity of Eq. (114) now follows from

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \left[T^t\right]_{x,x_1} = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \left(P(x) + \Delta^{(t)}_{x,x_1}\right) = P(x) + \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \Delta^{(t)}_{x,x_1}. \quad (122)$$

However since

$$\left|\sum_{t=1}^{n} \Delta^{(t)}_{x,x_1}\right| \leq \sum_{t=1}^{n} \left|\Delta^{(t)}_{x,x_1}\right| \leq b\sum_{t=1}^{n} e^{-t/\tau} \quad (123)$$

and $b\sum_{t=1}^{n} e^{-t/\tau}$ is a convergent series as $n \to \infty$, Eq. (114) is satisfied. **QED**

We now show on the basis of a simple example that if T is ergodic but not regular, Eq. (114) is still valid. Consider $N_s = 2$ and

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad T^{2n+1} = T. \quad (124)$$

T is ergodic but not regular. It is cyclic and in contrast to the regular transition matrices the $\lim_{t\to\infty} T^t$ does not exist. However, since

$$TP = P \;\text{ with }\; P = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \tag{125}$$

and for even and odd values of n

$$\frac{1}{2n}\sum_{t=1}^{2n} T^t = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad \frac{1}{2n+1}\sum_{t=1}^{2n} T^t = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} + \frac{1}{2n+1}T, \tag{126}$$

Eq. (114) holds. For further reading and a more precise and mathematical oriented discussion of Markov chains, the reader is referred to [56].

A.4 Construction of the transition matrix T

Having defined T, we now have to construct it explicitly. Let $T^0_{y,x}$ be the probability of proposing a move from x to y and $a_{y,x}$ the probability of accepting it. $1 - a_{y,x}$ corresponds to the probability of rejecting the move. T_0 is required to satisfy Eq. (109). Since in general we want to propose moves that change the initial configuration, $T^0_{x,x} = 0$. With $a_{y,x}$ and $T^0_{y,x}$ we build $T_{y,x}$ with

$$T_{y,x} = \begin{cases} T^0_{y,x} a_{y,x} & \text{if } \; y \neq x \\ \sum_{\substack{z \\ z\neq x}} T^0_{z,x}\,(1 - a_{z,x}) & \text{if } \; y = x \end{cases} \tag{127}$$

Clearly $T_{y,x}$ satisfies Eq. (109). To satisfy the stationarity, we impose the detailed balance condition to obtain the equality

$$T^0_{y,x}\, a_{y,x}\, P_x = T^0_{x,y}\, a_{x,y}\, P_y\,. \tag{128}$$

Let us set

$$a_{y,x} = \mathcal{F}\left(\frac{T^0_{x,y} P_y}{T^0_{y,x} P_x}\right) \tag{129}$$

with $\mathcal{F} :]0:\infty[\to [0,1]$. Since

$$a_{x,y} = \mathcal{F}\left(\frac{T^0_{y,x} P_x}{T^0_{x,y} P_y}\right) = \mathcal{F}\left(\frac{1}{\frac{T^0_{x,y} P_y}{T^0_{y,x} P_x}}\right), \tag{130}$$

the detailed balance condition reduces to

$$\frac{\mathcal{F}(Z)}{\mathcal{F}(1/Z)} = Z \;\text{ where }\; Z = \frac{T^0_{x,y} P_y}{T^0_{y,x} P_x}. \tag{131}$$

There are many possible choices. The *Metropolis algorithm* is based on the choice

$$\mathcal{F}(Z) = \min(Z, 1)\,. \tag{132}$$

Thus, one proposes a move from $\mathbf{x}$ to $\mathbf{y}$ and accepts it with probability $Z = \frac{T^0_{\mathbf{x},\mathbf{y}} P_{\mathbf{y}}}{T^0_{\mathbf{y},\mathbf{x}} P_{\mathbf{x}}}$. In the practical implementation, one picks a random number r in the interval $[0:1]$. If $r < Z$ ($r > Z$) one accepts (rejects) the move. Alternative choices of $\mathcal{F}(Z)$ are for example

$$\mathcal{F}(Z) = \frac{Z}{1+Z} \tag{133}$$

which is referred to as the heat-bath method.
That the so-constructed T matrix is ergodic depends upon the choice of T^0. In many cases, one will wish to combine different types of moves to achieve ergodicity. For a specific move i, we construct $T^{(i)}$ as shown above so that $T^{(i)}$ satisfies conditions (109) and (112). The moves may be combined in two ways

$$T = \sum_i \lambda_i T^{(i)}, \quad \sum_i \lambda_i = 1 \tag{134}$$

which is referred to as random updating since one picks with probability λ_i the move $T^{(i)}$. Clearly, T equally satisfies (109) and (112), and the moves have to be chosen appropriately to satisfy the ergodicity condition. Another choice is sequential upgrading. A deterministic ordering of the moves is chosen to obtain

$$T = \prod_i T^{(i)}. \tag{135}$$

This choice does not satisfy detailed balance but does satisfy stationarity (111) as well as (109). Again ergodicity has to be *checked* on a case to case basis.
The observable O may now be estimated with

$$\langle O \rangle_P \approx \frac{1}{N} \sum_{t=1}^{N} O(\mathbf{x}_t). \tag{136}$$

The required value of N depends on the autocorrelation time of the observable O

$$C_O(t) = \frac{\frac{1}{N}\sum_{s=1}^{N} O(\mathbf{x}_s) O(\mathbf{x}_{s+t}) - \left(\frac{1}{N}\sum_{s=1}^{N} O(\mathbf{x}_s)\right)^2}{\frac{1}{N}\sum_{s=1}^{N} O(\mathbf{x}_s)^2 - \left(\frac{1}{N}\sum_{s=1}^{N} O(\mathbf{x}_s)\right)^2}. \tag{137}$$

One expects $C_O(t) \sim e^{-t/\tau_O}$, where τ_O corresponds to the MC time scale on which memory of the initial configuration is lost. Hence, to obtain meaningful results, $N \gg \tau_O$. Note that one should also take into account a *warm up* time by discarding at least the first τ_O configurations in the MC sequence. Naively, one would expect $\tau_O = \tau$. However, this depends on the overlap of the observable with the slowest mode in the MC dynamics, which relaxes as $e^{-t/\tau}$. In particular in a model with spin rotation symmetry the slowest mode may correspond to the rotation of the total spin. In this case, observables that are invariant under a spin rotation will not be affected by the slowest mode of the MC dynamics. Hence, in this case $\tau_O < \tau$.
We now consider the error estimation. To apply the central limit theorem, we need a set of independent estimates of $\langle O \rangle_P$. This may be done by regrouping the data into *bins* of size $n\tau_O$.

$$\tilde{O}_n(t) = \frac{1}{n\tau_O} \sum_{s=1}^{n\tau_O} O(\mathbf{x}_{(t-1)n\tau_O+s}) \tag{138}$$

with $t = 1 \cdots N/(n\tau_O)$. If n is large enough (i.e. $n \approx 10 - 20$) then $\tilde{O}_n(t)$ may be considered as an independent estimate, and the error is given by

$$\sigma_n = \sqrt{\frac{1}{M}\left(\frac{1}{M}\sum_{t=1}^{M}\tilde{O}_n(t)^2 - \left(\frac{1}{M}\sum_{i=1}^{M}\tilde{O}_n(t)\right)^2\right)} \tag{139}$$

where $M = N/(n\tau_O)$. If n is large enough, the error σ_n should be independent of n.

A.5 One-dimensional Ising model

We conclude this appendix with an example of a Monte Carlo simulation with error analysis for the one-dimensional Ising model

$$H(\{\sigma\}) = -J\sum_{i=1}^{L}\sigma_i\sigma_{i+1} \quad , \quad \sigma_{L+1} = \sigma_1 \tag{140}$$

where $\sigma_i = \pm 1$. This model is easily solved exactly with the transfer matrix method and thus produces a useful testing ground for the MC approach. In particular at zero temperature, a phase transition to a ferromagnetically ordered phase ($J > 0$) occurs [57]. Spin-spin correlations are given by

$$g(r) = \frac{1}{L}\sum_{i=1}^{L}\langle\sigma_i\sigma_{i+r}\rangle \;\text{ with }\; \langle\sigma_i\sigma_{i+r}\rangle = \frac{\sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})}\sigma_i\sigma_{i+r}}{\sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})}} \tag{141}$$

where β corresponds to the inverse temperature.
We now construct the transition matrix T corresponding to a random single site updating scheme. For an L-site chain, we denote the spin configuration by

$$x = (\sigma_{x,1}, \sigma_{x,2}, \cdots, \sigma_{x,L}). \tag{142}$$

The transition matrix reads

$$T = \frac{1}{L}\sum_{i=1}^{L}T^{(i)} \;\text{ with }\; T^{(i)}_{x,y} = T^{(i),0}_{x,y}a_{x,y} \;\text{ and}$$

$$T^{(i),0}_{x,y} = \begin{cases} 1 & \text{if } x = (\sigma_{y,1}, \cdots, -\sigma_{y,i}, \cdots \sigma_{y,L}) \\ 0 & \text{otherwise} \end{cases} \tag{143}$$

The above corresponds to choosing a site i randomly, changing the orientation of the spin ($\sigma_i \to -\sigma_i$), and accepting the move with with probability $a_{x,y}$ corresponding to a heat-bath or Metropolis algorithm. The MC time unit corresponds to a single sweep, meaning that L sites are randomly chosen before a measurement is carried out.
Fig. 3 plots the autocorrelation time for $g(r = L/2)$ on an $L = 24$ site lattice at $\beta J = 1$ and $\beta J = 2$. From Fig. 3a one can extract the autocorrelation time: $\tau_O \approx 11, 54$ for $\beta J = 1, 2$ respectively. Fig. 3b plots the error as a function of bin size in units of the τ_O (see Eq. (139)). As one can see, $n \approx 10$ is sufficient to get a reliable estimate of the error.

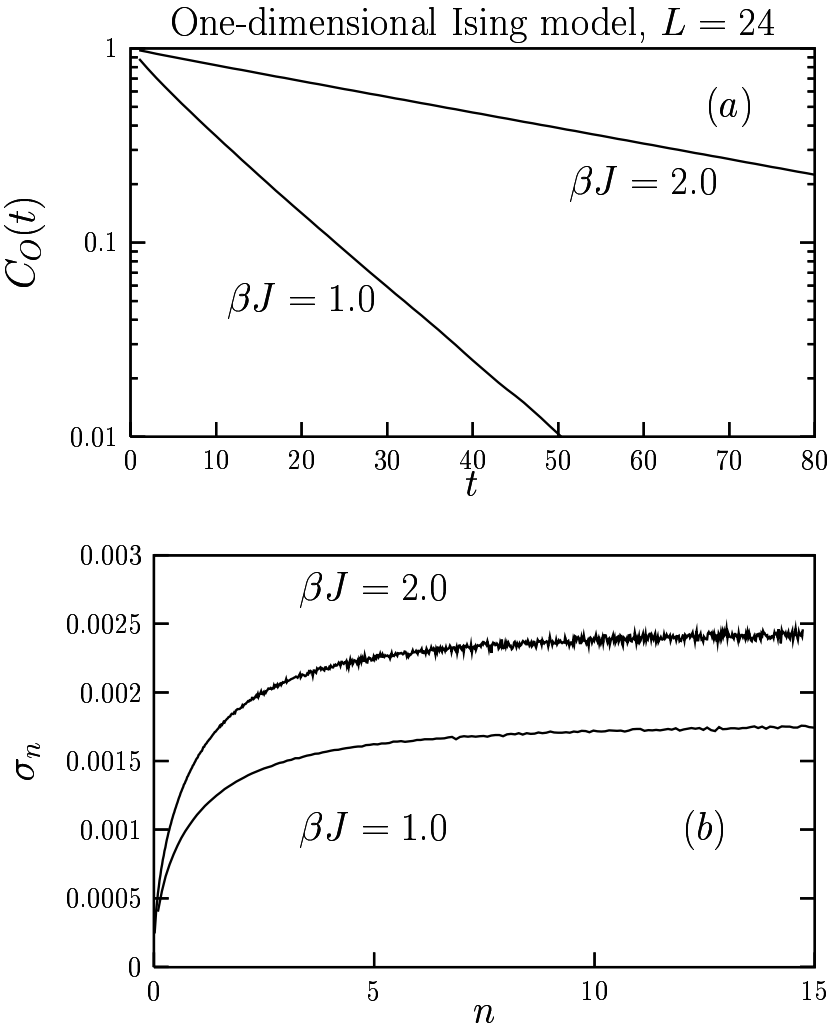

Fig. 3: *One dimensional Ising model on an L=24 site lattice. (a) Autocorrelation time (see Eq. (137)) for* $g(r = L/2)$*. The time unit corresponds to a single sweep. (b) Estimate of the error (see Eq. (139)). Here,* n *corresponds to the size of the bins in units of the autocorrelation time.* $n \sim 10$ *is sufficient to obtain a reliable estimate of the error. After* 2×10^6 *sweeps, our results yield* $g(r = L/2) = 0.076 \pm 0.0018$ *and* 0.909 ± 0.0025 *for* $\beta t = 1$ *and* 2 *respectively. The exact result is* $g(r = L/2) = 0.0760$ *and* 0.9106 *at* $\beta t = 1$ *and* 2 *respectively.*

B Proof of the determinant identity

In this section, a general determinant identity is proven [32] that can be used to derive Wick's theorem for contributions of a configuration C_n to physical observables. Let us define the vectors $\mathbf{u_i}$, $\mathbf{v_i} \in \mathbb{C}^m$ and the numbers $\alpha_{ij} \in \mathbb{C}$. Further, let $A \in \mathbb{C}^{m\times m}$ be a matrix of rank m. We define the non-singular matrices $\mathbf{M_n} \in \mathbb{C}^{(m+n)\times(m+n)}$ and $\mathbf{A_{ij}} \in \mathbb{C}^{(m+1)\times(m+1)}$ by

$$\mathbf{M_n} = \begin{pmatrix} \mathbf{A} & \mathbf{u_1} & \dots & \mathbf{u_n} \\ \mathbf{v_1}^T & \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v_n}^T & \alpha_{n1} & \dots & \alpha_{nn} \end{pmatrix}, \quad \mathbf{A_{ij}} = \begin{pmatrix} \mathbf{A} & \mathbf{u_j} \\ \mathbf{v_i}^T & \alpha_{ij} \end{pmatrix}. \tag{144}$$

With these definitions, the following determinant identity holds

$$\det \mathbf{M_n} \, (\det \mathbf{A})^{n-1} = \det \begin{pmatrix} \det \mathbf{A_{11}} & \dots & \det \mathbf{A_{1n}} \\ \vdots & \ddots & \vdots \\ \det \mathbf{A_{n1}} & \dots & \det \mathbf{A_{nn}} \end{pmatrix}. \tag{145}$$

The identity can be proven by induction in n. It is trivial for $n = 1$, so we have to start with $n = 2$, where we have to show

$$\frac{\det \mathbf{M_2}}{\det \mathbf{A}} = \frac{\det \mathbf{A_{11}}}{\det \mathbf{A}} \frac{\det \mathbf{A_{22}}}{\det \mathbf{A}} - \frac{\det \mathbf{A_{12}}}{\det \mathbf{A}} \frac{\det \mathbf{A_{21}}}{\det \mathbf{A}}. \tag{146}$$

For the following calculations, we introduce several vectors

$$\mathbf{u^1_{ij}} = \begin{pmatrix} \mathbf{u_j} \\ \alpha_{ij} - 1 \end{pmatrix}, \; \mathbf{v^2_{ij}} = \begin{pmatrix} \mathbf{v_i} \\ 0 \end{pmatrix}, \; \mathbf{u^2} = \mathbf{v^1} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \in \mathbb{C}^{m+1}. \tag{147}$$

$$\mathbf{u^1_M} = \begin{pmatrix} \mathbf{u}_2 \\ \alpha_{12} \\ \alpha_{22} - 1 \end{pmatrix}, \mathbf{v^2_M} = \begin{pmatrix} \mathbf{v}_2 \\ \alpha_{21} \\ 0 \end{pmatrix}, \mathbf{u^2_M} = \mathbf{v^1_M} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \in \mathbb{C}^{m+2}. \tag{148}$$

Let us define the *expanded* matrix $\mathbf{C}_{\text{ex}}$ of a square matrix $\mathbf{C}$ as the matrix C expanded by one row and one column containing a unit vector

$$\mathbf{C}_{\text{ex}} = \begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}. \tag{149}$$

As a last definition, we introduce the abbreviation $b_{ij} = \mathbf{v_i}^T \mathbf{A}^{-1} \mathbf{u_j}$. Using these notations, we can write the matrices $\mathbf{A_{ij}}$ as

$$\mathbf{A_{ij}} = \mathbf{A}_{\text{ex}} + \mathbf{u^1_{ij}} \mathbf{v^1}^T + \mathbf{u^2} \mathbf{v^2_{ij}}^T. \tag{150}$$

To calculate the determinant $\det \mathbf{A_{ij}}$, we use the matrix determinant lemma $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^T) = (1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}) \det \mathbf{A}$, yielding

$$\frac{\det \mathbf{A_{ij}}}{\det \mathbf{A}_{\text{ex}}} = \left[1 + \mathbf{v^2_{ij}}^T (\mathbf{A}_{\text{ex}} + \mathbf{u^1_{ij}} \mathbf{v^1}^T)^{-1} \mathbf{u^2}\right] (1 + \mathbf{v^1}^T \mathbf{A}_{\text{ex}}^{-1} \mathbf{u^1_{ij}}). \tag{151}$$

The inverse matrix of $(\mathbf{A}_{\mathrm{ex}} + \mathbf{u}^{\mathbf{1}}_{\mathbf{ij}} \mathbf{v}^{\mathbf{1}\,T})$ can be obtained from the Sherman-Morrison formula, and a tedious calculation making use of the special form of the vectors and matrices gives the result

$$\frac{\det \mathbf{A_{ij}}}{\det \mathbf{A}} = \alpha_{ij} - b_{ij}. \tag{152}$$

From this, the right-hand side of Eq. (146) can be easily obtained. For the left-hand side, we have to perform an analogous calculation using the decomposition of the matrix $\mathbf{M_2}$

$$\mathbf{M_2} = \mathbf{A_{11}}_{\mathrm{ex}} + \mathbf{u}^{\mathbf{1}}_{\mathbf{M}} \mathbf{v}^{\mathbf{1}}_{\mathbf{M}}{}^{T} + \mathbf{u}^{\mathbf{2}}_{\mathbf{M}} \mathbf{v}^{\mathbf{2}}_{\mathbf{M}}{}^{T}. \tag{153}$$

Again, we apply the matrix determinant lemma two times and insert the Sherman-Morrison formula to calculate the inverse matrix of $(\mathbf{A_{11}}_{\mathrm{ex}} + \mathbf{u}^{\mathbf{1}}_{\mathbf{M}} \mathbf{v}^{\mathbf{1}}_{\mathbf{M}}{}^{T})$. Simplifying the result as far as possible, we finally arrive at

$$\frac{\det \mathbf{M_2}}{\det \mathbf{A}} = (\alpha_{11} - b_{11})(\alpha_{22} - b_{22}) - (\alpha_{12} - b_{12})(\alpha_{21} - b_{21}). \tag{154}$$

If we compare (154) with (152), it is clear that Eq. (146) holds.
We now assume that for a certain value $n \in \mathbb{N}$ Eq. (145) holds. For $n+1$, we can cast the matrix $\mathbf{M_{n+1}}$ in a form where we can make use of Eq. (145) holding for n

$$\mathbf{M_{n+1}} = \begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{u}}_{\mathbf{2}} & \dots & \tilde{\mathbf{u}}_{\mathbf{n+1}} \\ \tilde{\mathbf{v}}_{\mathbf{2}}^{T} & \alpha_{2,2} & \dots & \alpha_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{v}}_{\mathbf{n}}^{T} & \alpha_{n,2} & \dots & \alpha_{n,n+1} \\ \tilde{\mathbf{v}}_{\mathbf{n+1}}^{T} & \alpha_{n+1,2} & \dots & \alpha_{n+1,n+1} \end{pmatrix}, \tag{155}$$

where we have introduced the new matrix $\tilde{\mathbf{A}}$ and the vectors $\tilde{\mathbf{u}}_{\mathbf{i}}$ and $\tilde{\mathbf{u}}_{\mathbf{j}}$ with

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & \mathbf{u_1} \\ \mathbf{v_1}^{T} & \alpha_{11} \end{pmatrix}, \quad \tilde{\mathbf{u}}_{\mathbf{i}} = \begin{pmatrix} \mathbf{u_i} \\ \alpha_{1i} \end{pmatrix}, \quad \tilde{\mathbf{v}}_{\mathbf{i}} = \begin{pmatrix} \mathbf{v_i} \\ \alpha_{i1} \end{pmatrix}. \tag{156}$$

Further, we need the matrices $\tilde{\mathbf{A}}_{\mathbf{ij}}$ defined analogously to (144)

$$\tilde{\mathbf{A}}_{\mathbf{ij}} = \begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{u}}_{\mathbf{j}} \\ \tilde{\mathbf{v}}_{\mathbf{i}}^{T} & \alpha_{ij} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{u_1} & \mathbf{u_j} \\ \mathbf{v_1}^{T} & \alpha_{11} & \alpha_{1j} \\ \mathbf{v_i}^{T} & \alpha_{i1} & \alpha_{ij} \end{pmatrix}. \tag{157}$$

With these definitions, and with the abbreviations $a_{ij} = \det \mathbf{A_{ij}}$ and $\tilde{a}_{ij} = \det \tilde{\mathbf{A}}_{\mathbf{ij}}$, we are now able to apply Eq. (145) holding for n

$$\det \mathbf{M_{n+1}} \left(\det \tilde{\mathbf{A}}\right)^{(n-1)} = \det \begin{pmatrix} \tilde{a}_{2,2} & \dots & \tilde{a}_{2,n+1} \\ \vdots & \ddots & \vdots \\ \tilde{a}_{n+1,2} & \dots & \tilde{a}_{n+1,n+1} \end{pmatrix}. \tag{158}$$

For $\tilde{a}_{ij}$, we make use of Eq. (145) with $n = 2$, which we have proved above

$$\tilde{a}_{ij} = \frac{1}{\det \mathbf{A}} \left(a_{11} a_{ij} - a_{i1} a_{1j}\right). \tag{159}$$

Inserting this result in (158) yields a determinant with entries of the form $a_{11}a_{ij} - a_{i1}a_{1j}$. We make use of the multi linearity of the determinant to decompose this expression and we obtain a sum of determinants with prefactors of the form a_{ij}. Eliminating zero contributions, the resulting expression corresponds precisely to the Laplace expansion of a larger determinant, and we finally obtain

$$\det \mathbf{M_{n+1}} \det \mathbf{A}^n = \det \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n+1} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n+1,1} & a_{n+1,2} & \dots & a_{n+1,n+1} \end{pmatrix}. \tag{160}$$

This is the identity (145) for $n+1$. Hence, we have derived the determinant identity for $n+1$ using only the identity for n and $n = 2$. By induction, the identity (145) therefore holds for every $n \in \mathbb{N}$, as it is trivial for $n = 1$.

References

[1] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011)

[2] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)

[3] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)

[4] F. Assaad and H. Evertz in *Computational Many-Particle Physics* ed. by H. Fehske, R. Schneider, and A. Weiße, Lecture Notes in Physics, Vol. 739 (Springer, 2008) p. 277

[5] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A.I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004)

[6] F.F. Assaad and T.C. Lang, Phys. Rev. B **76**, 035116 (2007)

[7] P. Werner and A.J. Millis, Phys. Rev. Lett. **99**, 146404 (2007)

[8] F.F. Assaad, Phys. Rev. B **78**, 155124 (2008)

[9] M. Hohenadler, F.F. Assaad, and H. Fehske, Phys. Rev. Lett. **109**, 116407 (2012)

[10] M.Z. Hasan and C.L. Kane, Rev. Mod. Phys. **82**, 3045 (2010)

[11] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011)

[12] C.L. Kane and E.J. Mele, Phys. Rev. Lett. **95**, 146802 (2005)

[13] M. Hohenadler and F.F. Assaad, J. Phys.: Condens. Matter **25**, 143201 (2013)

[14] M. Hohenadler, T.C. Lang, and F.F. Assaad, Phys. Rev. Lett. **106**, 100403 (2011)

[15] M. Hohenadler and F.F. Assaad, Phys. Rev. B **85**, 081106 (2012)

[16] F. Goth, D.J. Luitz, and F.F. Assaad, Phys. Rev. B **88**, 075110 (2013)

[17] M. Weber, M. Hohenadler, and F.F. Assaad, Phys. Rev. B **89**, 205125 (2014)

[18] F. Goth and F.F. Assaad, Phys. Rev. B **85**, 085129 (2012)

[19] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)

[20] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008)

[21] A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. B **75**, 045118 (2007)

[22] J.L. Smith and Q. Si, Phys. Rev. B **61**, 5184 (2000)

[23] D. Rost, F. Assaad, and N. Blümer, Phys. Rev. E **87**, 053305 (2013)

[24] E.F. Huffman and S. Chandrasekharan, Phys. Rev. B **89**, 111101 (2014)

[25] A.N. Rubtsov, V.V. Savkin, and A.I. Lichtenstein, Phys. Rev. B **72**, 035122 (2005)

[26] S.M.A. Rombouts, K. Heyde, and N. Jachowicz, Phys. Rev. Lett. **82**, 4155 (1999)

[27] E. Gull, P. Werner, O. Parcollet, and M. Troyer, Europhys. Lett. **82**, 57003 (2008)

[28] P. Werner, A. Comanac, L. de' Medici, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **97**, 076405 (2006)

[29] P.W. Anderson, Phys. Rev. **124**, 41 (1961)

[30] A.C. Hewson: *The Kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, 1997)

[31] J.W. Negele and H. Orland: *Quantum Many body systems* (Addison-Wesley, Redwood City, 1988)

[32] D.J. Luitz and F.F. Assaad, Phys. Rev. B **81**, 024509 (2010)

[33] D.J. Luitz, F.F. Assaad, T. Novotný, C. Karrasch, and V. Meden, Phys. Rev. Lett. **108**, 227001 (2012)

[34] K. Mikelsons, A. Macridin, and M. Jarrell, Phys. Rev. E **79**, 057701 (2009)

[35] J.E. Hirsch and R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986)

[36] J. Yoo, S. Chandrasekharan, R.K. Kaul, D. Ullmo, and H.U. Baranger, J. Phys. A **38**, 10307 (2005)

[37] M. Hohenadler, Z.Y. Meng, T.C. Lang, S. Wessel, A. Muramatsu, and F.F. Assaad, Phys. Rev. B **85**, 115132 (2012)

[38] C. Wu and S.-C. Zhang, Phys. Rev. B **71**, 155115 (2005)

[39] D.J. Luitz: *Numerical methods and applications in many fermion systems*, PhD thesis, Universität Würzburg Astronomie (2013) `http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:20-opus-75927`

[40] K. Haule, Phys. Rev. B **75**, 155113 (2007)

[41] N. Parragh, A. Toschi, K. Held, and G. Sangiovanni, Phys. Rev. B **86**, 155158 (2012)

[42] A.M. Läuchli and P. Werner, Phys. Rev. B **80**, 235117 (2009)

[43] H. Shinaoka, M. Dolfi, M. Troyer, and P. Werner, arxiv:1404.1259 (2014)

[44] H. Hafermann, K.R. Patton, and P. Werner, Phys. Rev. B **85**, 205106 (2012)

[45] P. Werner and A.J. Millis, Phys. Rev. Lett. **99**, 126405 (2007)

[46] P. Werner, E. Gull, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **101**, 166405 (2008)

[47] N. Parragh, G. Sangiovanni, P. Hansmann, S. Hummel, K. Held, and A. Toschi, Phys. Rev. B **88**, 195116 (2013)

[48] J. Werner and F.F. Assaad, Phys. Rev. B **88**, 035113 (2013)

[49] J. Werner and F.F. Assaad, Phys. Rev. B **89**, 245119 (2014)

[50] P.W. Anderson, P.A. Lee, M. Randeria, T.M. Rice, N. Trivedi, and F.C. Zhang, Journal of Physics: Condens. Matter **16**, R755 (2004)

[51] M. Greiner, O. Mandel, T. Esslinger, T.W. Hansch, and I. Bloch, Nature **415**, 39 (2002)

[52] B. Sipos, A.F. Kusmartseva, A. Akrap, H. Berger, L. Forro, and E. Tutis, Nature Mater. **7**, 960 (2008)

[53] R.P. Hardikar and R.T. Clay, Phys. Rev. B **75**, 245103 (2007)

[54] M. Hohenadler and F.F. Assaad, J. Phys.: Condens. Matter **25**, 014005 (2013)

[55] B. Efron: *The jackknife, the bootsrap and other resampling plans*, CBMS-NSF conference series in applied mathematics (J. W. Arrowsmith, Bristol, 1982)

[56] G.J. Kemeny and J. Snell: *Finite Markov Chains*, The university series in undergraduate mathematics (Van Nostrand, New York, 1960)

[57] R.J. Baxter: *Exactly solved models in statistical mechanics* (Academic Press Limited, London, 1989)