

## Quantum Monte Carlo

QMC could be called application of Monte Carlo to Quantum many body systems (of bosons and fermions).

There are very powerful techniques available for bosonic many-body systems (like spin systems) but not so much success in fermionic systems. The reason is the so called **fermionic minus sign**. This issue of fermionic statistics has not been solved yet and techniques currently on the market most often "sample" the minus sign and estimate the error. And the error grows as the temperature is decreased.

Having the above mentioned minus sign problem in mind, QMC is still one of the most powerful techniques available for many-body systems. There are cases where even for fermions, the minus-sign problem does not appear. One such case is **Hirsch-Fye algorithm** for **quantum impurity** system which we will implement. Minus sign appears when more complicated atom or more atoms are considered.

Varios Quantum Monte Carlo techniques developed over the past with common denominator: Importance Monte-Carlo sampling.

- **Variational Monte Carlo** : for finding the ground state of quantum Hamiltonian. The ground state wave function is parametrized and the Metropolis algorithm is used to minimize the total energy.
- **Diffusion Monte Carlo** : Uses diffusion type of equation in combination with random walk to estimate the ground state wave function of many-body system
- **Path integral Monte Carlo and Determinantal Monte Carlo**: excited states also accessible therefore finite temperatures and response functions at finite frequencies accessible. The idea is to rewrite the problem in Feynman path integral formulation and compute the multidimensional integrals using Monte Carlo importance sampling. We will show the technique on Hirsch-Fye for quantum impurity.
- **Diagrammatic and Continuous time quantum Monte Carlo**: it samples in configuration space of Feynman diagrams. The partition function is divided into exactly solvable part (not necessarily quadratic - Wick's theorem not necessary) and the rest. This latter part is expanded in Taylor series. The resulting diagrams are sampled by Monte Carlo importance sampling.

# 1 What do we plan to cover in this class?

- Determinantal QMC with example of Hirsch-Fye algorithm.
- Continuous time QMC algorithm (expanding the action in terms of hybridization strength).

## Similarities and difference between the two algorithms

### 1.1 Determinantal QMC

Determinantal QMC samples in the space of Slater determinants. They are enumerated by ising-like spin configurations. A configuration in Markov chain is an ising-like spin configuration, denoted by  $\{\phi\}$  in this chapter.

The basic idea of the determinantal QMC is sketched below.

The partition function and average of any physical observable can be expressed (in the Feynman path integral formulation) by

$$Z = \int \mathcal{D}[\psi^\dagger \psi] e^{-S_0 - \Delta S} \quad (1)$$

$$\langle A \rangle = \frac{1}{Z} \int \mathcal{D}[\psi^\dagger \psi] e^{-S_0 - \Delta S} \quad (2)$$

In Determinantal QMC,  $S_0$  is the quadratic part of the action (*needs to be exactly solvable and needs to obey Wick's theorem!*) and  $\Delta S$  is the interacting part (usually quartic - Coulomb interaction).

First, the discrete Hubbard-Stratonovich transformation is used to decouple the quartic term in the action (Fermions then interact through ising spins rather than directly with instantaneous interaction - analogy with virtual photons and Coulomb interaction).

$$\Delta S[\psi] \rightarrow \widetilde{\Delta S}[\psi, \phi] \quad (3)$$

Here  $\phi$  stand for the ising like spin rather than bosonic field.

The advantage of the transformation is that the action is quadratic in fermionic operators.

The price we pay is that the Hilberts space is heavily enlarged, i.e.,

$$Z = \int \mathcal{D}[\phi] \int \mathcal{D}[\psi^\dagger \psi] e^{-S_0 - \Delta S[\psi, \phi]} \quad (4)$$

We need to sum over all fermionic paths and also over all ising configurations.

In the next step, we integrate out fermions  $\psi$  and as a result we get

$$Z = \int \mathcal{D}[\phi] \det G^{-1}(\{\phi\}) \quad (5)$$

$$\langle A \rangle = \frac{1}{Z} \int \mathcal{D}[\phi] \det G^{-1}(\{\phi\}) A(\{\phi\}) \quad (6)$$

Here  $\{\phi\}$  stand for an ising configuration. The interacting problem is cast into a form of classical problem of ising spins. There are however infinite number of spins. For each small time interval, we need an ising spin and for each degree of freedom (site index, electron spin, or band index) we need ising spins.

The Monte Carlo algorithm is used to sample over all possible ising configurations. The importance sampling is however different than in classical case. The weight in the classical case is  $e^{-\beta E_{\{\phi\}}}$  where  $E_{\{\phi\}}$  is the energy of the ising configuration and  $\beta = 1/T$ . In determinantal QMC, the weight of the configuration is

$$\det G^{-1}(\{\phi\}).$$

## 1.2 Diagrammatic and Continuous time QMC

Again we start with the Feynman path integral formulation of a general interacting problem

$$Z = \int \mathcal{D}[\psi^\dagger \psi] e^{-S_0 - \Delta S} \quad (7)$$

$$\langle A \rangle = \frac{1}{Z} \int \mathcal{D}[\psi^\dagger \psi] e^{-S_0 - \Delta S} \quad (8)$$

We expand the action in power series to get the series of Feynman diagrams

$$Z = \sum_k \int \mathcal{D}[\psi^\dagger \psi] e^{-S_0} \frac{(-1)^k}{k!} (\Delta S)^k \quad (9)$$

Here is an important difference between the two QMC algorithms. In this expansion,  $S_0$  does not need to be quadratic in Fermionic operators. It needs to be exactly solvable but not necessary quadratic (an example is an atom).

(Strictly speaking, if  $S_0$  is not quadratic, the resulting power series is not a series of Feynman diagrams in terms of  $\langle \psi \psi^\dagger \rangle$ . This series can however still be thought as a series of Feynman diagrams in another representation.)

Monte Carlo sampling is used to sample over all possible diagrams. (there are usually  $(2k)!$  terms at order  $k$  because different order of times leads to different diagrams).

- We do not introduce time discretization, hence the name continuous time Monte Carlo.
- We do not increase the Hilbert space by using spins or other Hubbard Stratonovich fields.
- The Markov chain does not sample over ising configurations but rather over diagrams in above power series.

The weight that corresponds to the particular diagrams is again proportional to its contribution to partition function  $Z$ , i.e.,

$$Weight[Diagram] = \int \mathcal{D}[\psi\psi^\dagger] e^{-S_0} \frac{(-1)^k}{k!} (\Delta S)^k \quad (10)$$

Here a typical contribution to  $(\Delta S)^k$  contains a product of  $2k$  fermionic operators. Each permutation of these operators leads to a distinct diagram which needs to be sampled with Monte Carlo importance sampling.

## Determinantal QMC and Hirsch-Fye

Original derivation used Hamiltonian formulation rather than path integral approach. We will follow the original derivation.

The derivation is technically involved but the algorithm is simple to implement.

When necessary, we will think in terms of quantum impurity problem, however, this is not really necessary since the derivation is very general and is (in practically the same form) used in many determinantal QMC's.

The Hamiltonian for quantum impurity is

$$H = \sum_s \epsilon_0 c_{0\sigma}^\dagger c_{0\sigma} + U n_{0\uparrow} n_{0\downarrow} + \sum_{p>0\sigma} [V_{0p} c_{0\sigma}^\dagger c_{p\sigma} + V_{0p}^* c_{p\sigma}^\dagger c_{0\sigma}] + \sum_{p>0,\sigma} \epsilon_p c_{p\sigma}^\dagger c_{p\sigma} \quad (11)$$

First two terms are onsite atomic terms, last corresponds to the infinite band of electrons which are not interacting, and the third term couples the atom with the band of electrons.



Only the second term is **non-quadratic** - Coulomb repulsion. Without this term, the problem is exactly solvable.

We will need the solution of the non-interacting ( $U = 0$ ) case. It is a matter of simple matrix inversion to show that the **impurity** Green's function in case of  $U = 0$  is

$$G^0 = (\omega - \epsilon_0 - \sum_{p>0} \frac{V_{0p}^* V_{p0}}{\omega - \epsilon_p})^{-1} \quad (12)$$

We proceed with a general derivation of determinantal QMC. First we need to separate interacting part from non-interacting part of Hamiltonian

$$H = H^0 + H^i$$

In case of Quantum impurity, the above terms are explicitly

$$H^0 = \sum_{p\sigma} (\epsilon_p + \delta_{p0} \frac{1}{2} U) c_{p\sigma}^\dagger c_{p\sigma} + \sum_{p>0, \sigma} [V_{0p} c_{0\sigma}^\dagger c_{p\sigma} + V_{0p}^* c_{p\sigma}^\dagger c_{0\sigma}] \quad (13)$$

$$H^i = U [n_{0\uparrow} n_{0\downarrow} - \frac{1}{2} (n_{0\uparrow} + n_{0\downarrow})] \quad (14)$$

Taking small time step in imaginary time

$$e^{-\beta H} = e^{-\Delta\tau H} e^{-\Delta\tau H} \dots e^{-\Delta\tau H}; \quad \Delta\tau L = \beta$$

and taking into account the identity

$$e^{\Delta\tau(A+B)} = e^{\Delta\tau A} e^{\Delta\tau B} + O(\Delta\tau^2[A, B]) \quad (15)$$

we can performed "Trotter-Suzuki decomposition" of trace

$$Z = \text{Tr}(e^{-\beta H}) = \text{Tr}\left[\prod_{l=0}^{L-1} e^{-\Delta\tau(H^0 + H^i)}\right] \approx \text{Tr}\left[\prod_{l=0}^{L-1} e^{-\Delta\tau H^0} e^{-\Delta\tau H^i}\right] + O(\Delta\tau^2 U) \quad (16)$$

This is the *only* approximation in the Determinantal QMC. One can often calculate with few decreasing  $\Delta\tau$ 's and estimate the limit  $\Delta\tau \rightarrow 0$ . Typically, one takes  $\Delta\tau \leq \sqrt{0.25/U}$  and therefore  $L \geq \beta\sqrt{4U}$ .

The second important step is the discrete Hubbard-Stratonovich decomposition of the interaction term

$$e^{-\Delta\tau U[n_{0\uparrow}n_{0\downarrow} - \frac{1}{2}(n_{0\uparrow} + n_{0\downarrow})]} = \frac{1}{2} \sum_{s=\pm 1} e^{\lambda s(n_{0\uparrow} - n_{0\downarrow})} \quad (17)$$

To check that the above decoupling truly works, we can check how the two operators work on every state in the Hilbert space of the atom

	<i>left term</i>	<i>right term</i>	
$ 0\rangle$	1	1	
$ \uparrow\rangle$	$e^{\Delta\tau U/2}$	$\frac{1}{2}(e^\lambda + e^{-\lambda})$	(18)
$ \downarrow\rangle$	$e^{\Delta\tau U/2}$	$\frac{1}{2}(e^\lambda + e^{-\lambda})$	
$ \uparrow\downarrow\rangle$	1	1	

It follows that  $e^{\Delta\tau U/2} = \cosh \lambda$  and we arrive at

$$Z = \text{Tr} \left[ \prod_{l=0}^{L-1} e^{-\Delta\tau H^0} \frac{1}{2^L} \sum_{s_0, s_1, \dots, s_{L-1}} e^{\lambda s_l (n_{0\uparrow} - n_{0\downarrow})} \right] \quad (19)$$

or

$$Z = \frac{1}{2^L} \sum_{\{s\}} \text{Tr} \left[ \prod_{l=0}^{L-1} e^{-\Delta\tau H^0(l) + \lambda s_l (n_{0\uparrow}(l) - n_{0\downarrow}(l))} \right] \quad (20)$$

Note that in the last step we combined two exponents  $e^{-\Delta\tau H^0} e^{\lambda s_l (n_{0\uparrow} - n_{0\downarrow})}$  into

~~$e^{-\Delta\tau H^0 + \lambda s_l (n_{0\uparrow} - n_{0\downarrow})}$ , which is allowed because we are taking small time-step, and~~

because  $\lambda \approx \sqrt{\Delta\tau U}$  is also small, hence we are making the same Trotter error as before in discretizing the action.

We mapped the many-body interacting problem to a non-interacting (quadratic in  $c, c^\dagger$ ) problem. The price we payed is the enlarged Hilbert space with  $2^L$  Ising spins.

We can also evaluate the Green's function or any correlation function in the same way

$$G_{ij}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z} \frac{1}{2^L} \sum_{\{s\}} \text{Tr} \left[ T_\tau c_i(\tau_{l_1}) c_j^\dagger(\tau_{l_2}) \prod_{l=0}^{L-1} e^{-\Delta\tau H^0 + \lambda s_l (n_{0\uparrow} - n_{0\downarrow})} \right] \quad (21)$$

Note that usual definition of the Green's function has minus sign

$G(\tau, \tau') = -\langle T_\tau c(\tau) c^\dagger(\tau') \rangle$  and is different from what QMC community uses. Be *carefull* with the minus sign when reading QMC literature.

If the temperature is high or  $U$  is small, the number of time-slices  $L$  can be taken to be small. In this case, we can evaluate the above summation over Ising-spins exactly and we do not need Monte-Carlo. This is called Gray code enumeration and is very straightforward to implement.

The Monte Carlo sampling is used because the phase space of Ising spins  $2^L$  is too large in almost all interesting cases. The Ising configurations  $\{s\} \equiv (s_0, s_1, s_2, \dots, s_{L-1})$  are visited with Metropolis or Heat-Bath algorithm. For the problem on the lattice (not impurity but lattice) the equations and algorithm is the same. Just decoupling of interaction terms needs to be performed on each site and the number of configurations is than  $2^{LN}$  where  $N$  is number of interacting sites.

The derivation below is pretty tedious but is just rewriting the non-interacting problem in a form convenient for use with computer.

At fixed Ising configuration we want to evaluate

$$\text{Tr}[e^{-\Delta\tau H_{L-1}^0 + \lambda s_{L-1}(n_{0\uparrow}(\tau_{L-1}) - n_{0\downarrow}(\tau_{L-1}))} \dots e^{-\Delta\tau H_0^0 + \lambda s_0(n_{0\uparrow}(\tau_0) - n_{0\downarrow}(\tau_0))}] \quad (22)$$

Note here the order of the terms: The earlier times must appear first due to the time ordering  $T_\tau$  operator!

Since we can write  $H^0$  in the form

$$H^0 = \sum_{ij\sigma} c_{i\sigma}^\dagger h_{ij}^0 c_{j\sigma} \quad (23)$$

the above trace at fixed Ising configuration becomes

$$Z_{\{s\}} = \text{Tr} \left[ \prod_{l=0}^{L-1} e^{-\Delta\tau \sum_{ij\sigma} c_{i\sigma}^\dagger(\tau_l) [h_{l\sigma}]_{ij} c_{j\sigma}(\tau_l)} \right] \quad (24)$$

where

$$[h_{l\sigma}]_{ij} = h_{ij}^0 - \frac{\lambda}{\Delta\tau} s_l \sigma \delta_{i=j=0} \quad (25)$$

To calculate the above trace, we will first rotate the problem to diagonal form and then rotate the result back to the original base. Since the Hamiltonian is quadratic (one particle problem) the diagonal form is obtained by diagonalizing the matrix  $h$  ( $hA = A\varepsilon$  or more precisely)

$$h_{(i\sigma),(j\sigma)} A_{(j\sigma),\alpha} = \varepsilon_\alpha A_{(i\sigma),\alpha} \quad (26)$$

Here we combined orbital and position index ( $i\sigma$ ) to transform into new orbital-position index  $\alpha$ . The transformation can be performed for each time slice  $l$  to obtain  $\varepsilon_\alpha(l)$ . The

transformation of operator is  $c_{i\sigma} = \sum_{\alpha} A_{(i\sigma),\alpha} c_{\alpha}$  and therefore one has

$$Z_{\{s\}} = \text{Tr} \left[ \prod_{l=0}^{L-1} e^{-\Delta\tau \sum_{\alpha} c_{\alpha}^{\dagger}(\tau_l) \varepsilon_{\alpha}(l) c_{\alpha}(\tau_l)} \right] \quad (27)$$

Since we have non-interacting problem, we can sum for each state  $\alpha$  separately

$$Z_{\{s\}} = \prod_{\alpha} \text{Tr}_{\alpha} \left[ \prod_{l=0}^{L-1} e^{-\Delta\tau \varepsilon_{\alpha}(l) n_{\alpha}(\tau_l)} \right] \quad (28)$$

where  $\text{Tr}_{\alpha}$  means trace at fixed  $\alpha$ . For fermions,  $n_{\alpha}$  at fixed  $\alpha$  can take only two values 0 or 1 therefore

$$Z_{\{s\}} = \prod_{\alpha} \text{Tr}_{\alpha} \left[ \prod_{l=0}^{L-1} e^{-\Delta\tau \varepsilon_{\alpha}(l) n_{\alpha}(\tau_l)} \right] = \prod_{\alpha} \left( 1 + \prod_{l=0}^{L-1} e^{-\Delta\tau \varepsilon_{\alpha}(l)} \right) \quad (29)$$

The last step is discretized version of the more familiar equation

$\text{Tr} [T_{\tau} e^{-\int_0^{\beta} d\tau \varepsilon_{\alpha} n_{\alpha}}] = (1 + e^{-\int_0^{\beta} d\tau \varepsilon_{\alpha}}) = (1 + e^{-\beta \varepsilon_{\alpha}})$ . *This is because  $n_{\alpha}$  can be either zero for all times or 1 for all times, since there is no term in the Hamiltonian that would change the occupation with time.*

Now we can define diagonal matrix  $(\tilde{B}_l)_{\alpha\alpha} \equiv e^{-\Delta\tau\varepsilon_\alpha(l)}$  and write the above equation in base-invariant form

$$Z_{\{s\}} = \text{Det}(1 + \tilde{B}_{L-1}\tilde{B}_{L-2}\cdots\tilde{B}_0) \quad (30)$$

where Det is taken over the space and orbital index  $\alpha$ . The matrix  $\tilde{B}$  can be transformed back to the original representation of index  $i, j$ :

$B = (A\tilde{B}A^\dagger) = e^{-\Delta\tau h(l)} = e^{-\Delta\tau h^0} e^{\lambda s_l \sigma \delta_{i=j=0}}$ , to give the result

$$Z_{\{s\}} = \text{Det}(1 + B_{L-1}B_{L-2}\cdots B_0) \quad (31)$$

We finally arrived at expression, which is basis independent

$$Z = \sum_{\{s\}} \frac{1}{2^L} \text{Det}(1 + B_{L-1}B_{L-2}\cdots B_0) \quad (32)$$

$$B_l = e^{-\Delta\tau h^0} e^{\lambda s_l \sigma \delta_{i=j=0}} \quad (33)$$

In the literature, the same derivation is usually based on the identity:

$$\text{Tr}_{c_i^\dagger, c_i} \left[ e^{-\sum_{ij} c_i^\dagger A_{ij} c_j} e^{-\sum_{ij} c_i^\dagger B_{ij} c_j} e^{-\sum_{ij} c_i^\dagger C_{ij} c_j} \right] = \text{Det}[1 + e^{-A} e^{-B} e^{-C}] \quad (34)$$



Starting with Eq. 24, we could take  $A = \Delta\tau h(l = 0)$ ,  $B = \Delta\tau h(l = 1)$ ,  $C = \Delta\tau h(l = 2)$ , etc... and we would then immediately arrive at Eq. 31 by using identity Eq. 34.

We need similar expression for the Green's function. The derivation goes along the same lines. Similar to the partition function at fixed Ising configuration  $Z_{\{s\}}$ , we define Green's function at fixed Ising configuration  $g_{\{s\}}(i\sigma), (j\sigma)$  so that the electron Green's function is

$$G = \frac{\sum_{\{s\}} g_{\{s\}} Z_{\{s\}}}{\sum_{\{s\}} Z_{\{s\}}}; \quad \{s\} \text{ Ising configuration} \quad (35)$$

Green's function at fixed Ising configuration is

$$g_{(i\sigma), (j\sigma)}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z_{\{s\}}} \text{Tr} \left[ T_{\tau} c_{i\sigma}(\tau_{l_1}) c_{j\sigma}^{\dagger}(\tau_{l_2}) \prod_{l=0}^{L-1} e^{-\Delta\tau H^0 + \lambda s_l (n_{0\uparrow} - n_{0\downarrow})} \right] \quad (36)$$

We can again go to the diagonal representation where Hamiltonian is diagonal therefore the

Green's function is also diagonal and takes the form

$$g_{\alpha\alpha}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z_{\{s\}}} \text{Tr} \left[ T_{\tau} c_{\alpha}(\tau_{l_1}) c_{\alpha}^{\dagger}(\tau_{l_2}) \prod_{l=0}^{L-1} e^{-\Delta\tau \sum_{\alpha} \varepsilon_{\alpha}(l) c_{\alpha}^{\dagger}(\tau_l) c_{\alpha}(\tau_l)} \right] \quad (37)$$

The time ordering is very important here. Suppose  $l_1 > l_2$  and if we define the following operator operator  $U_l \equiv e^{-\Delta\tau \varepsilon_{\alpha}(l) n_{\alpha}(\tau_l)}$  we have

$$g_{\alpha\alpha}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z_{\{s\}}} \text{Tr} [U_{L-1} \cdots U_{l_1} c_{\alpha}(\tau_{l_1}) U_{l_1-1} \cdots U_{l_2} c_{\alpha}^{\dagger}(\tau_{l_2}) U_{l_2-1} \cdots U_0] \quad (38)$$

Cyclic permutation of the trace gives

$$g_{\alpha\alpha}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z_{\{s\}}} \text{Tr} [U_{l_2-1} \cdots U_0 U_{L-1} \cdots U_{l_1} c_{\alpha}(\tau_{l_1}) U_{l_1-1} \cdots U_{l_2} c_{\alpha}^{\dagger}(\tau_{l_2})] \quad (39)$$

Since we have one particle problem, we need to consider only two possibilities: state  $\alpha$  is either empty or occupied. The only nonzero contribution is obviously when  $\alpha$  is empty

$$g_{\alpha\alpha}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z_{\{s\}}} \langle 0 | U_{l_2-1} \cdots U_0 U_{L-1} \cdots U_{l_1} c_{\alpha}(\tau_{l_1}) U_{l_1-1} \cdots U_{l_2} c_{\alpha}^{\dagger}(\tau_{l_2}) | 0 \rangle \quad (40)$$

which gives

$$g_{\alpha\alpha}(\tau_{l_1}, \tau_{l_2}) = \frac{1}{Z_{\{s\}}} [\tilde{B}_{l_1-1} \cdots \tilde{B}_{l_2}] = \frac{\tilde{B}_{l_1-1} \cdots \tilde{B}_{l_2}}{(1 + \tilde{B}_{L-1} \cdots \tilde{B}_0)} \quad (41)$$

In the original non-diagonal representation, Green's function takes the form

$$g(\tau_{l_1}, \tau_{l_2}) = \begin{cases} l_1 > l_2 & B_{l_1-1} \cdots B_{l_2} [1 + B_{L-1} \cdots B_0]^{-1} \\ l_1 < l_2 & -B_{l_1-1} \cdots B_0 B_{L-1} \cdots B_{l_2} [1 + B_{L-1} \cdots B_0]^{-1} \end{cases} \quad (42)$$

Note here the minus sign in the second line, which comes from the fermionic commutation rules.

This complicated matrix can be written in much more elegant way as an inversion of a large matrix of size  $L * N \times L * N$ , where  $L$  is number of times slices and  $N$  is the number of orbitals times number of sites. We combined position-orbital index with spin index for convenience to have one index  $(i\sigma)$ . However, all matrices  $g, B, \dots$  are diagonal in spin index therefore we can work with smaller matrix of size (position\*orbital) and not (position\*orbital\*spin) and use  $B^\sigma$  instead of  $B$ .

Matrix  $g_{l_1, l_2}$  is

$$g = \begin{pmatrix} 1 & 0 & \cdots & 0 & B_{L-1} \\ -B_0 & 1 & \cdots & \cdots & 0 \\ 0 & -B_1 & 1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & -B_{L-2} & 1 \end{pmatrix}^{-1} \quad (43)$$

and, at the same time we also have

$$Z_{\{s\}} = \text{Det}[g^{-1}] \quad (44)$$

Perhaps the best way to check the above identity is to examine a few lowest order terms. The non-trivial  $3 \times 3$  matrix gives

$$\begin{pmatrix} 1 & 0 & B_2 \\ -B_0 & 1 & 0 \\ 0 & -B_1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} (1 + B_2 B_1 B_0)^{-1} & -B_2 B_1 (1 + B_0 B_2 B_1)^{-1} & -B_2 (1 + B_1 B_0 B_2)^{-1} \\ B_0 (1 + B_2 B_1 B_0)^{-1} & (1 + B_0 B_2 B_1)^{-1} & -B_0 B_2 (1 + B_1 B_0 B_2)^{-1} \\ B_1 B_0 (1 + B_2 B_1 B_0)^{-1} & B_1 (1 + B_0 B_2 B_1)^{-1} & (1 + B_1 B_0 B_2)^{-1} \end{pmatrix} \quad (45)$$

which shows the right structure for 3-time slices. Note that  $B$ 's are itself matrices in orbital, spin, (position) indices. Determinant of the above  $3 \times 3$  matrix is  $\text{Det}[1 + B_2 B_1 B_0]$  which is compatible with the second identity.

Let's repeat the results obtained so far

$$G = \frac{\sum_{\{s\}} \text{Det}[g^{-1}]g}{\sum_{\{s\}} \text{Det}[g^{-1}]} \quad (46)$$

$$g^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & B_{L-1} \\ -B_0 & 1 & \cdots & \cdots & 0 \\ 0 & -B_1 & 1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & -B_{L-2} & 1 \end{pmatrix} \quad (47)$$

$$B_l = e^{-\Delta\tau h^0} e^{\lambda s_l \sigma \delta_{i=j=0}} \quad (48)$$

These are general results of determinantal QMC, and for general case we just do not have  $\delta_{i=j=0}$  in the last term, but rather such decoupled term exists for each site.

## Hirsch-Fye

In the specific case of a quantum impurity, we can make a few more steps. The size of the matrix  $g$  is now  $L * N \times L * N$  where  $L$  is the number of time slices and  $N$  is the proportionant to the number of sites (including the non-interacting conduction band sites). As usual, it is possible to integrate out the degrees of freedom that correspond to the non-interacting part of the system (the conduction band) and work with the interacting sites only. The great advantage is that the conduction band can then be treated as an infinite band (with infinite number of sites).

Hirsch and Fye pointed out that there exists something like a Dyson equation which connects Green's functions of different Ising configurations.

We defined matrix  $B$  above as

$$(B_l)_{(i\sigma),(j\sigma)} = (e^{-\Delta\tau h^0})_{ij} e^{\lambda s_l \sigma \delta_{i=j=0}}$$

It is convenient to define another matrix of size  $N \times N$ , which is identity everywhere except on impurity site  $i = j = 0$  where it is diagonal and takes the form  $e^{\lambda s_l \sigma}$ . Let's call this matrix  $e^{V_l}$ :

$$(e^{V_l})_{(i\sigma),(j\sigma)} = \delta_{i,j} + \delta_{i=j=0} (e^{\lambda s_l \sigma} - 1) \quad (49)$$

Matrix  $B_l$  therefore becomes

$$B_l = e^{-\Delta\tau h^0} e^{V_l} \quad (50)$$

Let's construct  $L * N \times L * N$  matrix  $e^V$  and multiply it with  $g^{-1}$

$$g^{-1} e^{-V} = \begin{pmatrix} 1 & 0 & \cdots & 0 & B_{L-1} \\ -B_0 & 1 & \cdots & \cdots & 0 \\ 0 & -B_1 & 1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & -B_{L-2} & 1 \end{pmatrix} \begin{pmatrix} e^{-V_0} & 0 & 0 & 0 & 0 \\ 0 & e^{-V_1} & 0 & 0 & 0 \\ 0 & 0 & e^{-V_2} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & e^{-V_{L-1}} \end{pmatrix}$$

$$= \begin{pmatrix} e^{-V_0} & 0 & \cdots & 0 & B_{L-1} e^{-V_{L-1}} \\ -B_0 e^{-V_0} & e^{-V_1} & \cdots & \cdots & 0 \\ 0 & -B_1 e^{-V_1} & e^{-V_2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & e^{-V_{L-2}} & 0 \\ \cdots & \cdots & \cdots & -B_{L-2} e^{-V_{L-2}} & e^{-V_{L-1}} \end{pmatrix}$$

It is important to realize that the **off-diagonal** elements of the resulting matrix **do not depend on Ising spin configuration** since  $B_l e^{-V_l} = e^{-\Delta\tau h^0(l)} e^{V_l} e^{-V_l} = e^{-\Delta\tau h^0(l)}$ .

It is then simple to see that the following matrix equation is valid

$$g^{-1} e^{-V} - g'^{-1} e^{-V'} = e^{-V} - e^{-V'} \quad (53)$$

where  $(g, V)$  and  $(g', V')$  correspond to any two (different) Ising configurations.



The above equation (53) is important because: if we know  $g$  for certain configuration and we obtain by Markov chain simulation new  $V'$ , we can get  $g'$ . The big advantage of using equation (53) is that one can isolate the impurity site and monitor what happens with the impurity site only. One can thus avoid using large  $N * L \times N * L$  matrices, instead one can manipulate  $L \times L$  matrices.

It is instructive to rewrite the above equation in the form

$$g'^{-1} - g^{-1} = (1 - g^{-1})(1 - e^{V'-V}) \quad (54)$$

or

$$g' - g = (g - 1)(e^{V'-V} - 1)g'. \quad (55)$$

As we noted before  $e^V$  is unity for each but impurity site. The term  $(e^{V'-V} - 1)$  therefore vanishes for each but impurity site. The change of the green's function on the impurity site can hence be computed without the knowledge of non-impurity sites, i.e.,

$$(g')_{00} - (g)_{00} = ((g)_{00} - 1)(e^{V'-V} - 1)_{00}(g')_{00} \quad (56)$$

It is therefore sufficient to monitor the impurity only in the Markov chain simulation.

Slightly different form of Eq. (55) is actually used in computation

$$\mathcal{A} g' = g \quad (57)$$

with

$$\mathcal{A} = 1 + (1 - g)(e^{V' - V} - 1) \quad (58)$$

The important point is that

$$\text{Det}(g) = \text{Det}(g') \text{Det}(\mathcal{A})$$

or

$$\text{Det}(g'^{-1}) = \text{Det}(g^{-1}) \text{Det}(\mathcal{A})$$

with  $\text{Det}(\mathcal{A}) = \text{Det}(\mathcal{A}_{00}) \times 1$ , where 00 means the impurity site. Note that  $\mathcal{A}$  is not just unity outside the impurity, as it takes the form

$$\mathcal{A}_{i0} = -g_{i0}(e^{V' - V} - 1)_{00} \quad (59)$$

for  $i \neq 0$ . However,  $\mathcal{A}_{0j} = 0$  for  $j \neq 0$ , hence the determinant expansion by minors gives

$$\mathcal{A}_{00} - \mathcal{A}_{10}0 + \mathcal{A}_{20}0 + \cdots = \mathcal{A}_{00}$$

## The CT-auxiliary field MC

While Hirsch-Fye algorithm is not much used anymore in practice, its close cousin (CTAUX) is used a lot for solving large cluster problems when Coulomb repulsion is very simple (only Hubbard  $U$  or Slater- $F^0$ ).

The CTAUX arrives at almost identical equations as in the case of determinantal MC, however, it has continuous time evolution, which makes it free from time-discretization systematic error.

We will derive this algorithm with the action formulation, which is a faster way of getting to the same equations. We write

$$\begin{aligned}
 S = & \int_0^\beta d\tau \int_0^\beta d\tau' \sum_\sigma c_\sigma^\dagger(\tau) [-G_0^{-1}(\tau - \tau')] c_\sigma(\tau') + K \\
 & + \int_0^\beta \left[ -\frac{K}{\beta} + U n_\uparrow(\tau) n_\downarrow(\tau) - U \frac{n_\uparrow(\tau) + n_\downarrow(\tau)}{2} \right]
 \end{aligned} \tag{60}$$

The partition function is then

$$Z = \int \mathcal{D}[c^\dagger c] e^{-S_0} \exp \left( \frac{K}{\beta} \int_0^\beta d\tau \left[ 1 - \frac{U\beta}{K} \left( n_\uparrow n_\downarrow - \frac{n_\uparrow + n_\downarrow}{2} \right) \right] \right) \tag{61}$$

Now we expand it in powers of the interaction  $\Delta S$ , to obtain

$$\begin{aligned}
 Z = & \int \mathcal{D}[c^\dagger c] e^{-S_0} \sum_{k=0}^{\infty} \left(\frac{K}{\beta}\right)^k \frac{1}{k!} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \cdots \int_0^\beta d\tau_k \\
 & \times \prod_{i=1}^k \left[ 1 - \frac{U\beta}{K} \left( n_\uparrow(\tau_i) n_\downarrow(\tau_i) - \frac{n_\uparrow(\tau_i) + n_\downarrow(\tau_i)}{2} \right) \right]
 \end{aligned} \tag{62}$$

The crucial step is now the discrete Hubbard-Stratonovich transformation, which is slightly different than in the usual determinantal MC

$$1 - \frac{\beta U}{K} \left( n_\uparrow(\tau) n_\downarrow(\tau) - \frac{n_\uparrow(\tau) + n_\downarrow(\tau)}{2} \right) = \frac{1}{2} \sum_{s=\pm 1} e^{\lambda s (n_\uparrow - n_\downarrow)} \tag{63}$$

with  $\cosh \lambda = 1 + \frac{\beta U}{2K}$ .

With this transformation, we can rewrite the action as

$$Z = \int \mathcal{D}[c^\dagger c] e^{-S_0} \sum_{k=0}^{\infty} \left(\frac{K}{2\beta}\right)^k \frac{1}{k!} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \cdots \int_0^\beta d\tau_k \times \prod_{i=1}^k \sum_{s_i=\pm 1} e^{\lambda s_i (n_\uparrow(\tau_i) - n_\downarrow(\tau_i))} \quad (64)$$

or

$$Z = \sum_{k=0}^{\infty} \left(\frac{K}{2\beta}\right)^k \frac{1}{k!} \sum_{\{s\}_k} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \cdots \int_0^\beta d\tau_k Z_k[\{s\}_k, \{\tau\}_k] \quad (65)$$

where

$$Z_k[\{s\}_k, \{\tau\}_k] = \int \mathcal{D}[c^\dagger c] e^{-S_0} e^{\sum_{l=1}^k \lambda s(\tau_l) (n_\uparrow(\tau_l) - n_\downarrow(\tau_l))} \quad (66)$$

Now we recognize that  $Z_k[\{s\}_k, \{\tau\}_k]$  corresponds exactly to our determinantal MC action (see Eq. 20), in which the time discretization was chosen to be fixed, with  $k$  equal time slices. We then recognize that our previous derivation can be used to compute the partition function.

More precisely, the partition function evaluates to

$$Z_k[\{s\}_k, \{\tau\}_k] = \text{Det} (1 + B_k B_{k-1} \cdots B_1) \quad (67)$$

with

$$B_l = e^{-\Delta\tau_l h^0 + \lambda s_l \sigma_z \delta_{i=j=0}} \quad (68)$$

Compare this to Eqs. 32 and 33.

While the similarity between this expression and determinantal MC is close, the simulation is quite different. In CTAUX one needs to allow for variable number of time slices, and arbitrary distance between these times. This is the essence of continuous time Monte Carlo, namely, that the time discretization is not being fixed at the outset, but one allows for an arbitrary configuration of times being sampled.

Lets go back to Hirsh-Fye and summarize the equations we derived

$$G_\sigma = \frac{\sum_{\{s\}} \text{Det}[g_\uparrow^{-1}] \text{Det}[g_\downarrow^{-1}] g_\sigma}{\sum_{\{s\}} \text{Det}[g_\uparrow^{-1}] \text{Det}[g_\downarrow^{-1}]} . \quad (69)$$

$$\mathcal{A} g' = g \quad (70)$$

$$\mathcal{A} = 1 + (1 - g)(e^{V'} - 1) \quad (71)$$

$$e^{V_i} = e^{\lambda s_i \sigma} \quad (72)$$

The non-interacting Green's function  $G^0$  obtained from the original impurity Hamiltonian by switching off the interaction ( $U = 0$ ) was written in Eq. (12) and is

$$G^0 = (\omega - \epsilon_0 - \sum_{p>0} \frac{V_{0p}^* V_{p0}}{\omega - \epsilon_p})^{-1}$$

We can get Green's function  $g$  at any Ising configuration  $(s_0, s_1, \dots, s_{L-1})$  from  $G^0$  using Eq. (61) since  $G^0$  corresponds to the case  $e^V = 1$ .

Note that all equations above are written for impurity site only and  $g_\sigma$  are just matrices in time index (not in site index).  $\mathcal{A}$  and  $V$  are also matrices in imaginary time.

We will use two types of updates for Green's function  $g$ : single-spin flip and clean update.

If Ising configurations  $\{s\}'$  and  $\{s\}$  differ in only one spin (lets call it  $s_p$ ),  $g'$  can be

obtained from  $g$  without the inversion of the  $L \times L$  matrix  $\mathcal{A}$ .

Matrix  $e^{V'-V}$  is

$$e^{V'-V} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & e^{\lambda\sigma(s_{p'}-s_p)} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \quad (73)$$

or in short  $(e^{V'-V} - 1)_{ll'} = a\delta_{l=l'=p}$  with  $a \equiv (e^{\lambda\sigma(s_{p'}-s_p)} - 1)$

From Eq. 62, we can obtain  $\mathcal{A}$  which is  $\mathcal{A}_{ll'} = \delta_{ll'} + a(\delta_{lp} - g_{lp})\delta_{l'p}$ . The inversion of matrix  $\mathcal{A}$  can be explicitly write down and takes the form

$$(\mathcal{A}^{-1})_{ll'} = \delta_{ll'} + \frac{a}{1 + a(1 - g_{pp})} (g_{lp} - \delta_{lp})\delta_{l'p} \quad (74)$$

In matrix form it takes the form

$$\mathcal{A} = I - a \begin{pmatrix} 0 & \cdots & g_{0p} & \cdots & 0 \\ 0 & \cdots & g_{1p} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & g_{pp} - 1 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & g_{L-1p} & \cdots & 0 \end{pmatrix} \rightarrow \mathcal{A}^{-1} = I + \frac{a}{1 + a(1 - g_{pp})} \begin{pmatrix} 0 & \cdots & g_{0p} & \cdots & 0 \\ 0 & \cdots & g_{1p} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & g_{pp} - 1 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & g_{L-1p} & \cdots & 0 \end{pmatrix} \quad (75)$$



This identity can be derived from a more general ShermanMorrison formula, which reads

$$(A + u \otimes v^T)^{-1} = A^{-1} - \frac{A^{-1}u \otimes v^T A^{-1}}{1 + v^T \cdot A^{-1} \cdot u} \quad (76)$$

If we take  $A = I$ ,  $u_l = a(\delta_{lp} - g_{lp})$  and  $v_l^T = \delta_{lp}$  we arrive at the Eq. 65.

The update operation for single spin flip is therefore  $g' = \mathcal{A}^{-1}g$ :

$$g_{l_1 l_2}' = g_{l_1 l_2} + \frac{a}{1 + a(1 - g_{pp})} (g_{l_1 p} - \delta_{l_1 p}) g_{p l_2} \quad (77)$$

This is simple rank-1 update implemented in BLAS routine "dgerr" ( $A = A + \alpha x * y'$ ) and the QMC code can be made very efficient just by using this blas routine. No further optimization of the QMC code is necessary in this case, since most of the time is spend in blas routine and therefore one gets theoretical performance (number of operations equal to the number of clock ticks of the processor).

We will also need the determinant of  $\mathcal{A}$  since this will give us the transition probability. It is simple to see that

$$\text{Det}\mathcal{A} = 1 + (1 - g_{pp})a \quad (78)$$

Due to numerical roundoff error, it is advisable to make a "clean" update from time to time.

It this case we just use Eqs. (61),(62) to get from non-interacting  $G^0$  to  $g$  of the current Ising configuration

$$\mathcal{A} = 1 + (1 - G^0)(e^V - 1) \quad (79)$$

$$\text{or } \mathcal{A}_{l_1 l_2} = \delta_{l_1 l_2} (1 + (e^{\lambda \sigma_{s_{l_2}}} - 1)) - G_{l_1 l_2}^0 (e^{\lambda \sigma_{s_{l_2}}} - 1)$$

How is Monte Carlo sampling used in this approach?

From Eq. 60, we see that  $G$  is weighted average of  $g$  over all possible Ising configurations

$$G_\sigma = \frac{\sum_{\{s\}} \rho g_\sigma}{\sum_{\{s\}} \rho} \quad (80)$$

$$\rho = \text{Det}[g_\uparrow^{-1}] \text{Det}[g_\downarrow^{-1}] \quad (81)$$

The simplest QMC step is a spin flip of randomly chosen Ising spin. From analogy with the classical Ising system it is clear that each configuration can be reached from any configuration in finite number of steps (*connectedness*).

The acceptance probability will be handled by Metropolis algorithm

$$A(\{s\} \rightarrow \{s'\}) = \min \left[ \frac{\rho[\{s'\}]}{\rho[\{s\}]}, 1 \right]$$

Here we see the "fermionic minus sign" problem. If it happens that the above product of determinants is negative, we can not use it as probability density. The problem is at present unsolved. However, it turns out that for single impurity problem, determinants are always positive. This was rigorously proved very recently.

For single spin flip, the ratio of distributions  $\rho'/\rho$  can be further simplified. It is clear from

Eq. (61) that

$$\text{Det}[g'^{-1}]/\text{Det}[g^{-1}] = \text{Det}\mathcal{A} = 1 + (1 - g_{pp})(e^{\lambda\sigma(s_p' - s_p)} - 1)$$

and therefore

$$\rho'/\rho = [1 + (1 - g_{pp}^{\uparrow})(e^{\lambda(s_p' - s_p)} - 1)][1 + (1 - g_{pp}^{\downarrow})(e^{-\lambda(s_p' - s_p)} - 1)] \quad (82)$$

## The algorithm for Hirsch-Fye simulation is

- Read input non-interacting Green's function  $G^0$ . Make a matrix of  $L \times L$  out of  $G^0$  by  $G^0(l_1, l_2) = G^0(\tau_{l_1} - \tau_{l_2})$ . Be careful about antiperiodicity of  $G^0(\tau + \beta) = -G^0(\tau)$ .
- Choose starting Ising configuration (if "good" configuration is not known from some previous run, start with completely random configuration)
- Calculate  $g$  with "clean update" from  $G^0$  :  $g_\sigma = \mathcal{A}^\sigma G_\sigma^0$  with

$$\mathcal{A}_{l_1 l_2}^\sigma = \delta_{l_1 l_2} e^{\lambda \sigma s_{l_2}} - G_{l_1 l_2}^0 (e^{\lambda \sigma s_{l_2}} - 1)$$

- Repeat these steps until enough measurements are collected ( $\sim 10^5 L$ )
  - Choose imaginary time slice  $p$  randomly and accept spin flip at  $p$  with transition probability according to Metropolis ( $A(s \rightarrow s') = \min[\rho'/\rho, 1]$ ). We have

$$\rho'/\rho = [1 + (1 - g_{pp}^\uparrow)(e^{\lambda(s_p' - s_p)} - 1)][1 + (1 - g_{pp}^\downarrow)(e^{-\lambda(s_p' - s_p)} - 1)]$$

- Calculate new  $g'$  from  $g$  by

$$g_{l_1 l_2}^{\sigma'} = g_{l_1 l_2}^\sigma + \frac{a}{1 + a(1 - g_{pp}^\sigma)} (g_{l_1 p}^\sigma - \delta_{l_1 p}) g_{p l_2}^\sigma$$

Use rank-1 update BLAS routine "dgerr".

- If the number of accepted single spin flip steps exceeds certain number ( $\sim 100$ ), make clean update again.
- If the distribution reaches equilibrium - number of all steps exceeds number of warmup steps ( $\sim 100L$ ), start collecting measurements.
- Save new measurement (of  $g$ ) after correlation time has passed ( $\sim 3L$ ).
- Calculate averages and standard deviation of correlation functions (of  $G$ ).
- Calculate  $G(\tau)$  out of  $L \times L$  matrix  $G(l_1, l_2)$ .

## Grouping of measurements into bins

The measurements in Markov chain are highly correlated. Even though measurements are recorded only after few hundred spin-flips, the correlations between measurements are still very large. As a result, the standard deviation calculated with the familiar formula

$$\sigma^2 = \frac{1}{N_d - 1} (\langle G^2 \rangle - \langle G \rangle^2) \quad (83)$$

is orders of magnitude smaller than the actual sampling error. There is a simple way to improve the estimate of the QMC error. The idea is to group large number of successive measurements into one bin and then use the above formula for standard deviation of bins. A good choice for bin size is  $N_b \sim 1000$  and one needs of the order of 100 bins to estimate the error and covariant matrix.

To determine the best bin size one can monitor standard deviation as a function of  $N_b$ . It turns out that it increases with increasing bin-size and saturates for large  $N_b$ . The optimal  $N_b$  can be chosen close to saturation of  $\sigma^2$ .

## Fourier transformation and analytic continuation

The input  $G^0$  is usually in imaginary frequency  $i\omega_n$  and one needs to make inverse Fourier transform

$$G(\tau) = T \sum_{i\omega_n} G(i\omega_n) e^{-i\omega_n \tau}.$$

The problem is that Green's function falls off as  $1/(i\omega_n)$  and therefore the above sum converges very badly. The usual trick is to subtract  $1/(i\omega_n)$  term from  $G$  and treat the tail analytically. One therefore evaluates

$$G(\tau) = T \sum_{i\omega_n} \left( G(i\omega_n) - \frac{1}{i\omega_n} \right) e^{-i\omega_n \tau} + T \sum_{i\omega_n} \frac{1}{i\omega_n} e^{-i\omega_n \tau} \quad (84)$$

We will show below that the second term is equal to  $-1/2$ . One can further simplify the sum by realizing that the real part of  $G(i\omega_n)$  is even function of  $i\omega_n$  and imaginary part is odd. We have

$$G(\tau) = \sum_{i\omega_n} \left( G'(i\omega_n) + i \left[ G''(i\omega_n) + \frac{1}{\omega_n} \right] \right) e^{-i\omega_n \tau} - \frac{1}{2} \quad (85)$$

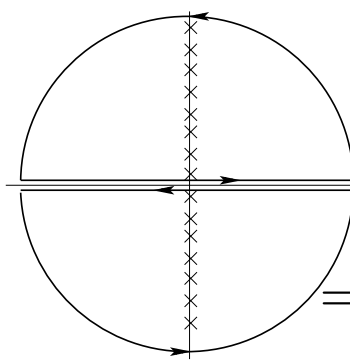
$$= \sum_{i\omega_n > 0} 2 \left[ G'(i\omega_n) \cos \omega_n \tau + \left( G''(i\omega_n) + \frac{1}{\omega_n} \right) \sin \omega_n \tau \right] - \frac{1}{2} \quad (86)$$



Here we want to show that

$$T \sum_{i\omega_n} \frac{1}{i\omega_n} e^{-i\omega_n \tau} = -1/2.$$

We derive this as a special case of a standard contour integration in complex plane.



$$T \sum_{i\omega_n} g(i\omega_n) e^{-i\omega_n \tau} = \oint \frac{dz}{2\pi i} \begin{cases} f(-z) e^{-z\tau} g(z) \\ -f(z) e^{-z\tau} g(z) \end{cases} \quad (87)$$

$$= \begin{cases} 0 < \tau < \beta & \int_{-\infty}^{\infty} \frac{d\omega}{\pi} f(-\omega) e^{-\omega\tau} g''(\omega + i\delta) \\ -\beta < \tau < 0 & - \int_{-\infty}^{\infty} \frac{d\omega}{\pi} f(\omega) e^{-\omega\tau} g''(\omega + i\delta) \end{cases} \quad (88)$$

In special case of  $g(i\omega) = \frac{1}{i\omega}$  we have  $g''(\omega) = -\pi\delta(\omega)$  and the results is  $-1/2$  for  $\tau > 0$  and  $1/2$  for  $\tau < 0$ .

Usually, one needs Green's function in imaginary (or real) frequency rather than imaginary time. The Fourier transformation (analytic continuation) is required in this case. This is very troublesome step and needs a lot of technical effort.

It amounts to solve

$$G(i\omega) = \int_0^\beta d\tau e^{i\omega\tau} G(\tau) \quad (89)$$

$$G(\tau) = - \int d\omega f(-\omega) e^{-\omega\tau} A(\omega) \quad (90)$$

The second equation connects  $G(\tau)$  with the spectral function. It can be derived as a special case of the above contour integral. Inserting for  $g$  Green's function and taking into account  $A(\omega) = -\frac{1}{\pi} G''(\omega)$  we arrive at (90). We need the inverse of the equation (90): to obtain  $A(\omega)$  from  $G(\tau)$ . This is very nontrivial since the kernel of the above integral equation is highly singular because it cutts-off high frequencies.

For Fourier transformation, the FFT is useless because Green's function in  $\omega$  is not periodic function but rather falls of as  $1/\omega$ . The trouble is caused by high-frequencies again since this information is not very well captured in noisy  $G(\tau)$ .

The simplest approach to Fourier transformation is by assuming linear interpolation of  $G(\tau)$  between two successive imaginary time slices. The result is

$$G'(i\omega_n) = \frac{1}{\omega_n} \left( \sin(\omega_n \tau_{l+1}) G_{l+1} - \sin(\omega_n \tau_l) G_l + \frac{1}{\omega_n} \frac{G_{l+1} - G_l}{\tau_{l+1} - \tau_l} [\cos(\omega_n \tau_{l+1}) - \cos(\omega_n \tau_l)] \right) \quad (91)$$

$$G''(i\omega_n) = \frac{1}{\omega_n} \left( \cos(\omega_n \tau_l) G_l - \cos(\omega_n \tau_{l+1}) G_{l+1} + \frac{1}{\omega_n} \frac{G_{l+1} - G_l}{\tau_{l+1} - \tau_l} [\sin(\omega_n \tau_{l+1}) - \sin(\omega_n \tau_l)] \right) \quad (92)$$

The problem of this approach is that frequencies larger than  $\omega > \frac{1}{\Delta\tau}$  are completely wrong and one gets nonphysical oscillations with frequency  $\frac{1}{\Delta\tau}$ .

The standard approach is to spline  $G(\tau)$  by cubic spline function. Since than one knows the function analytically (at least locally) one can compute Fourier transformation analytically and arrives at somewhat more complicated but similar equation than Eq. (92).

The standard approach for analytic continuation is so-called Maximum Entropy Method. The naive way to solve for  $A(\omega)$  in Eq. (90) would be to minimize the distance between the QMC measurements  $\bar{G}_l$  and  $G(\tau_l)$  obtained from  $A$  with the help of equation (90)

$$\chi^2 = \sum_l (G(\tau_l) - \bar{G}_l)^2 \quad (93)$$

$$G(\tau_l) = - \int d\omega f(-\omega) e^{-\omega \tau_l} A(\omega) \quad (94)$$

and minimization determines  $A(\omega)$ . However, the problem is that the kernel in Eq. (90) is singular and  $G(\tau)$  is completely insensitive to high-frequencies in  $A(\omega)$ . Even worse is the fact that QMC measurements have a lot of noise and this kind of minimization would overfit QMC noise and the resulting  $A(\omega)$  would have weird unphysical oscillations.

The idea of Maximum entropy method is to add a damping term to  $\chi$  in order not to overfit the QMC noisy data. The functional to be minimized is

$$\chi^2[A] = \sum_l (G(\tau_l) - \bar{G}_l) [C^{-1}]_{ll'} (G(\tau_{l'}) - \bar{G}_{l'}) + \alpha S[A] \quad (95)$$

with  $S[A]$  being entropy like term which does not allow  $A(\omega)$  to be far from default model  $m(\omega)$  (usually taken to be a constant)

$$S[A] = \int d\omega [A(\omega) - m(\omega) - A(\omega) \log A(\omega)/m(\omega)] \quad (96)$$

Here  $C$  is the covariant matrix and appears because measurements are not statistically independent, and  $\alpha$  is parameter to be varied. In historical MEM,  $\alpha$  is chosen to be equal to number of real frequency points of  $A(\omega)$ . More elaborated schemes are currently used in literature. See M. Jarrell, and J. E. Gubernatis, 1996, Phys. Rep. 269, 133 (1996).

## Multiband models

The algorithm can be easily generalized to more complicated impurity with some more degrees of freedom. As long as interaction is of the density-density type  $U nn$ , the decoupling is straightforward and the algorithm is free from minus-sign problem.

Below, the case of two bands is sketched.

The Hamiltonian for the  $SU(N)$  problem is

$$H = \sum_q \epsilon_{q0} c_{0q}^\dagger c_{0q} + \frac{1}{2} \sum_{q \neq q'} U n_{0q} n_{0q'} + \sum_{p > 0q} [V_{0p} c_{0q}^\dagger c_{pq} + V_{0p}^* c_{pq}^\dagger c_{0q}] + \sum_{p > 0, q} \epsilon_p c_{pq}^\dagger c_{pq} \quad (97)$$

where  $q$  runs over  $N$  spin-orbitals for  $SU(N)$  case. The special  $SU(2)$  case was addressed above. In the two band case, the interaction term can be written in the form

$$H^i = U_1(n_1 n_2 - \frac{1}{2}(n_1 + n_2)) + U_2(n_1 n_3 - \frac{1}{2}(n_1 + n_3)) + U_3(n_1 n_4 - \frac{1}{2}(n_1 + n_4)) \\ + U_4(n_2 n_3 - \frac{1}{2}(n_2 + n_3)) + U_5(n_2 n_4 - \frac{1}{2}(n_2 + n_4)) + U_6(n_3 n_4 - \frac{1}{2}(n_3 + n_4))$$

There are  $N_f \equiv N(N-1)/2$  interaction terms and each term needs one Ising spin. The number of Ising spins is therefore  $LN_f$ . The Hubbard-Stratonovich decoupling at each

time slice in the two band case is

$$\begin{aligned}
 e^{-\Delta\tau H^i} &= \frac{1}{2^6} \sum_{s_1, \dots, s_6} \exp [\lambda_1 s_1 (n_1 - n_2) + \lambda_2 s_2 (n_1 - n_3) + \lambda_3 s_3 (n_1 - n_4) \\
 &\quad + \lambda_4 s_4 (n_2 - n_3) + \lambda_5 s_5 (n_2 - n_4) + \lambda_6 s_6 (n_3 - n_4)] \\
 &= \frac{1}{2^{N_f}} \sum_{s_1, \dots, s_{N_f}} e^{\sum_{ij} \lambda_i s_i f_{ij} n_j} \tag{100}
 \end{aligned}$$

with  $2 \cosh \lambda_i = e^{-\Delta\tau U_i}$ ,  $N_f = 6$  and  $f_{ij}$  a matrix

$f_{ij}$	1	2	3	4
1	1	-1	0	0
2	1	0	-1	0
3	1	0	0	-1
4	0	1	-1	0
5	0	1	0	-1
6	0	0	1	-1

(101)

For  $SU(N)$ , all  $U_i$  are equal to  $U$ . The Hirsch-Fye equations are almost unchanged from  $SU(2)$  case with the only important difference that the number of Ising spins over which one needs to sample is increased from  $L$  to  $N_f L$ .

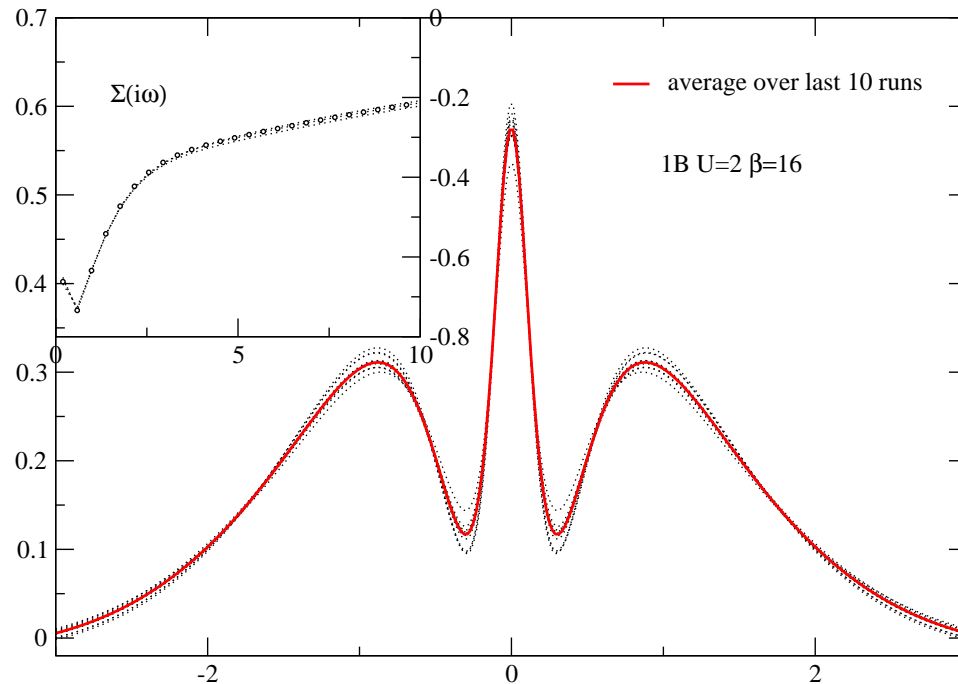


Figure 1: The spectral functions of the DMFT solution for the Hubbard model on the Bethe lattice. Last 10 DMFT iterations are plotted and the red curve shows the average spectral function, averaged over last 10 DMFT steps. The inset shows the imaginary axis self-energy which is very smooth and precise.

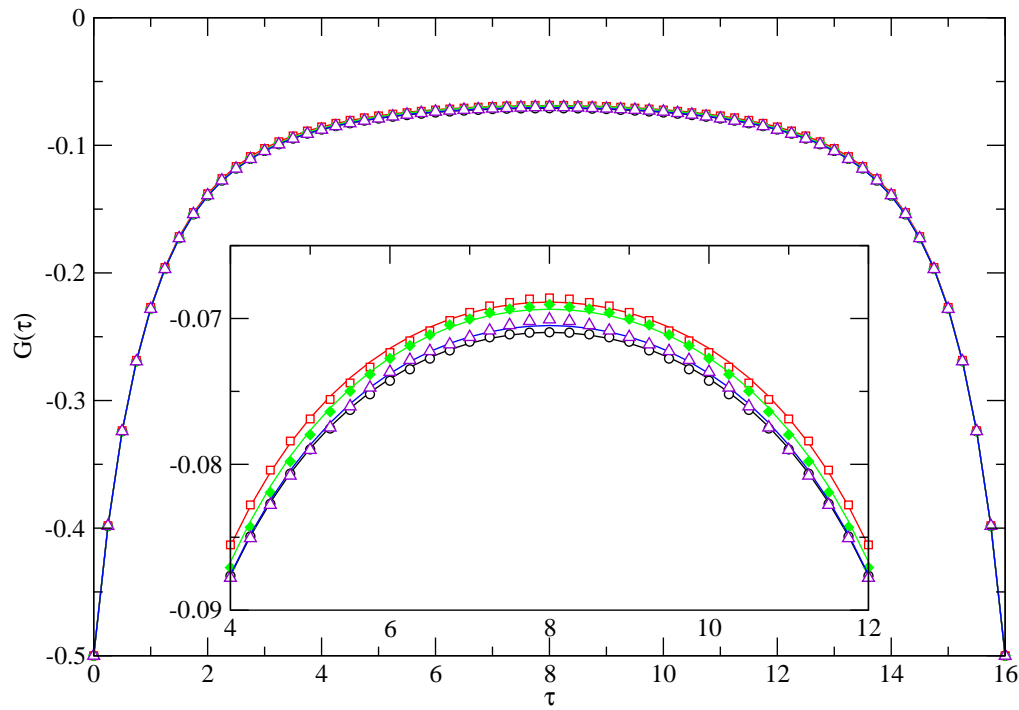


Figure 2: QMC imaginary time Green's function  $G(\tau)$  for few latest DMFT iterations. QMC never converges to extremely high accuracy because of QMC noise (not perfect statistics). And here one can see few latest iterations how they change in imaginary time and how good the fit of 8 SVD-functions is. The symbols stand for QMC data and lines connect the interpolation with those SVD functions.



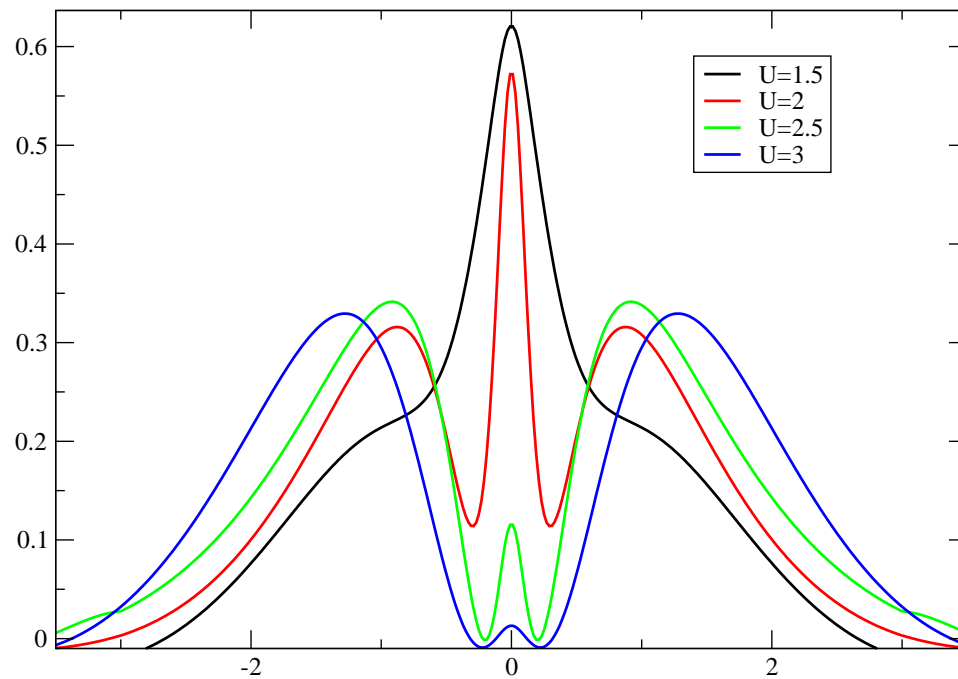


Figure 3: Metal insulator transition within DMFT. Here  $\beta = 16$  therefore the coexistence region is not detected.

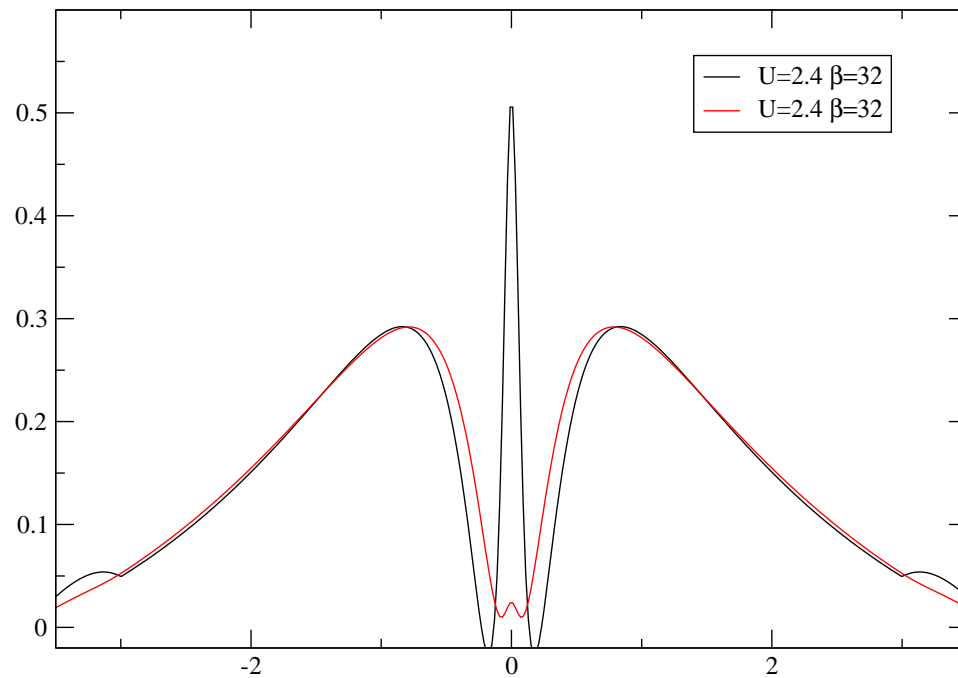


Figure 4: This is a low temperature run with  $\beta = 32$  and 128 time-slices showing the coexistence of solutions at  $U = 2.4$ .

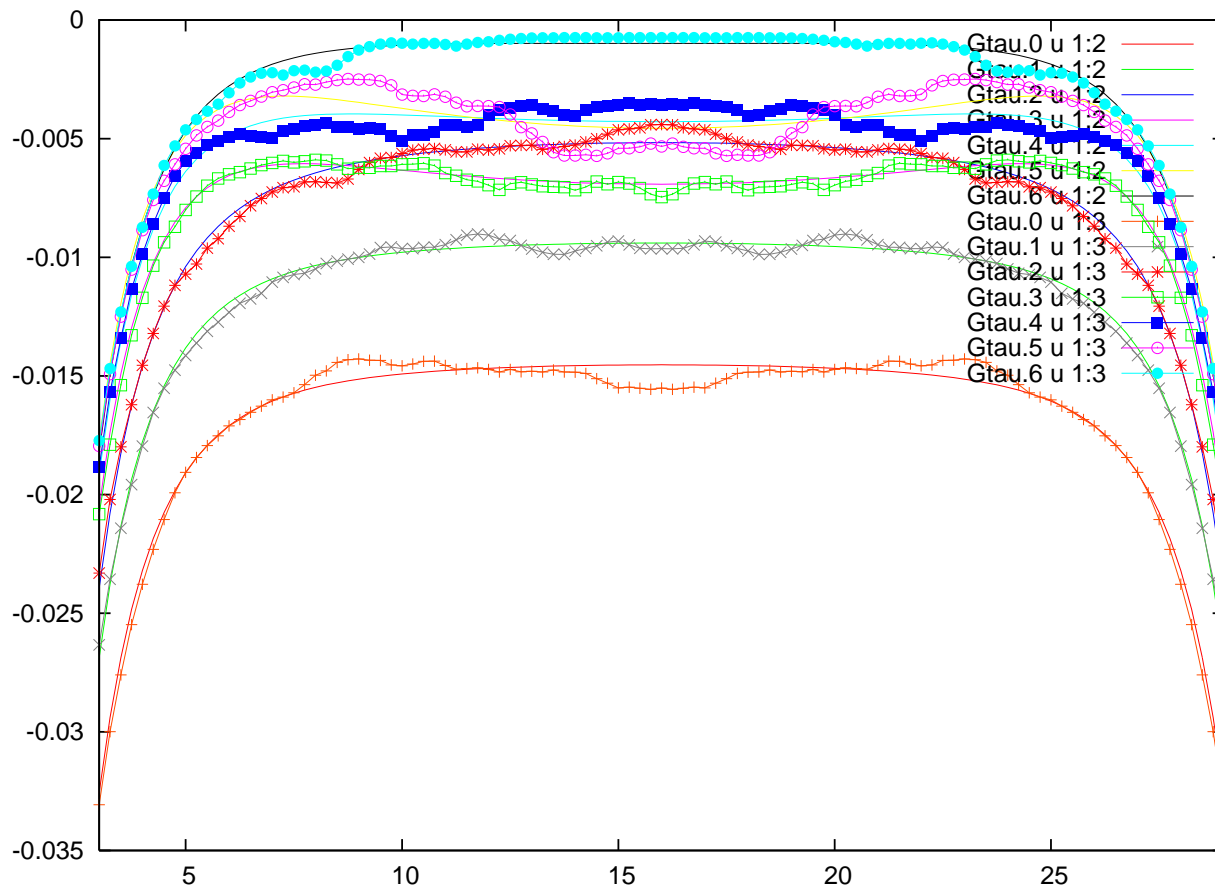


Figure 5: This is an example of low temperature run where one has 128 time slices and if one zooms in the central region, it is obvious that the oscillations of QMC data are pretty bad.