COMPUTATIONAL QUANTUM MAGNETISM

Boulder 2003 Summer School

Lecture Three- Determinant Quantum Monte Carlo

A. Introduction

In the preceding lecture, we emphasized that the world-line QMC algorithm, which is very effective for quantum spin and boson systems (in the absence of frustration) fails for fermions in more than one dimension because it is impossible to make the matrix elements, and hence the weight for the world-line configurations, positive. Fortunately, there is an alternative, the determinant QMC method. In general, this approach also has a sign problem, but there are certain important special cases, such as the half-filled Hubbard model, where the sign problem does not occur. Even when the sign problem arises, it becomes a problem at relatively low temperature, and one can often extract interesting physics before the method breaks down.

The determinant QMC method has certain characteristic features. Unlike monte carlo methods which have a local action and whose update time per degree of freedom is independent of lattice size, the fermion action is non-local. Updating one of the field variables takes a cpu time which grows as the square of the lattice size N (number of electrons). Hence to update all the field variables is a process which scales as the cube of N. Current simulations are of several hundred electrons an order of magnitude larger than possible with exact diagonalization. Because the fermion Green's function is the central quantity in the method, determinant QMC also allows detailed contact with analytic many body methods.

While the determinant QMC method has not seen the refinements that world-line QMC has undergone in the last several years, there is one very important development, which is its application to dynamical mean field theory (DMFT). DMFT is an approach to interacting electron systems which allows the electron self-energy Σ to fluctuate in imaginary time (that is, Σ retains its full frequency dependence), while ignoring spatial fluctuations (momentum dependence). While it initially was applied to model Hamiltonians, DMFT is increasingly being integrated into electronic structure calculations within the local density approximation (LDA). The use of determinant QMC in the context of "LDA+DMFT" is clearly one of the most important frontier areas of QMC at the moment.

Finally, it is also worth noting that determinant QMC is the general methodology behind Lattice Gauge Theory, although there are, of course, significant differences in the details.

B. Multidimensional Gauusian Integration

The equations involved in determinant QMC bear many similarities with multidimensional Gaussian integrals. Reviewing these identities will help provide an intuitive feel for the formulae of determinant QMC, within a familiar context.

The generalization of the familiar one dimensional Gaussian integral,

$$\int_{-\infty}^{+\infty} dx \, e^{-ax^2} = \frac{\sqrt{\pi}}{a}, \tag{1}$$

to many dimensions is,

$$Z = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1 dx_2 \ldots dx_N \, e^{-\vec{x}A\vec{x}^T} = \frac{\pi^{n/2}}{\sqrt{\det A}}. \tag{2}$$

Here $\vec{x}$ is an N dimensional vector of real numbers and A is a real, symmetric, N dimensional matrix. I have used the notation Z for the integral to emphasize that it would be the partition function for a set of classical variables whose action is given by the quadratic form $\vec{x}A\vec{x}^T$.

We also know how to do these integrals when the integrand includes factors of x_i.

$$\langle x_i x_j \rangle = Z^{-1} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1 dx_2 \ldots dx_N \, x_i x_j \, e^{-\vec{x}A\vec{x}^T} = \frac{1}{2}[A^{-1}]_{ij} \tag{3}$$

Again, the notation $\langle x_i x_j \rangle$ emphasizes a possible statistical mechanical interpretation of the ratio of integrals.

Further factors of x_i in the integrand generate expressions which are similar in form to 'Wick's Theorem', which tells us that contractions of products of many fermion operators can be expressed as sums of products of contractions taken two operators at a time, in all possible permutations.

$$\begin{aligned}\langle x_i x_j x_k x_l \rangle &= Z^{-1} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty} dx_1 dx_2 \ldots dx_N x_i x_j x_k x_l e^{-\vec{x}A\vec{x}^T} \\ &= \frac{1}{4}([A^{-1}]_{ij}[A^{-1}]_{kl} + [A^{-1}]_{ik}[A^{-1}]_{jl} + [A^{-1}]_{il}[A^{-1}]_{jk})\end{aligned} \tag{4}$$

While it is possible to do these integrals with arbitrary polynomials as part of the integrand, they cannot be done when a quartic term appears in the exponential. We shall see shortly the analogies of these various statements for traces over fermion Hamiltonians.

C. Basic Formalism of Determinant QMC

In solving the Hubbard model we want to evaluate expressions like

$$\begin{aligned}\langle A \rangle &= Z^{-1} \operatorname{Tr}[Ae^{-\beta H}] \\ Z &= \operatorname{Tr} e^{-\beta H}\end{aligned} \tag{5}$$

As we emphasized in lecture I, the "Tr" is a trace over the 4^N dimensional Hilbert space, where N is the number of sites: The exact diagonalization method explicitly contructs the matrix for H in an occupation number basis.

In analogy with multidimensional Gaussian integration, we can do such traces if they are over *quadratic* forms of fermion operators. Suppose

$$H = (c_{1\sigma}^\dagger \quad c_{2\sigma}^\dagger \quad . \quad .) \begin{pmatrix} h_{11} & h_{12} & . & . \\ h_{21} & h_{22} & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix} \begin{pmatrix} c_{1\sigma} \\ c_{2\sigma} \\ . \\ . \end{pmatrix}. \tag{6}$$

Here h is an NxN matrix. The identity is,

$$Z = \operatorname{Tr} e^{-\beta H} = \det[I + e^{-\beta h}]. \tag{7}$$

Note that while the original "Tr" is over a quantum mechanical 4^N dimensional Hilbert space, the "det" is a usual determinant of NxN matrices. "I" is the N dimensional identity matrix and "h" is the matrix of *numbers* entering the definition of H. It is worth emphasizing that because we are taking the trace over the full 4^N dimensional Hilbert space, we are including states of all occupation numbers. The determinant QMC method, as formulated here, works in the grand canonical ensemble. Particle density is controlled by changing the chemical potential. This contrasts with our formulation of the world line method which works in the canonical ensemble (fixed particle number).

It is trivial to check that Eq. 7 holds for a single fermion degree of freedom, with Hamiltonian $H = \epsilon c^\dagger c$. There are two states in the Hilbert space and

$$Z = \langle 0|e^{-\beta \epsilon c^\dagger c}|0\rangle + \langle 1|e^{-\beta \epsilon c^\dagger c}|1\rangle = 1 + e^{-\beta \epsilon}. \tag{8}$$

More generally (e.g. for more than one fermion degree of freedom) Eq. 7 can be verified by going to the basis where h is diagonal. However, the equations are best derived by employing the techniques of Grassman integration.

There is a more general identity. If one has a *set* of quadratic forms $l = 1, 2, \ldots L$

$$H(l) = (c_{1\sigma}^\dagger \quad c_{2\sigma}^\dagger \quad . \quad .) \begin{pmatrix} h(l)_{11} & h(l)_{12} & . & . \\ h(l)_{21} & h(l)_{22} & . & . \\ . & . & . & . \\ . & . & . & . \end{pmatrix} \begin{pmatrix} c_{1\sigma} \\ c_{2\sigma} \\ . \\ . \end{pmatrix}, \tag{9}$$

then,

$$Z = \operatorname{Tr} \left[e^{-\Delta\tau H(1)} e^{-\Delta\tau H(2)} \ldots e^{-\Delta\tau H(L)} \right] = \det\left[I + e^{-\Delta\tau h(1)} e^{-\Delta\tau h(2)} \ldots e^{-\Delta\tau h(L)} \right]. \tag{10}$$

Here I have changed the prefactor in the exponential from β to $\Delta\tau$ for reasons which will soon be clear. It is also true that,

$$\begin{aligned} G_{ij} = \langle c_{i\sigma} c_{j\sigma}^\dagger \rangle &= Z^{-1} \operatorname{Tr} \left[c_{i\sigma} c_{j\sigma}^\dagger \, e^{-\Delta\tau H(1)} e^{-\Delta\tau H(2)} \ldots e^{-\Delta\tau H(L)} \right] \\ &= \left[I + e^{-\Delta\tau h(1)} e^{-\Delta\tau h(2)} \ldots e^{-\Delta\tau h(L)} \right]_{ij}^{-1}. \end{aligned} \tag{11}$$

The "fermions Greens function" is just an appropriate matrix element of the inverse of the NxN matrix whose determinant gives the partition function. These formulae also are best verified by the techniques of Grassmann integration.

3

The above formulae describe how to perform traces over quadratic forms of fermion degrees of freedom. Unfortunately, the Hubbard Hamiltonian has an interaction term $Un_{i\uparrow}n_{i\downarrow} = Uc_{i\uparrow}^{\dagger}c_{i\uparrow}c_{i\downarrow}^{\dagger}c_{i\downarrow}$ which is quartic in the fermion operators. To handle such terms, we employ the (discrete) Hubbard–Stratonovich transformation,

$$e^{-U\Delta\tau(n_{\uparrow}-\frac{1}{2})(n_{\downarrow}-\frac{1}{2})} = \frac{1}{2}e^{-\frac{U\Delta\tau}{4}}\sum_{S}e^{\lambda S(n_{\uparrow}-n_{\downarrow})} \tag{12}$$

Here $\cosh\lambda = e^{U\Delta\tau/2}$, and S is an Ising variable which can take on the two values $S = \pm1$. This identity can be verified by explicitly enumerating the 4 possible choices for $n_{\uparrow}, n_{\downarrow}$.

Now we divide $\beta = L\Delta\tau$ and employ the Trotter decomposition. As in the discussion of world line QMC in lecture II, this allows us to isolate different pieces of the Hamiltonian. But rather than introducing complete sets of states, our approach here is to employ the Hubbard-Stratonovich transformation. We write $H = K + V$ where K contains all the one–body pieces and V the on–site Hubbard interaction. Then,

$$Z = \text{Tr}\,e^{-\beta H} = \text{Tr}\,[\,e^{-\Delta\tau K}\,e^{-\Delta\tau V}\,e^{-\Delta\tau K}\,e^{-\Delta\tau V}\,\ldots]. \tag{13}$$

The $e^{-\Delta\tau K}$ are quadratic in the fermion operators. For each factor of the L terms $e^{-\Delta\tau V}$ above, we introduce N Hubbard–Stratonovich fields, one for each of the spatial sites where we have an on–site interaction to decouple. The Hubbard–Stratonovich field $S(i, l)$ therefore has two indices, space i and imaginary–time l. Now the $e^{-\Delta\tau V(l)}$ are also quadratic in the fermion operators. We put an argument l on V to emphasize that while the K are all identical, the $V(l)$ contain different Hubbard–Stratonovich fields on the different imaginary time slices.

Applying the preceding identities allows the analytic evaluation of the trace,

$$Z = \sum_{S(il)} \det M_{\uparrow}\,\det M_{\downarrow}. \tag{14}$$

We get a determinant for each of the two spin species. The quantum partition function has now been expressed to a *classical* monte carlo problem: We need to sum over the possible configurations of the real, classical, variables $S(i, l)$ with the "Boltzmann weight" which is the product of the two fermion determinants. Note that as in world-line QMC, the classical variable to be summed over has an additional index l labeling imaginary time.

We will write down the explicit forms of the matrices for the one–dimensional Hubbard model.

$$M_{\sigma} = I + e^{-k}e^{-v_{\sigma}(1)}e^{-k}e^{-v_{\sigma}(2)}\ldots e^{-k}e^{-v_{\sigma}(L)}. \tag{15}$$

Here

$$k = -\Delta\tau \begin{pmatrix} \mu & t & 0 & 0 & . & . \\ t & \mu & t & 0 & . & . \\ 0 & t & \mu & t & . & . \\ 0 & 0 & t & \mu & . & . \end{pmatrix} \tag{16}$$

4

and

$$v_\sigma(l) = \lambda\sigma \begin{pmatrix} S(1,l) & 0 & 0 & 0 & . & . \\ 0 & S(2,l) & 0 & 0 & . & . \\ 0 & 0 & S(3,l) & 0 & . & . \\ 0 & 0 & 0 & S(4,l) & . & . \end{pmatrix} \tag{17}$$

Note that the matrices $v_\sigma(l)$ for different spin species σ differ only in the sign of the coupling to the Hubbard–Stratonovich field.

In higher dimensions, k picks up additional off–diagonal "bands", but the $v_\sigma(l)$ remain diagonal.

D. Summary of Determinant QMC Algorithm

The most primitive determinant QMC code then proceeds as follows:
(1.) Initialize all the Hubbard–Stronovich variables (for example, set them up randomly).
(2.) Compute the matrices M_σ and their determinants.
(3.) Change one or more of the Hubbard–Stratonovich fields, and compute the new matrices M'_σ and their determinants.
(4.) Throw a random number $0 < r < 1$ and "accept" the new configuration with probability min(1,detM'/det$M$), the usual Metropolis algorithm.
(5.) Repeat (1–4). Measurements of the Greens function G are obtained by accumulating M_{ij}^{-1}. Other observables like magnetic susceptibility by appropriate products of matrix elements of M^{-1}.

E. Subtleties and "Tricks of the Trade"

While the above formulae allow you to write a "bare–bones" determinant QMC algorithm, there are a number of refinements which are essential for a "real working code".
(1.) The algorithm, as stated, scales in CPU time as $N^4 L$. The reason is that re–evaluating the determinant of M' takes N^3 operations, and we must do that NL times to sweep through all the Hubbard–Stratonovich variables (if, as is typically done, we change just one at a time). This scaling can be reduced to $N^3 L$. (In what follows I will drop the spin indices.) The idea is to write $M' = M + dM$ and the ratio of determinants as,

$$\det M'/\det M = \det(M^{-1} M') = \det(M^{-1}(M + dM)) = \det(I + G\,dM), \tag{18}$$

with the definition $G = M^{-1}$. It turns out that dM is very simple because when a Hubbard–Stratonovich field is flipped, a single diagonal entry in $v(l)$ changes. Because dM is sparse, the evaluation of $\det(I + G\,dM)$ takes a cpu time independent of N and L! The update decision is no more costly than for a local weight as in a typical classical monte carlo.

However, we need to know $G = M^{-1}$ for this calculation, and once the Hubbard-Stratonovich field change is made, one needs to update G. This updating G does not take N^3 iterations, as one might expect of a matrix inversion, but can be done in only

N^2 operations, again as a result of the simplicity of the change dM. The relevant identity which relates the new $G' = (M + dM)^{-1}$ to the old $G = M^{-1}$ is an application of the "Sherman–Morrison" formula given, for example, in Press's 'Numerical Recipes'. In short, steps (3–4) above are replaced by: (3'–4'.) Change one of the Hubbard–Stratonovich fields, Throw a random number $0 < r < 1$ and "accept" the new configuration with probability min(1,det(I+$G\,dM$)). Recompute G if move accepted. There is a savings of one power of the lattice size in computation time over the primitive approach.

(2.) The matrix G needed in this refinement of the algorithm is precisely the usual equal time fermion Greens function (Eq. 11), familiar from many-body physics, which determines all the interesting physics of the Hamiltonian. Thus measurements are "free"– the matrix inverse used in the updates is the quantity needed for construction of all observables.

(3.) More complicated two–body measurements are made using "Wick's theorem". For example to measure the transverse spin spin correlation $\langle c_{i\downarrow}^\dagger c_{i\uparrow} c_{j\uparrow}^\dagger c_{j\downarrow} \rangle$ one accumulates the product of matrix elements $G_{\uparrow ij} G_{\downarrow ji}$. This is similar to expressions for higher order expectation values in Gaussian integration.

(4.) It is also possible to measure correlation functions with non–zero imaginary time separation, but this requires considerably more work. Analytic continuation of such correlations is required to get the dynamical response. That is quite difficult.

(5.) The big product of matrices required in constructing M is numerically unstable at low temperatures and strong couplings. That is, the product has a very high ratio of largest to smallest eigenvalue. Special "stabilization" is required to do the matrix manipulations. While these add to the complexity of the code, they however have no content in the sense that all the above equations are valid, it is just a question of how best to multiply matrices on a machine of finite precision.

(6.) The determinants of the matrices can go negative. This is called the "fermion sign problem." The sign problem does not occur for certain special cases. For example, if U is negative (the "attractive" Hubbard model), the individual determinants can go negative, but the matrices are always equal and hence the determinant appears as a perfect square. This is a consequence of the fact that the appropriate Hubbard–Stratonovich transformation couples S to the charge $n_\uparrow + n_\downarrow$ as opposed to the spin as given in Eq. 12 for the repulsive model. If U is positive but the chemical potential $\mu = U/2$ ("half–filling") one is also okay. The matrices are not identical in this case, but the determinants are nevertheless related by a positive factor, that is, they again have the same sign, so their product is always positive. Some types of randomness are also acceptable. It is okay for the hoppings t and interactions U to depend on the link or site. These statements are demonstrated by various particle–hole transformations on the Hamiltonian.

(7.) Alternate Hubbard–Stratonovich transformations are possible. One can couple more symmetrically to the spin, that is not single out the z component Or, one can couple to pair creation operators. So far, all such alternatives give a worse sign problem than the transformation Eq. 12.

(8.) Very similar "ground state" determinant simulations exist which work at $T = 0$ and in the canonical ensemble.

F. What QMC Simulations Can Do

The state of the art of QMC simulations, in the absence of a sign problem, are studies of several hundred electrons down to temperatures of $\beta t = 10 - 20$. In terms of temperature and bandwidth, this means T of roughly $1/100$ of the bandwidth $W = 8t$ of the 2-d Hubbard model. This is plenty cold enough to see well developed magnetic correlations. For typical parameters, $t = 1, U = 4$ one chooses $\Delta\tau = 1/8$ so these beta values correspond to roughly $L = 100$, and the simulation involves approximately 10^4 Hubbard–Stratonovich variables.

In cases where one has a sign problem, βt is limited to 4–5. This is, unfortunately, not low enough in temperature to make conclusive statements about certain important problems, perhaps most prominently the question of the existence of long range $d-$wave superconducting correlations in the Hubbard model away from half-filling.

G. Some Results

One of the things that we looked at with exact diagonalization in lecture I was the development of the magnetic moment $\langle m_z^2 \rangle$ at half-filling as the temperature decreased and doubly occupied and empty sites were frozen out. This quantity is shown in Fig. 1 on a 6x6 lattice for different interaction strengths U. We see the local moment begin to develop from its uncorrelated value $\langle m_z^2 \rangle = \frac{1}{2}$ at a temperature set by U, and then saturate at low T. The local moment does not reach $\langle m_z^2 \rangle = 1$ at $T = 0$ because significant quantum fluctuations allow doubly occupied and empty sites to occur even in the ground state. However, as U increases, these fluctuations are suppressed and the moment becomes better and better formed. $\langle m_z^2 \rangle$ also makes a further small adjustment at low T, which is due to the onset of long range magnetic order.

Another quantity we examined in lecture I was the specific heat of the Heisenberg model. We also discussed the mapping between the Hubbard and Heisenberg models at large U. This connection is re-emphasized in Fig. 2 which shows that the low temperature peak in the specific heat of the Hubbard model can be mapped onto that of the Heisenberg model with $J = 4t^2/U$.

The near-neighbor spin correlation are shown in Fig. 3. The magnetic structure factor $S(\pi, \pi) = \frac{1}{N}\sum_{ij}\langle S_{z,i}S_{z,j}\rangle$ sums these correlations over the whole lattice. It is found that $S(\pi, \pi)$ grows linearly with N at low T, indicating that the correlations extend over the whole lattice.

Finally, in Fig. 4 we show the density of states at $\omega = 0$ for the half-filled Hubbard model at different values of U. The suppression of $N(\omega = 0)$ at low T and large U is a signature of the presence of an insulating gap caused by the on-site repulsion.
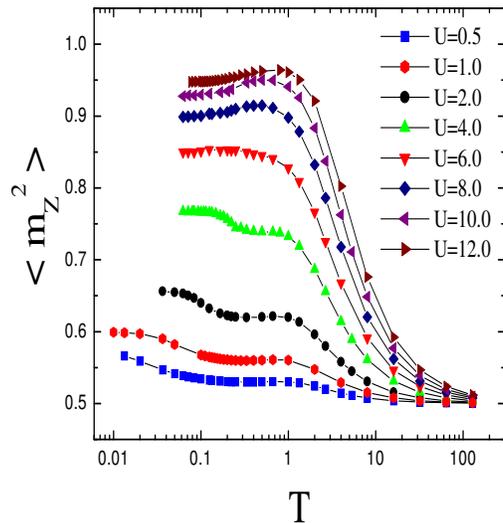
FIG. 1: The local moment $\langle m_z^2 \rangle$ as a function of temperature for different interaction strengths U and lattice size 6x6. The lattice is half-filled.
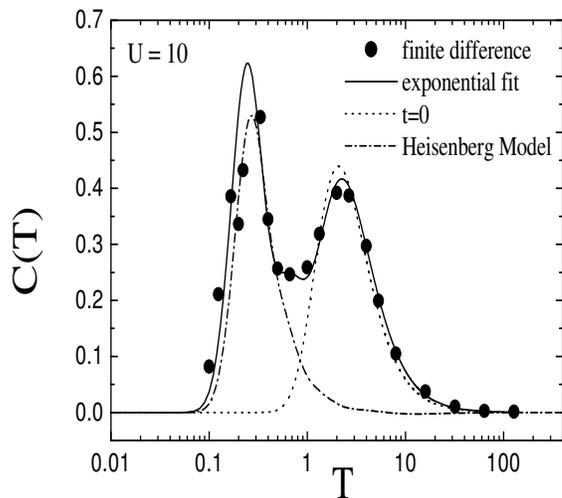

FIG. 2: The specific heat of the Hubbard model for $U = 10t$. There is a high temperature peak which is fit well by considering a single site Hubbard model ($t = 0$) and a low temperature peak which agrees well with the Heisenberg model with $J = 4t^2/U = 0.4$.

H. Conclusions

Determinant QMC is the method of choice for simulating interacting electron Hamiltonians in more than one dimension. One can easily study problems with several hundred particles, an order of magnitude greater than with exact diagonalization, and often large
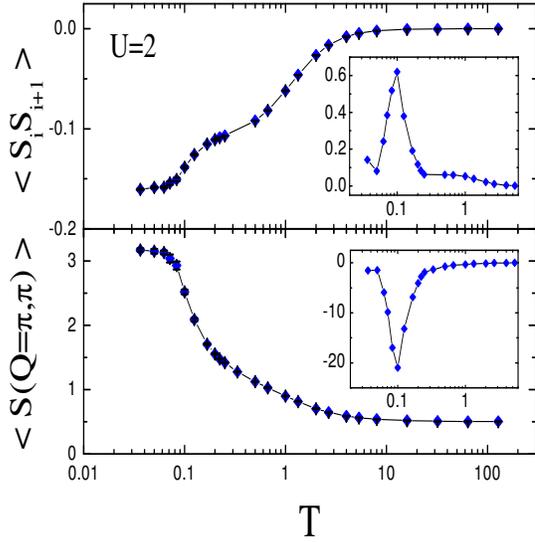
The near neighbor spin correlations and magnetic structure factor of the half-filled Hubbard model at $U = 2t$.
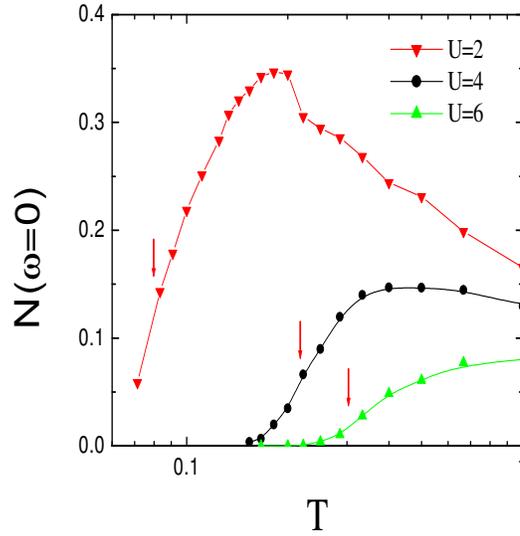
FIG. 3:

FIG. 4: The density of states at $N(\omega = 0)$. As T is lowered, a Mott-Hubbard gap opens up. The half-filled Hubbard model is insulating.

enough to make compelling finite size scaling analysis. The sign problem is much less serious than in world-line QMC, and, for the repulsive Hubbard model, one can go to temperatures on the order of at least $W/30$ where W is the bandwidth. For special cases like the attractive Hubbard model or the repulsive model at half-filling, there is no sign problem, and the ground state properties can be obtained.

Algorithm development in determinant QMC currently focusses on applications to

DMFT, where the Hubbard-Stratonovich field is allowed to fluctuate only in imaginary time. A number of questions are being actively explored in this field: How does one incorporate more complex (e.g. Hund's rules) interactions into simulation which include multiple orbitals? Can one re-introduce some degree of spatial fluctuations?

I. Acknowledgements