STOCHASTIC DIAGONALIZATION 

HANS DE RAEDT[1], WERNER FETTES[2], AND
KRISTEL MICHIELSEN[1]

[1] *Institute for Theoretical Physics and Materials Science Centre*
University of Groningen, Nijenborgh 4
NL-9747 AG Groningen, The Netherlands
E-mail: deraedt@phys.rug.nl, kristel@phys.rug.nl

[2] *Fakultät Physik, Universität Regensburg*
D-93040 Regensburg, Germany
E-mail: Werner.Fettes@physik.uni-regensburg.de

1. Introduction

The most direct approach to calculate the physical properties of a quantum system is to solve the eigenvalue problem of the Hamiltonian H. In particular the ground state properties of a quantum system can be computed from the solution of the eigenvalue problem

$$H|\Phi\rangle = E_1|\Phi\rangle \quad , \tag{1}$$

where E_1 denotes the smallest eigenvalue of the "matrix" H, the Hamiltonian of the system, and $|\Phi\rangle$ represents the corresponding eigenvector.

A critical factor for this approach to be useful in practice is the amount of memory needed to characterize a state of the system. In a classical N-particle system a many-particle state can be specified in terms of the position and momentum of each of the particles, i.e. $2dN$ numbers are sufficient to uniquely characterize a state (d denotes the dimension of the space in which the particles move). In quantum mechanics this is not the case: A state of the same system is described by a wave function which in general is a linear combination of all the allowed "classical" states. Thus instead of $\mathcal{O}(N)$ numbers, we need to specify all the coefficients of this linear combination, i.e. M numbers where M denotes the dimension of the matrix H representing the Hamiltonian. In general $M \gg \mathcal{O}(N)$ as will be clear from the following, rather generic, example.

Consider a lattice model of L sites, filled with $L/2$ electrons with spin up and $L/2$ electrons with spin down. Taking into account the fermion character of the electrons simple counting shows that

$$M = \binom{L}{L/2}^2 \quad , \tag{2}$$

which for large L ($L \geq 16$ will do) can be approximated using Stirling's formula

to give $M \approx 2^{2L+2}/2\pi L$, demonstrating that M increases exponentially with $2L$. Using 8 bytes/floating point number the estimated amount of memory we need to store a single eigenvector is given by

$$\mathcal{M} \approx \frac{2^{2L-25}}{2\pi L} \text{ Gbyte} \quad . \tag{3}$$

From (3) it follows that $\mathcal{M} \approx 1$Gbyte if $L = 16$, $\mathcal{M} \approx 10^9$Gbyte if $L = 32$, and $\mathcal{M} \approx 10^{28}$Gbyte if $L = 64$. Clearly any method that requires storage of the full matrix (i.e. $\mathcal{M}^2$ Gbyte will be of very limited use (as far as the range of system sizes that can be studied is concerned) to solve models for interacting fermions.

Although our estimate of the required amount of memory is somewhat crude (it does not incorporate reductions due to the use of symmetries) it gives a feeling for the kind of systems that can be solved by conventional, sparse matrix eigenvalue techniques such as the Lanczos [1, 2, 3], the (inverse) power [1, 2], or (generalized) Davidson method [4, 5]: $L = 16$ is within reach [6 − 8] but $L = 32$ is not.

Usually one is not content with the solution of a many-body problem describing a few particles only: The system-size dependence of the physical properties needs to be studied in order to get insight into the collective behavior of the particles. Unfortunately, as the many-particle system becomes larger, M grows exponentially fast and it is not a matter of just waiting for the next generation of computers to become affordable. If we want to solve these monster eigenvalue problems for values of M that our computers cannot manage we have to adopt another strategy. In fact we have only one option: We have to make the fundamental hypothesis that of all the M possible configurations (states) sampling only a small fraction M_I ($M_I \ll M$) of states will suffice to compute the ground-state properties to the desired accuracy. Searching the very large number of M states for the M_I important states may be viewed as a problem of importance sampling. However, the probability of a state to occur, i.e. its contribution to the ground state, is not known until we actually solve the eigenvalue problem.

In classical equilibrium statistical mechanics one faces a similar problem: The probability for a configuration is $p_j \equiv e^{-\beta E_j} / \sum_j e^{-\beta E_j}$ where E_j is the energy corresponding to the configuration j. The partition function $Z \equiv \sum_i e^{-\beta E_i}$ is, in general, unknown and hence so is p_j. Any Markov process which has $\{p_j\}$ as its limit distribution can be used to generate the "important" configurations, i.e. those that give the largest contributions to Z. The Metropolis Monte Carlo (MMC) method [9, 10, 11] is the most widely used algorithm implementing this idea but other simulation techniques such as Molecular Dynamics or Langevin Dynamics can be used as well. The MMC method uses the ratio p_i/p_j to determine the transition probability for the underlying stochastic process. Crucial thereby is that in forming the ratio, the unknown partition function drops out.

The apparent similarity between quantum and classical problems can be exploited to reformulate the calculation of the lowest eigenvalue or thermal expectation values of a quantum system in terms of a Markov process on the space of states. Various standard methods of linear algebra have a stochastic counterpart. The power method is at the heart of the Diffusion Monte Carlo technique which uses a stochastic process to sample powers of the matrix H [12]. The Green Function Monte Carlo technique performs the inverse iteration steps by solving this linear equation by a stochastic method [12]. Quantum statistical problems can be

recasted into a "classical" form by means of the Feynman path integral [13, 14] or, more generally by invoking a Trotter-Suzuki formula [14−19]. The Markov process will properly sample the important contributions to the ground state provided the elements of the stochastic matrix, defining the Markov process, correspond to the matrix elements of a judiciously chosen function $f(H)$ of the Hamiltonian H. However, for many problems of interest constructing this correspondence seems extremely hard: The elements of the stochastic matrix, being probabilities, have to be positive but, more as a general rule than an exception, the matrix elements of $f(H)$ may be negative and positive (except for $f(H) \propto 1$).

The fundamental difficulty of constructing the appropriate stochastic matrix is called the minus-sign problem in Quantum Monte Carlo (QMC) simulations [20]. It results from the choice of the representation used to calculate the matrix elements of $f(H)$, in combination with the desire or need to use a Markov process as a vehicle to search for the important states. Accordingly the minus-sign problem should also be present in cases where there are no fermionic degrees of freedom and indeed, there are ample examples that show it is, including systems with only one degree of freedom [18]. The analogy with classical statistical mechanics breaks down completely if the minus-sign problem is present. In practice this implies that a quantity of interest might be a sum of a positive and a negative contribution which, unfortunately, nearly cancel each other. Extremely good statistics and accuracy may be required to obtain meaningful results.

This lecture is devoted to a method, called stochastic diagonalization [21, 22], that is free of minus-sign problems by construction. It is fundamentally different from QMC methods in that it uses a random process, based on orthogonal instead of stochastic matrices, to collect the important contributions to the ground state.

2. Minus-sign Problem in Quantum Monte Carlo Methods

From mathematical point of view all QMC methods [20] have in common that at some point they evaluate, using a sampling technique, matrix elements of a function $f(H)$ of the Hamiltonian H that has the structure of a matrix product. For the Diffusion QMC method

$$f_{DQMC}(H) = (1 - \tau H)^m \quad , \tag{4}$$

where τ is some control parameter and m is related to the number of steps in the random walk. The Green Function Monte Carlo method uses

$$f_{GFMC}(H) = [(\omega + H)^{-1}]^m \quad , \tag{5}$$

where ω is a real number, the purpose of which is discussed below. Path-integral or Trotter-Suzuki-based techniques push the product-formula structure of (4) and (5) a level further by approximating the function $f_{PI}(H) = e^{-\beta H}$ by an ordered product of exponents, e.g.

$$f_{PI}(H) \approx \left(e^{-\beta(H-A)/m)} e^{-\beta A/m)} \right)^m \quad . \tag{6}$$

where A denotes a contribution to H.

It is now instructive to ask the question under which conditions these QMC

techniques will **not** suffer from the minus-sign problem. For the Diffusion QMC Method (4) a sufficient condition is given by (almost trivial)

Theorem 1. A sufficient condition for $\langle\phi|(1-\tau H)|\psi\rangle$ to possess positive matrix elements is given by $\langle\phi|H|\psi\rangle \leq 0$ for all $|\phi\rangle$,$|\psi\rangle$.

For the Green Function Monte Carlo method (5) it is clear that if a real number ω can be found such that the matrix $(\omega+H)^{-1}$ has all positive elements it can, in principle, be used to define a Markov process [23]. A sufficient condition follows from

Theorem 2. If ω is taken such that $\omega + H$ is a positive definite matrix with the property that $\langle\phi|H|\psi\rangle \leq 0$ for all $|\phi\rangle \neq |\psi\rangle$ then $(\omega + H)^{-1}$ has only positive elements.

The proof of theorem 2 can be found in [22, 24]. Finally for QMC Methods (6) that are based on $e^{-\beta H}$ (or approximations thereof) we have

Theorem 3. The necessary and sufficient condition for $\langle\phi|e^{-\tau H}|\psi\rangle$ to be positive for all $\tau > 0$ is $\langle\phi|H|\psi\rangle \leq 0$ for all $|\phi\rangle \neq |\psi\rangle$.

The proof of theorem 3 can be found in [22, 24]. Obviously, for all three QMC methods the sufficient condition is the same.

It is easy to see that the sufficient condition puts rather strong constraints on the choice of the representation that is used to actually evaluate the matrix elements. Due to the product-formula structure of the algorithms the product of all (positive or negative) factors might still be positive, for instance because the total number of negative factors is always even. An example of such a system is the two-dimensional Heisenberg model on a bipartite lattice [14].

The sufficient condition for not having minus-sign problems seems rather restrictive. There are a number of examples where, at first glance, the necessary condition is not fulfilled but where there are no minus-sign problems. This, in all cases that we know of, is due to the presence of symmetries that allow us to reverse the sign of the non-diagonal elements of H by changing the representation of the states.

Usually it is expedient to carry out analytically as many integrations (or summations) as possible, for instance by introducing auxillary degrees of freedom, completing the square and performing the resulting Gaussian integrals [18]. In some cases, e.g. some electron-phonon models [25], the minus-sign problem disappears completely while in other cases (e.g. the two-dimensional Hubbard model) it is absent for particular values of the model parameters [18]. We are not aware of general results on the presence or absence of the minus-sign problem in QMC methods that use auxillary fields. For fermion systems this amounts to a study of one or more determinants of non-symmetric real matrices [18, 26], a non-trivial problem in itself.

3. Stochastic Diagonalization

If the dimension of the Hilbert space is so large that it is no longer possible to store even a single vector, standard diagonalization methods cannot be used to solve the Schrödinger equation. Then the only way to proceed is to make the basic assumption (i.e. the fundamental hypothesis mentioned above) that of

the whole, large set of basis vectors $\{|\phi_j\rangle\,;\, j = 1, \ldots, M\}$ spanning the Hilbert space, only a relatively small portion is *important* for the computation of physical properties. The stochastic diagonalization (SD) algorithm implements this idea in the following manner [21, 22]. Instead of using the sparseness of the matrix, it is assumed that the solution itself is "sparse" in the sense that only a small fraction of the elements of the eigenvector, corresponding to the smallest eigenvalue, is *important*.

As the ground state can be written as a linear combination of all the basis states

$$|\Phi\rangle = \sum_{j=1}^{M} a_j |\phi_j\rangle \quad , \tag{7}$$

we can, at least in principle, rearrange the terms in this sum so that the ones with the largest amplitude are in front:

$$|\Phi\rangle = \sum_{j=1}^{M} a_{Pj} |\phi_{Pj}\rangle \quad . \tag{8}$$

Here P denotes the permutation of the set $\{1, \ldots, M\}$ such that $|a_{Pj}| \geq |a_{P(j+1)}|$. Keeping only the first $M_I = M_{Important}$ terms we have

$$|\Phi\rangle \approx |\widetilde{\Phi}\rangle = \sum_{j=1}^{M_I} a_{Pj} |\phi_{Pj}\rangle \quad . \tag{9}$$

From Poincaré's theorem [1, 2] it follows that $E \leq \widetilde{E} = \langle\widetilde{\Phi}|H|\widetilde{\Phi}\rangle / \langle\widetilde{\Phi}|\widetilde{\Phi}\rangle$ demonstrating that keeping the M_I important states gives an upperbound to the ground-state energy.

If the basic premise on the existence of important states holds we may expect that $M_I \ll M$. In practice the choice of the M_I that will give satisfactory accuracy will depend on the actual choice of the basis vectors (i.e. the representation used) and on the model itself.

Up to now, we argued as if we already know the permutation P that re-shuffles the expansion (8) in terms of the chosen basis states but in fact we don't know P nor do we know the coefficients a_{Pj}. The essence of the SD algorithm is that it finds P and the coefficients a_{Pj} simultaneously through a combination of plane (Jacobi) rotations and matrix inflation. As will be clear from the theory given below, this algorithm cannot suffer from the minus-sign problems by construction.

3.1. THEORY: PRELIMINARIES

We start by introducing some notations. The projection of a real and symmetric matrix H on the subspace spanned by the $n \leq M$ orthonormal states (vectors) $S^{(n)} \equiv \{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$ will be denoted by $H^{(n)}$, matrix elements by $H^{(n)}_{i,j} = \langle\phi_i|H|\phi_j\rangle$ and the eigenvalues $E^{(n)}_i$ of $H^{(n)}$ are assumed to be ordered such that $E^{(n)}_1 \leq E^{(n)}_2 \leq \ldots \leq E^{(n)}_n$. Eventually the superscript n will keep track of the number of important states and therefore also of the size of the matrix $H^{(n)}$. Evidently $H^{(M)} = H$ and $E^{(M)}_i = E_i$. We will use H and $H_{i,j}$ to denote the full matrix and the corresponding matrix elements respectively. Without loss of generality we may assume that in each row (or column) $i = 1, \ldots, M$ there is at least one non-diagonal matrix element ($H_{i,j}$) that differs from zero, i.e. $\sum_{j\neq i} |H_{i,j}| > 0$ for $1 \leq i \leq M$. Otherwise the matrix would decompose into blocks of smaller matrices and the determination of the smallest eigenvalue amounts to the calculation of the smallest eigenvalue of each block. The transpose of a matrix A is denoted by A^T. The norm of a vector $x = (x_1, \ldots, x_n)$ will be denoted by $\|x\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$ and $\|A\|$ stands for the spectral norm, i.e. the square root of the largest eigenvalue of the matrix $A^T A$ [1].

A plane rotation involving states $|\phi_{i_{n,k}}\rangle$ and $|\phi_{j_{n,k}}\rangle$ ($1 \leq i_{n,k} < n$, $i_{n,k} < j_{n,k} \leq n$) is represented by a $n \times n$ orthogonal matrix $U^{(n,k)} = U^{(n,k)}(i_{n,k}, j_{n,k}, c_{n,k}, s_{n,k})$ which, in block matrix form, can be written as

$$U^{(n,k)}(i_{n,k}, j_{n,k}, c_{n,k}, s_{n,k}) = \begin{array}{c} \begin{array}{ccccccc} 1 & & i_{n,k} & & j_{n,k} & & n \end{array} \\ \begin{pmatrix} 1 & & & & & & \\ & \cdots & & & & & \\ & & c_{n,k} & \cdots & s_{n,k} & & \\ & & \vdots & 1 & \vdots & & \\ & & -s_{n,k} & \cdots & c_{n,k} & & \\ & & & & & \cdots & \\ & & & & & & 1 \end{pmatrix} \end{array} . \quad (10)$$

In (10) all diagonal elements are unity except for the two elements $c_{n,k}$ in columns $i_{n,k}$ and $j_{n,k}$. All non-diagonal elements are zero except the two elements $-s_{n,k}$ and $s_{n,k}$. The subscript k will be used as a running index of the plane rotations for fixed dimension n. This admittedly complicated notation is necessary to avoid ambiguities in the interpretation of the symbols. The product of a sequence of plane rotations will be denoted by $\mathcal{U}^{(n,m)} = U^{(n,1)} \ldots U^{(n,m)}$. We adopt the convention that the order in which plane rotations are applied corresponds to the value of k, i.e. first $U^{(n,1)}$, then $U^{(n,2)}$ and so on. The transformed matrix is given by $H^{(n,m)} = \left[\mathcal{U}^{(n,m)}\right]^T H^{(n)} \mathcal{U}^{(n,m)}$. Note that the label n only determines the dimension of the matrices and that it puts no restriction on m.

Plane rotations will be determined by the following elementary result.

Lemma 1. The eigenvalues $\lambda_1 \leq \lambda_2$ of a real and symmetric matrix $A = \begin{pmatrix} x & y \\ y & z \end{pmatrix}$ where $x \leq z$ and $y \neq 0$ satisfy $\lambda_1 < x \leq z < \lambda_2$.

The eigenvalues of A are $\lambda_1 = x - ty < x$ and $\lambda_2 = z + ty > z$ and the orthogonal matrix

$$U = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \quad , \tag{11}$$

with $c = 1/\sqrt{1+t^2}$, $s = t/\sqrt{1+t^2}$ and $t = 2y/(z - x + \sqrt{(z-x)^2 + 4y^2})$ (i.e. $|t| \leq 1$), diagonalizes the matrix A, i.e.

$$U^T A U = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad . \tag{12}$$

The strict inequality $\lambda_1 < x$ is essential for the importance sampling algorithm.

3.2. THEORY: MODIFIED JACOBI METHOD

One strategy to compute the ground state would be to transform the matrix H as

$$U^T H U = \begin{pmatrix} E_1 & 0^T \\ 0 & \widetilde{H} \end{pmatrix} \quad . \tag{13}$$

We modify the cyclic Jacobi method [1] to accomplish this. Let us assume for a moment that $H_{1,1} \leq H_{j,j}$, $2 \leq j \leq M$. We will remove this restriction later. If, instead of considering all pairs (i,j), the plane rotations involve pairs $(1,2), \ldots, (1,M)$ only, then

$$\widehat{H}^{(M)} \equiv \lim_{m \to \infty} H^{(M,m)} = \begin{pmatrix} E_i & 0^T \\ 0 & \widetilde{H} \end{pmatrix} \quad , \tag{14}$$

where $E_i = E_i^{(M)}$ is one of the eigenvalues. The proof of (14) is straightforward. According to Lemma 1, application of a plane rotation involving a pair $(1,j)$ strictly reduces $H_{1,1}^{(M,m)}$, i.e. $H_{1,1}^{(M,m+1)} < H_{1,1}^{(M,m)}$ for $m > 0$ and since $E_1^{(M)} \leq H_{1,1}^{(M,m)}$ the sequence $\{H_{1,1}^{(M,m)}\}$ is monotonically decreasing and bounded from below. Thus $\lim_{m\to\infty} H_{1,1}^{(M,m)} = \widehat{E}$ exists. Furthermore $\lim_{m\to\infty} H_{1,j}^{(M,m)} = 0$ for all $j \in \{2, \ldots, M\}$. To prove this assume the contrary, i.e. $\lim_{m\to\infty} H_{1,j}^{(M,m)} \neq 0$ for at least one $j \in \{2, \ldots, M\}$. Then, according to Lemma 1, a plane rotation involving the pair $(1,j)$ would reduce the $(1,1)$ element, in contradiction with the assumption that the monotonically decreasing sequence $\{H_{1,1}^{(M,k)}\}$ converges to $\widehat{E}$. Moreover, since $\lim_{m\to\infty} H_{1,j}^{(M,m)} = 0$ for all $j \in \{2, \ldots, M\}$, $\widehat{E}$ is an eigenvalue. Hence $\widehat{E} = E_i^{(M)}$ for some $i \in \{1, \ldots, M\}$. This completes the proof that this variant of the Jacobi method isolates an eigenvalue.

The eigenvector corresponding to $E_i^{(M)}$ is given by $|\Phi_i^{(M)}\rangle$ with

$$|\Phi_i^{(n)}\rangle \equiv \lim_{m\to\infty} \sum_{j=1}^{n} \mathcal{U}_{j,i}^{(n,m)} |\phi_j\rangle \quad ; \quad 1 \leq n \leq M \quad , \tag{15}$$

i.e. the i-th column vector of $\mathcal{U}^{(n,m)}$ in the basis $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$.

3.3. THEORY: MATRIX INFLATION

The modified Jacobi method isolates an eigenvalue and yields the corresponding eigenvector but at this point it is not known which eigenvalue it will find. In order to obtain the smallest eigenvalue we combine the modified Jacobi method with a matrix inflation procedure. The latter will turn out to be essential to determine which states are important and which are not.

The theoretical justification of the method is by induction. Consider the submatrix

$$H^{(n)} = \begin{pmatrix} H_{1,1} & H_{1,2} & \ldots & H_{1,n} \\ H_{1,2} & H_{2,2} & \ldots & H_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{1,n} & H_{2,n} & \ldots & H_{n,n} \end{pmatrix} \quad , \tag{16}$$

and assume that application of the modified Jacobi scheme reduces $H^{(n)}$ to the form

$$\widehat{H}^{(n)} \equiv \lim_{m\to\infty} H^{(n,m)} = \begin{pmatrix} E_1^{(n)} & 0^T \\ 0 & \widetilde{H} \end{pmatrix} \quad , \tag{17}$$

where $E_1^{(n)}$ is the smallest eigenvalue of $H^{(n)}$. This assumption is trivially satisfied for $n = 1$. We now inflate the matrix $H^{(n)}$ by adding the $(n+1)$-th row and column. Apply to $H^{(n+1)}$ the sequence of plane rotations that transforms $H^{(n)}$ to the form (17) and obtain

$$\widehat{H}^{(n+1)} = \begin{pmatrix} \widehat{\mathcal{U}}^{(n)} & 0 \\ 0 & 1 \end{pmatrix}^T H^{(n+1)} \begin{pmatrix} \widehat{\mathcal{U}}^{(n)} & 0 \\ 0 & 1 \end{pmatrix} \quad , \tag{18a}$$

$$= \begin{pmatrix} E_1^{(n)} & 0 & \ldots & 0 & \alpha_1^{(n+1)} \\ 0 & \widetilde{H}_{2,2} & \ldots & \widetilde{H}_{2,n} & \alpha_2^{(n+1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \widetilde{H}_{2,n} & \ldots & \widetilde{H}_{n,n} & \alpha_n^{(n+1)} \\ \alpha_1^{(n+1)} & \alpha_2^{(n+1)} & \ldots & \alpha_n^{(n+1)} & H_{n+1,n+1} \end{pmatrix} \quad , \tag{18b}$$

where $\widehat{\mathcal{U}}^{(n)} = \lim_{m\to\infty} \mathcal{U}^{(n,m)}$, $\alpha_j^{(n+1)} = \lim_{m\to\infty} \alpha_j^{(n+1,m)}$ for $j = 1, \ldots, n$,

$$\alpha_j^{(n+1,m)} = \left(\begin{pmatrix} \mathcal{U}^{(n,m)} & 0 \\ 0 & 1 \end{pmatrix}^T H^{(n+1)} \begin{pmatrix} \mathcal{U}^{(n,m)} & 0 \\ 0 & 1 \end{pmatrix} \right)_{j,n+1} \quad , \tag{19a}$$

and

$$\mathcal{U}^{(n,m)} \equiv \begin{pmatrix} \widehat{\mathcal{U}}^{(n-1)} & 0 \\ 0 & 1 \end{pmatrix} \mathcal{U}^{(n,m)} \quad ; \quad \widehat{\mathcal{U}}^{(1)} = 1 \quad . \tag{19b}$$

Here use has been made of the symmetry of H and the fact that the plane rotations in (19) do not affect the matrix elements in column $n+1$.

Let us now assume that

$$\alpha_1^{(n+1)} \neq 0 \quad . \tag{20}$$

We will discuss the case $\alpha_1^{(n+1)} = 0$ in more detail below. According to Lemma 1, a single plane rotation involving the pair $(1, n+1)$ will lead to a reduction of the $(1,1)$ element of $H^{(n+1)}$ provided

$$E_1^{(n)} \leq H_{n+1,n+1} \quad , \tag{21}$$

a restriction to be removed later. With $x = E_1^{(n)}$, $y = \alpha_1^{(n+1)}$, and $z = H_{n+1,n+1}$, Lemma 1 gives

$$H^{(n+1,1)} = \begin{pmatrix} \lambda_1 & \beta^T & 0 \\ \beta & \widetilde{H} & \widetilde{\gamma} \\ 0 & \widetilde{\gamma}^T & \lambda_2 \end{pmatrix} \quad , \tag{22}$$

with $\beta^T = -s_{n+1,1} \left(\alpha_2^{(n+1)}, \ldots, \alpha_n^{(n+1)} \right)$ and $\lambda_1 < E_1^{(n)}$. Invoking the separation theorem [1] gives $E_1^{(n+1)} \leq E_1^{(n)}$ and hence

$$E_1^{(n+1)} \leq \lambda_1 < E_1^{(n)} \quad . \tag{23}$$

If $\beta = 0$ we have $E_1^{(n+1)} = \lambda_1$. In general $\beta \neq 0$ but we already showed that in the modified Jacobi method, the $(1,1)$ element monotonically decreases and converges to an eigenvalue. According to inequality (23), application of the modified Jacobi strategy to the matrix (22) will yield the smallest eigenvalue of $H^{(n+1)}$, i.e. $\lim_{m\to\infty} H_{1,1}^{(n+1,m)} = E_1^{(n+1)}$. Then, returning to (18) with n replaced by $n+1$, the whole procedure can be repeated. This completes the proof that the method will isolate the smallest eigenvalue of H.

Summarizing: The calculation starts by diagonalizing the 2×2 matrix. Then one row and column is added to the matrix and the modified Jacobi method is employed to compute the smallest eigenvalue of the 3×3 matrix. This step is repeated, yielding the smallest eigenvalue of a 4×4 matrix, 5×5 matrix, and so on.

We now review the assumptions made. Restriction (21) (which includes the condition $H_{1,1} \leq H_{j,j}$) is trivially removed. If this condition is not satisfied, application of the permutation

$$P = \begin{pmatrix} 1 & \ldots & n+1 & \ldots & M \\ n+1 & \ldots & 1 & \ldots & M \end{pmatrix} \quad , \tag{24}$$

will bring the matrix in the desired form, without loosing numerical stability. In

practice this is a trivial operation.

At each inflation step ($n \to n+1$) we might have $\alpha_1^{(n+1)} = 0$. Then the arguments that were used to prove convergence to the smallest eigenvalue cannot be used because the inequality $\lambda_1 < E_1^{(n)}$ does not hold. If the matrix is block diagonal, i.e. $\alpha_j^{(n+1)} = 0$ for all $j \in \{1, \ldots, n\}$, it is clear that we have to compute the lowest eigenvalue of each block. However this case cannot occur because we assumed that there is at least one non-zero off-diagonal matrix element in each column (or row) and the application of orthogonal transformations does not change this property. The process of clearing a matrix element on the first row and inflating the matrix may "accidentally" lead to $\alpha_1^{(n+1)} = 0$, some rather exotic examples being given in ref. [22]. As long as $n+1 < M$ there is no immediate danger for the method to break down. If there exists a permutation of the columns (and rows) $n+1$ and n' ($n+1 < n' \leq M$) that yields $\alpha_1^{(n+1)} \neq 0$, we perform this permutation (in theory, not in practice of course) and continue as usual. However, if $n+1 = M$ or if there does not exist such a permutation then the method has isolated an eigenvalue but there is no guarantee that it is the smallest. In this case the matrix has been reduced to the block-diagonal form and we have no other option than to repeat the procedure, i.e. isolate the smallest eigenvalue, for the remaining $(M-1) \times (M-1)$ block matrix. However, according to the hypothesis made in the introduction, the number of important states M_I is assumed to be a small fraction of M. Hence $n \leq M_I \ll M$ and the case $\alpha_1^{(p)} = 0$ with $n < p < M$ will hardly occur in practice.

3.4. THEORY: IMPORTANCE SAMPLING ALGORITHM

The theoretical method can be turned into a useful importance sampling algorithm by a few minor modifications. The order to annihilate the off-diagonal elements of the first row (and column) is fully determined by our desire to efficiently isolate an eigenvalue. Accordingly, the pair $(1, j)$ and $(j, 1)$ is chosen such that $|H_{1,j}^{(n,m)}| = \max_{i>1} |H_{1,i}^{(n,m)}|$.

The first modification, identical to the one made in the case of the Jacobi method [1], is to limit the number of plane rotations for fixed n by introducing the threshold $\epsilon_R^{(n,m)} > 0$. Rotations will be carried out if $|H_{1,i}^{(n,m)}| \geq \epsilon_R^{(n,m)}$ for any $i = 2, \ldots, n$ or, in different words, until the size of all off-diagonal elements on the first row becomes smaller than the threshold $\epsilon_R^{(n,m)}$. Keeping $\epsilon_R^{(n,m)}$ fixed the transformed matrix reads

$$H^{(n,m)} = \begin{pmatrix} E_1^{(n,m)} & \delta^{(n,m)^T} \\ \delta^{(n,m)} & \widetilde{H} \end{pmatrix} \quad , \tag{25}$$

where $E_1^{(n)} \leq E_1^{(n,m)}$ and $\delta_i^{(n,m)} = H_{1,i}^{(n,m)}$ for $i > 1$. Invoking the monotonicity theorem [1, 2, 22], yields

$$E_1^{(n,m)} - E_1^{(n)} \leq \|\delta^{(n,m)}\| = \sqrt{\sum_{i>1} \left(H_{1,i}^{(n,m)}\right)^2} < \sqrt{n}\epsilon_R^{(n,m)} \quad . \tag{26}$$

The second modification concerns the inflation step, providing the criterion to decide which states are important and which are not. Again we proceed by induction. Assume the number of important states is n. We pick a trial state $|\widehat{\phi}\rangle$ from the set of $M-n$ remaining states, for instance randomly. Recall that there must be at least one non-zero element in the new row and column. We temporarily set $|\phi_{n+1}\rangle = |\widehat{\phi}\rangle$, compute $\alpha_1^{(n+1,m)}$ and the corresponding change of the $(1,1)$ element (see Lemma 1)

$$\Delta_{n+1}^{(n+1,m)} = \frac{2\left(H_{1,n+1}^{(n,m)}\right)^2}{\Delta + \sqrt{\Delta^2 + 4\left(H_{1,n+1}^{(n+1,m)}\right)^2}} \quad , \tag{27}$$

where $\Delta = H_{n+1,n+1}^{(n+1,m)} - H_{1,1}^{(n+1,m)}$.

If $\Delta_{n+1}^{(n+1,m)} \geq \epsilon_A^{(n+1,m)}$ the trial state $|\widehat{\phi}\rangle$ is considered to be important and is added to the set of states. Clearly the threshold $\epsilon_A^{(n+1,m)} > 0$ will control the importance sampling process. We set $|\phi_{n+1}\rangle = |\widehat{\phi}\rangle$ and

$$U^{(n+1,k)} = \begin{pmatrix} U^{(n,k)} & 0 \\ 0 & 1 \end{pmatrix} \quad ; \quad k = 1, \ldots, m \quad . \tag{28}$$

Unlike in the previous sections of this chapter, the plane rotation index is not reset to its initial value $m = 1$ when we inflate the matrices. Annihilation of the matrix element $\alpha_1^{(n+1,m)} = 0$ determines the new rotation matrix $U^{(n+1,m+1)}$. We finally replace n by $n+1$, m by $m+1$ and continue. If $\Delta_{n+1}^{(n+1,m)} < \epsilon_A^{(n+1,m)}$ the trial state is rejected and a new trial state $|\widehat{\phi}\rangle$ is generated. If $\alpha_1^{(n+1,m)} = 0$ the trial state is always rejected since $\Delta_{n+1}^{(n+1,m)} = 0$.

In order to isolate the smallest eigenvalue the reduction has to be large enough. A sufficient condition can be derived by repeating the steps that led to (18). In place of (18) we now have

$$H^{(n+1,m)} = \begin{pmatrix} E_1^{(n,m)} & \delta_2^{(n,m)} & \ldots & \delta_n^{(n,m)} & \alpha_1^{(n+1,m)} \\ \delta_2^{(n,m)} & \widetilde{H}_{2,2} & \ldots & \widetilde{H}_{2,n} & \alpha_2^{(n+1,m)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \delta_n^{(n,m)} & \widetilde{H}_{2,n} & \ldots & \widetilde{H}_{n,n} & \alpha_n^{(n+1,m)} \\ \alpha_1^{(n+1,m)} & \alpha_2^{(n+1,m)} & \ldots & \alpha_n^{(n+1,m)} & H_{n+1,n+1} \end{pmatrix} \quad , \tag{29}$$

because now m is finite. Annihilating $\alpha_1^{(n+1,m)}$ leads to a matrix of the form (22). From (26) it is clear that if $\lambda_1 \leq E_1^{(n,m)} - \|\delta^{(n,m)}\|$, we will have $\lambda_1 < E_1^{(n)}$.

Repeating the reasoning that follows (23) establishes that if $\lambda_1 \leq E_1^{(n,m)} -$

$\|\delta^{(n,m)}\|$ the inflation step will guarantee convergence to the smallest eigenvalue. The condition for isolating the smallest eigenvalue follows from $\lambda_1 \leq E_1^{(n,m)} - \|\delta^{(n,m)}\|$ and $E_1^{(n,m)} - \lambda_1 \geq \epsilon_A^{(n,m)}$ and is given by $\epsilon_A^{(n,m)} \geq \|\delta^{(n,m)}\|$. As is usual with this kind of theoretical analysis, the bounds on the maximum reduction of the smallest eigenvalue may be too weak and strict use of (29) may have a negative impact on the performance of the algorithm. The construction of the importance sampling algorithm and the mathematical proof that it yields the smallest eigenvalue of H have now been completed.

To summarize, the SD algorithm looks like this:

Initialize data structures
do
  if {Maximum of absolute value of off-diagonal elements of the first row smaller than threshold for rejecting plane rotations}
  then
    Generate a new trial state
      if {No important state has been found}
        then
          Reduce the threshold(s)
        else
          Inflate the matrix
        end if
  else
    Annihilate the pair of off-diagonal elements with the largest absolute value by performing a plane rotation
  end if
end do

3.5. COMPUTATION OF PHYSICAL PROPERTIES

Assuming the ground state has been found, either in exact form by e.g. the Lanczos method or in the variational sense through the SD algorithm, the calculation of expectation values of physical quantities may become a non-trivial computational problem if the matrix representing the observable is not diagonal in the basis $\{|\phi_j\rangle\,;\, j = 1, \ldots, M\}$ that was used to represent the Hamiltonian. Indeed, if A denotes the physical observable, the expectation value of A is given by

$$\langle A \rangle = \langle \Phi | A | \Phi \rangle = \sum_{i,j=1}^{M_I} a_{Pi} a_{Pj} \langle \phi_i | A | \phi_j \rangle \quad , \tag{30}$$

showing that in general it will take $\mathcal{O}(M_I^2)$ operations to carry out this computation. The calculation of the ground-state energy itself does not require extra work because (the approximation to) it is known at each stage of the SD process [22]. However for large M_I the calculation of certain expectation values, e.g. the reduced two-fermion density matrix (see below), might take a substantial amount of CPU time.

Table 1. Comparison between ground state energies of the 2D tt'-Hubbard model as obtained by exact diagonalization (ED), stochastic diagonalization (SD) and projector quantum Monte Carlo (PQMC) for a $L_x = L_y = 4$ lattice. The density of electrons with spin-up or spin-down is 5/16. For $U < 0$, $t' = 0$ whereas for $U > 0$, $t' = -0.22t$.

$U/\|t\|$	ED	SD	$PQMC$
-6	-2.458782	-2.4568	-2.460 ± 0.004
-4	-2.045849	-2.0453	-2.045 ± 0.002
-2	-1.731689	-1.7316	-1.731 ± 0.003
2	-1.230034	-1.2300	-1.231 ± 0.001
4	-1.126160	-1.1261	-1.125 ± 0.003
6	-1.058717	-1.0581	-1.061 ± 0.005

3.6. COMPARISON WITH OTHER METHODS

We have tested the SD algorithm by comparing the results of SD calculations to results obtained by other, more established methods. In Table 1 we collect some data, obtained by various numerical methods, for the two-dimensional (2D) tt'-Hubbard model [27, 28]. In reciprocal space the model Hamiltonian reads

$$H^{Hub} = \sum_k \sum_{\sigma=\uparrow,\downarrow} \epsilon_k \, c^+_{k,\sigma} c_{k,\sigma} + \frac{U}{L} \sum_{k,p,q} c^+_{k+q,\uparrow} c^+_{p-q,\downarrow} c_{p,\downarrow} c_{k,\uparrow} \quad , \tag{31}$$

where $\epsilon_k = -2t(\cos(2\pi k_x/L_x) + \cos(2\pi k_y/L_y)) - 4t' \cos(2\pi k_x/L_x)\cos(2\pi k_y/L_y)$, $k = (k_x, k_y)$, U is the on-site Coulomb interaction, L_x (L_y) denotes the number of lattice points in the x-(y-)direction, $L \equiv L_x \cdot L_y$, and t (t') is the nearest (next-nearest) neighbour hopping-integral. From Table 1 it is clear that the SD algorithm is working properly in this case. In Table 2 we compare results of a fixed-node approximation approach [29] and SD calculations for a small two and three-dimensional Hubbard model ($t' = 0$, $\epsilon_k = -2t(\cos(2\pi k_x/L_x) + \cos(2\pi k_y/L_y) + \cos(2\pi k_z/L_z))$, $k = (k_x, k_y, k_z)$). The number of states collected in these SD calculations was limited to 10 and 600 for the first and second (and third) row respectively, the corresponding dimension of the Hilbert space being 36 and 4900. Compared to mean-field approximations the fixed-node approximation yields a significant improvement for the upper-bound to the ground-state energy of the Hubbard model [29] but is clearly not as accurate as the SD method.

In Table 3 we present a selection of results of SD and other calculations on some typical quantum chemistry problems [30]. The SD results using $M_I = 40000$ compare favourably with the results of the full Configuration Interaction (CI) calculation [5]. These and other [30] results demonstrate that the SD algorithm can be quite effective in reducing the number of basis states while giving up little in terms of accuracy.

4. Application: Off-Diagonal Long-Range Order

In boson systems Bose-Einstein condensation is characterized by the existence of

Table 2. Comparison between ground state energies per site of the two and three-dimensional Hubbard models as obtained from a fixed-node aproximation (FN) [29], exact diagonalization (ED) and stochastic diagonalization (SD). N denotes the number of electrons with spin-up and spin-down. The difference between the FN results of the second and third row stems from the choice of the trial state [29].

Lattice	N	$U/\|t\|$	FN	ED	SD
2×2	2	1	-3.3172	-3.3408	-3.3408
$2 \times 2 \times 2$	4	10	-2.6507	-2.8652	-2.8634
$2 \times 2 \times 2$	4	10	-2.6382	-2.8652	-2.8634

Table 3. The energy E and correlation energy E_{corr}, both in hartree of the Mg atom as obtained by SD and various standard quantum chemistry methods. M_I denotes the number of basis states (determinants with D_{2h} symmetry) that was used.

Method	E	E_{corr}	M_I
SCF	-199.585212	0.0000	1
RASSCF	-199.615701	-0.0305	4
CISD	-199.721386	-0.1362	2960
CISDT	-199.722039	-0.1368	102928
CISDTQ	-199.726256	-0.1410	1964232
SD	-199.724237	-0.1409	2000
SD	-199.726164	-0.1410	40000
CI [5]	-199.7263	-0.1411	2538603250

Off-Diagonal Long-Range Order (ODLRO) in the reduced single-particle density matrix [31, 32]. Yang has shown that the concept of ODLRO can also be used to characterize the superconducting state of fermion systems [33]. Under certain simplifying assumptions, ODLRO implies the existence of the Meissner effect and magnetic flux quantization [34 – 36]. ODLRO in the reduced n-particle density matrix implies ODLRO in the reduced m-particle density matrices for all $m > n$ [33].

For a fermion system the one-particle reduced density matrix cannot exhibit ODLRO [33]. Therefore we will compute the largest eigenvalue λ_0 of the reduced two-body density matrix [33]

$$\rho_{r,s} \equiv \rho(i,j,\sigma;k,l,\sigma') = \langle c^+_{i,\sigma} c^+_{j,-\sigma} c_{l,-\sigma'} c_{k,\sigma'} \rangle \quad , \tag{32}$$

where $r = (i,j,\sigma)$ and $s = (k,l,\sigma')$ and $c^+_{i,\sigma}$ and $c_{i,\sigma}$ are the creation and annihilation operators, respectively, for a fermion with spin $\sigma = \uparrow, \downarrow$ at the generalized site index i. As is evident from the spin labels in (32) we only consider singlet pairing. For conciseness we will from now on use the term ODLRO, always refering to ODLRO in the two-body density matrix.

The eigenvector of the two-body density matrix, corresponding to λ_0, contains

all the information about the type of pairing, including all exotic forms of pairing [37 − 39]. For instance, in the case of pure s-wave pairing, for all $r = (i, i, \sigma)$ the elements of the eigenvector are non-zero and the same whereas they are zero for all $r = (i, j, \sigma)$, $i \neq j$. In general, having obtained this eigenvector, it is a simple matter to identify the dominant pairing mechanism.

One could also use flux quantization [33, 40] or the superfluid density [41] as criteria to look for superconductivity in a particular model. Exact calculations for free fermions show that for systems of sizes accessible to numerical simulations, both these quantities display a very strong size dependence, making them less suited for our purpose [25, 42].

The evaluation of the two-particle density matrix (32) can be time consuming. The number of operations in the algorithm that we use to compute all entries of this matrix scales with $L^4 M_I$. For most of the systems that we have studied $M_I = \mathcal{O}(10^5)$, and the CPU time required to set up the two-particle density matrix is substantial.

There is ODLRO in a fermion system if the largest eigenvalue λ_0 of the $2L^2 \times 2L^2$ matrix $\rho_{r,s}$ grows linearly with the size of the system (assuming the density of particles is kept constant) [33]. Accordingly, a plot of λ_0 versus the system size will reveal whether or not the system exhibits ODLRO. It is also of interest to compute the on-site (s-wave) pairing correlation function

$$P_0 \equiv \frac{1}{L} \sum_{i,j} \langle c^+_{i,\uparrow} c^+_{i,\downarrow} c_{j,\downarrow} c_{j,\uparrow} \rangle \quad . \tag{33}$$

As the contributions to P_0 appear on the diagonal of the two-particle density matrix (32) we must have $P_0 \leq \lambda_0$, an inequality that is never violated by our numerical data. From (33) it is clear that there is ODLRO if $P_0 \propto L$ for large L, i.e. ODLRO of the on-site type.

A simple check on the numerical results is that they should satisfy the rigorous bound [33]

$$\lambda_0 \leq L \frac{n(2-n)}{2} + n \quad ; \quad nL \quad \text{even} \quad , \tag{34}$$

where $n = L^{-1} \sum_{i,\sigma} \langle n_{i,\sigma} \rangle$ denotes the density of particles. All our numerical results are in concert with (34).

On the one hand the computational effort required to compute the ground state energy and the reduced two-particle density matrix grows (exponentially) fast with the system size. On the other hand it is crucial to have data for significantly different system sizes in order for the plot of λ_0 versus L to be of any use at all. With this in mind it is of interest to start searching for ODLRO in one-dimensional (1D) systems [42]. SD results for two-dimensional models can be found in refs. [43, 44].

Although in a 1D model there can be no ODLRO at non-zero temperature in the strict sense [45], at $T = 0$ there can be ODLRO even in a 1D system. As the numerical method we employ is designed to compute the ground-state properties we may expect to find in our data clear signals for ODLRO whenever it is there. Due to the quantum fluctuations there can at most be "quasi" ODLRO in 1D systems with short-range interactions: The pairing correlation functions exhibit a

slow (power-law) decrease for large distances, resulting in a sublinear dependence of λ_0 on L.

In our numerical work on Hubbard-like models we adopt periodic boundary conditions. Our SD codes work either with the real-space or Fourier space representation and can take advantage of the spatial and spin symmetries of the model. Most of the data presented below have been obtained from runs that use all obvious tricks to reduce the size of the Hilbert space. For many of the systems studied, the calculations were carried out using both representations, providing a highly non-trivial consistency check. Occasionally some runs have been repeated without the use of symmetries. For small systems, the results of the SD calculations have been compared against those obtained from exact diagonalization and, as expected on theoretical grounds, no differences were found.

4.1. BCS REDUCED HAMILTONIAN

From pedagogical viewpoint it is important to have at least one example for which it is known that the system supports ODLRO. Such an example is provided by the Hamiltonian

$$H = -t \sum_{\langle i,j \rangle} \sum_{\sigma=\uparrow,\downarrow} \left(c^+_{i,\sigma} c_{j,\sigma} + c^+_{j,\sigma} c_{i,\sigma} \right) - \frac{|U|}{L} \sum_{i,j} c^+_{i,\uparrow} c^+_{i,\downarrow} c_{j,\downarrow} c_{j,\uparrow} \quad , \tag{35}$$

where $c^+_{i,\sigma}$ and $c_{i,\sigma}$ are the creation and annihilation operators, respectively, for a fermion with spin $\sigma =\uparrow,\downarrow$ at the site (or orbital) i and the sum over $\langle i,j \rangle$ is over distinct pairs of nearest neighbor lattice sites on a chain of length L. t is the hopping parameter and U is the pairing interaction. A variational, BCS-like treatment of (35) yields the exact solution [46], hence the name "BCS reduced Hamiltonian". As ODLRO is a characteristic feature of the BCS wave function [33] any numerical method that solves (35) should be able to reproduce this feature.

SD results [42] for the ground-state energy per site E/L, the on-site pairing correlation function P_0 and the largest eigenvalue λ_0 of the two-particle density matrix as a function of system size for half-filled rings are shown in Fig.1. For small system sizes E/L increases with L. For $L \geq 14$ the L-dependence of the ground-state energy is no longer visible on the scale used in Fig.1. For $6 \leq L < 22$ the largest eigenvalue λ_0 of the two-particle density matrix grows linearly with L, as expected since the system described by Hamiltonian (35) exhibits ODLRO [46]. For larger system sizes λ_0 decreases, indicating that the number of important states M_I that has been collected is too small for these system sizes. The number of important states M_I collected by the SD algorithm, working in the Fourier space representation, varies from $M_I \approx 6$ for $L = 4$ to $M_I \approx 100000$ for $L \geq 22$. The dimension of the Hilbert space varies from $M = 36$ for $L = 4$ to $M \approx 10^{14}$ for $L = 26$. The behavior of P_0 and λ_0 as a function of system size is identical, as expected in this case. Hence, the ODLRO exhibited by the system is mainly of the on-site (s-wave) pairing type.

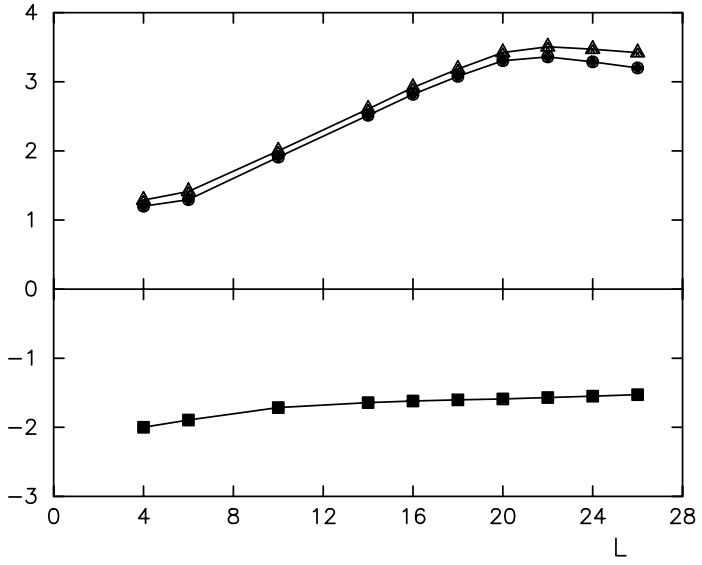

Fig.1. Ground-state energy per site E/L, on-site pairing correlation function P_0 and largest eigenvalue λ_0 of the reduced two-particle density matrix as a function of system size L for the BCS reduced Hamiltonian for $t = 1$, $U = -4$ and $n = 1$. Squares: E/L; bullets: P_0; triangles: λ_0. The lines are guides to the eye.

4.2. HUBBARD MODEL

The Hubbard model is the generic model for the description of electron correlations in narrow energy-band systems [47] and, because of its apparent simplicity, is often the model of choice for numerical work on correlated electron systems. The Hamiltonian of the Hubbard model reads

$$H^{Hub} = -t \sum_{\langle i,j \rangle} \sum_{\sigma=\uparrow,\downarrow} \left(c^+_{i,\sigma} c_{j,\sigma} + c^+_{j,\sigma} c_{i,\sigma} \right) + U \sum_i c^+_{i,\uparrow} c^+_{i,\downarrow} c_{i,\downarrow} c_{i,\uparrow} \quad , \qquad (36)$$

where U is the on-site Coulomb interaction.

In Fig.2 we present results for the attractive ($U < 0$) and repulsive ($U > 0$) Hubbard model for the case of a three-quarter filled band [42]. For $U > 0$, λ_0 does not increase with the system size. Hence the 1D three-quarter filled repulsive Hubbard model does not show ODLRO, as expected. For large negative U ($U = -4$ for example) λ_0 grows with L. This points to ODLRO. For $U = -0.2$ there is no noticeable increase of λ_0 with L. From a BCS treatment of the attractive Hubbard model [42] it follows that for $U = -0.2$, the size of an electron pair is much larger than the length of the rings we have studied with the SD method while for $U = -4$, the size of an electron pair is approximately one lattice site. Hence, due to these finite-size effects for small negative U, our numerical results cannot show the characteristic signal of ODLRO.

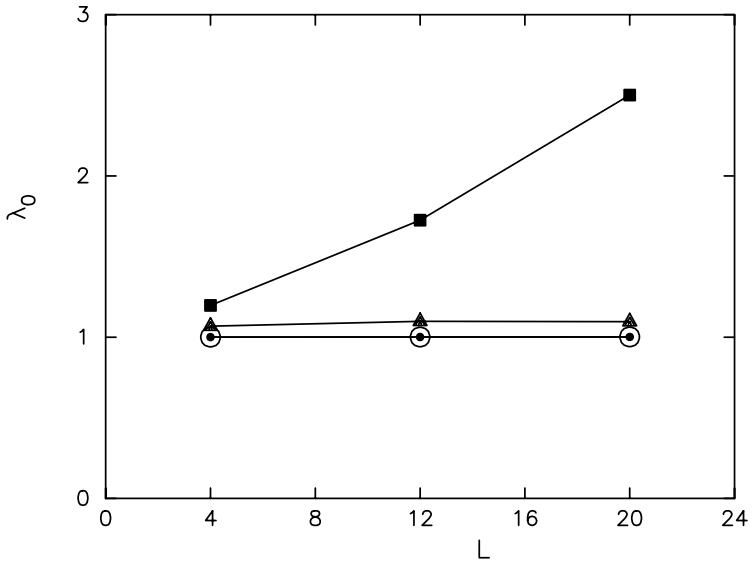

Fig.2. Largest eigenvalue λ_0 of the reduced two-particle density matrix as a function of system size L for the Hubbard model for $t = 1$ and $n = 1.5$. Squares: $U = -4$; bullets: $U = -0.2$; circles: $U = 0.2$; triangles: $U = 4$. The lines are guides to the eye.

4.3. HUBBARD MODEL WITH CORRELATED HOPPING

The tight binding Hamiltonian (for a single band) as derived by Hubbard contains several different types of interactions [47]. The Hubbard integrals $U = (ii|1/r|ii)$ (on-site) and $V = (ij|1/r|ij)$ (inter-site) set the strength of the interactions between electrons at the same site and neighboring sites, respectively. The correlated hopping amplitude $\Delta t = (ii|1/r|ij)$ describes the interaction between an electron hopping between two neighboring sites i and j and another electron localized either on site i or j, hence the name bond-charge site-charge interaction. The integral $X = (ii|1/r|jj)$ represents the interaction between electrons on the same bond. Here we will consider the case $V = X = 0$. Then the Hamiltonian reads [47]

$$H = H^{Hub} + \Delta t \sum_{<i,j>} \sum_{\sigma} \left(n_{i,-\sigma} + n_{j,-\sigma}\right) \left(c^{+}_{i,\sigma} c_{j,\sigma} + c^{+}_{j,\sigma} c_{i,\sigma}\right) \quad . \qquad (37)$$

This model was first studied by Caron and Pratt using a self-consistent cluster treatment [48]. For $\Delta t = t$ the exact ground state of the model at half-filling (including V) is known, for any dimension and a wide range of model parameters [49, 50] and in one dimension the model has been solved exactly away from half-filling [50 − 52]. In more than one dimension the qualitative form of the ground-state phase diagram for $\Delta t = t$ is basically the same as that of the ground-state phase diagram in one dimension although the exact location of all phase boundaries cannot be determined [52]. Exact diagonalization for chains up to 12 sites [53] and weak-coupling continuum-limit calculations [54] provide additional information on

(part of) the ground-state phase diagram.

It has been suggested that the correlated hopping interaction is essential for the occurrence of superconductivity [55, 56]. The Hubbard model with correlated hopping can be viewed as an effective one-band model for the CuO_2-planes of the cuprate superconductors $[57 - 59]$. The ground state of model (37) contains η-pairs and the η-paired states exhibit ODLRO [51, 52, 60]. However the presence of ODLRO in the η-paired states is not a sufficient condition for the existence of superconductivity [61].

Adding spin-flip hopping processes, it is possible to obtain the static and dynamic properties of the model and a complete picture of the full $(n, \Delta t/t, U/t)$ phase diagram $[62 - 69]$. For $\Delta t = t$ the qualitative form of the ground-state phase diagram is similar to the ground-state phase diagram of model (37) and the dimensionality of the lattice does not play an important role. From the phase diagram it follows that for $\Delta t = t$, model (37) with spin-flip hopping processes exhibits a continuous Mott metal-insulator transition at $n = 1$, $U = 4d|t|$ where d is the lattice dimensionality. For $0 < \Delta t < t$ model (37) with spin-flip hopping processes has a discontinuous metal-insulator transition at half-filling.

For $\Delta t = t$, $n = 1$, $U < -4t$ and zero temperature the BCS treatment yields the exact ground state with an on-site pairing correlation function given by [42] $P_0 = (L+1)/4$ showing that there is ODLRO in this case. Although at first sight there may be a flow of particles because $t \neq 0$, closer inspection reveals that the current operator acting on the ground state (with ODLRO) is identically zero, hence the ground state state is not superconducting [42]. This is due to the choice $\Delta t = t$ which implies conservation of local pairs of particles.

SD results for $\Delta t = t$ and $n = 1$ for rings of various lengths (results not shown) indicate that for $U > U_c$ the ground-state energy is zero and that no on-site electron pairs are formed. For $U < -U_c$ all electrons are paired, the pairs are static and the ground-state energy is equal to the number of pairs times U. For $L = 6$, $U_c = 3.5$; for $L = 10$, $U_c = 3.9$ and for rings with fourteen or more sites $U_c = 4$. These SD results [42] are in perfect agreement with the analytical results obtained in the thermodynamic limit $[49 - 52]$.

In Figs.3,4 we present SD results for the ground-state energy per site E/L, the on-site pairing correlation function P_0 and the largest eigenvalue λ_0 of the two-particle density matrix for the three-quarter filled Hubbard model with $\Delta t = 0.4$ and $t = 1$. The number of important states M_I collected by the SD algorithm, working in the Fourier space representation, varies from $M_I = 4$ for $L = 4$ to $M_I \approx 192000$ for $L = 28$. In the latter case the dimension of the Hilbert space $M \approx 1.4 \times 10^{12}$, so that $M_I \ll M$ indeed.

For $U = -1$ and $4 \leq L < 28$ the ground-state energy is almost constant, as shown in Fig.3. The largest eigenvalue λ_0 of the two-particle density matrix increases with L, indicating that the system exhibits ODLRO. The on-site pairing correlation function P_0 also increases with L but is significantly smaller than λ_0. Hence, the ODLRO is not of the pure on-site (s-wave) type. Analysis of the eigenvector of the two-body density matrix, corresponding to λ_0, shows that the ODLRO is mainly of the extended s-wave type.

For $U = 1$ (see Fig.4) the behavior of E/L, P_0 and λ_0 as a function of L is qualitatively the same as for $U = -1$ and $U = 0$ [42]. There is ODLRO, mainly of the extended s-wave type, and in a parameter regime where there is no special symmetry in the model and for which the continuum theory [54] does

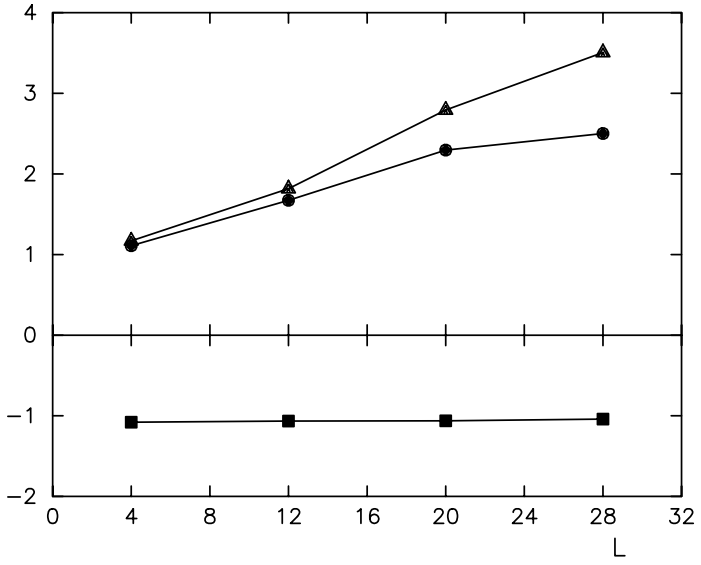

Fig.3. Ground-state energy per site E/L, on-site pairing correlation function P_0 and largest eigenvalue λ_0 of the reduced two-particle density matrix as a function of system size L for the Hubbard model with correlated hopping for $t = 1$, $\Delta t = 0.4$, $U = -1$ and $n = 1.5$. Squares: E/L; bullets: P_0; triangles: λ_0. The lines are guides to the eye.

not apply. This demonstrates, without invoking bosonization techniques or BCS-like arguments, that correlated hopping terms can lead to ODLRO in a system of electrons with a repulsive on-site interaction U. Since for small U we find ODLRO in the Hubbard model with correlated hopping whereas for the standard Hubbard model we do not find ODLRO, it seems that the correlated hopping interaction not only favors the formation of pairs but also reduces the size of the electron pairs.

To summarize: Our results suggest that the (repulsive) Hubbard model, supplemented with correlated hopping terms, can exhibit Off-Diagonal Long-Range Order for a wide range of model parameters.

5. Basis Set Optimization

The SD algorithm gives us a systematic, mathematically correct, recipe to collect the M_I most important contributions to the ground state wave function *for a fixed set of basis states* $\{|\phi_1\rangle, \ldots, |\phi_M\rangle\}$. As can be expected on general grounds, the convergence of the results with M_I depends on the particular choice of the basis states one makes. As discussed above, for the Hubbard-like models, our SD codes operate with states built from simple single-particle states, i.e. $|\psi_i\rangle = c_i^+|0\rangle$ or $|\varphi_k\rangle = c_k^+|0\rangle$ respectively (whenever possible we will suppress the spin labels from now on). It is obvious that the former will lead to poor performance if the ground state has an extended structure (in real space) and that the latter is not well-suited in cases where spin-up and spin-down electrons form localized (in real

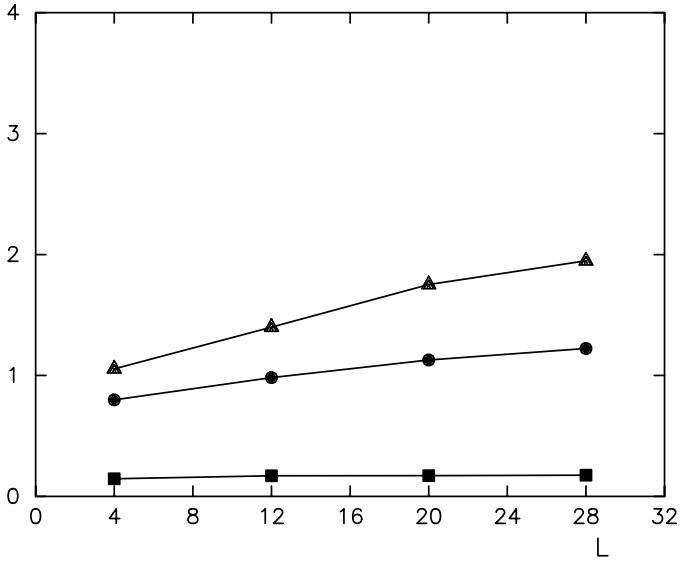

Fig.4. Ground-state energy per site E/L, on-site pairing correlation function P_0 and largest eigenvalue λ_0 of the reduced two-particle density matrix as a function of system size L for the Hubbard model with correlated hopping for $t = 1$, $\Delta t = 0.4$, $U = 1$ and $n = 1.5$. Squares: E/L; bullets: P_0; triangles: λ_0. The lines are guides to the eye.

space) pairs. Therefore it is to be expected that in situations where the many-body ground state describes an extended state of localized pairs of electrons (as in the case of ODLRO), the number of important states might be reduced if the SD would be carried out using single-particle states that are adaptive, in the sense that they can smoothly interpolate between the two extreme cases. In quantum chemistry problems one faces the same problem: To start say a calculation one first has to make a choice for the single-particle orbitals. In most cases the problem is "solved" through a combination of a lot of experience and knowledge about the problem at hand.

In this section we discuss an attempt to approach the problem of optimizing the single-particle basis states in a systematic manner [28]. As a toy model we will use the Hubbard model. For simplicity we only will consider N-particle states $|\phi_n\rangle$ that take the form of a Slater determinant, i.e.

$$|\phi_n\rangle = \sum_{i_j=1}^{K} a^{(n)}(i_1,1)a^{(n)}(i_2,2)\ldots a^{(n)}(i_N,N)c_{i_1}^{+}c_{i_2}^{+}\ldots c_{i_N}^{+}|0\rangle \quad , \qquad (38)$$

where the number of single-particle states (= number of lattice sites in the case of the Hubbard model, ignoring spin) is denoted by K.

It is convenient to arrange all the coefficients $a^{(n)}(i,j)$ of the complete set of K single-particle states into a $K \times K$ matrix

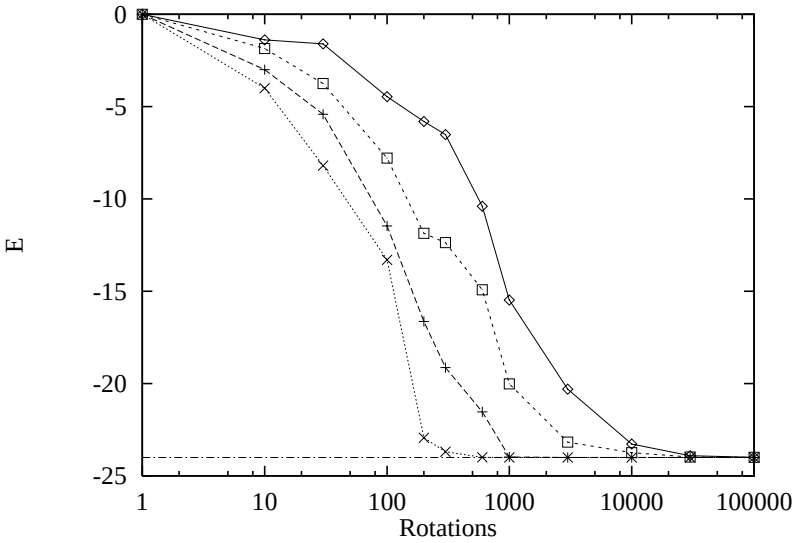

Fig.5. The ground state energy E as a function of the number of plane rotations, as obtained by optimizing one Slater determinant for the case of a 4×4 Hubbard model with $N_\uparrow = N_\downarrow = 5$, $t = 1$ and $U = 0$. The single-particle wave functions used to construct the initial many-body wave function are $|\psi_i\rangle = c_i^+|0\rangle$ where i refers to the i-th lattice site. The various symbols correspond to different optimization strategies (see text). The exact result $E_1 = -24$ is indicated by dashed-dotted line. Other lines are guides to the eye.

$$A^{(n,K)} = \begin{pmatrix} a^{(n)}(1,1) & a^{(n)}(1,2) & \dots & a^{(n)}(1,K) \\ a^{(n)}(2,1) & a^{(n)}(2,2) & \dots & a^{(n)}(2,K) \\ \vdots & \vdots & \vdots & \vdots \\ a^{(n)}(K,1) & a^{(n)}(K,2) & \dots & a^{(n)}(K,K) \end{pmatrix} \quad . \tag{39}$$

The N-particle state (38) is build from the first N columns of $A^{(n,K)}$. As the calculation of matrix elements $\langle\phi_{n'}|H|\phi_n\rangle$ (for details see [28]) simplifies considerably if the single-particle states are orthonormal [28] we will from now on assume that the initial $A^{(n,K)}$ is a unitary matrix, i.e. $A^{(n,K)}(A^{(n,K)})^\dagger = 1$ [1]. For example, adopting the real-space approach corresponds to working with the *fixed* matrix

$$A^{(n,K)} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 1 \end{pmatrix} \quad . \tag{40}$$

Let us now consider simple and efficient methods to optimize the sub-matrix $A^{(n,N)}$ in the sense that it minimizes the current approximation to the ground state energy. One simple way to transform $A^{(n,K)}$ is to perform the operation [28]

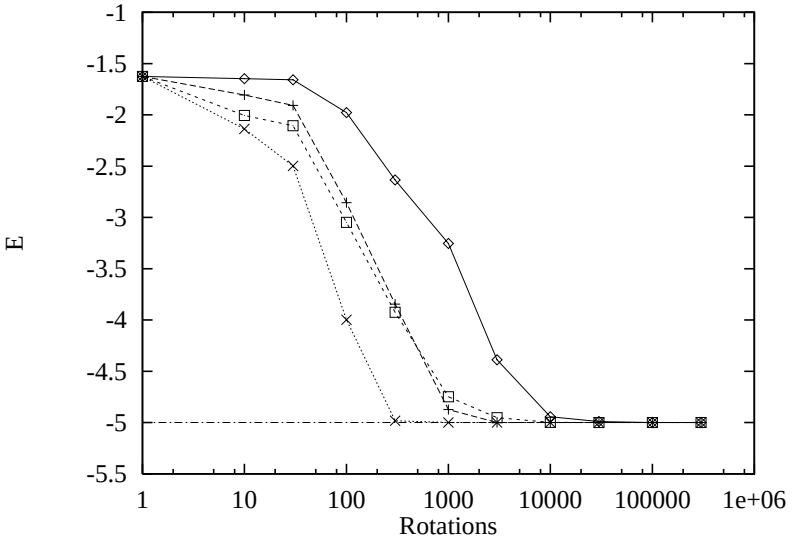

Fig.6. The ground state energy E as a function of the number of plane rotations as obtained by optimizing one Slater determinant for the case of a 4×4 Hubbard model with $N_\uparrow = N_\downarrow = 5$, $t = 0$ and $U = -1$. The single-particle wave functions used to construct the initial many-body wave function are $|\varphi_k\rangle = c_k^+|0\rangle$ where k denotes a wave vector of the reciprocal lattice. The various symbols correspond to different optimization strategies (see text). The exact result $E_1 = -5$ is indicated by the dashed-dotted line. Other lines are guides to the eye.

$$A^{(n,K)} \leftarrow A^{(n,K)} U^{(p,q)} \quad , \tag{41}$$

where $U^{(p,q)}$ is the plane rotation matrix (10) of size $K \times K$. As a (Slater) determinant does not change if we replace a row (column) by a linear combination of rows (columns) it is sufficient to let $p \leq N$, and $N < q \leq K$. In practice (41) tells us how to replace the columns p and q by a linear combination of them. Alternatively we can "mix" two arbitrary elements of each of the column vectors of $A^{(n,K)}$ by computing [28]

$$A^{(n,K)} \leftarrow U^{(p,q)} A^{(n,K)} \quad . \tag{42}$$

Both (41) and (42) preserve the unitary character of $A^{(n,K)}$ and can be used to define a "dynamics" for changing Slater determinants [70]. Including the electron spin requires a minor extension: We only have to replace each $|\phi_n\rangle$ by a product of two similar functions, one for the electrons with spin up and one for the electrons with spin down [28]. For details on the calculation of matrix elements of the type $\langle \phi_{n,\uparrow} \otimes \phi_{n,\downarrow} | H | \phi_{n',\uparrow} \otimes \phi_{n',\downarrow} \rangle$ see Ref. [28].

In general there are several possibilities for optimizing the single-particle states. Not only do we have the choice between (41) and (42) but we can also decide (not) to use the same plane rotation to the spin-up and spin-down part of the wave func-

Table 4. The ground state energy of the 4×4 Hubbard model with $N_\uparrow = N_\downarrow = 5$ and $t = 1$ as obtained by various methods. E_1^{Exact}: Exact diagonalisation. $E_1^{(k)}$: SD + basis-set optimization using $|\varphi_k\rangle = c_k^+|0\rangle$ as initial single-particle states. $E_1^{(i)}$: SD + basis-set optimization using $|\psi_i\rangle = c_i^+|0\rangle$ as initial single-particle states. E_1^{SD}: SD. E_1^{HF}: Hartree-Fock. In all SD calculations the number of many-body wave functions $M_I = 200$. A greedy algorithm was used to optimize the single-particle states, using 10000 plane rotations for each new many-body state added.

U	E_1^{Exact}	$E_1^{(k)}$	$E_1^{(i)}$	E_1^{SD}	E_1^{HF}
6	−18.35837	−18.09710	−17.93031	−17.41384	−14.62500
4	−19.58094	−19.50157	−19.42179	−19.24434	−17.75000
2	−21.37695	−21.36579	−21.34991	−21.32807	−20.87500
0	−27.12500	−27.12500	−27.12500	−27.12500	−27.12500
−2	−27.70702	−27.69305	−27.68078	−27.64086	−27.12500
−4	−32.73360	−32.51795	−32.45901	−32.05033	−30.25000
−6	−39.34051	−38.19426	−38.27180	−36.85642	−33.37500

tion. Furthermore we have the option to use different methods for chosing the angle of the plane rotations. In Figs.5,6 we present some results for a very simple case that illustrate the effect of using different optimization strategies. We use only one wave function $|\phi_{1,\uparrow} \otimes \phi_{1,\downarrow}\rangle$ and minimize

$$\langle \phi_{1,\uparrow} \otimes \phi_{1,\downarrow}|H|\phi_{1,\uparrow} \otimes \phi_{1,\downarrow}\rangle \quad . \tag{43}$$

For the case of the Hubbard model with $U = 0$ (Fig.5) and $t = 0$ (Fig.6) the initial state was taken to be the ground states for $t = 0$ (Fig.5) and $U = 0$ (Fig.6) respectively, i.e. the most unfavourable initial state.

The results represented by diamonds have been obtained by using dynamics (42) with p, q, and the rotation angle all chosen randomly. At each step we change either the spin-up or the spin-down part of the wave function. Data marked by squares has been obtained by applying the same (random) plane rotation to both spin components of the wave function. With some additional effort for each pair (p,q) it is possible to find the angle that yields the maximum decrease of the energy. From Figs.5,6 it is clear that application of the "best" plane rotation ($\times$ and $+$) seems the most effective of the four strategies used. However the calculation of the optimal angle, although not complicated, takes additional CPU time that compensates for the reduction of the number of rotations that results from it (for details see [28]).

In the case of an attractive interaction ($U < 0$) between electrons with different spin, performing the same rotation on both components simultaneously ($\times$) instead of rotating the components separately ($+$) yields a substantial improvement. For $U > 0$ the opposite behavior is found (not shown) [28]. Of course this behavior is closely linked to the difference in physical behavior. The main point in the present context is that the optimization procedure changes the state such that it describes the physics as good as possible.

The next step is to combine the procedure of optimizing the single-particle

wave functions with matrix-inflation. In principle this can be done along the lines described above but in practice a number of technical but nevertheless important complications arise [28]. For a detailed description of various optimization strategies and a lot of results see [28]. Some results of SD calculations with and without optimized single-particle states are given in Table 4. It is clear that for a fixed number of important states M_I, the optimization can bring substantial improvements over the standard SD, especially in the intermediate coupling regime ($U \approx 4|t|$).

6. Outlook

Stochastic diagonalization (SD) is an importance sampling method that does not suffer from the minus-sign problem. It is complementary to existing Quantum Monte Carlo methods that compute ground state properties of quantum many-body systems. The potential of the SD approach has been demonstrated in applications to quantum chemistry problems and to models of strongly correlated electron systems. The results of a first attempt to combine ideas of SD and basis-set optimization suggest that this might be a fruitful direction for further research.

7. Acknowledgements

This work has been supported by the Human Capital and Mobility program of the EEC, the "Deutsche Forschunggemeinschaft (DFG)", and the "Stichting Nationale Computer Faciliteiten (NCF)".

8. References

1. Wilkinson, J.H. (1965) *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford.
2. Parlett, B.N. (1981) *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs.
3. Cullum, J.K. and Willoughby, R.A. (1985) *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, Birkäuser, Boston .
4. Davidson, E.R. (1975) The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices, *J. Comp. Phys.* **17**, 87 – 94.
5. Olsen, J., Jørgensen, P., and Simons, J. (1990) Passing the one-billion limit in full configuration interaction (FCI) calculations, *Chem. Phys. Lett.* **169**, 463 – 472.
6. Dagotto, E. and Moreo, A. (1985) Improved Hamiltonian variational technique for lattice models, *Phys. Rev.* **D31**, 865 – 870.
7. Gagliano, E., Dagotto, E., Moreo, A., and Alcaraz, F. (1986) Correlation functions of the antiferromagnetic Heisenberg model using a modified Lanczos method, *Phys. Rev.* **B34**, 1677 – 1682.
8. Fano, G., Ortolani, F., and Parola, A. (1992) Electron correlations in the two-dimensional Hubbard model – A group-theoretical and numerical study, *Phys. Rev.* **B46**, 1048 – 1060.
9. Hammersley, J.M. and Handscomb, D.C. (1964) *Monte Carlo Methods*, Methuen, London.
10. Binder, K. (1979) Introduction: Theory and Technical Aspects of Monte Carlo Simulations, in K. Binder (ed.), *Monte Carlo Methods in Statistical Physics – Topics in Current Physics 7*, Springer, Berlin, pp. 1 – 45.
11. Binder, K. and Stauffer, D. (1984) A simple introduction to Monte Carlo simulation and some specialized topics, in K. Binder (ed.), *Applications of the Monte Carlo Methods in*

Statistical Physics – Topics in Current Physics 36, Springer, Berlin, pp. 1 – 36.
12. Schmidt, K.E and Kalos, M.H. (1984) Few- and many-fermion problems, in K. Binder (ed.), *Applications of the Monte Carlo Methods in Statistical Physics – Topics in Current Physics 36*, Springer, Berlin, pp. 125 – 143.
13. Feynman, R.P. and Hibbs, A.R. (1965) *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York .
14. De Raedt, H. and Lagendijk, A. (1985) Monte Carlo simulation of quantum statistical lattice models, *Phys. Rep.* **127**, 233 – 307.
15. Suzuki, M., Miyashita, S., and Kuroda, A. (1977) Monte Carlo simulation of quantum spin systems I, *Prog. Theor. Phys.* **58**, 1377 – 1387 .
16. Suzuki, M. (1985) Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics, *J. Math. Phys.* **26**, 601 – 612.
17. Suzuki, M. (1986) General review of quantum statistical Monte Carlo methods, in M. Suzuki (ed.), *Quantum Monte Carlo Methods – Solid State Sciences 74*, Springer, Berlin, pp. 2 – 22.
18. De Raedt, H. and von der Linden, W. (1992) Quantum Lattice problems, in K. Binder (ed.), *Monte Carlo Methods in Condensed Matter Physics – Topics in Applied Physics 71*, Springer, Berlin, pp. 249 – 284.
19. De Raedt, H. (1996) Quantum Theory, in K. Binder and G. Ciccotti (eds.), *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, SIF, Bologna, pp. 401 – 442.
20. We exclude from this discussion techniques that implement the variational principle based on a pre-defined form of a trial wave function. In this lecture we will concentrate on methods that can, be it only in principle, yield the exact answer if enough CPU time is available.
21. De Raedt, H. and von der Linden, W. (1992) Monte-Carlo diagonalization of many-body problems – Applications to fermion systems, *Phys. Rev.* **B45**, 8787 – 8790.
22. De Raedt, H. and Frick, M. (1993) Stochastic diagonalization, *Phys. Rep.* **231**, 107 – 149.
23. Hetherington, J.H. (1984) Observations on the statistical iteration of matrices, *Phys. Rev.* **A30**, 2713 – 2719.
24. Bellman, R. (1960) *Introduction to Matrix Analysis*, Maple Press, York.
25. Michielsen, K. and De Raedt, H. (1997) Quantum molecular dynamics study of the Su-Schrieffer-Heeger model, *Z. Phys.* **B103**, 391 – 400 .
26. De Raedt, H., Michielsen, K., and van Dijk, L. (1997) Rigorous Bounds on the Free Energy of Electron-Phonon Models, *Int. J. Mod. Phys.* **B11**, 1591 – 1605.
27. Fettes, W., Morgenstern, I., and Husslein, T. (1997) Parallelization of the exact diagonalization of the t-t' Hubbard model, *Comp. Phys. Comm.* **106**, 1 – 9.
28. Fettes, W. (1998) Supraleitung im tt'-Hubbard-Modell und in BCS-reduzierten Hubbard-Modellen, Ph.D Thesis, University of Regensburg.
29. ten Haaf, D.F.B., van Bemmel, H.J.M., van Leeuwen, J.M.J., van Saarloos, W., and Ceperley, D.M. (1995) Low-temperature behavior of the large-U Hubbard model from high-temperature expansions, *Phys. Rev.* **B51**, 353 – 367.
30. Visscher, L., De Raedt, H., and Nieuwpoort, W.C. (1994) A new configuration selection method for configuration-interaction calculations, *Chem. Phys. Lett.* **227**, 327 – 336.
31. Penrose, O. (1951) On the Quantum Mechanics of Helium II, *Phil. Mag.* **42**, 1373 – 1377.
32. Penrose, O. and Onsager, L. (1956) Bose-Einstein condensation and liquid helium, *Phys. Rev.* **104**, 576 – 584.
33. Yang, C.N. (1962) Concept of off-diagonal long-range order and the quantum phases of liquid He and of superconductors, *Rev. Mod. Phys.* **34**, 694 – 704.
34. Sewell, G.L. (1990) Off-diagonal long-range order and the Meissner effect, *J. Stat. Phys.*

61, 415 – 422.
35. Nieh, H.T., Su, G., and Zhao, B.-H. (1995) Off-diagonal long-range order – Meissner effect and flux-quantization, *Phys. Rev.* **B51**, 3760 – 3764.
36. Au,C. and Zhao, B.-H. (1995) From ODLRO to the Meissner effect and flux-quantization, *Phys. Lett.* **A209**, 235 – 240.
37. Yang, C.N. (1989) η-pairing and off-diagonal long-range order in a Hubbard model, *Phys. Rev. Lett.* **63**, 2144 – 2147.
38. Betsuyaku, H. (1991) η-pairing and superconductivity in the negative-U Hubbard model, *Phys. Rev.* **B44**, 871 – 874.
39. Singh, R.R.P. and Scalettar, R.T. (1991) Exact demonstration of η-pairing in the ground-state of an attractive-U Hubbard-model, *Phys. Rev. Lett.* **66**, 3203 – 3204.
40. Byers, N. and Yang, C.N. (1961) Theoretical considerations concerning quantized magnetic flux in superconduction cylinders, *Phys. Rev. Lett.* **7**, 46 – 49.
41. Fisher, M.E., Barber, M.N., and Jasnow, D. (1973) Helicity modulus, superfluidity, and scaling in isotropic systems, *Phys. Rev.* **A8**, 1111 – 1124.
42. Michielsen, K. and De Raedt, H. (1997) Off-Diagonal Long-Range Order in Generalized Hubbard Models, *Int. J. Mod. Phys.* **B11**, 1311 – 1335.
43. Husslein, T., Fettes, W., and Morgenstern, I. (1997) Comparison of calculations for the Hubbard model obtained with quantum-Monte-Carlo, exact, and stochastic diagonalization, *Int. J. Mod. Phys.* **C8**, 397 – 415.
44. Fettes, W., Morgenstern, I., and Husslein, T. (1997) Application of the SD technique for solving a BCS-reduced Hubbard like Hamiltonian, *Int. J. Mod. Phys.* **C8**, 1037 – 1061.
45. Hohenberg, P.C. (1967) Existence of long-range order in one and two dimensions, *Phys. Rev.* **158**, 383 – 386.
46. Bursill, R.J. and Thompson, C.J. (1993) Rigorous treatment of the BCS model of superconductivity, *J. Phys. A: Math. Gen.* **26**, 769 – 786.
47. Hubbard, J. (1963) Electron correlations in narrow energy bands, *Proc. Roy. Soc. London* **A276**, 238 – 257.
48. Caron, L.G. and Pratt Jr., G.W. (1968) Correlation and magnetic effects in narrow energy bands II, *Rev. Mod. Phys.* **40**, 802 – 806.
49. Strack, R. and Vollhardt, D. (1993) Hubbard-model with nearest-neighbor and bond-charge interaction – Exact ground-state solution in a wide-range of parameters, *Phys. Rev. Lett.* **70**, 2637 – 2640.
50. Ovchinnikov, A.A. (1994) Metal-inulator-transition in the generalized Hubbard-model, *J. Phys.: Condens. Matter* **6**, 11057 – 11069.
51. Arrachea, L. and Aligia, A.A. (1994) Exact solution of a Hubbard chain with bond-charge interaction, *Phys. Rev. Lett.* **73**, 2240 – 2243.
52. Schadschneider, A. (1995) Superconductivity in an exactly solvable Hubbard-model with bond-charge interaction, *Phys. Rev.* **B51**, 10386 – 10391.
53. Arrachea, L., Aligia, A.A., Gagliano, E., Hallberg, K., and Balseiro, C. (1994) Superconducting correlations in Hubbard chains with correlated hopping, *Phys. Rev.* **B50**, 16044 – 16051.
54. Japaridze, G. and Müller-Hartmann, E. (1994) Electrons with correlated hopping in one dimension, *Ann. Physik* **3**, 163 – 180.
55. Hirsch, J.E. (1989) Bond-charge repulsion and hole superconductivity, *Physica* **C158**, 326 – 336.
56. Hirsch, J.E. (1989) Coulomb attraction between Bloch electrons, *Phys. Lett.* **A138**, 83 – 87.
57. Schüttler, H.B. and Fedro, A.J. (1992) Copper-oxygen charge excitations and the effective

single-band theory of cuprate superconductors, *Phys. Rev.* **B45**, 7588 – 7591.
58. Simón, M.E., Baliña, M., and Aligia A.A. (1993) Effective one-band Hamiltonian for cuprate superconductor metal-insulator-transition, *Physica* **C206**, 297 – 304.
59. Simón, M.E. and Aligia, A.A. (1993) Brinkman-Rice transition in layered perovskites, *Phys. Rev.* **B48**, 7471 – 7477.
60. de Boer, J., Korepin, V.E., and Schadschneider, A. (1995) η-pairing as a mechanism of superconductivity in models of strongly correlated electrons, *Phys. Rev. Lett.* **74**, 789 – 792.
61. Arrachea, L., Aligia, A.A., and Gagliano, E. (1996) Anomalous flux quantization in a Hubbard ring with correlated hopping, *Phys. Rev. Lett.* **76**, 4396 – 4399.
62. Michielsen, K., De Raedt, H., and Schneider, T. (1992) Metal-insulator-transition in a generalized Hubbard-model, *Phys. Rev. Lett.* **68**, 1410 – 1413.
63. Michielsen, K. (1993) Metal-insulator transitions and strong electron correlations, *Int. J. Mod. Phys.* **B7**, 2571 – 2653.
64. de Vries, P., Michielsen, K., and De Raedt, H. (1993) Gaps in densities of states of two Hubbard-like models, *Phys. Rev. Lett.* **70**, 2463 – 2466.
65. de Vries, P., Michielsen, K., and De Raedt, H. (1993) The simplified Hubbard model in one and two dimensions – Thermodynamic and dynamic properties, *Z. Phys.* **B92**, 353 – 362.
66. Michielsen, K., De Raedt, H., Schneider, T., and de Vries, P. (1994) Finite-temperature phase-transition in the Montorsi-Rasetti model, *Europhys. Lett.* **25**, 599 – 604.
67. de Vries, P., Michielsen, K., and De Raedt, H. (1994) Single-particle self-energy and optical conductivity of the simplified Hubbard model, *Z. Phys.* **B95**, 475 – 479.
68. Michielsen, K. (1994) Bond-charge site-charge interaction and metal-insulator transitions, *Phys. Rev.* **B50**, 4283 – 4291.
69. Michielsen, K. and De Raedt, H. (1994) First-order phase transitions in the Montorsi-Rasetti model, *Phys. Rev.* **E50**, 4371 – 4379.
70. This can be useful in Car-Parinello algorithms, eliminating the need for re-orthogonalization of the wave functions after a MD step.