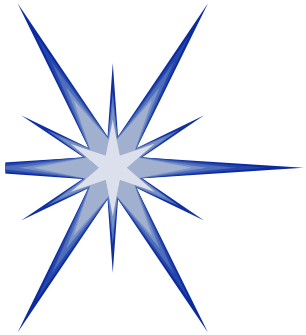
Stochastic Diagonalization

H. De Raedt

Institute for Theoretical Physics
and Materials Science Centre
University of Groningen,
The Netherlands

Work done in collaboration with:

W.Fettes, (Regensburg, D)
K. Michielsen, (Groningen, NL)



Outline

- **Introduction**
  - **Formulation of the problem**
- **Minus-sign problem in Quantum Monte Carlo methods**
  - **General aspects only**
- **Stochastic diagonalization**
  - **Theory**
  - **Comparison with other (QMC) methods**
    - **2D *tt'* Hubbard model**
    - **Quantum Chemistry Problems**
- **Application:**
  - **Off-diagonal long-range order, superconductivity**
- **Basis-set optimization**
- **Conclusions**

Formulation of the problem

- **Quantum statistical mechanics, (lattice) QED, (lattice) QCD:**

$$\langle X \rangle = \frac{\mathrm{Tr} e^{-\beta H} X}{\mathrm{Tr} e^{-\beta H}} = \frac{\sum_{\{\psi\},\{\varphi\}} \langle \varphi | e^{-\beta H} | \psi \rangle \langle \psi | X | \varphi \rangle}{\sum_{\{\psi\}} \langle \varphi | e^{-\beta H} | \varphi \rangle}$$

  - ***H*** **: Hamiltonian**
  - ***X*** **: Observable**
  - $\{\psi\},\{\varphi\}$ **: Complete sets of states**
  - $\beta = 1/k_B T$ **: Inverse temperature**
- **Quantum dynamics:**

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = H |\Psi(t)\rangle \quad \Leftrightarrow \quad |\Psi(t)\rangle = e^{-itH/\hbar} |\Psi(0)\rangle$$

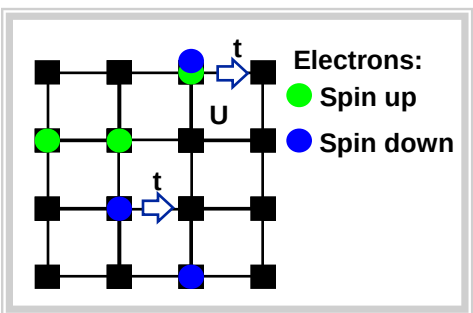
- **Ground-state (*T*=0) properties:**

$$H |\Phi_1\rangle = E_1 |\Phi_1\rangle$$

  - E_1 **is the smallest eigenvalue of** ***H***

Formulation of the problem

- **For numerical purposes, a state of the quantum system, i.e. the wave function, is represented by a vector of length *M* (the dimension of the Hilbert space) and the Hamiltonian *H* is a hermitian matrix of dimension *M* x *M* .**
- **We know what we have to do: "Just" solve the Schrödinger equation** $H|\Phi\rangle = E|\Phi\rangle$
  - **A standard problem of linear algebra**
    - **Example: The Hubbard-Gutzwiller model (1963)**



- **Describes fermions (electrons) that hop from one lattice point to another and that feel each other via an on-site repulsion (*U>0*).**

Formulation of the problem

- **Count the number of possible ways to distribute *N* electrons with spin up and $N_\downarrow$ electrons with spin down over *L* lattice sites:**

$$M = \begin{pmatrix} L \\ N_\uparrow \end{pmatrix} \mathrm{x} \begin{pmatrix} L \\ N_\downarrow \end{pmatrix}$$

- **For $N = N_\downarrow = L/2$:** $M \approx \dfrac{2^{2L+2}}{2\pi L}$

$$Memory \approx \frac{2^{2L-25}}{2\pi L} \text{ Giga bytes}$$

L	$N_\uparrow = N_\downarrow$	M	Memory in Gb	Approx. cost
16	8	2×10^8	2	$\$ 10^3$
36	18	8×10^{19}	6×10^{11}	$\$ 10^{14}$
64	32	3×10^{36}	3×10^{28}	$\$ 10^{20}$

Very Large Matrices !

Formulation of the problem

- **In general, there is no way to solve these monster eigenvalue problems using perturbative or conventional numerical (Eispack, Lanczos, Davidson, ...) methods.**
- **Key issue**

How to solve (eigenvalue) problems involving extremely large matrices ?

- **General strategy: Trade memory for CPU time**
  - **Quantum Monte Carlo (QMC) Methods**
  - **Stochastic Diagonalization**

Quantum Monte Carlo Methods

- **Stochastic methods:**
  - **Use much less than *M* storage elements but use much more CPU time**
  - **Example: A QMC method based on**

$$\lim_{m\to\infty}(H-E_{shift})^m\left|\Phi_{initial}\right\rangle=\left|\Phi_1\right\rangle$$

    - **Linear algebra: Power method**

$$(H-E_{shift})^m\left|\Phi_{initial}\right\rangle=\sum_{i=1}^{M}(E_i-E_{shift})^m\left|\varphi_i\right\rangle\left\langle\varphi_i\right|\Phi_{initial}\rangle$$

$$\xrightarrow{m\to\infty}\max_i(E_i-E_{shift})^m\left|\varphi_i\right\rangle\left\langle\varphi_i\right|\Phi_{initial}\rangle\propto\left|\Phi_1\right\rangle$$

      - **for a "good" choice of E_{shift}**
  - **Physical quantities are related to matrix elements:**

$$\left\langle\Phi_i\right|(H-E)^m\left|\Phi_f\right\rangle=\sum_{\{\Phi_k\}}\left\langle\Phi_i\right|H-E\left|\Phi_1\right\rangle\left\langle\Phi_1\right|H-E\left|\Phi_2\right\rangle\cdots\left\langle\Phi_{m-1}\right|H-E\left|\Phi_f\right\rangle$$

    - **Takes approximately *mM* operations**
    - **Does not help much if we would like to sum all terms**

Quantum Monte Carlo Simulation: General Strategy

- **Importance sampling**
  - **Basic assumption *(hope)* : If the problem can be solved at all, it must be such that not all the states of the system are equally important. Using only a few of them should be sufficient to compute the physical properties with reasonable accuracy.**
    - **In practice a "few " can still mean 10^4 - 10^8 states**
    - **No a-priori justification in the case of quantum systems**
    - **OK for classical statistical problems**
  - **How to find these important states ?**
    - **Random search gives bad results**

    - **Metropolis Monte Carlo method, molecular dynamics, Langevin dynamics, ...**

Importance sampling

- **Importance sampling:**
  - **Search for *important* states uses**

    $$\rho = \rho(\{\Phi_k\}) = \langle \Phi_0 | F | \Phi_1 \rangle \langle \Phi_1 | F | \Phi_2 \rangle \cdots \langle \Phi_{m-1} | F | \Phi_m \rangle$$

    as a "density" distribution for the states $\{\Phi_k\}$
    - ***F* depends on the application**
- **Classical statistical mechanics**

  $$F_{Boltzmann}(H) = e^{-\beta H} \quad ; \quad H \in \Re$$
- **Diffusion QMC (power method):**

  $$F_{DQMC}(H) = (1 - \tau H)^m$$
- **Green Function QMC (inverse iteration:**

  $$F_{GFMC}(H) = [(E_{shift} + H)^{-1}]^m$$
- **Path Integral MC (Trotter-Suzuki product formulae)**

  $$F_{PI}(H = A + B) = \left(e^{-\beta A/m} e^{-\beta B/m}\right)^m$$

Quantum Monte Carlo Methods: General Strategy

- **All importance sampling methods require a probability distribution**
  - **Only positive functions qualify as probability distributions**
- **Is $\rho > 0$ for all possible states ?**
  - **In classical mechanics (Boltzmann factor): ALWAYS**
  - **In quantum mechanics: Almost NEVER**
    - **"Theorems": If all the off-diagonal elements of the matrix representing the Hamiltonian are negative then $\rho > 0$**
      - **H. De Raedt and M . Frick, Phys. Rep. 231, 107-149 (1993)**
      - **The reverse is NOT true but it is save to expect it is.**

Minus Sign Problem

Theorems on the positivity of $F(H)$

- **Classical statistical mechanics**

  $$F_{Boltzmann}(H) = e^{-\beta H} > 0 \quad ; \quad H \in \Re$$

- **DQMC:** $F_{DQMC}(H) = (1-\tau H)^m$

  If $\langle\varphi|H|\psi\rangle \le 0$ for all $|\varphi\rangle, |\psi\rangle$

  then $\langle\varphi|(1-\tau H)|\psi\rangle \ge 0$

- **GQMC:** $F_{GFMC}(H) = [(E_{shift} + H)^{-1}]^m$

  If $\omega + H$ is positive definite and $\langle\varphi|H|\psi\rangle \le 0$

  for all $|\varphi\rangle \ne |\psi\rangle$ then $\langle\varphi|(\omega + H)^{-1}|\psi\rangle > 0$

- **Trotter-Suzuki:** $F_{PI}(H = A + B) = \left(e^{-\beta A/m} e^{-\beta B/m}\right)^m$

  If $\langle\varphi|X|\psi\rangle \le 0$ for all $|\varphi\rangle \ne |\psi\rangle$

  then $\langle\varphi|e^{-\tau X}|\psi\rangle \ge 0$ for all $\tau > 0$

- **Auxilary-field QMC:**

  $$F \propto \det(1 + e^{A_1} \ldots e^{A_m}) = ?$$

Quantum Monte Carlo Minus Sign Problem

- **Why should negative contributions bother us at all ?**
  - **Importance sampling without (mathematical) justification**
    - **Trick: Transfer sign to measurement**
      - **H. De Raedt and A. Lagendijk, Phys. Rev. Lett. 46, 77 (1981)**

$$\langle A \rangle = \mathrm{Tr}\rho A = \mathrm{Tr} \mid \rho \mid sign(\rho) A = \frac{\langle sign(\rho) A \rangle_{|\rho|}}{\langle sign(\rho) \rangle_{|\rho|}}$$

  - **Simulation methods suffer from statistical errors**
    - **This trick fails if $\langle sign(\rho) \rangle_{|\rho|} \to 0$ which is usually the case if**
      - **the number of factors in the product formula increases, or**
      - **the system size increases, or**
      - **the temperature decreases, or**
      - **...**

Minus Sign Problem: Simple Example

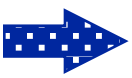
- **The minus sign problem can also appear when there are no fermions !**
- **Example: A spin-1/2 model**

$$-\beta H = a\sigma^x - b\sigma^y = \begin{pmatrix} 0 & a+ib \\ a-ib & 0 \end{pmatrix}$$

Generic QMC approach: Rewrite the partition function using the Trotter-Suzuki formula:

$$Z = \lim_{m\to\infty} Z_m \quad ; \quad Z_m = \mathrm{Tr}\left(e^{a\sigma^x/m} e^{-b\sigma^y/m}\right)^m$$

$$1)\ Z_1 = \cosh a \cosh b\, \mathrm{Tr}(1 + i\sigma^z \tanh a \tanh b)$$



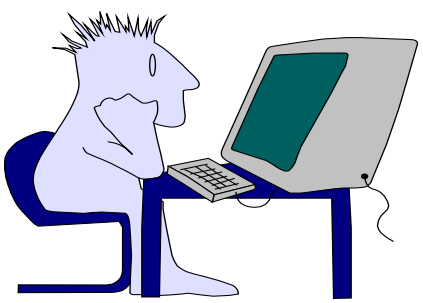
None of the contributions is REAL !

$$2)\ \text{Change of representation :} \quad -\beta H = a\sigma^x - b\sigma^z$$

$$Z_m = \sum_{\{s_j=\pm 1\}} \prod_{j=0}^{m-1} \left[\delta_{s_j,s_{j+1}} + (1-\delta_{s_j,s_{j+1}}) \tanh a\right] e^{-bs_j/m} \cosh\frac{a}{m}$$

Quantum Monte Carlo Methods: Summary

- **Quantum Monte Carlo methods work well if there is no minus sign problem**
- **The minus sign problem is due to the use of a *product formula* in combination with an *importance sampling algorithm***
- **Use of importance sampling seems unavoidable (just too many possibilities)**
- **Without using some product formula it seems very difficult to compute physical quantities without making (uncontrolled) approximations**



Stochastic Diagonalization

- **Method to compute the smallest eigenvalue and corresponding eigenvector of very (very) large matrices.**
  - **Basic idea: Assume that of all possible states of the quantum system, only a relatively small fraction is "important"**
    - **The same assumption is at the heart of all Quantum Monte Carlo techniques**
    - **How to find these important states ?**
      - **HDR & W. von der Linden (1992):**

$$|\Phi\rangle = \sum_{j=1}^{M} a_j |\phi_j\rangle = \sum_{j=1}^{M} a_{Pj} |\phi_{Pj}\rangle \quad ; \quad |a_{Pj}| \geq |a_{P(j+1)}|$$

$$|\Phi\rangle \approx |\tilde{\Phi}\rangle = \sum_{j=1}^{M_{imp}} a_{Pj} |\phi_{Pj}\rangle \quad ; \quad M_{imp} = \#\ \text{Important states}$$

$$E \leq \tilde{E} = \frac{\langle \tilde{\Phi} | H | \tilde{\Phi} \rangle}{\langle \tilde{\Phi} | \tilde{\Phi} \rangle}$$

Stochastic Diagonalization

- **Lemma: If** $U^{-1}\begin{pmatrix} a & b \\ b^* & c \end{pmatrix}U \to \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix}$ **and if**

$a \le c$ **and** $|b| \neq 0$ **then** $\boxed{a' < a}$

- **Step 1: Modified Jacobi method**

$$U_{1,k}^{-1}\begin{pmatrix} a & b_1 & \cdots & b_k & \cdots & b_M \\ b_1^* & \bullet & \bullet & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ b_k^* & \bullet & \bullet & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ b_M^* & \bullet & \bullet & \bullet & \bullet & \bullet \end{pmatrix}U_{1,k} \to \begin{pmatrix} a' & b_1' & \cdots & 0 & \cdots & b_M' \\ b_1'^* & \bullet & \bullet & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ b_M'^* & \bullet & \bullet & \bullet & \bullet & \bullet \end{pmatrix} \Rightarrow a' < a$$

  - **Repeat until** $\to \begin{pmatrix} E_0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \bullet & \bullet & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet & \bullet & \bullet \end{pmatrix}$

- **Step 2: Use the difference** $\boxed{a - a' > 0}$ **to set up an importance sampling scheme**

Stochastic Diagonalization

- **Use plane rotations (2x2 matrices) to inflate the matrix and to simultaneously isolate the lowest eigenvalue**

$$\left(E_0^{(0)}\right) \rightarrow \begin{pmatrix} E_0^{(1)} & * \\ * & \bullet \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} E_0^{(2)} & * & * \\ * & \bullet & \bullet \\ * & \bullet & \bullet \end{pmatrix} \rightarrow \cdots \rightarrow \begin{pmatrix} E_0 & 0 & \cdots & 0 \\ 0 & \bullet & \bullet & \bullet \\ \vdots & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix}$$

- **How to pick "candidates" ?**
  - **Use** $\langle \varphi_{Pi} | H | \varphi_{Trial} \rangle \neq 0$
- **Why does the (1,1) element converges to the ground state energy ?**

Stochastic Diagonalization

- **Cauchy theorem:**

$$H = \begin{pmatrix} H^{(k)} & X^T \\ X & Y \end{pmatrix} \quad ; \quad k = 1, \ldots, M-1$$

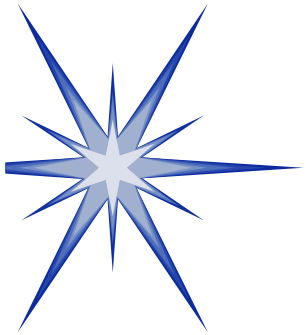
$$E_j^{(k)} : \quad j\text{-th eigenvalue of } H^{(k)}$$

Then :

$$E_j \le E_j^{(k)} \le E_{M-k+j}$$

- **Poincaré separation theorem: Relates eigenvalues of a (*k*-1) x (*k*-1) matrix to the eigenvalues of a *k* x *k* matrix**

$$E_1^{(k)} \le E_1^{(k-1)} \le E_2^{(k)} \le \ldots \le E_{k-1}^{(k-1)} \le E_k^{(k)}$$



SD: Algorithm

- **Initialize data structures**
- **Do**
- **If (Maximum of absolute value of off-diagonal elements of the first row < 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

SD : Practical aspects

- **Necessary condition to be useful in practice: The calculation of matrix elements of H should be (very) fast.**
  - **Each attempt to find a new important state requires the calculation of the matrix element**

$$H^{trial}_{1,n+1} = ((U^{(1)} \ldots U^{(p)})^T H U^{(p)} \ldots U^{(1)})_{1,n+1}$$

$$= \sum_{j=1}^{n} (U^{(1)} \ldots U^{(p)})^T_{1,j} H_{j,n+1}$$

    - **Calculation of the product of plane rotation matrices $U^{(q)}$ is fast**
    - **Calculating $H_{i,n+1}$ takes most of the CPU time**
      - **Can be done in parallel**

Computation of Physical Properties

- **At *T*=0 the expectation value of a physical observable is given by**

$$\langle A \rangle = \langle \Phi | A | \Phi \rangle = \sum_{i,j=1}^{M} a_i a_j \langle \phi_i | A | \phi_j \rangle$$

$$\approx \frac{\sum_{i,j=1}^{M_{imp}} a_{Pi} a_{Pj} \langle \phi_{Pi} | A | \phi_{Pj} \rangle}{\sum_{i=1}^{M_{imp}} a_{Pi}^2}$$

  - **Takes of the order of M_{imp} x M_{imp} operations**
    - **For $M_{imp} \sim 10^5 - 10^6$ computing a physical quantity may take a substantial amount of CPU time**

Example: Repulsive Hubbard Model

- **2D Hubbard model with nearest and next-nearest neighbor hopping**

$$H = -t \sum_{<i,j>\sigma=\uparrow,\downarrow} \sum \left(c^{+}_{i,\sigma} c_{j,\sigma} + c^{+}_{j,\sigma} c_{i,\sigma} \right) + U \sum_{i} c^{+}_{i\uparrow} c^{+}_{i\downarrow} c_{i\downarrow} c_{i\uparrow}$$

$$-t' \sum_{<<i,j>>} \sum_{\sigma=\uparrow,\downarrow} \left(c^{+}_{i,\sigma} c_{j,\sigma} + c^{+}_{j,\sigma} c_{i,\sigma} \right)$$

$$= \sum_{k} \sum_{\sigma=\uparrow,\downarrow} \varepsilon_k c^{+}_{k,\sigma} c_{k,\sigma} + \frac{U}{L} \sum_{k,p,q} c^{+}_{k+q,\uparrow} c^{+}_{p-q,\downarrow} c_{p,\downarrow} c_{k,\uparrow}$$

- **Ground state energy per site of 10 electrons on a 4x4 lattice filled as obtained by exact diagonalization (ED), stochastic diagonalization (SD), and projector quantum Monte Carlo (PQMC).**

$t'/\lvert t\rvert$	$U/\lvert t\rvert$	ED	SD	$PQMC$
0	-6	-2.458782	-2.4568	-2.460 ± 0.004
0	-4	-2.045849	-2.0453	-2.045 ± 0.002
0	-2	-1.731689	-1.7316	-1.731 ± 0.001
-0.22	2	-1.230034	-1.2300	-1.231 ± 0.001
-0.22	4	-1.126160	-1.1261	-1.125 ± 0.003
-0.22	6	-1.058717	-1.0581	-1.061 ± 0.005

Example: Quantum Chemistry

- **Energy and correlation energy (both in hartree) of the Mg atom as obtained by SD and various standard quantum chemistry methods. M_I denotes the number of the basis states (Slater determinants with D_{2h} symmetry) that has been used**

Method	E	E_{corr}	M_I
SCF	−199.585212	0	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**	−199.7263	−0.1411	2538603250

  - **J.Olsen et al., Chem. Phys. Lett. 169, 463 (1990)**

Stochastic Diagonalization versus Fixed-Node Diffusion Monte Carlo

- **Comparison between the ground state energy per site as obtained from a fixed-node approximation (FN), exact diagonalization (ED) and stochastic diagonalization (SD). The difference between the FN results of the second and third row stems from the choice of the trial state.**

Lattice	$U/\|t\|$	FN*	ED	SD	M
2x2 $N_\uparrow = N_\downarrow = 2$	1	-3.3172	-3.3408	$-3.3408^{(a)}$	36
2x2x2 $N_\uparrow = N_\downarrow = 4$	10	-2.6597	-2.8652	$-2.8634^{(b)}$	4900
2x2x2 $N_\uparrow = N_\downarrow = 4$	10	-2.6382	-2.8652	$-2.8634^{(b)}$	4900

- **(A) : D.F.B. ten Haaf *et al.*, Phys. Rev. B (1995)**
  - **(a) : Using < 100 s CPU-time on an IBM Thinkpad 755CD collecting 10 states, without using symmetry**
  - **(b) : Using < 1000 s CPU-time on an IBM Thinkpad 755CD collecting 600 states, without using symmetry**

Stochastic Diagonalization: Main features

- **Exact results (but requires much more effort than standard methods) for small systems because then we can put $M_{imp} = M$ from the start**
- **The "best" variational wave function that can be build from M_{imp} basis states**
- **NO minus sign problems**
- **"Quality" of the result depends on the choice of the representation (basis states)**
  - **Use real- and momentum space representation (very hard for QMC)**
  - **Checks: Compare with exact diagonalization (small systems) and Projector Quantum Monte Carlo**

Application: Off-Diagonal Long-Range Order

- **Can a model system of (strongly) interacting electrons exhibit superconductivity ?**
  - **No degrees of freedom other than those of the electrons**
  - **Dominant interaction between the electrons is repulsive**
- **How would theoreticians know the system is in the superconducting state ?**
  - **Criteria:**
    - **Infinite electrical conductivity**

    AND
    - **Off-Diagonal Long-Range order OR**
    - **Flux Quantization OR**
    - **Superfluid density**

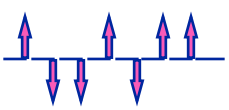
Off-Diagonal Long-Range Order (ODLRO)

- **Penrose & Onsager (1951) :**
  - **Properties of the reduced one-particle density matrix can be used to "detect" Bose-Einstein condensation**
- **Yang (1962) :**
  - **Reduced two-particle density matrix can be used to "detect" superconductivity**
    - **"Detect" : ODLRO in the two-particle density matrix ⇒ superconductivity (zero resistance & Meissner effect)**
      - **Sewell (1990); Nieh, Su, Zhao (1995): ODLRO (+ ...) ⇒ Meissner effect**

ODLRO

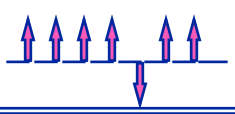
- **Ising Spin model**
- **Phase transition**
  - $T > T_c$ **: There is little correlation between spins on different sites**
  - $T < T_c$ **: Strong correlation between spins on different sites**
- **Quantitative measure of correlation:**

$$C(j) = \frac{1}{L}\sum_{i=1}^{L}\langle S_i S_{i+j}\rangle$$



$$T > T_c \Rightarrow C(j) \approx e^{-j/\lambda} \Rightarrow \sum_{j=1}^{L} C(j) = a$$

$$T < T_c \Rightarrow C(j) \approx 1 \Rightarrow \sum_{j=1}^{L} C(j) = a + bL$$



- **Fermions**
- **Normal ⇔ super-conductor phase transition: Pairs of electrons**
  - **Yang: No long-range correlation in one-particle density matrices**
  - **Correlation functions or reduced two-particle density matrix**

$$C(j) = \frac{1}{L}\sum_{i=1}^{L}\left\langle c^{+}_{i,\uparrow} c^{+}_{i+k,\downarrow} c_{i+k+j,\downarrow} c_{i+j,\uparrow}\right\rangle$$

  - ***k*** **determines the "kind" of electron pairing**
    - ***k*****=0 ⇨ on-site *s*-wave pairing**
- **ODLRO :** $\sum_{j=1}^{L} C(j) = a + bL$

ODLRO: Full two-particle density matrix

- **Yang: Compute all entries of the reduced two-particle density matrix**

$$A(n,m) = \left\langle c^{+}_{i,\sigma} c^{+}_{i',-\sigma} c_{j',-\sigma'} c_{j,\sigma'} \right\rangle \quad ; n = (i, i', \sigma)\ ,\ m = (j,\ j', \sigma')$$

  - **Compute *A(n,m)* : $2L^2$ x $2L^2$ matrix**
  - **Find the largest eigenvalue λ_0 of *A***
  - **If $\lambda_0 = a + bL$ then there is ODLRO**
  - **The eigenvector corresponding to λ_0 contains the information on the kind of pairing**
- **Analogy with magnetic moments:**
  - **Diagonalize $A(n,m) = \left\langle S_n S_m \right\rangle$ by Fourier transformation**
    - **q=0 ⇨ ferromagnetic**
    - **q=π ⇨ antiferromagnet**

BCS reduced Hamiltonian

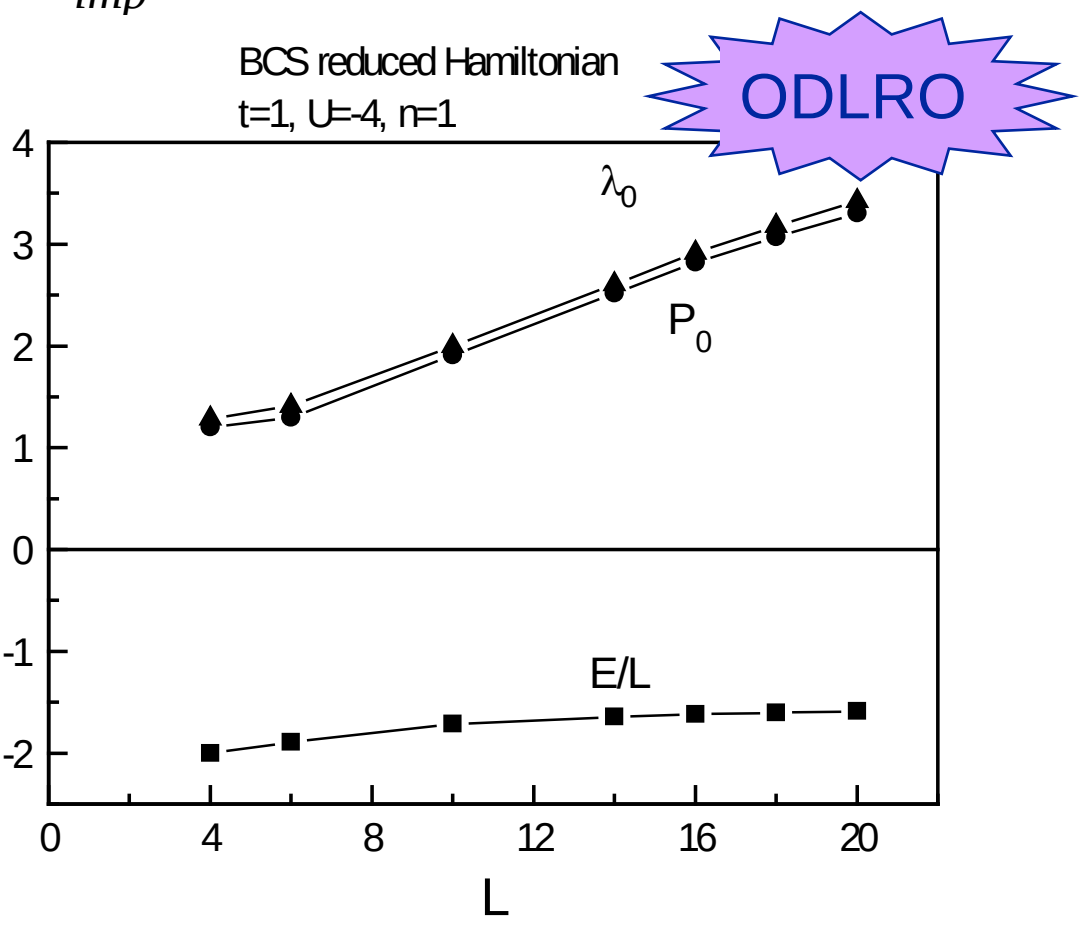
$$H_{BCS} = -t \sum_{<i,j>} \sum_{\sigma=\uparrow,\downarrow} \left(c^{+}_{i,\sigma} c_{j,\sigma} + c^{+}_{j,\sigma} c_{i,\sigma} \right) - \frac{|U|}{L} \sum_{i,j} c^{+}_{i\uparrow} c^{+}_{i\downarrow} c_{j\downarrow} c_{j\uparrow}$$

$$= \sum_{k} \sum_{\sigma=\uparrow,\downarrow} \varepsilon_k c^{+}_{k,\sigma} c_{k,\sigma} - \frac{|U|}{L} \sum_{k,p} c^{+}_{k,\uparrow} c^{+}_{-k,\downarrow} c_{-p,\downarrow} c_{p,\uparrow}$$

- **For the BCS reduced Hamiltonian the BCS treatment yields the exact solution (in the thermodynamic limit)**
  - **R.J. Bursill and C.J. Thompson (1993)**
- **This system has to exhibit ODLRO**
  - **the BCS wave function has ODLRO build in**

BCS reduced Hamiltonian

SD: $M_{imp} \leq 100000$



On-site pairing correlation function

$$P_0 = \frac{1}{L}\sum_{i,k} < c^+_{i\uparrow} c^+_{i\downarrow} c_{k\downarrow} c_{k\uparrow} >$$

$$P_0 \leq \lambda_0$$

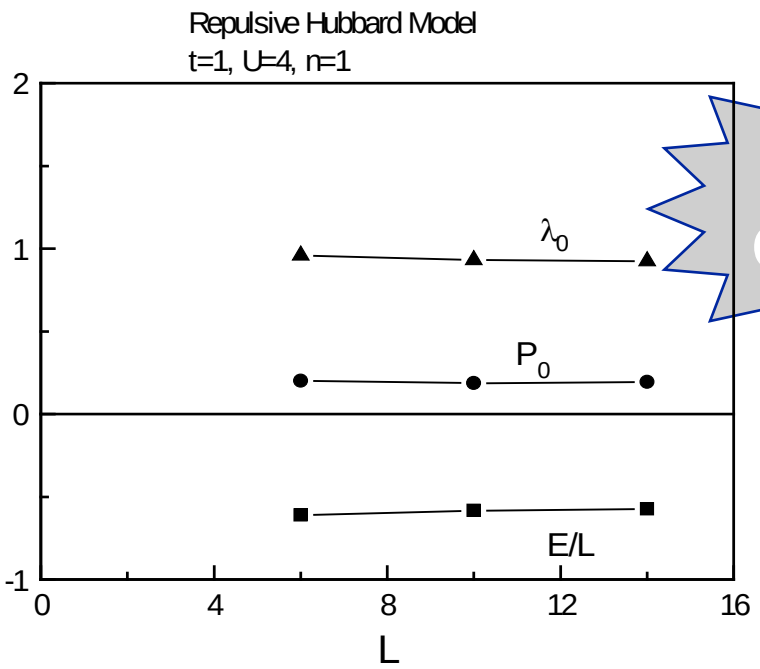
Repulsive Hubbard Model

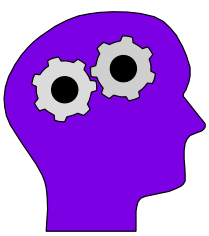
$$H = -t \sum_{<i,j>} \sum_{\sigma=\uparrow,\downarrow} \left(c^{+}_{i,\sigma} c_{j,\sigma} + c^{+}_{j,\sigma} c_{i,\sigma} \right) + |U| \sum_{i} c^{+}_{i\uparrow} c^{+}_{i\downarrow} c_{i\downarrow} c_{i\uparrow}$$

$$= \sum_{k} \sum_{\sigma=\uparrow,\downarrow} \varepsilon_k c^{+}_{k,\sigma} c_{k,\sigma} + \frac{|U|}{L} \sum_{k,p,q} c^{+}_{k+q,\uparrow} c^{+}_{p-q,\downarrow} c_{p,\downarrow} c_{k,\uparrow}$$

- **BCS treatment: No ODLRO**
- SD: $M_{imp} \leq 198000$

$$M \leq 1.2 \times 10^{6}$$





Repulsive Hubbard Model with Correlated Hopping

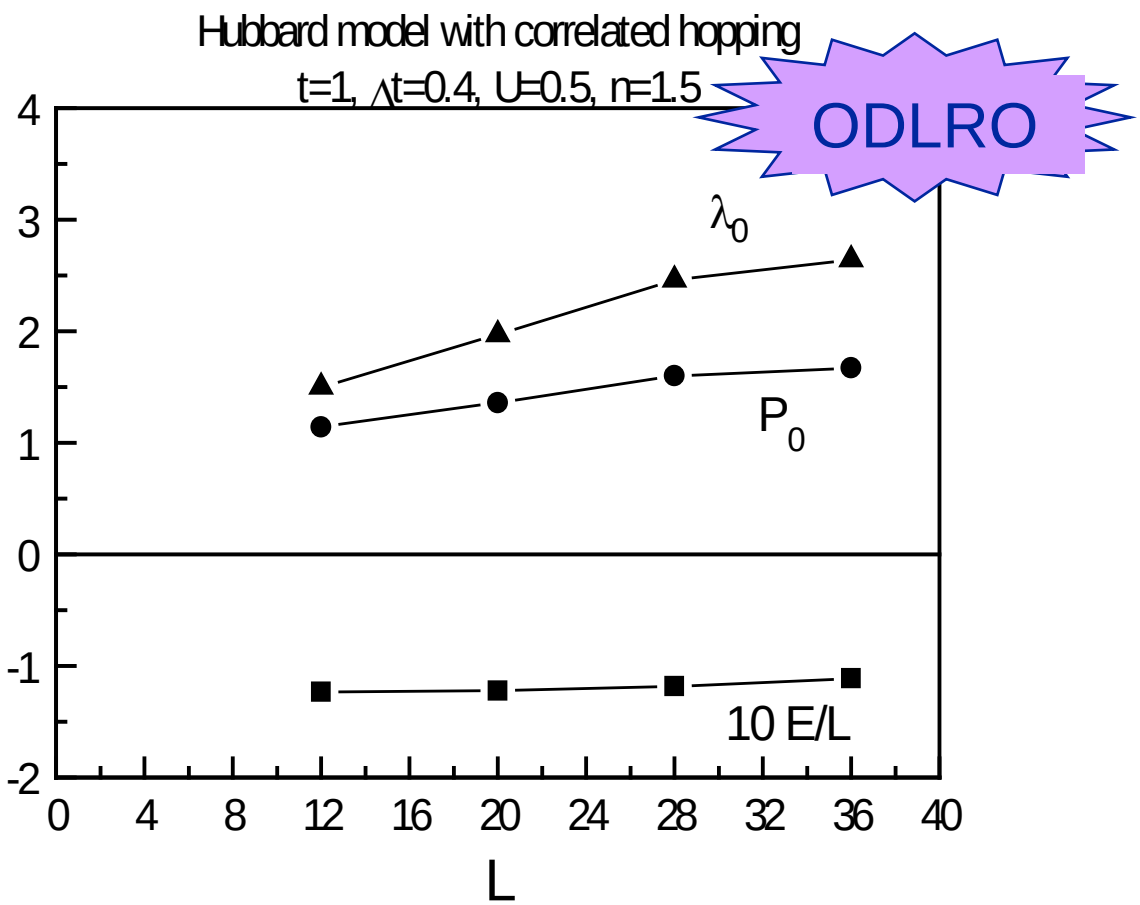
$$H = -t\sum_{<i,j>\sigma=\uparrow,\downarrow}\sum\left(c^{+}_{i,\sigma}c_{j,\sigma} + c^{+}_{j,\sigma}c_{i,\sigma}\right) + U\sum_{i} n_{i\uparrow}n_{i\downarrow}$$
$$+\Delta t\sum_{<i,j>\sigma=\uparrow,\downarrow}\sum\left(n_{i,-\sigma} + n_{j,-\sigma}\right)\left(c^{+}_{i,\sigma}c_{j,\sigma} + c^{+}_{j,\sigma}c_{i,\sigma}\right)$$
$$= \sum_{k}\sum_{\sigma=\uparrow,\downarrow}\varepsilon_{k}c^{+}_{k,\sigma}c_{k,\sigma}$$
$$+\frac{1}{L}\sum_{k,p,q}\left[U - \frac{\Delta t}{t}(\varepsilon_{k} + \varepsilon_{k+q})\right]c^{+}_{k+q,\uparrow}c^{+}_{p-q,\downarrow}c_{p,\downarrow}c_{k,\uparrow}$$

- **Interaction may be "attractive" for some (*U,t,Δt,n*)**
- **Hirsch (1989): Basic model for (high-T_c) superconductivity (based on BCS treatment)**
- **Exact solution for (*U,t=Δt,n=1*): ODLRO for $U<-4|t|$ but NO superconductivity**
- **Japaridze & Müller-Hartmann (1994): Continuum limit + Bosonization: ODLRO for $U \ll \Delta t \ll t$, $n_0 < n < n_1$**

Repulsive Hubbard Model with Correlated Hopping

- SD: $M_{imp} \leq 192000$

$$M \leq 8.7 \times 10^{19}$$



Application: Summary

- **Off-Diagonal Long-Range Order (implies superconductivity ?) in models of interacting fermions.**
  - **The repulsive Hubbard model supplemented with correlated hopping may exhibit Off-Diagonal Long-Range Order, depending on the choice of ($U,t,\Delta t,n$)**
  - **Building the many-body states from single-particle states in either the real-space or k-space representation is far from optimal for this kind of problem**

Basis set optimization

- **SD can use “any” representation**
  - **if $\langle\varphi|H|\psi\rangle$ can be calculated efficiently**
  - **why not use basis sets that are “more sophisticated” ?**
- **An attempt motivated by our study of ODLRO**
  - **real-space representation is adequate to describe e.g. local pairs**
  - **k-space is the representation of choice to describe e.g. extended structure**
- **Can we let the algorithm find an optimum that interpolates between these two extremes ?**
  - **basis set optimization (cfr. Quantum chemistry)**

Optimization of the single-particle states

- **Example: Fermions on a lattice (atomic orbitals)**
- **Many-body basis states are build from single-particle states**
  - **Real-space :** $|\varphi_i\rangle = c_i^+|0\rangle$
  - **Fourier-space :** $|\varphi_k\rangle = c_k^+|0\rangle$
  - **Many-body state (Slater determinant):**

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

  - **Matrix of coefficients:**

$$A^{(n,K)} = \begin{pmatrix} a^{(n)}(1,1) & a^{(n)}(1,2) & \cdots & a^{(n)}(1,N) & \cdots & a^{(n)}(1,K) \\ a^{(n)}(2,1) & a^{(n)}(2,2) & \cdots & a^{(n)}(2,N) & \cdots & a^{(n)}(2,K) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a^{(n)}(K,1) & a^{(n)}(K,2) & \cdots & a^{(n)}(K,N) & \cdots & a^{(n)}(K,K) \end{pmatrix}$$

    - **e.g. “real space” :**

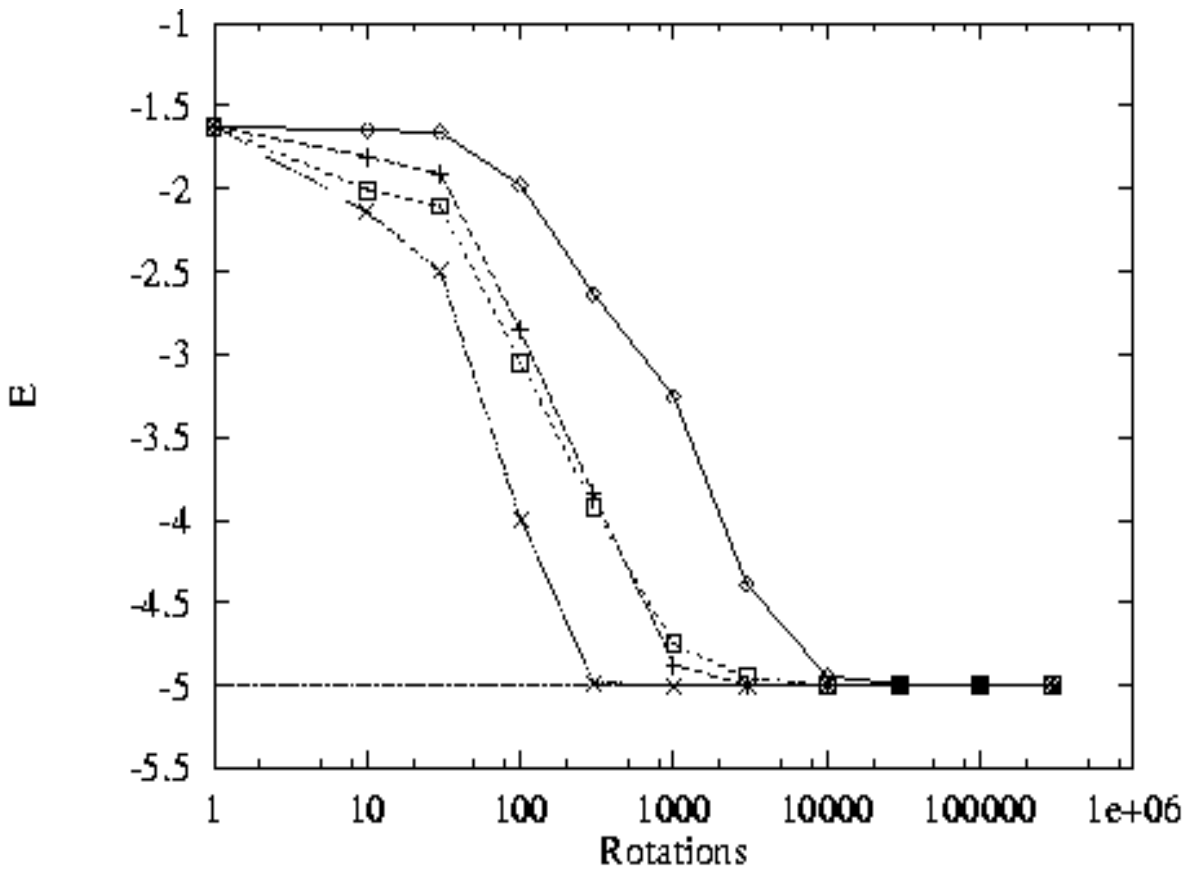
$$A^{(n,K)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Optimization of the single particle states

- **Change the matrix $A^{(n,K)}$ using "simple" moves**
  - **keep single-particle states orthonormal**
    - **Simplifies calculation of matrix elements of *H***
    - **Use plane-rotations:** $U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$
      - **Update two rows (columns) *p*,*q* at a time**
  - **Mix columns:** $A^{(n,K)} \leftarrow A^{(n,K)} U^{(p,q)} \quad ; \quad 1 \le p \le N < q \le K$
  - **Mix rows:** $A^{(n,K)} \leftarrow U^{(p,q)} A^{(n,K)} \quad ; \quad p,q = 1,K$
  - **Preserves unitary character of $A^{(n,K)}$ and defines a "dynamics" for changing Slater determinants**

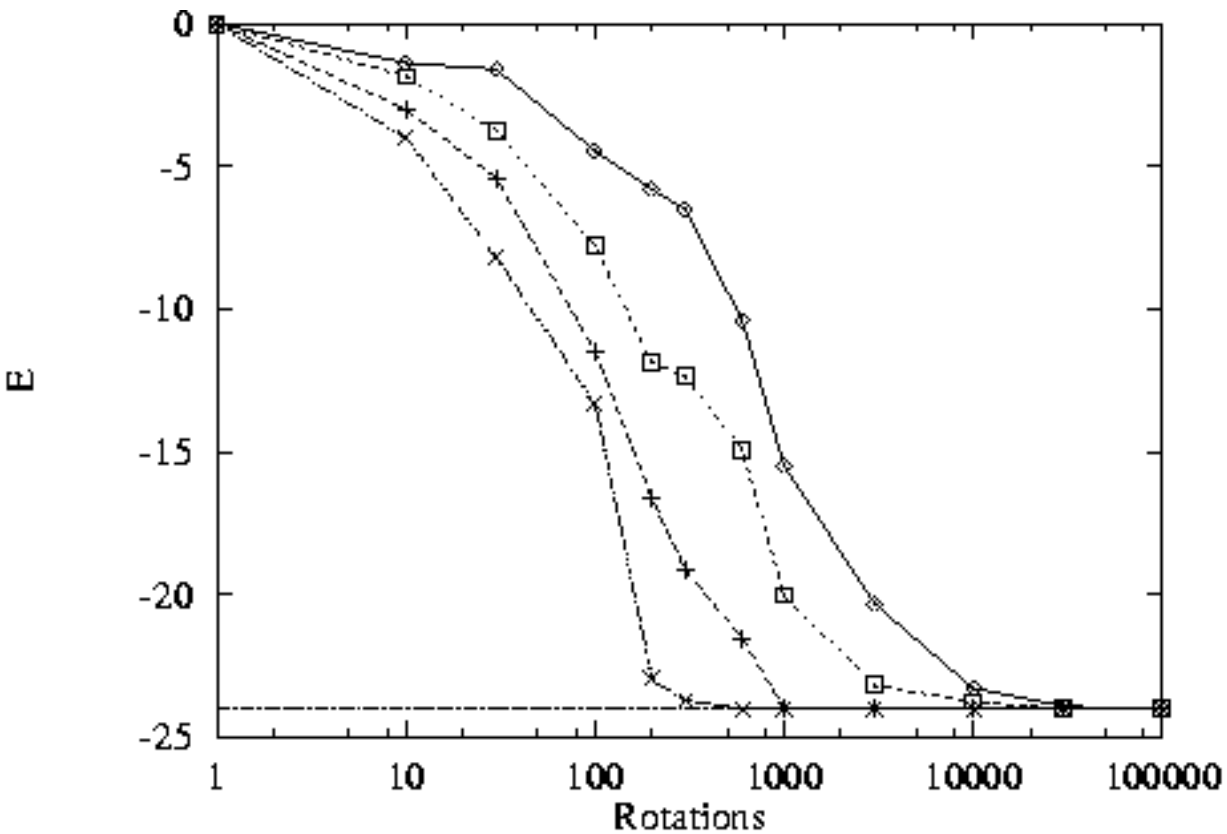
Basis-set optimization: Example

- **The energy *E* as a function of the number of rotations as obtained by optimizing one Slater determinant for the case of a 4x4 Hubbard model ($N_\uparrow = N_\downarrow = 5\,, t = 0\,, U = -1$). The single-particle states used to construct the initial many-body wave function are $|\varphi_k\rangle = c_k^+ |0\rangle$. Various symbols correspond to different optimization strategies. The exact result is indicated by the dashed-dotted line.**



Basis-set optimization: Example

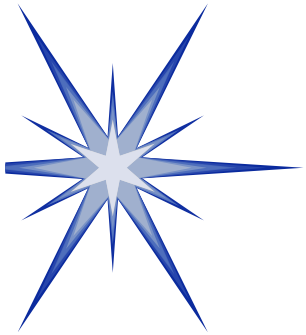
- **The energy *E* as a function of the number of rotations as obtained by optimizing one Slater determinant for the case of a 4x4 Hubbard model ($N_\uparrow = N_\downarrow = 5, t = 1, U = 0$). The single-particle states used to construct the initial many-body wave function are $|\varphi_i\rangle = c_i^+ |0\rangle$. Various symbols correspond to different optimization strategies. The exact result is indicated by the dashed-dotted line.**



Basis-set optimization: Example

- **The ground state energy of the 4x4 Hubbard model ($N_\uparrow = N_\downarrow = 5\,, t = 1$) as obtained by various methods:**
  - E_1^{Exact} **: Exact diagonalization**
  - $E_1^{(k)}$ **: SD + basis-set optimization using** $|\varphi_k\rangle = c_k^+|0\rangle$
  - $E_1^{(i)}$ **: SD + basis-set optimization using** $|\varphi_i\rangle = c_i^+|0\rangle$
  - E_1^{SD} **: SD**
  - E_1^{HF} **: Hartree Fock**
  - **In all SD calculations the limit to the number of important states was set to 200: $M_I = 200$.**
  - **A greedy algorithm was used to optimize the single-particle states, using 10000 plane rotations for each 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	−14.625
4	−19.58094	−19.50157	−19.42179	−19.24434	−17.750
2	−21.37695	−21.36576	−21.34991	−21.32807	−20.875
0	−27.12500	−27.12500	−27.68078	−27.64086	−27.125
−2	−27.70702	−27.69305	−27.68078	−27.64086	−27.125
−4	−32.73360	−32.51795	−32.45901	−32.05033	−30.250
−6	−39.34051	−38.19426	−38.27180	−36.85642	−33.375



Summary

- **General aspects of the sign problem in Quantum Monte Carlo methods**
- **Stochastic diagonalization method**
- **Application: Off-Diagonal Long-Range Order (⇦ superconductivity) in models of interacting fermions.**
  - **The repulsive Hubbard model supplemented with correlated hopping may exhibit Off-Diagonal Long-Range Order, depending on the choice of (*U*,*t*,*Δt*,*n*)**
- **Basis-set optimization**