Reprinted from

JOURNAL OF THE PHYSICAL SOCIETY OF JAPAN

FULL PAPER

The Fermionic Minus-Sign Problem for the Ground State, Revisited with Higher Order Suzuki–Trotter Decompositions

Xiaoxing WANG and Hans-Georg MATUTTIS

J. Phys. Soc. Jpn. **81** (2012) 104002

Journal of the Physical Society of Japan **81** (2012) 104002 **FULL PAPERS**
http://dx.doi.org/10.1143/JPSJ.81.104002

The Fermionic Minus-Sign Problem for the Ground State, Revisited with Higher Order Suzuki–Trotter Decompositions

Xiaoxing WANG* and Hans-Georg MATUTTIS

Department of Mechanical Engineering and Intelligent Systems, The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan

(Received April 25, 2012; accepted August 6, 2012; published online September 19, 2012)

A theorem by Yoshida which states that the discretization error for the Suzuki–Trotter (ST) decomposition is one order smaller than for the corresponding approximated Hamiltonian has long been overlooked with grave consequences for the accuracy of algorithms which use decomposition schemes. In our error analysis for the ground state energy of the Hubbard Hamiltonian by the projector quantum Monte Carlo method, we have used various orders of ST-decompositions, including for the first time in this field, pseudo-symplectic methods. We show that higher order does not necessarily imply better convergence, and identify the condition to obtain good convergence with higher order ST-decompositions. For the first order ST-decomposition, the ground state energy does not converge to the values for numerical diagonalization, in agreement with Yoshida's theorem, which may have caused confusion in connection with the "fermionic sign problem". We show how the use of the sign of the fermion determinant is in fact a reweighing method which for the ground state energy violates some basic Monte Carlo properties (the Gaussian distribution of observables), in contrast to sampling without sign. Whether sampling with or without the sign deviates less from the exact ground state energy depends on the system parameters (interaction, filling) and on the test wave function, so sampling with sign gives not necessarily better or more physical than sampling without sign, a result which can be supported by arguments on the error compensation. We discuss the implications for related methods.

KEYWORDS: Hubbard model, projector Quantum Monte Carlo, Hubbard–Stratonovich transformation, higher order Suzuki–Trotter decomposition, minus-sign problem

1. Introduction

The inception of the Quantum Monte Carlo (QMC) method via the Suzuki–Trotter (ST) decomposition[1,2] has opened new possibilities to describe many body systems with controlled numerical approximations. Nevertheless, the occurrence of negative transition properties in fermionic systems gave rise to what became labeled as "minus-sign problem", which has hampered the development and use of determinant quantum Monte Carlo (DQMC) algorithms, both for the ground state[3] and finite temperature[4] for the better part of the last twenty years. Only when negative transition probabilities were also found in spin-glasses,[5] it became clear that the "minus-sign" was not necessarily something related to fermions. Therefore, some researchers have argued that the occurrence of the minus-sign problem originates rather from the combination of the ST-decomposition and the choice of the representation for the states of the system and the Monte Carlo sampling.[6,7] Due to the lack of alternative approaches to "1st or 2nd order ST-decomposition, with Hubbard–Stratonovich transformation", it has been hardly possible to discriminate universal behavior and artifacts of the traditional combination of approaches. To overcome this gridlock, we apply both "symmetric" and "asymmetric" higher order ST-decompositions to answer the question whether the problem is physical at all or purely an artifact of the methods employed up to now. Higher order ST-decompositions have largely been shunned in the field of QMC (with one exception[5]), as they have the odium of being cumbersome and numerically expensive to implement.[8] Such "symplectic" decompositions have attracted recent interest in the numerical analysis community for use with ordinary differential equations to preserve the "area/volume" in the phase space.[9,10] The connection between ST-decompositions and symplectic integrators runs deep: The popular Verlet–Störmer[11] scheme can be derived from the ST-decomposition of second order.[12] Higher order methods, with smaller truncation error should cast a light on what really happens in the lower order methods. We simulate the two-dimensional Hubbard Hamiltonian

$$\begin{aligned}\hat{H} &= \hat{K} + \hat{U} \\ &= -t \sum_{\langle ij\rangle,\sigma} (c^{\dagger}_{i,\sigma} c_{j,\sigma} + \text{h.c.}) + U \sum_i n_{i,\uparrow} n_{i,\downarrow},\end{aligned} \tag{1}$$

where the operator $c^{\dagger}_{i,\sigma}$ ($c_{i,\sigma}$) creates (annihilates) a fermion of spin $\sigma = \uparrow$ or $\downarrow$ at site i. $n_{i,\sigma} = c^{\dagger}_{i,\sigma} c_{i,\sigma}$ is the number operator, t is the hopping parameter for nearest neighbors $\langle ij\rangle$, and U represents the strength of the on-site interaction with $t = 1$ as the energy unit. Even with the very recent computing power of the Earth Simulator (there may be larger machines in the top500-list, but whether their performance as a whole has been measured, or only been "extrapolated" based on the performance of a smaller configuration, is in most cases not documented[13]), exact diagonalization is hardly possible beyond 25 sites.[14,15] Therefore, the importance of QMC methods (see refs. 8 and 16 for overviews and references therein) lies in the fact that they are both practicable and their errors can at least in principle be controlled.

1.1 *ST-decompositions*

The operator exponential $e^{-\theta\hat{H}} = e^{-\theta(\hat{K}+\hat{U})}$ with parameter θ is split into M ST-slices $e^{-\tau(\hat{K}+\hat{U})}$ with $\tau = \theta/M$

$$e^{-\tau(\hat{K}+\hat{U})} = e^{-\tau\hat{K}} e^{-\tau\hat{U}} + O(\tau^2), \tag{2}$$

in first order and in second order as

*E-mail: wang@matuttis.mce.uec.ac.jp

$$e^{-\tau(\hat{K}+\hat{U})} = e^{-\tau\hat{K}/2}e^{-\tau\hat{U}}e^{-\tau\hat{K}/2} + O(\tau^3). \tag{3}$$

To obtain p-th order variants, a single slice of the operator exponential $e^{-\tau(\hat{K}+\hat{U})}$ is divided into $M_{\rm sub}$ substeps as

$$e^{-\tau(\hat{K}+\hat{U})} = \prod_{n=1}^{M_{\rm sub}} e^{-\gamma_n\tau\hat{K}}e^{-\eta_n\tau\hat{U}} + O(\tau^{p+1}),$$
$$\sum_{n=1}^{M_{\rm sub}}\gamma_n = \sum_{n=1}^{M_{\rm sub}}\eta_n = 1, \tag{4}$$

i.e., for $M_{\rm sub} = 4$, we have

$$e^{-\gamma_1\tau\hat{K}}e^{-\eta_1\tau\hat{U}}e^{-\gamma_2\tau\hat{K}}e^{-\eta_2\tau\hat{U}}e^{-\gamma_2\tau\hat{K}}e^{-\eta_1\tau\hat{U}}e^{-\gamma_1\tau\hat{K}}, \tag{5}$$

with coefficients γ_n and η_n for, e.g., Candy & Rozmus$_4$ or Chambers$_{6A}$ in Table A·I, where we have summarized the coefficients for all methods we are going to use. Coefficients in quadruple precision are available elsewhere.[17–19] ST-decompositions can be either symmetric or asymmetric with respect to ST-time. For conventional ST-decompositions for orders higher than 3, it has been proven that there must be some negative substeps.[20] Such decompositions are derived from zeroing higher order commutators in the Baker–Campbell–Hausdorff formula, while for weakly-coupled systems in celestial mechanics, "pseudo-symplectic" decompositions have been developed for Hamiltonians $\hat{H} = \hat{A} + \epsilon\hat{B}$ with small perturbation parameters ϵ. The elimination of the terms of $O(\epsilon\tau^{p-1})$ imposes additional constraints,[21–23] but allows to derive decompositions with purely positive ST-substeps γ_n and η_n. We also try out these pseudo-symplectic decompositions, as they are up to now untapped resources in the field of QMC. In the following, we indicate the methods in Table A·I by the conventional or author's name, the order as a following index, as well as a letter if there are several methods of the same order by the same author, i.e., "Yoshida$_{6A}$".

1.2 *Projector quantum Monte Carlo (PQMC)*

The ground state can be obtained from the exponential of the Hamiltonian $\hat{H}$ acting on a trial function as $e^{-\theta\hat{H}}|T\rangle$. Our choice is to let an exponential of ST-length θ act on bra and on ket, so the total length of our ST-product is 2θ. Expansion in the energy representation yields

$$\begin{aligned} e^{-\theta\hat{H}}|T\rangle &= e^{-\theta\hat{H}}\sum_m \langle m|T\rangle\cdot|m\rangle \\ &= e^{-\theta E_0}\langle 0|T\rangle\cdot|0\rangle \\ &\quad + e^{-\theta E_0}\sum_{m>0} e^{-\theta(E_m-E_0)}\langle m|T\rangle\cdot|m\rangle, \end{aligned} \tag{6}$$

with eigenvalues $E_0 < E_1 < \cdots < E_m$ and eigenvectors $|0\rangle, |1\rangle, \ldots, |m\rangle$. The contributions for states energetically higher than the ground state vanish exponentially with θ and the corresponding contributions are exponentially suppressed. An exponential $\exp(-\theta\hat{H})|T\rangle$ used to obtain a ground state is sometimes also called a "filter".[24] The interaction $\hat{U}$ in $e^{-\alpha\hat{U}}$ is decoupled using the discrete Hubbard–Stratonovich (HS) transformation.[25,26] We restrict ourselves to the repulsive model with $U > 0$, and abbreviate the ST-substeps as $\alpha = \eta_j\tau$ (or $\alpha = \gamma_i\tau$, see §2.3). For $\alpha \geq 0$, we obtain

$$\begin{aligned} e^{-\alpha U n_{i,\uparrow}n_{i,\downarrow}} &= \frac{1}{2}\sum_{s_i=\pm1}\exp\left[\lambda s_i(n_{i,\uparrow}-n_{i,\downarrow}) - \frac{\alpha U}{2}(n_{i,\uparrow}+n_{i,\downarrow})\right] \\ &= \frac{1}{2}\sum_{s_i=\pm1}\exp\left[\left(\lambda s_i - \frac{\alpha U}{2}\right)n_{i,\uparrow}\right] \\ &\quad\times\exp\left[\left(-\lambda s_i - \frac{\alpha U}{2}\right)n_{i,\downarrow}\right], \end{aligned} \tag{7}$$

and for negative substeps $\alpha < 0$,

$$\begin{aligned} e^{-\alpha U n_{i,\uparrow}n_{i,\downarrow}} &= \frac{1}{2}\sum_{s_i=\pm1}\exp\left[\lambda s_i(n_{i,\uparrow}+n_{i,\downarrow}-1)\right. \\ &\quad\left. - \frac{\alpha U}{2}(n_{i,\uparrow}+n_{i,\downarrow}-1)\right] \\ &= \frac{1}{2}\sum_{s_i=\pm1}\exp\left[\left(\lambda s_i - \frac{\alpha U}{2}\right)\left(n_{i,\uparrow}-\frac{1}{2}\right)\right] \\ &\quad\times\exp\left[\left(\lambda s_i - \frac{\alpha U}{2}\right)\left(n_{i,\downarrow}-\frac{1}{2}\right)\right], \end{aligned} \tag{8}$$

with $\lambda = \operatorname{arccosh}[\exp(|\alpha U|/2)]$ and auxiliary HS fields $s_i \pm 1$ for each lattice site i over the whole ST-time. For higher order ST-decompositions, the HS configurations are not only independent for all ST-slices, but also for all substeps. For a given ST-slice l_n, we denote the entire configuration of all HS spins s_i on sites L by Ω. The resulting diagonal exponential matrix $V_{l_n,\sigma}(\Omega)$ with electron spin σ leads to potentials which fluctuate in space and ST-time during the MC process. The ST-product is then

$$B_{l_n,\sigma}(\Omega) = \prod_{n=1}^{M_{\rm sub}} e^{-\gamma_n\tau K}e^{-\eta_n\tau V_{l_n,\sigma}(\Omega)}. \tag{9}$$

In PQMC, the wave functions $|\Psi\rangle$ and $\langle\Phi|$ for M ST-slices are computed from the trial wave function $|T\rangle$ as

$$|\Psi\rangle = \prod_{l=M/2+1}^{M}\prod_{n=1}^{M_{\rm sub}} B_{l_n,\sigma}(\Omega)|T\rangle, \tag{10}$$

$$\langle\Phi| = \langle T|\prod_{l=1}^{M/2}\prod_{n=1}^{M_{\rm sub}} B_{l_n,\sigma}(\Omega). \tag{11}$$

For N up and N down electrons with the system of L sites, $|T\rangle$, $|\Psi\rangle$, and $|\Phi\rangle$ are $L \times N$ matrices. The statistical weight for a particular set of HS configurations is given by the fermion determinant for each spin direction σ

$$W_\sigma(\Omega) = \det\langle\Phi|\Psi\rangle. \tag{12}$$

The Metropolis-transition probability is

$$P_{\rm trans} = P_\uparrow P_\downarrow = \frac{|W_\uparrow(\Omega^{\rm new})W_\downarrow(\Omega^{\rm new})|}{|W_\uparrow(\Omega^{\rm old})W_\downarrow(\Omega^{\rm old})|}, \tag{13}$$

while we have used the heat-bath transition probability

$$P_{\rm trans} = \frac{P_\uparrow P_\downarrow}{1 + P_\uparrow P_\downarrow}. \tag{14}$$

The Green's function G_σ is obtained from the wave functions in eqs. (10) and (11) as

$$G_\sigma = |\Psi\rangle(\langle\Phi|\Psi\rangle)^{-1}\langle\Phi|. \tag{15}$$

1.3 *The discretization error in the approximated Hamiltonian*

The following discussion is independent of the nature of

the quantum particles, and applies to all Hamiltonians, and probably to a large class of path integral methods, too. PQMC-methods must give the right answer for the ground state energy, in agreement with exact diagonalization methods, within the numerical and statistical errors of the method. Unfortunately, a proper error analysis of the DQMC-procedures has never taken place. For a ST-stepsize τ, the approximation to the operator exponential $\exp(\theta H)$ is correct up to $O(\tau^p)$. Correspondingly, the truncation error is of order $O(\tau^{p+1})$. The more interesting question is: How accurate is such an approximation for the Hamiltonian? Regarding the wide-spread use of first order decomposition, the tacit assumption among many researchers may be that the Hamiltonian is approximated with the same order as its exponential. Unfortunately, this is not the case. The most convenient approach which is common in numerical analysis for similar cases is that of "backward error analysis": Instead of computing how the actual result $H + \epsilon$ deviates from the exact result H ("forward error analysis"), one figures out for which problem H' the given approach is exact, and then obtains the error as $\epsilon = H - H'$. Surprisingly, for classical Hamiltonians $H = H_A + H_B$, it turns out[27] that for the application of the first order decomposition of $\exp(\theta H)$, one has

$$H' = H_A + H_B + \frac{\tau}{2}\{H_A, H_B\} + O(\tau^2), \tag{16}$$

with the Poisson brackets $\{\cdot\}$. Let us flavor this result explicitly in words: Results obtained for a first order approximation of the exponential of a Hamiltonian are of zero order in the Hamiltonian! In classical physics, the solution of Newton's equation of motion in zero order means that the dynamics is lost, the result converges towards the equilibrium position.[28] And a zero order approximation is zero order, no matter how small τ is taken. For the second order ST-decompositions,

$$H' = H_A + H_B + \frac{\tau^2}{12}\left\{\{H_A, H_B\}, H_B + \frac{1}{2}H_A\right\} + O(\tau^4), \tag{17}$$

is obtained[27] where the approximation of the Hamiltonian is of first order. At least for second order ST-decompositions, the error in the approximated Hamiltonian can be controlled by the ST-stepsize τ. Nevertheless, if the simulation contains a programming mistake of even only "half" a ST-slice multiplied at the wrong place, this will make the algorithm first order. For quantum Hamiltonians, it should be enough to replace the Poisson brackets by the commutator. For general $p+1$-st order ST-decompositions, Yoshida[29] has shown that the error is of $p+1$-st order, so the Hamiltonian is approximated in order p. Of course, the occurrence of the Poisson brackets in the error term of the Hamiltonian leads to very system-dependent deviations from the exact results. Depending on the sign of the Poisson brackets in eqs. (16) and (17), the exact value may be approached from below or from above for decreasing τ, so there are no bounds for ST-approximations, unlike e.g., the Ritz-variational method ("always above the exact value").

1.4 Negative weight sampling and minus-sign problem

In eq. (13), the sign of the product of two determinants $W_\uparrow W_\downarrow$ may be negative, and it has become a custom[30] to factor in this sign not in the transition probabilities, but in the computation of expectation values as

$$[A] = \frac{\langle AS\rangle}{\langle S\rangle} = \frac{A^+ - A^-}{S^+ - S^-}, \tag{18}$$

where $\langle\cdot\rangle$ denotes the sampling average during the simulation and $[\cdot]$ denotes the thermal expectation value. A^+ is the average of configurations with positive weight, and A^- with negative weight. $S^\pm$ is the likelihood for positive and negative sign, respectively. When $\langle S\rangle$ is close to zero, $[A]$ becomes undefined or fluctuates large, which is called the minus-sign problem. One purpose of this work is to compare the quality of averages $[A]$ with conventional averages sampled without sign $\langle A\rangle$.

2. Higher Order Decompositions

We run the simulations for 4×4 lattice systems with periodic boundary conditions in FORTRAN and reference codes in MATLAB. If not mentioned otherwise, a single run is thermalized with $N_{\text{thermal}} = 1 \times 10^5$ sweeps, after that measurements are taken in every of the following $N_{\text{meas}} = 3 \times 10^5$ sweeps. If not indicated otherwise, we sample 100 independent HS configurations from runs with different MC seeds. Error-bars are computed as standard deviations of the results of these 100 configurations. Where no error-bars are shown they are smaller than symbols. Where appropriate, we included in many cases the Monte Carlo convergence of the data. As a projection parameter, where not mentioned otherwise, we use $\theta = 4$. The stabilization is carried out at every $\Delta\tau = 0.5$ ST-time. Unlike in the finite temperature algorithm, even at half-filling the minus-sign problem cannot be excluded in PQMC unless a proper trial wave function is chosen. We mostly used dimerized trial wave functions, i.e., the lowest eigenvectors of the hopping matrix resp. its exponential where every second matrix element is chosen to be larger than t.[8] For 1st and 2nd order ST-decompositions, this trial wave function does not have any minus-sign problem at half-filling. Test runs with plane waves are also performed for some cases. The matrix inversion uses the LU-decomposition from LAPACK. For the stabilization, we have used QR-decomposition with pivoting after Householder. Other techniques gave equivalent results. The coefficients of all decomposition methods used are given in the Appendix in Table A·I. After some familiarisation, we tended to use the pseudo-symplectic methods to test assumptions, as they were the computationally cheapest approaches (minimal discretization error with acceptance rates nearly as high as the lower order methods) and tried out the other methods afterwards. We use multi-spin updates where a certain percentage of the whole HS configuration is changed, not like single-spin updates in which single HS spins are flipped randomly on a given ST-slice in sequential.[16] We prefer multi-spin updates as we feel qualms about the ergodicity of the algorithm.[31] Per MC step, 10^{-3} of the HS spins are reversed and the acceptance of this reversal is accepted with the heat-bath probability in eq. (14). For $\theta = 4$ and $\tau = 1/8$, for one subslice this means that about two spins are modified per trial, and four spins for two subslices. There may be more efficient possibilities for

Table I. Comparison of the ground state energy from a dimerized trial wave function sampled with sign according to eq. (18) and without sign as well as the average sign. The methods marked with ⋆ have coefficients of absolute values $|\gamma_n|$ and $|\eta_n|$ larger than 1. For methods of order four and higher, only for pseudo-symplectic methods the $|\gamma_n|$ and $|\eta_n|$ sum up to 1. $N_\uparrow = N_\downarrow = 5$ with $U = 4$ and $\tau = 1/8$.

Method	p	$\langle E\rangle$ sign ignored	$[E]$ with sign	Average sign $\langle S\rangle$	$\sum_n \lvert\gamma_n\rvert$ for $\hat{K}$	$\sum_n \lvert\eta_n\rvert$ for $\hat{U}$
Exact		-1.2238				
Suzuki–Trotter$_1$	1	-1.2418 ± 0.0008	-1.2420 ± 0.0008	0.9954 ± 0.0004	1.0	1.0
McLachlan$_2$	2	-1.2229 ± 0.0008	-1.2232 ± 0.0008	0.9954 ± 0.0004	1.0	1.0
Suzuki–Trotter$_2$	2	-1.2237 ± 0.0008	-1.2240 ± 0.0008	0.9955 ± 0.0004	1.0	1.0
McLachlan$_3$	3	-1.222 ± 0.001	-1.225 ± 0.001	0.952 ± 0.001	1.3760	1.3760
Ruth$_3$	3	-1.205 ± 0.001	-1.225 ± 0.002	0.502 ± 0.004	1.0833	2.3333
Blanes & Moan$_4$	4	-1.201 ± 0.001	-1.223 ± 0.002	0.425 ± 0.005	1.1562	2.4007
Calvo & Sanz-Serna$_4$	4	-1.222 ± 0.001	-1.224 ± 0.001	0.9735 ± 0.0009	1.2811	1.2418
Candy & Rozmus$_4$⋆	4	-1.167 ± 0.002	-1.6 ± 2.3	0.002 ± 0.002	1.7024	4.4048
Chambers$_{4A}$	4	-1.2239 ± 0.0009	-1.2242 ± 0.0009	0.9953 ± 0.0004	1.0	1.0
Chambers$_{4B}$	4	-1.2240 ± 0.0007	-1.2243 ± 0.0007	0.9949 ± 0.0004	1.0	1.0
McLachlan$_4$	4	-1.2234 ± 0.0009	-1.2242 ± 0.0009	0.9838 ± 0.0008	1.4496	1.1716
Suzuki fractal$_4$	4	-1.204 ± 0.002	-1.227 ± 0.002	0.462 ± 0.004	1.4869	2.3159
McLachlan$_5$	5	-1.198 ± 0.001	-1.221 ± 0.002	0.440 ± 0.004	1.4606	2.3833
Tselios$_5$⋆	5	-1.225 ± 0.004	-1.4 ± 0.3	-0.000 ± 0.001	1.9717	10.8147
Blanes & Moan$_{6A}$	6	-1.217 ± 0.001	-1.223 ± 0.001	0.860 ± 0.003	2.0118	1.6592
Blanes & Moan$_{6B}$	6	-1.200 ± 0.002	-1.225 ± 0.003	0.372 ± 0.004	1.2061	2.4935
Chambers$_{6A}$	6	-1.2229 ± 0.0008	-1.2231 ± 0.0008	0.9949 ± 0.0004	1.0	1.0
Chambers$_{6B}$	6	-1.2252 ± 0.0008	-1.2257 ± 0.0008	0.9941 ± 0.0005	1.0	1.0
Yoshida$_{6A}$⋆	6	-1.158 ± 0.003	-0.8 ± 0.6	0.002 ± 0.001	2.8842	5.7107
Laskar & Robutel$_{8A}$	8	-1.2236 ± 0.0008	-1.2240 ± 0.0008	0.9945 ± 0.0005	1.0	1.0
Yoshida$_{8D}$⋆	8	-0.878 ± 0.005	-2.0 ± 1.2	-0.000 ± 0.004	10.3491	15.2548

the sampling for higher order methods according to the width of the ST-slices, but that is something we will address in the future.

Production runs were performed on AMD and SUN processors and were consistent with tentative tests in quadruple processors on DEC ALPHA and INTEL processors.

We have focused on the computation of the ground state energy as the most crucial quantity for zero temperature. According to the Hohenberg–Kohn theorem for the (correct) ground state energy there is a unique electron density distribution. Therefore, if the QMC simulation reproduces the energy correctly, it can also be expected that all other observables can be computed in a reliable fashion.

2.1 $N_\uparrow = N_\downarrow = 5$ *with* $U = 4$

For the system with 10 electrons ($N = N_\uparrow = N_\downarrow = 5$) at Hubbard interaction $U = 4$, the average sign is approximately 1 as the ground state of the kinetic energy is non-degenerate,[3,32] corresponding to a closed shell system. We have implemented the exact diagonalization and the ground state energy per lattice site $E_{\text{exact}} = -1.2238$, which is consistent with ref. 33.

In Table I, we show the comparison of the ground state energy sampled with sign [eq. (18)] and without sign as well as the average sign. The last two columns are the absolute sum of the coefficients for the ST-substeps from eq. (4) for the kinetic $\sum_n |\gamma_n|$ and interaction $\sum_n |\eta_n|$ terms. Test runs with plane waves are also performed and the results are similar. We have performed the same simulations on MATLAB and we obtained the similar results. In our simulations, there are some higher order ST-decompositions which fail to give "good" average signs, namely Ruth$_3$, Blanes & Moan$_4$, Candy & Rozmus$_4$, Suzuki fractal$_4$, McLachlan$_5$, Tselios$_5$, Blanes & Moan$_{6B}$, Yoshida$_{6A}$, and Yoshida$_{8D}$. For all methods with reasonable average signs, the energies are the same for sampling with and without sign, consistent with the value for the exact energy, except for Suzuki–Trotter$_1$. For Suzuki–Trotter$_1$, while the results for large ST-stepsize are consistent on different architectures and compilers, when we refine the ST-stepsize, the results converge to different energies on different computer architectures. For the best accuracy we have (10 byte extended precision on AMD with PGI-FORTRAN compiler) the ground state energy does not converge to the exact value (Table II and Fig. 1). For SUN processors with its native FORTRAN compiler and unidentifiable precision (nominally 8 byte double precision, but how many digits are lost internally due to optimization is not clear), the energy values show a drift towards higher than the exact values, while on INTEL processors with MATLAB, the values are in between the SUN and AMD values. As the numerical error is the sum of rounding and discretization errors, for zero order algorithms, the increase of the rounding error is proportional to the timestep M and the reduction of the discretization error is proportional to $1/M$, see Fig. 2. Accordingly, the error is not minimal for the smallest timestep. Due to this interference of errors, deviation of the energy from the exact ground state value on the AMD machines is consistent with the theorem in §1.3, while the better agreement on SUN and INTEL is spurious, as there is no control or information whether proper rounding or only truncation happens. The average sign $\langle S\rangle$ is subtly larger on AMD than on SUN processor. Depending on the ST-stepsize τ, there are

Table II. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign and taking the sign into account with different values of τ for $\theta = 4$ by Suzuki–Trotter$_1$ on AMD (upper part) and SUN processors (lower part).

Processor	ST-stepsize τ	$\langle E \rangle$ sign ignored	$[E]$ with sign	Average sign $\langle S \rangle$
	Exact	-1.2238		
AMD 10 byte	1/8	-1.2420 ± 0.0006	-1.2423 ± 0.0006	0.9960 ± 0.0002
	1/16	-1.2292 ± 0.0008	-1.2296 ± 0.0007	0.9949 ± 0.0003
	1/64	-1.2257 ± 0.0007	-1.2264 ± 0.0007	0.9939 ± 0.0006
	1/128	-1.2248 ± 0.0009	-1.2251 ± 0.0009	0.9943 ± 0.0004
	1/256	-1.2254 ± 0.0009	-1.2258 ± 0.0009	0.9938 ± 0.0005
SUN 8 byte	1/8	-1.2418 ± 0.0008	-1.2420 ± 0.0008	0.9954 ± 0.0004
	1/16	-1.2294 ± 0.0006	-1.2299 ± 0.0006	0.9946 ± 0.0003
	1/64	-1.2234 ± 0.0008	-1.2238 ± 0.0008	0.9952 ± 0.0004
	1/128	-1.2231 ± 0.0009	-1.2236 ± 0.0009	0.9936 ± 0.0005
	1/256	-1.2233 ± 0.0008	-1.2238 ± 0.0008	0.9932 ± 0.0006

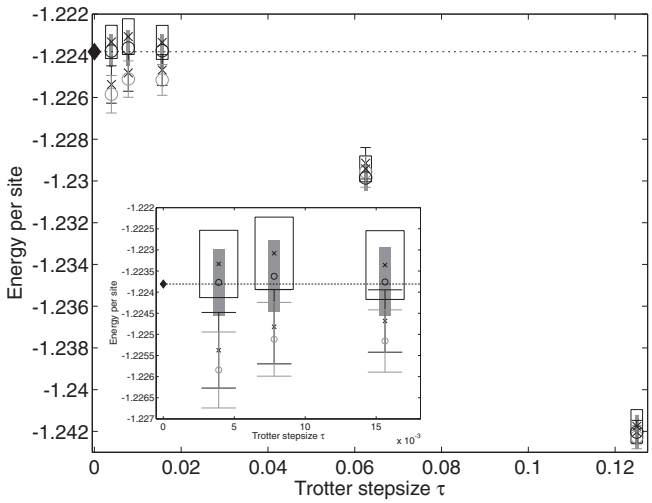

Fig. 1. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs with different values of τ for $\theta = 4$ by Suzuki–Trotter$_1$ in AMD processor (10 byte extended precision, $\times$ solid line: without sign, ○ gray line: with sign) and in SUN processor (8 byte double precision, $\times$ box-frame: without sign, ○ gray straight-bar: with sign). To the value for the exact energy (◆), a dotted line is added to guide the eye.

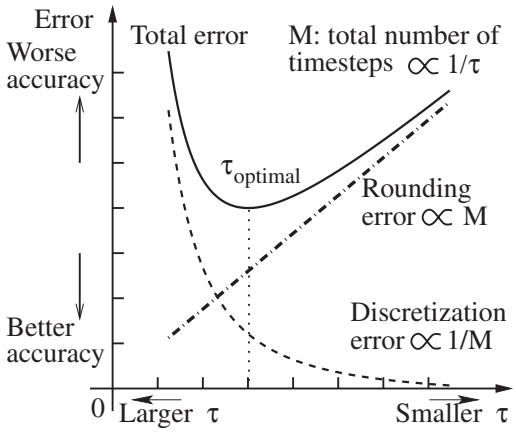

Fig. 2. Schematic representation for the total error as the sum of the discretization error and rounding error for increasing the number of steps, i.e., decreasing stepsize. Depending on the implementation (rounding or truncation), the rounding error may be negative, not positive as in the sketch.

between 6 and 15% more negative configurations on SUN than on AMD. We consider this as an additional indication (other than the number of bytes) that the rounding is more accurate on AMD, as rounding errors should increase the number of negative configurations. Other effects as reason for the deviation (too small projection parameter etc.) can be excluded as the energy is too low and the convergence of the higher order methods for the same parameters is unimpaired. This should serve as a warning that reducing the ST-step in first order methods does not guarantee convergence. Neither first nor second order fitting of the data in Fig. 1 would reliably lead to the minimal deviation from the exact energy for $\tau \to 0$. For stepsizes larger than 1/64, no adverse effects were observed from SUN's double precision, and the results were consistent with the ones obtained from AMD's quadruple precision, so it seems it is really an accumulation of rounding errors.

The higher order Suzuki fractal decompositions[34,35] which were developed to effectively resolve the minus-sign problem[36,37] and help to clarify the mechanism of high-T_c superconductivity[20] have malign effects on the average sign compared to other methods. These methods have been applied primary in the frustrated spin systems.[38]

The pseudo-symplectic decompositions of Chambers$_{4A}$, Chambers$_{4B}$, Chambers$_{6A}$, Chambers$_{6B}$, and Laskar & Robutel$_{8A}$ which are all positive-definite decompositions give benign average signs [$\approx O(1)$], and the ground state energies from ignoring the sign are consistent with the exact value.

With the outcome in this section, we have seen which decompositions give a "good" sign and which give a "bad" sign. Therefore, for the cases where there are control data with "good" sign available, we can at the same time see the behavior of sampling with bad sign. In general, the deviation between the averages computed with and without sign become larger with decaying average sign. For some higher order ST-methods, the ground state energies with sign are far away from the exact value with large error-bars (marked with $\star$ in Table I). The ground state energy from ignoring the sign by Tselios$_5$ is consistent with the exact value while the value computed with sign is separated from the exact value by a large error-bar. In Fig. 3, one can see that the ground state energies for methods with malign signs (close

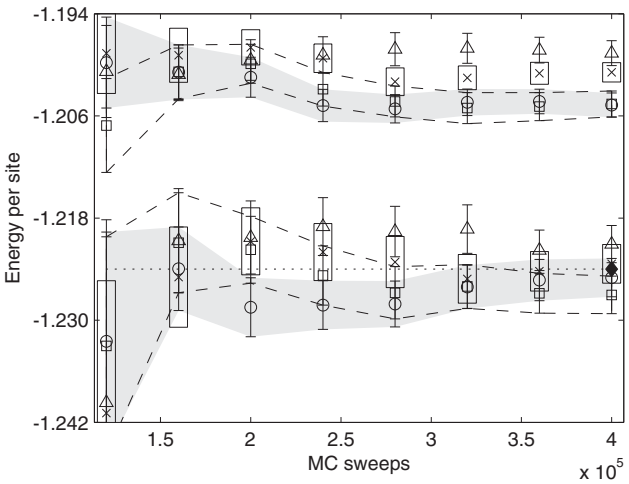

Fig. 3. Ground state energy vs MC sweeps for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign (upper data cloud) and taking the sign into account (lower data cloud) for the same configurations in $\tau = 1/8$ for $\theta = 4$ with Ruth$_3$ (○ with gray shadow), Blanes & Moan$_4$ (× with box-frame), Suzuki fractal$_4$ (□ with dashed envelope), and McLachlan$_5$ (△). To the value for the exact energy (◆), a dotted line is added to guide the eye.

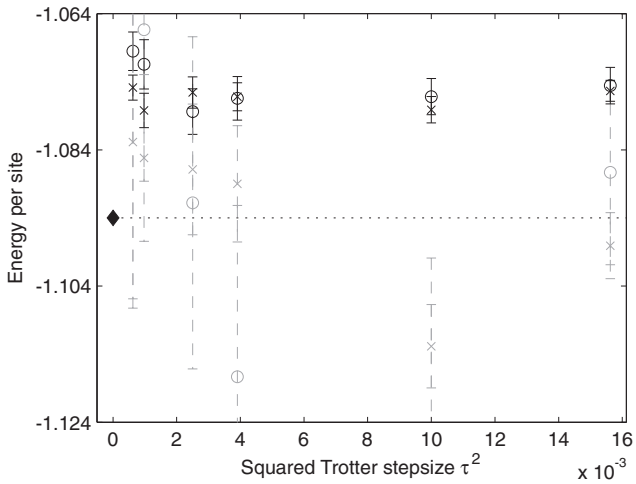

Fig. 4. Ground state energy for the 4×4 systems at $U = 8$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign (black line) and taking the sign into account (gray dashed line) for the same configurations with different values of τ for $\theta = 8$ by Suzuki–Trotter$_2$ (×) and Chambers$_{4A}$ (○). To the value for the exact energy (◆), a dotted line is added to guide the eye.

to 0 so that the convergence is bad) in Table I are closer to the exact value than those from ignoring the sign. It will turn out that this is not universal, but depends on the filling.

When we look at the absolute sum of the coefficients over the kinetic $\sum_n |\gamma_n|$ and interaction $\sum_n |\eta_n|$ terms of the corresponding methods, we note that when $\sum_n |\eta_n| > 2$, the average sign deteriorates for Ruth$_3$, Blanes & Moan$_4$, Candy & Rozmus$_4$, Suzuki fractal$_4$, McLachlan$_5$, Tselios$_5$, Blanes & Moan$_{6B}$, Yoshida$_{6A}$, and Yoshida$_{8D}$. Moreover, we find that such a deterioration of the average sign is proportional to the magnitude of $\sum_n |\eta_n|$. This means that the average sign varies strongly with the symmetry of the position of the interaction term $\hat{U}$ in the ST-substeps, but not with the order. The determinant of a ST-product of commuting symmetric exponential (positive) matrices would be positive-definite, so the sign cannot be attributed to the ST-decomposition alone. Only the decoupling of the interaction with the HS-transformation leads to alternate multiplication of the exponentials for the kinetic term and varying diagonal matrices which leads to a loss of symmetry of the accumulated product and the subsequent decay of the sign. For Suzuki–Trotter$_{1,2}$, the acceptance rate for multi-spin updates is 45–46%. The acceptance rates for higher order pseudo-symplectic methods are nearly as high, for general higher order decompositions it is lower, and lowest are those for decompositions with malign signs.

From Table I we see that the ground state energy with sign is always lower than that from ignoring the sign. An exception is Yoshida$_{6A}$, for which very large outliers and abysmal sign in some runs leads to averaging "good" and "bad" configurations. We are pretty sure that we did not e.g., mistype the coefficients, since the outcome for QMC is consistent with our experiences in molecular dynamics simulations,[17] where Yoshida$_{6A}$ also yielded the largest outliers and consequently the most unreliable configuration averages. That indicates that not all ST-decompositions of high order are actually useful.

2.2 $N_\uparrow = N_\downarrow = 5$ *with* $U = 8$

For $U = 8$, we use a projection parameter with $\theta = 8$ to obtain better convergence. The thermalization is not as efficient as for $U = 4$ with multi-spin updates since the HS configuration space becomes larger with respect to the projection parameter θ for fixed τ. We therefore take an average of 50 HS configuration samples. In Fig. 4, we show the convergence to the ground state energy by Suzuki–Trotter$_2$ and Chambers$_{4A}$. We also tried to halve the flip probability in updating multi-HS spins to increase the acceptance rate for Chambers$_{4A}$. The deviations from the original flip-rate are negligible. For both methods the fluctuations for the computation with sign taken into account ($[A]$-averages) are larger than the $\langle A \rangle$ with ignoring the sign. In Fig. 2 of ref. 39, with a different trial wave function for various projection parameters and different stepsizes, both $[A]$- and $\langle A \rangle$-averages are consistent, and the convergence to the same energy value was obtained.

2.3 *Variants of higher order ST-decompositions*

Alternative simulations are obtained by simply exchanging the kinetic $\hat{K}$ and interaction $\hat{U}$ terms in eq. (4), so that

$$e^{-\tau(\hat{K}+\hat{U})} = \prod_{n=1}^{M_{\rm sub}} e^{-\gamma_n \tau \hat{U}} e^{-\eta_n \tau \hat{K}} + O(\tau^{p+1}). \qquad (19)$$

The ordering of γ_n and η_n may not be interchanged. We have interchanged $\hat{K}$ and $\hat{U}$ for Blanes & Moan$_4$, Blanes & Moan$_{6A}$, and Blanes & Moan$_{6B}$, though in the original paper[40] it was mentioned that for the problems considered the kinetic $\hat{K}$ and the interaction $\hat{U}$ were qualitatively different and not interchangeable. As our problems are different, we try the exchange anyway. In Table III the ground state energy are sampled with and without sign, the last two columns are the absolute sum of the coefficients over interaction $\sum_n |\gamma_n|$ and kinetic $\sum_n |\eta_n|$ terms. We clearly see that, expect for Blanes & Moan$_{6A}$, the average signs increase as the asymmetry of the ST-products seems to

Table III. Comparison of the ground state energy from a dimerized trial wave function sampled with sign according to eq. (18) and without sign as well as the average sign after interchanging the order of the kinetic term $\hat{K}$ and the interaction term $\hat{U}$ in the decomposition. $N_\uparrow = N_\downarrow = 5$ with $U = 4$ and $\tau = 1/8$.

Method	p	$\langle E\rangle$ sign ignored	$[E]$ with sign	Average sign $\langle S\rangle$	$\sum_n \lvert\gamma_n\rvert$ for $\hat{U}$	$\sum_n \lvert\eta_n\rvert$ for $\hat{K}$
Blanes & Moan$_4$	4	-1.201 ± 0.001	-1.223 ± 0.002	0.425 ± 0.005	2.4007	1.1562
", order reversed		-1.223 ± 0.001	-1.224 ± 0.001	0.983 ± 0.001	1.1562	2.4007
Suzuki fractal$_4$	4	-1.204 ± 0.002	-1.227 ± 0.002	0.462 ± 0.004	2.3159	1.4869
", order reversed		-1.220 ± 0.001	-1.225 ± 0.001	0.917 ± 0.002	1.4869	2.3159
Blanes & Moan$_{6A}$	6	-1.217 ± 0.001	-1.223 ± 0.001	0.860 ± 0.003	1.6592	2.0118
", order reversed		-1.211 ± 0.001	-1.224 ± 0.002	0.670 ± 0.004	2.0118	1.6592
Blanes & Moan$_{6B}$	6	-1.200 ± 0.002	-1.225 ± 0.003	0.372 ± 0.004	2.4935	1.2061
", order reversed		-1.223 ± 0.001	-1.224 ± 0.001	0.978 ± 0.001	1.2061	2.4935

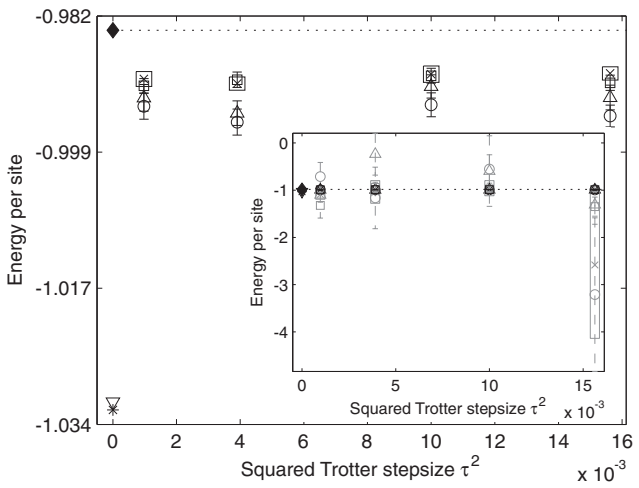

Fig. 5. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 7$ from Monte Carlo runs ignoring the sign (black line) and taking the sign into account (gray dashed line) for the same configurations with different values of τ for $\theta = 4$ by Ruth$_3$ (○), Chambers$_{4A}$ (□), McLachlan$_5$ (△), Laskar & Robutel$_{8A}$ (×). Added are the data for sampling without sign from ref. 41 extrapolated to $\tau \to 0$ with the initial trial wave function from unrestricted Hartree–Fock solution (∗) and paramagnetic state (▽). To the value for the exact energy (◆), a dotted line is added to guide the eye.

be affected less by the absolute sum of ST-substeps over the kinetic term, but over the interaction term. Similarly, the decay of the average signs are proportional to the absolute sum over interaction term, as for Blanes & Moan$_{6A}$, but not to the absolute sum over kinetic term unless the absolute sum over kinetic term starts to become too large. A symmetric product of symmetric positive-definite ST-slices would of course always have a determinant with positive sign. The decrease of the average sign can therefore be attributed to the asymmetry of the ST-product which accumulated during the multiplication of exponential operators.

2.4 $N_\uparrow = N_\downarrow = 7$ *with* $U = 4$

The average sign for the system with 14 electrons ($N = N_\uparrow = N_\downarrow = 7$) at Hubbard interaction $U = 4$ becomes prohibitively small where the ground state of the kinetic energy is degenerate,[3] corresponding to an open shell system with the exact ground state energy per lattice site $E_{\text{exact}} = -0.9840$.[33] In Fig. 5, we show the convergence of our ground state energy computed with and without the sign for symmetric and asymmetric decompositions. Symmetric pseudo-symplectic higher order methods (Chambers$_{4A}$ and Laskar & Robutel$_{8A}$) give much better accuracy than asymmetric methods (Ruth$_3$ and McLachlan$_5$) do. For Laskar$_{8A}$ at $\tau = 1/32$, the ground state energy with sign gives a miserable error-bar of 0.104, no wonder with an average sign of $\langle S\rangle = 0.014 \pm 0.003$. Elsewhere,[41] a ground state energy of $E_{\text{ig1}} = -1.032$ with unrestricted Hartree–Fock solution and $E_{\text{ig2}} = -1.031$ with paramagnetic state as an initial trial wave function were reported for sampling without sign (∗ and ▽ in Fig. 5) extrapolation to $\theta \to 0$, but our deviations from the exact value for sampling without sign are considerably smaller.

3. The "Sign Problem"

3.1 *Sampling "with sign"*

The argument that the "negative configurations" result from "fermion exchange" (virtual or otherwise) can be borrowed from the world-line simulation. It is both tempting and misleading: The permutation matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

has indeed a negative determinant. But in Quantum Monte Carlo, we never deal with matrices themselves, only with their exponential, and $\det[\exp(\tau P)]$ is of course positive, as its eigenvalues are the exponentials of the eigenvalues of τP. The sign of the fermion determinant in DQMC [see eq. (12) for the ground state and ref. 42 for finite temperature] is the product of the sign of all singular values of the $\langle\Phi|\Psi\rangle$, some of them are exponentially small, much smaller than the error-bounds on the original ST-product and especially of the approximated Hamiltonian according to eqs. (16) and (17). There is no rigorous mathematical argument conceivable to assign quantities or their sign a crucial role in the sampling process which to all intents and purposes is smaller than the discretization error in the simulation, but the sign of the determinant is exactly that.

A similar argument can be made for eq. (18): That $\langle AS\rangle$ cannot be factored into $\langle A\rangle\langle S\rangle$, so that $\langle S\rangle$ simply drops out implies a statistical correlation between A and S, as has been argued by one of the authors earlier.[39] If we would take such a correlation at face value, again it would hinge on the signs of the smallest singular values, smaller than the error in the decomposition. So that sampling with sign will give better observables than sampling without sign, we need a rather more "sizeable" effect.

As statistical weight in eq. (13) in DQMC the absolute value of the fermion determinant is chosen. Some articles refs. 43 and 44 give the formula without absolute values even when they state later that the absolute value must be taken. The actual sampling has to use the absolute value, everything else is impracticable, and, on top of that, negative probability measures would be at odds with Kolmogorov's axioms. For importance sampling, the transition probabilities are to be chosen so that the algorithm becomes ergodic, and then the averages are computed as simple averages over the Monte Carlo sweeps. As we have shown, the sign of the fermion determinant is not a physical quantity, as it depends at least on the specific ST-decomposition and, as we show later in §4 in case of PQMC on the test wave function. Methodologically, in the "minus-sign sampling" in eq. (18) the sign is reintroduced *ad hoc* after the sampling and the measurement, so it is actually a reweighing procedure, which must be justified by its statistical merits.

3.2 Distributions

Monte Carlo sampling is based on the law of large numbers, and accordingly, the sampled data, especially the energy data, should be Gaussian distributed.[45] In Fig. 3, the ground state energies computed with sign lie closer to the exact value than the values computed without sign. However, in Monte Carlo sampling, we should not only have an eye on the averages, but also on the distributions. For the sampling with sign, we have to subtract the bins of the samples for negative sign from those with positive sign. The distribution of the sampled values for the energy with the decomposition after Ruth$_3$ is Gaussian distributed when the sign is ignored (Fig. 6, above). The result with sign (Fig. 6, below) is a far cry from the Gaussians which are assumed for the sampling in Monte Carlo theory. Nevertheless, Gaussian distributions are a prerequisite for the importance sampling: We cannot just give it up on a whim to justify an *ad hoc* conceived post-processing method, so the shape of the distribution does not favor sampling with sign.

3.3 Electron densities

After the Hamiltonian has turned out to be rather less well preserved than its exponential in eqs. (16) and (17), it is time to ask how negative transition probabilities are obtained anyway. The Metropolis transition probability for single-spin flips is

$$P_\sigma = 1 + \delta_\sigma \cdot G_\sigma(l)_{i,i}, \tag{20}$$

$$\delta_\uparrow = \exp[\lambda(s_i^{\rm new}(l) - s_i^{\rm old}(l))] - 1 \tag{21}$$

$$\delta_\downarrow = \exp[-\lambda(s_i^{\rm new}(l) - s_i^{\rm old}(l))] - 1, \tag{22}$$

where $s_i^{\rm old}(l)$ is the old HS spin on the i-th site at the l-th ST-slice and $s_i^{\rm new}(l)$ is after the HS spin flip-trial. $n_{i,\sigma} = G_\sigma(l)_{i,i}$ is the i-th diagonal element of the Green's function at the l-th ST-slice, with $\lambda = \arccosh[\exp(|\alpha U|/2)]$ from eq. (7). For higher order ST-decompositions, the Green's functions for the substeps n are obtained by replacing l with l_n, and for negative substeps $\delta_\uparrow = \delta_\downarrow$ according to eq. (8).

For relatively small ST-substeps α with $U > 0$, and therefore reasonable αU, negative transition probabilities in eq. (20) can only be reached for electron densities $n_{i,\sigma}$ smaller than 0 ($\delta_\sigma > 0$) or larger than 1 ($\delta_\sigma < 0$), see Fig. 7. Of course, the meaning of the $n_{i,\sigma} < 0$ is not that of holes in a half-filled band, it is a negative density created in an empty state. Symmetric to the densities smaller than 0 in Fig. 8 which violate common sense, we see densities with $n_{i,\sigma} > 1$ which violate fermion statistics. For fermion systems, quantitative errors in the density are actually qualitative violations of the fermion principles, which affects all variable computations. When we look at the histogram for electron densities $n_{i,\sigma}$ at strong repulsion at half-filling in Fig. 8, we see that the higher the order of the ST-decomposition is, the more unphysical densities smaller than 0 and larger than 1 are suppressed. We conclude that the higher order methods in Table I give more accurate results than Suzuki–Trotter$_1$ for the energy, as there are less unphysical densities. Current symplectic ST-decomposition only conserve global properties of the total Hamiltonian and its exponential, they do not constrain individual electron densities to their physical values, though higher order methods reduce the number of negative densities.

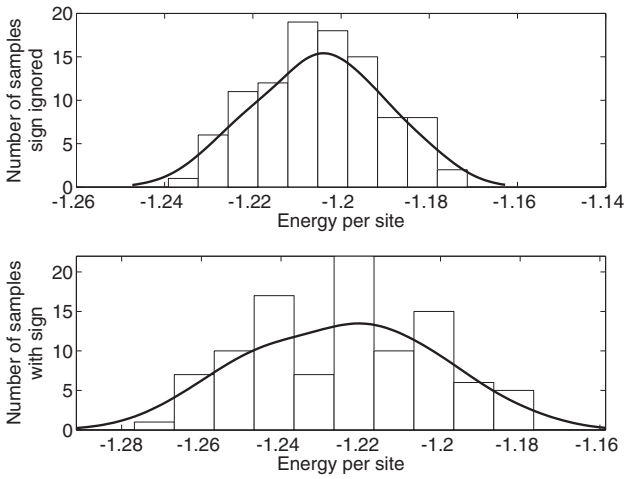

Fig. 6. Histogram for 100 HS configuration averages of the ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ from ignoring the sign (upper) and taking the sign into account (lower) for the same configurations in $\tau = 1/8$ with $\theta = 4$ by Ruth$_3$.

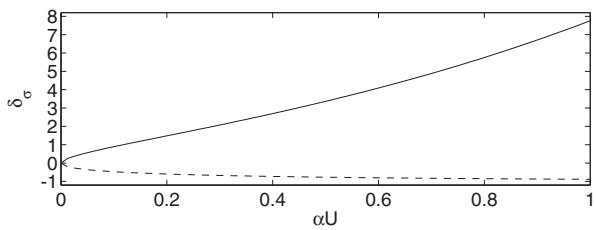

Fig. 7. Graph for $\delta_\uparrow, \delta_\downarrow$ in eqs. (21) and (22).

3.4 An alternative sampling scheme

The hope that the negative densities in PQMC stem from ST-slices "close" to the test wave function, which are not used for Monte Carlo sampling, but Fig. 8 tells otherwise: For anti-ferromagnetic ordering, obviously the observable computation itself is affected. This error will not "go away" by averaging: Unphysical spikes in the density will influence the results for Fourier transform and density correlation functions, and, moreover, unphysical diagonal elements of

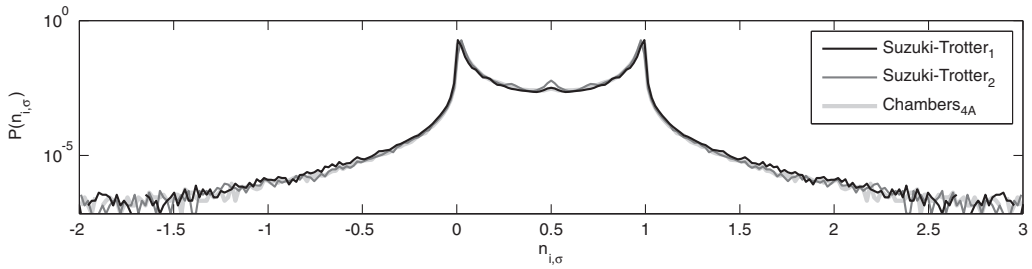

Fig. 8. Histogram for single densities $n_{i,\sigma}$ for Suzuki–Trotter$_1$ (thin black line), Suzuki–Trotter$_2$ (gray line), and Chambers$_{4A}$ (fat light gray line) for 15 configurations of 3×10^4 measurements, each for 16 sites, for spin up and down each, for $U = 8$ and half-filling, normalized with the total number of samples. Some isolated densities in $[3, 4]$ and $[-2, -3]$ for Suzuki–Trotter$_1$ are not shown.

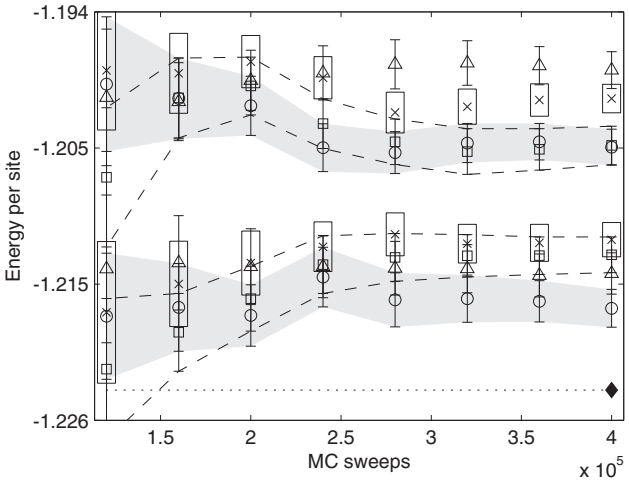

Fig. 9. Ground state energy vs MC sweeps for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign without density cutoff (upper data cloud) and with density cutoff (lower data cloud) for the same configurations in $\tau = 1/8$ for $\theta = 4$ with Ruth$_3$ (○ with gray shadow), Blanes & Moan$_4$ (× with box-frame), Suzuki fractal$_4$ (□ with dashed envelope), and McLachlan$_5$ (△). To the value for the exact energy (◆), a dotted line is added to guide the eye.

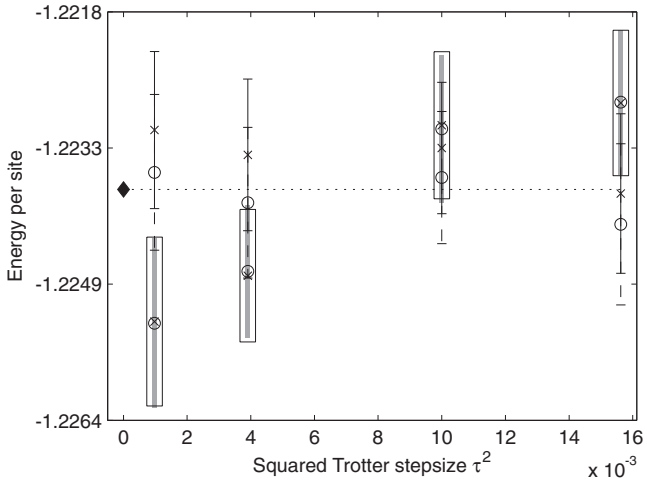

Fig. 10. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ for different values of τ for $\theta = 4$ by Chambers$_{4A}$ without density cutoff (× solid line: without sign, ○ dashed line: with sign) and with density cutoff (× box-frame: without sign, ○ gray straight-bar: with sign). To the value for the exact energy (◆), a dotted line is added to guide the eye.

the Green's function $n_{i,\sigma} = G_\sigma(l)_{i,i}$ may indicate that the off-diagonal elements $G_\sigma(l)_{i,j}$ are also unreliable: Fourier transforms and higher order correlation functions may be even farther off. As we have seen that ST-methods do not constrain individual electron densities, so we can try to eliminate unphysical densities "by hand". We can contrive a new "constrained sampling" scheme where we continue to use the weight function after eq. (14) to obtain the MC time sequence of HS configurations for the sake of ergodic sampling, but we allow only those densities for sampling where all densities fulfill the fermion condition $0 \leq n_{i,\sigma} \leq 1$ and throw away all other HS configurations which lead to unphysical $n_{i,\sigma} > 1$ or $n_{i,\sigma} < 0$ for each spin direction σ during the sample procedure. In Fig. 9, it can be seen that the deviations from the exact value when ignoring the sign becomes smaller compared to direct sampling indiscriminate of the densities. In Fig. 10, we compare the ground state energy with traditional sampling and sampling with density cutoff ignoring the sign and taking the sign into account for the symmetric Chambers$_{4A}$. By using the sampling with density cutoff, the ground state with sign and without sign lie at the same value within the statistical error-bar. Tests with other methods such as Chambers$_{6B}$ and Laskar & Robutel$_{8A}$ yielded similar results. For the asymmetric decomposition Ruth$_3$ shown in Fig. 11, the ground state energy of ignoring the sign by using the sampling with density cutoff becomes much closer to the exact value while the values with sign start to deviate more from the exact value. Results with McLachlan$_5$ are similar. Even though the unphysical densities are relatively rare in Fig. 8, the deviations are significant, so large, that the distance of PQMC-data for the ground state energy and the exact diagonalization is larger than the error-bar. That is the effect of "unphysical densities", that a few measurements can create a significant drift in the data. In other words, when we do not allow unphysical electron densities on those slices which enter the sampling (mind that there may still be unphysical densities hidden away from these slices), the values for the ground state energy computed with and without sign get closer to each other. So there is a strong suspicion that the deviations of the samples with and without signs stem from weird correlations between the configurations with negative sign and their unphysical densities. The minus-sign sampling in eq. (18) will only eliminate unphysical electron densities if either the up- or the down-spin determinants are affected. If both up- and down-determinants are negative, the total

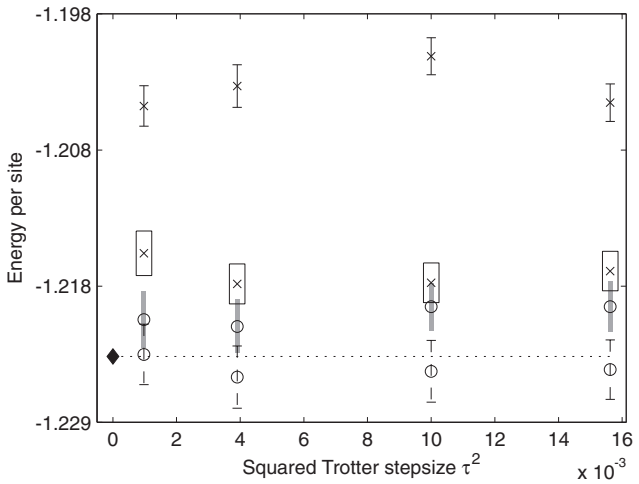

Fig. 11. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ for different values of τ for $\theta = 4$ by Ruth$_3$ without density cutoff ($\times$ solid line: without sign, ○ dashed line: with sign) and with density cutoff ($\times$ box-frame: without sign, ○ gray straight-bar: with sign). To the value for the exact energy (◆), a dotted line is added to guide the eye.

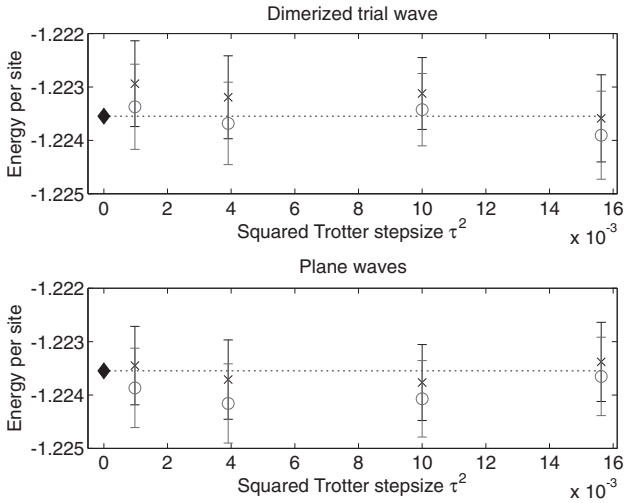

Fig. 12. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign ($\times$) and taking the sign into account (○ gray) for the same configurations with different values of τ for $\theta = 4$ by Chambers$_{4A}$. Dimerized trial state (upper) and plane waves (lower) are used as initial trial wave functions. To the value for the exact energy (◆), a dotted line is added to guide the eye.

configuration is positive, and there is no compensation and the two unphysical densities will affect the observable computation.

In the presence of errors, e.g., due to lack or faulty stabilization, too large ST-stepsize, too many unphysical densities, dysfunctional matrix inversions,[39] beneficial effects of the inclusion of the sign can be expected up to a certain point. The noise will affect the determinant so that the contaminated configurations will be distributed with equal probability into the bin with positive and that with negative sign. Nevertheless, if both up- and down-determinant become negative, the contaminated data are sampled as "positive configuration", and no compensation will occur. The minus-sign sampling bloats the error-bars. Historically, large deviations between diagonalization values and QMC simulations also arose from erroneously reported data: We were notified[46] that exact diagonalization results in ref. 47, had been obtained probably under some symmetry assumptions, so that the exact diagonalization results for the ground state energy with the full matrix are lower by 10% and more. Therefore, also reference data from exact diagonalization may be unreliable, so that we computed our own data. In Table I, for Suzuki–Trotter$_1$ the deviation from the exact energy are hopelessly too large both for sampling with and without sign, but that is no contradiction to the validity of the PQMC-method, only a proof that the commutator in eq. (16) will not oblige to vanish with Suzuki–Trotter$_1$. It is first order ST-decompositions which lead to uncontrolled approximations, not the sampling ignoring the sign.

4. Effect of the Trial Wave Functions

When we want to analyze the errors of our PQMC-method, a question which has not been touched is the influence of the test wave function. Suitable properties have to be postulated in the ansatz in eqs. (10) and (11), and their violation may lead to systematical errors which cannot be analyzed via MC error-bars. It has been known[3,41] for considerable time that there are effects of the test wave function: Such effects of the wave function are again an argument against the "reality" of the minus-sign: As the results of the PQMC should be independent of $|T\rangle$ for "reasonable" test wave functions, a quantity like the sign which varies strongly with $|T\rangle$ has no place in a controllable numerical approximation, and variations due to the test wave functions are systematic errors. For a given ST-decomposition, different averages in the sign then becomes a mere correlation between HS configurations and the respective test wave function. The test wave function is not a physical quantity anyway, as the result should be independent of it, so correlations from it are not physically meaningful, either.

4.1 $N_\uparrow = N_\downarrow = 5$ *with* $U = 4$

In Fig. 12, we plotted the ground state energy by Chambers$_{4A}$ for different values of the Trotter stepsizes τ with dimerized and plane waves as trial wave functions to show that the convergence to the same energy value is obtained. Our results for both ignoring the sign and taking the sign into account agree with other references.[48]

4.2 $N_\uparrow = N_\downarrow = 5$ *with* $U = 8$

In Figs. 13 and 14, we show the energy convergence of our ground state energy computed with and without the sign along the MC sweeps by Suzuki–Trotter$_2$ and Chambers$_{4A}$ for dimerized and plane waves, respectively. For both methods, the values computed with sign are closer to the exact value whereas their fluctuations are much influenced by the different trial wave functions which is in contrast to the values computed without sign.

4.3 $N_\uparrow = N_\downarrow = 7$ *with* $U = 4$

In Figs. 15 and 16, we show the energy convergence of our ground state energy computed with and without the sign along the MC sweeps by Suzuki–Trotter$_2$ and Chambers$_{4A}$ for dimerized and plane waves, respectively. In contrast to the closed shell system, for an open shell system, the values

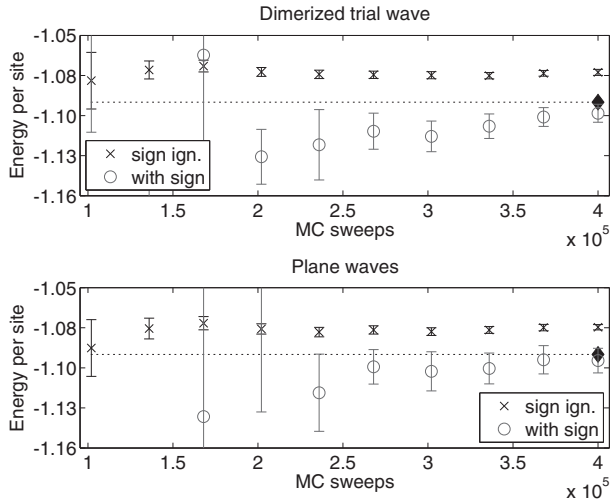

Fig. 13. Ground state energy for the 4×4 systems at $U = 8$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign ($\times$) and taking the sign into account (○ gray) for the same configurations in $\tau = 1/16$ with $\theta = 8$ by Suzuki–Trotter$_2$ computed from dimerized trial wave functions (upper) and plane waves (lower). To the value for the exact energy (◆), a dotted line is added to guide the eye.

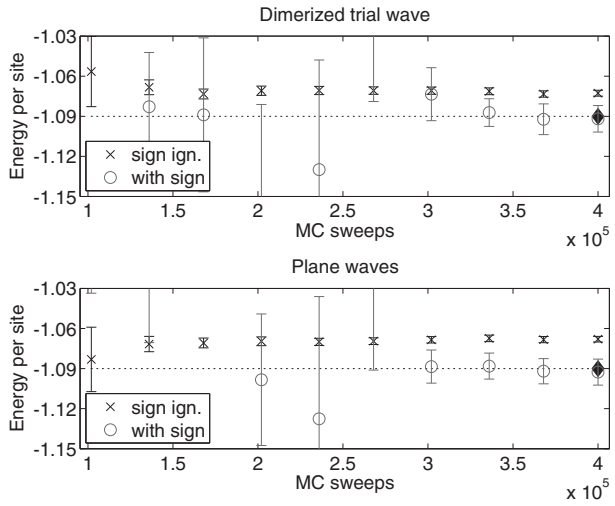

Fig. 14. Ground state energy for the 4×4 systems at $U = 8$ with $N_\uparrow = N_\downarrow = 5$ from Monte Carlo runs ignoring the sign ($\times$) and taking the sign into account (○ gray) for the same configurations in $\tau = 1/16$ with $\theta = 8$ by Chambers$_{4A}$ computed from dimerized trial wave functions (upper) and plane waves (lower). To the value for the exact energy (◆), a dotted line is added to guide the eye.

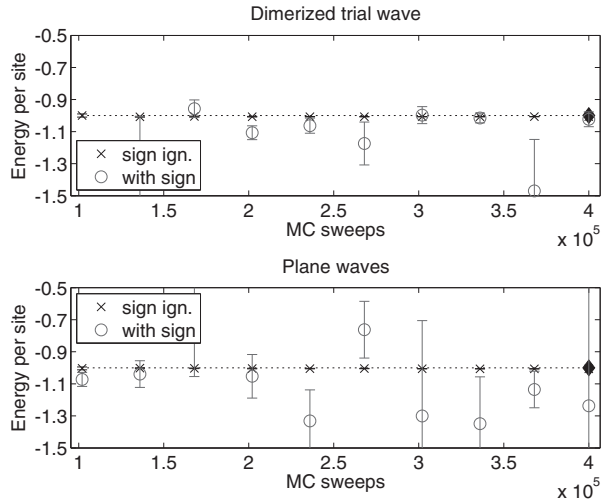

Fig. 15. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 7$ from Monte Carlo runs ignoring the sign ($\times$) and taking the sign into account (○ gray) for the same configurations in $\tau = 1/16$ with $\theta = 4$ by Suzuki–Trotter$_2$ computed from dimerized trial wave functions (upper) and plane waves (lower). To the value for the exact energy (◆), a dotted line is added to guide the eye.

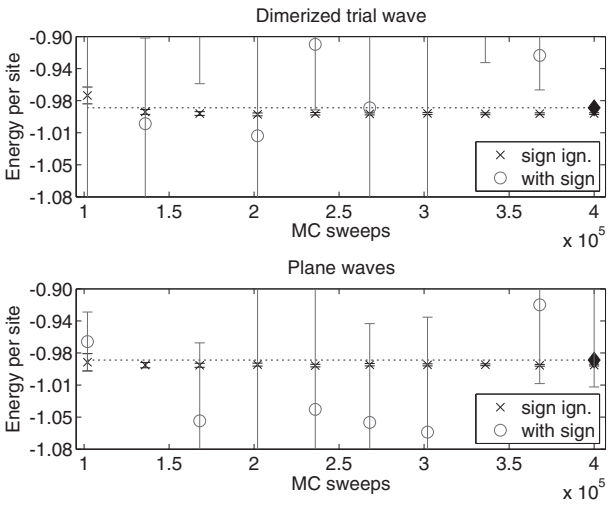

Fig. 16. Ground state energy for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 7$ from Monte Carlo runs ignoring the sign ($\times$) and taking the sign into account (○ gray) for the same configurations in $\tau = 1/16$ with $\theta = 4$ by Chambers$_{4A}$ computed from dimerized trial wave functions (upper) and plane waves (lower). To the value for the exact energy (◆), a dotted line is added to guide the eye.

computed without sign by both methods with different trial wave functions are closer to the exact value while the values computed with sign are distant from the exact value and again have large fluctuations. As shown in Fig. 17, the average sign is much influenced by the initial trial wave functions where the dimerized trial wave function has benign average signs than those from plane waves for both methods.

4.4 $N_\uparrow = N_\downarrow = 7$ *with* $U = 8$

That in §4.1 and §4.3 the data with $U = 4$ are closer to the values with exact diagonalization for sampling without sign, and in §4.2 the data with $U = 8$ are closer for sampling with sign should not create the impression that sampling with sign is beneficial for strong interaction. For an open shell system with $N_\uparrow = N_\downarrow = 7$ at $U = 8$, the results with ignoring the sign are closer, see Fig. 18. With the given wave functions, it depends on the physical parameters whether sampling with or without sign gives results closer to exact diagonalization.

4.5 *Alternative trial wave functions*

We have made comparisons with other trial wave functions (eigenvectors to the largest eigenvalues of the exponential of the matrix of the kinetic energy $\hat{K}$, as well as antiferromagnetic states) for selected parameters. The results were within the limits of the dimerized and plane waves shown here.

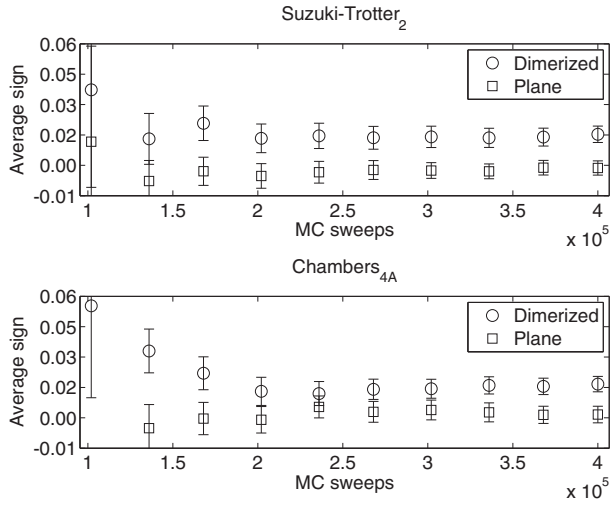

Fig. 17. Average sign for the 4×4 systems at $U = 4$ with $N_\uparrow = N_\downarrow = 7$ from Monte Carlo runs for the same configurations in $\tau = 1/16$ with $\theta = 4$ by Suzuki–Trotter$_2$ (upper) and Chambers$_{4A}$ (lower) computed from dimerized trial wave functions (○) and plane waves (□).

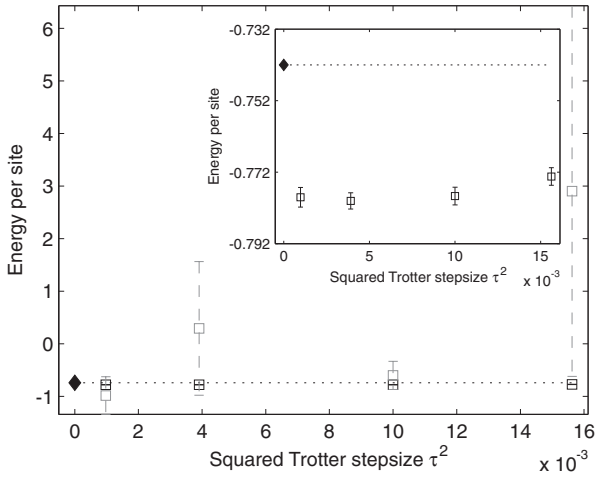

Fig. 18. Ground state energy for the 4×4 systems at $U = 8$ with $N_\uparrow = N_\downarrow = 7$ from Monte Carlo runs ignoring the sign (black line) and taking the sign into account (gray dashed line) for the same configurations with different values of τ for $\theta = 8$ by Chambers$_{4A}$ (□). To the value for the exact energy (◆), a dotted line is added to guide the eye.

5. Relation with Other Methods

5.1 *Path integral methods*

In the field of path integral methods, what we call first order ST-decomposition is called[49] the "primitive approximation", introduced before Quantum Monte Carlo methods were thought of, see e.g., eq. (2-34) p. 38 in Feynman.[50] The error analysis for the approximated Hamiltonian in eqs. (16) and (17) is valid for these cases. The field of path integral methods lacks the self-confidence of the QMC community, and, for the evaluation of observables talks bashfully only about "estimators"[49,51,52] for the energy, etc. From eq. (16), one can deduce that this terminology spread because more often than not, the first order approximation to the partition function left a lot to be desired for the actual accuracy for which energies were evaluated when alternative methods were available for comparison.

5.2 *Langevin dynamics*

Our simulation has been performed exclusively with the PQMC-scheme with heat-bath dynamics. Nevertheless, one of the authors (H.-G. M) played around with heat-bath sampling modified by orders of magnitude of multiplicative noise, up to 5 to 8 decades, which is not so large, taking into account the magnitude of the fermion determinant. The results for the ground state energy did not change, after all, the noise was ergodic. Therefore, we think that our result may be also valid for simulations performed with Langevin dynamics.[53]

5.3 *World-line QMC*

Originally, the minus-sign sampling was introduced *ad hoc* for world-line simulation.[30] Many misleading conclusions about the DQMC come from indiscriminate adoptions of ideas and results in the world-line QMC[54] method (WLQMC), especially in two[55] and higher dimensions. The world-line algorithm is in principle a path integral method for lattices. Unfortunately, this incurs negative transition probabilities when fermions change their place, the mother of all minus-sign problems. A recent analysis[56] talks about "the" fundamental limitations to QMC simulations in general, without discriminating between world-line and determinant (auxiliary field) methods. First of all, WLQMC will in general be applied with first order ST-decompositions, else there would be a tremendous blow-up of the necessary number of ST-slices if disconnected-bond decompositions have to be performed in higher order. In that case, WLQMC suffers from the zero order accuracy of the approximated Hamiltonian in eq. (16) when observables are computed. While the fermion amplitudes can be constrained to physical densities between 0 and 1, the world-line approach has a more severe glitch even for ST-decompositions of higher order: The decoupling is done via the closure relation so that a complete sets of orthogonal single-particle-states is between the ST-slices to decouple the interaction. This gives access to world-lines of individual particles in contradiction to the fundamentals of fermion statistics as a theory for indistinguishable particles, and therefore such an approach should give errors beyond those of the decomposition. Only if several world-lines become connected in ST-time, the respective particles become indistinguishable again. That can probably not be realized with a ground state algorithm, as periodic boundary conditions in ST-time would be necessary, which necessitate finite temperature simulations. In DQMC, only one Green's function is accessible for the computation of observables, so there is no possibility to discriminate between individual electrons. We cannot exclude that for the world-line approach, the minus-sign weighing may have merits, though there are no bounds on an error which is due to the introduction of discriminable particles. Nevertheless, on the computational side, another problem of WLQMC is that it is far from clear how world-lines must be moved so that the configuration space is sampled evenly. The minus-sign problem there might as well be related to ergodicity problems, which is suggested by the impact different updating strategies[57] have on the magnitude of the sign.

5.4 *"Interpolation"*

World-line and DQMC algorithms are substantially different. Though Hirsch[42] proposed something what he chose to call an "interpolation" between QMC methods, it is only a coupling of the same system simulated with two different methods. The coupling of two models via a statistical weight does not prove any equivalence or relation between the two models. Neither can there be any conclusions be drawn about relation or equivalence from a coupling between different methods for the same system. The sign for the world-line approach has nothing to do with the sign in the DQMC, as can be seen just by counting the degrees of freedom: In the determinant approach, their number depends on the number of HS spins per site, while in the world-line approach they depend on the number of particles in the system.

5.5 *Grand-canonical simulations at finite temperature*

Many results and conclusions in this article on PQMC are also relevant for its finite temperature equivalent, the grand-canonical quantum Monte Carlo[4] (GQMC). First of all, there is also a fermion determinant, only its rank is equal to the number of sites, not to the number of electrons. The unphysical densities there have been noted earlier. In our experience,[58] the energies for low temperature GQMC match those for PQMC quite well, even with the tendency, though not the magnitude of the sign. Nevertheless, for the GQMC there is no effect of a test wave function. In the future, we will experiment with higher order ST-decompositions also with GQMC.

6. Conclusions

All things being equal, asymmetric ST-decompositions show smaller signs than symmetric decompositions, negative fermion determinants can therefore be attributed to the loss of symmetry in the product matrix of itself symmetric matrices. All things being equal, higher order ST-decompositions with large negative substeps give lower signs than those with smaller negative substeps, an issue which affects the convergence. Accordingly, the pseudo-symplectic decompositions without negative ST-substeps perform rather well, in the face of the fact that they are derived under additional assumptions.

The energies computed without sign are higher than with sign, exactly the opposite of what has been claimed in ref. 59: We find consistently smaller energies when we take the sign into account. Whenever we could control the parameters, deviations between the averages computed with and without sign were due to unphysical elements in the simulation, like asymmetric decompositions (§2.4) or unphysical electron densities (§3.3). Averages sampled without sign may be higher or lower than the exact values, and correspondingly the inclusion of the sign gives sometimes better, sometimes worse averages. The assertion in ref. 59: "It should be emphasized that the neglect of the sign of the fermion determinant is an uncontrolled approximation, and can lead to misleading results if the correction terms are not included". For the occurrence of negative configurations, unphysical fermion densities outside $[0, 1]$ are necessary. Therefore the unphysical approximation is the one which takes the sign due to such unphysical densities into account. The post-processing using the sign only enhances the fluctuations in the data to an extent that the distributions become non-Gaussian. We have at least as small, if not smaller errors than other publications: For the closed shell system, for methods with benign sign, sampling with and without sign was consistent. For open shell systems, sampling without sign was more accurate.

From the error compensation of the sampling with sign (numerical errors, due to lack or malfunction of numerical stabilizations or due to dysfunctional matrix inversions, error bounds in the approximated Hamiltonian or unphysical densities) works only if the error and the change in sign affects one of the two spin direction, and then only statistically. Unphysical densities are a problem which cannot be eliminated by better stabilizations, inversions or the use of quadruple precision: It will have to be seen in the future whether it is sufficient to use higher order methods, after all, in fluid dynamics, the use of 6th order CIP-methods (Constrained Interpolated Profile) is common. Nevertheless, it has to be seen whether alternatives to the conventionally used HS-transformation have to be found or maybe even "constrained ST-decomposition" should be derived which are able to fix the electron densities (and hopefully, all other entries in the Green's function) within the limits given by fermion statistics. Regarding how much the trial wave functions are responsible for the unphysical densities, comparisons with the finite temperature algorithm will have to be seen. Deviations between ground state energies sampled with different trial wave functions can be larger than the MC error-bars.

For any system, classical, fermionic or bosonic, the approximation of the Hamiltonian via ST-decompositions is one order lower than for its exponential. Our results in Table II show that this theorem is not just a pessimistic estimate, but there is indeed a lack of convergence for the ground state energy of the first order ST-decomposition for the ST-stepsize approaching zero: First order ST-approximations are therefore zero order computations of the observables and not controlled approximations. Many "results" and "methods" have been derived as path integral formulas via the "primitive approximation"[49] or via the "Baker–Hausdorff formula"[60] which are equivalent to Suzuki–Trotter$_1$, so that the accuracy of the results for the corresponding Lagrangian or Hamiltonian hinges on the actual disappearance of the commutator of the constituents which the authors could not deal with in the first place. It is rather mind-boggling to imagine the Sisyphean task which lies ahead of physics to figure out which results obtained since the inception of the path integral method[50] are correct, which are amendable and which are just irrelevant.

The requirement for the Markov probability of the Monte Carlo process is that the detailed balance is fulfilled, beyond that, the computation of the average is reduced to a simple average of the obtained observables. In this sense, a Metropolis weight or a heat-bath weight can be used, there is no need to manipulate the variables afterwards: Nobody would decorate the observables in conventional Monte Carlo with additional factors (quotient between the heat-bath and Metropolis weight or *vice versa*): The guarantee of the detailed balance guarantees the correctness of the Monte Carlo sampling. Though the formula for the sampling in

Table A·I. Coefficients for higher order ST-decompositions.

Method	Order p	Substeps $M_{\rm sub}$	γ_n	η_n
Suzuki–Trotter$_1$[1,2]	1	1	1.0	1.0
McLachlan$_2$[65,66]	2	2	$1-1/\sqrt{2}$	$1/\sqrt{2}$
			$1/\sqrt{2}$	$1-1/\sqrt{2}$
Suzuki–Trotter$_2$[67,68]	2	2	0.5	1.0
			0.5	0.0
McLachlan$_3$[65,66] AS	3	3	$\gamma_1=\eta_3$	$\eta_1=0.9196615230173999$
			$\gamma_2=\eta_2$	$\eta_2=1/(4\eta_1)-\eta_1/2$
			$\gamma_3=\eta_1$	$\eta_3=1-\eta_1-\eta_2$
Ruth$_3$[69] AS	3	3	7/24	2/3
			3/4	−2/3
			−1/24	1.0
Blanes & Moan$_4$[40]	4	7	0.0829844064174052	0.2452989571842710
			0.396309801498368	0.6048726657110800
			−0.0390563049223486	$1/2-(\eta_1+\eta_2)$
			$1-2(\gamma_1+\gamma_2+\gamma_3)$	η_3
			γ_3	η_2
			γ_2	η_1
			γ_1	0.0
Calvo & Sanz-Serna$_4$[70] AS	4	5	0.06175885813562632	0.2051776615422864
			0.3389780265536433	0.4030212816042146
			0.6147913071755776	−0.1209208763389141
			−0.1405480146593734	0.5127219331924131
			0.1250198227945261	0.0
Candy & Rozmus$_4$,[71]	4	4	$(2+2^{1/3}+2^{-1/3})/6$	$1/(2-2^{1/3})$
Forest & Ruth$_4$[72]			$(1-2^{1/3}-2^{-1/3})/6$	$1/(1-2^{2/3})$
			γ_2	η_1
			γ_1	0.0
Chambers$_{4A}$[21–23]	4	3	$(1-1/\sqrt{3})/2$	1/2
			$1/\sqrt{3}$	1/2
			γ_1	0.0
Chambers$_{4B}$[21–23]	4	3	0.0	1/6
			1/2	2/3
			1/2	η_1
McLachlan$_4$[65,66] AS	4	4	0.1344961992774311	0.5153528374311229
			−0.2248198030794208	−0.08578201941297364
			0.7563200005156683	0.4415830236164665
			0.3340036032863214	0.1288461583653842
Suzuki fractal$_4$[36,73,74]	4	6	0.2072453858971879	0.4144907717943757
			0.4144907717943757	0.4144907717943757
			−0.1217361576915636	−0.6579630871775028
			γ_3	η_2
			γ_2	η_1
			γ_1	0.0

Continued on next page.

fermionic QMC procedures is always written with the allegedly signed (positive and negative) probabilities as in eq. (18), the probabilities actually used to generate the Markov chain always use absolute values of the fermion determinants. Anything else makes no sense, as there is no mathematical theory of mixed negative and positive probabilities, or a theory of Markov processes with negative probabilities. An introduction of a weight for the sampling which is different from the weight which generates the probability distribution is unique to the field of Quantum Monte Carlo, and has no mathematical justification from Monte Carlo theory, on the contrary: Due to the exceptional treatment of using different probabilities for the generation of the Markov chain (absolute value of the fermion determinant) and the observable computation (determinant with sign), the probability distributions can deviate from the Gaussian shape, in contradiction to the law of large number itself. Originally, the introduction of the sign in the sampling for world-lines was *ad hoc*, the multi-particles states were claimed to be bosonic, and the sign was allegedly used to recover the "fermionic character". Mathematically, the multi-particle states in the world-line formalism are not

Continued.

Method	Order p	Substeps M_{sub}	γ_n	η_n
McLachlan$_5$[65] AS	5	6	0.1193900292875673	0.3398396258391100
			0.6989273703824752	−0.08860133690302732
			−0.1713123582716008	0.5858564768259621
			0.4012695022513534	−0.6030393565364919
			0.01070508184823598	0.3235807965546976
			0.05897962549803117	0.4423637942197495
Tselios$_5$[75] AS	5	7	0.4515650720436606	1.904232780508446
			−0.002625517726040550	−1.939586366441925
			−0.2887462490910128	0.3960766510231830
			0.4703720043422902	0.5133868104090695
			0.3704466763359328	−2.967739460604547
			0.1934796732533846	0.004177409528669316
			−0.1944916591582146	3.089452175577104
Blanes & Moan$_{6A}$[40]	6	15	0.0	0.03785931984061160
			0.09171915262446165	0.1026356331024350
			0.1839831700050060	−0.02586788826655870
			−0.05653436583288827	0.3142414030714470
			0.004914688774712854	−0.1301444595174150
			0.1437611271683580	0.1064177003695430
			0.3285676937468040	−0.008794243128510581
			−0.1964114664864542	0.2073050690568954
			γ_8	η_7
			γ_7	η_6
			γ_6	η_5
			γ_5	η_4
			γ_4	η_3
			γ_3	η_2
			γ_2	η_1
Blanes & Moan$_{6B}$[40]	6	12	0.04146499851826240	0.1232297759462710
			0.1981286719180670	0.2905537977995580
			−0.04000619210415330	−0.1270492126254170
			0.07525398430158070	−0.2463317610620750
			−0.01151138742068790	0.3572088727959280
			0.2366699247869311	0.2047770542914701
			γ_6	η_5
			γ_5	η_4
			γ_4	η_3
			γ_3	η_2
			γ_2	η_1
			γ_1	0.0

Continued on next page.

quantum states at all as different particles have different world-lines which can be discriminated in violation of the basic tenets of quantum mechanics and quantum statistics. Therefore, the whole argument for the introduction of the sign is invalid. It is the actual (positive) probability used in the Markov chain which determines the sampling, which also guarantees that the observables are Gaussian distributed if the detailed balance is fulfilled. Therefore, using a sign in eq. (18) "with hindsight" for a state which was generated independent of this sign is a misrepresentation of the Monte Carlo sampling process.

A different question is why the use of the absolute value of the determinant can generate the Markov chain at all. From the point of approximation via the ST-decomposition, it is clear that the sign of the determinant depends on the sign of all of the singular values in $\langle\Phi|\Psi\rangle$. Many of them are considerably smaller than the error which is in the Suzuki–Trotter decomposition itself: Letting a quantity used in sampling depend on something which is smaller than the noise in the sampling procedure makes no sense. Another argument concerns the trial wave function. The results in the filtering process should be independent of the trial wave function $|T\rangle$, and in particular, of the phases of the single particle waves ϕ_i it is composed of. If we change one of ϕ_i to $-\phi_i$, the sign of the determinant will be reversed, while the Green's function is invariant under this change: In the computation of eq. (15), the phase turns up twice, once in inverted part, once in the linear part, and both phase changes cancel. This independence is according to the principles of quantum mechanics. For the computation of the Markov

Continued.

Method	Order p	Substeps $M_{\rm sub}$	γ_n	η_n
Chambers$_{6A}$[22,23]	6	4	$(1-3/\sqrt{15})/2$	5/18
			$3/(2\sqrt{15})$	4/9
			γ_2	η_1
			γ_1	0.0
Chambers$_{6B}$[22,23]	6	4	0.0	1/12
			$(1-1/\sqrt{5})/2$	5/12
			$1/\sqrt{5}$	η_2
			γ_2	η_1
Yoshida$_{6A}$[76,77]	6	8	0.3922568052387786	0.7845136104775573
			0.5100434119184577	0.2355732133593581
			−0.4710533854097564	−1.177679984178871
			0.06875316825252009	1.315186320683911
			γ_4	η_3
			γ_3	η_2
			γ_2	η_1
			γ_1	0.0
Laskar & Robutel$_{8A}$[23]	8	5	0.06943184420297371	0.1739274225687269
			0.2605776340045981	0.3260725774312730
			0.3399810435848563	η_2
			γ_2	η_1
			γ_1	0.0
Yoshida$_{8D}$[76,77]	8	16	0.4574221231148958	0.9148442462297915
			0.5842687913980015	0.2536933365662113
			−0.5955794501471969	−1.444852236860605
			−0.8015464361143276	−0.1582406353680502
			0.8899492511272379	1.938139137622526
			−0.01123554767639245	−1.960610232975311
			−0.9289051917917571	0.1027998493917964
			0.9056264600895391	1.708453070787282
			γ_8	η_7
			γ_7	η_6
			γ_6	η_5
			γ_5	η_4
			γ_4	η_3
			γ_3	η_2
			γ_2	η_1
			γ_1	0.0

chain, the use of the absolute value again eliminates the phase. Only the sampling with sign will take into account such an artificial sign explicitly. We are therefore currently looking for systems where the sampling with sign leads to macroscopic deviations from the true values, in contrast to the diminutive differences we found in §4 which are smaller than the deviations obtained with different test wave functions.

Acknowledgments

The authors acknowledge the financial support of The Iwatani Naoji Foundation and East Asian Circle of Applied Technology Foundation. We are indebted to many participants at the QMC methods at work for novel phases of matter, Trieste, Jan. 23–Feb. 3, Italy 2012 for enlightening discussions, and especially thank Professor F. F. Assaad for the code to generalize the dimerized trial wave function as well as for his permission to use them in this paper.

Appendix: Coefficients for Higher Order ST-Decompositions

We show the coefficients γ_n and η_n for the various higher order ST-decompositions in Table A·I in alphabetical order. Yoshida$_{6A}$, of which the closed form is given by Koseleff,[61,62] is known as the best among the three possible solutions.[63] Yoshida$_{8D}$ is the optimal solution among the five possible solutions.[64] Note that the symmetric higher order ST-decompositions indicate $\gamma_n = \gamma_{M_{\rm sub}+1-n}$ and $\eta_n = \eta_{M_{\rm sub}-n}$ with $\eta_{M_{\rm sub}} = 0$ or $\eta_n = \eta_{M_{\rm sub}+1-n}$ and $\gamma_n = \gamma_{M_{\rm sub}+2-n}$ with $\gamma_1 = 0$. Antisymmetric decompositions are indicated by AS.

1) M. Suzuki: Prog. Theor. Phys. **56** (1976) 1454.
2) M. Suzuki: Commun. Math. Phys. **51** (1976) 183.
3) M. Imada and Y. Hatsugai: J. Phys. Soc. Jpn. **58** (1989) 3752.

4) J. E. Hirsch: Phys. Rev. B **31** (1985) 4403.
5) N. Hatano: Dr. Thesis, The University of Tokyo, Department of Physics (1992).
6) H. De Raedt and W. von der Linden: in *The Monte Carlo Method in Condensed Matter Physics*, ed. K. Binder (Springer, Heidelberg, 1992) Chap. 8, pp. 249–284.
7) M. Frick and H. De Raedt: Z. Phys. B **88** (1992) 173.
8) F. F. Assaad and H. G. Evertz: in *Computational Many Particle Physics*, ed. H. Fehske, R. Schneider, and A. Weiße (Springer, Heidelberg, 2008) Lecture Notes in Physics, Vol. 739, pp. 277–356.
9) E. Hairer, C. Lubich, and G. Wanner: *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations* (Springer, Heidelberg, 2006) 2nd ed., p. 149.
10) R. I. McLachlan and G. R. W. Quispel: J. Phys. A **39** (2006) 5251.
11) L. Verlet: Phys. Rev. **159** (1967) 98.
12) P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel: *Computational Molecular Dynamics: Challenges, Methods, Ideas* (Springer, Heidelberg, 1997) p. 301.
13) M. Müller: private communication.
14) S. Yamada, T. Imamura, and M. Machida: Proc. ACM/IEEE Superconducting Conf., 2005 (SC'05) p. 44.
15) S. Yamada, T. Imamura, T. Kano, and M. Machida: Proc. ACM/IEEE Supercomputing Conf., 2006 (SC'06).
16) W. von der Linden: Phys. Rep. **220** (1992) 53.
17) X. Wang and H.-G. Matuttis: Theor. Appl. Mech. Jpn. **59** (2010) 309.
18) M. Sofroniou and G. Spaletta: Opt. Methods Software **20** (2005) 597.
19) W. Kahan and R.-C. Li: Math. Comput. **66** (1997) 1089.
20) M. Suzuki: J. Math. Phys. **32** (1991) 400.
21) R. I. McLachlan: BIT Numer. Math. **35** (1995) 258.
22) J. E. Chambers and M. A. Murison: Astron. J. **119** (2000) 425.
23) J. Laskar and P. Robutel: Celestial Mech. Dyn. Astron. **80** (2001) 39.
24) J. W. Negele and H. Orland: *Quantum Many-Particle Systems* (Perseus Books, 1998) pp. 416–423.
25) J. E. Hirsch: Phys. Rev. B **28** (1983) 4059.
26) J. E. Hirsch: Phys. Rev. B **29** (1984) 4159.
27) P. Saha and S. Tremaine: Astron. J. **104** (1992) 1633.
28) S. H. Ng and H.-G. Matuttis: Theor. Appl. Mech. Jpn. **59** (2010) 323.
29) H. Yoshida: Celestial Mech. Dyn. Astron. **56** (1993) 27.
30) H. De Raedt and A. Lagendijk: Phys. Rev. Lett. **46** (1981) 77.
31) Y. Asai: Phys. Rev. B **62** (2000) 10674.
32) F. F. Assaad and D. Würtz: Z. Phys. B Condens. Matter **80** (1990) 325.
33) E. Dagotto, A. Moreo, F. Ortolani, D. Poilblanc, and J. Riera: Phys. Rev. B **45** (1992) 10741.
34) M. Suzuki: Phys. Lett. A **165** (1992) 387.
35) M. Suzuki: Physica A **205** (1994) 65.
36) M. Suzuki: Phys. Lett. A **146** (1990) 319.
37) M. Suzuki: Physica A **194** (1993) 432.
38) N. Hatano and M. Suzuki: Prog. Theor. Phys. **85** (1991) 481.
39) H.-G. Matuttis and N. Ito: J. Phys. Soc. Jpn. **70** (2001) 1519.
40) S. Blanes and P. C. Moan: J. Comput. Appl. Math. **142** (2002) 313.
41) N. Furukawa and M. Imada: J. Phys. Soc. Jpn. **60** (1991) 3669.
42) J. E. Hirsch: Phys. Rev. B **38** (1988) 12023.
43) R. Blankenbecler, D. J. Scalapino, and R. L. Sugar: Phys. Rev. D **24** (1981) 2278.
44) S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar: Phys. Rev. B **40** (1989) 506.
45) K. Binder and D. W. Heermann: *Monte Carlo Simulation in Statistical Physics: An Introduction* (Springer, Heidelberg, 1992) Springer Series in Solid-State Sciences, p. 16.
46) H. Shi: private communication.
47) H. De Raedt and W. von der Linden: Phys. Rev. B **45** (1992) 8787.
48) T. Husslein, W. Fettes, and I. Morgenstern: Int. J. Mod. Phys. C **8** (1997) 397.
49) D. M. Ceperley: Rev. Mod. Phys. **67** (1995) 279.
50) R. P. Feynman and A. R. Hibbs: *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965) 1st ed.
51) M. F. Herman, E. J. Bruskin, and B. J. Berne: J. Chem. Phys. **76** (1982) 5150.
52) M. Parrinello and A. Rahman: J. Chem. Phys. **80** (1984) 860.
53) S. Sorella, E. Tosatti, S. Baroni, R. Car, and M. Parrinello: Int. J. Mod. Phys. B **2** (1988) 993.
54) J. E. Hirsch, R. L. Sugar, D. J. Scalapino, and R. Blankenbecler: Phys. Rev. B **26** (1982) 5033.
55) I. Morgenstern and D. Würtz: Z. Phys. B Condens. Matter **70** (1988) 115.
56) M. Troyer and U.-J. Wiese: Phys. Rev. Lett. **94** (2005) 170201.
57) S. Chandrasekharan and U.-J. Wiese: Phys. Rev. Lett. **83** (1999) 3116.
58) H.-G. Matuttis: Dr. Thesis, Universität Regensburg, Department of Physics (1995).
59) E. Y. Loh, Jr., J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar: Phys. Rev. B **41** (1990) 9301.
60) J. B. Kogut: Rev. Mod. Phys. **51** (1979) 659.
61) P.-V. Koseleff: AAECC'10 Proc., Lect. Not. Comp. Sci., 1993.
62) P.-V. Koseleff: Fields Inst. Commun. **10** (1996) 103.
63) H. Kinoshita, H. Yoshida, and H. Nakai: Celestial Mech. Dyn. Astron. **50** (1991) 59.
64) R. I. McLachlan: SIAM J. Sci. Comput. **16** (1995) 151.
65) R. I. McLachlan and P. Atela: Nonlinearity **5** (1992) 541.
66) D. Okunbor and R. D. Skeel: Math. Comput. **59** (1992) 439.
67) M. Suzuki: J. Math. Phys. **26** (1985) 601.
68) H. De Raedt and B. De Raedt: Phys. Rev. A **28** (1983) 3575.
69) R. D. Ruth: IEEE Trans. Nucl. Sci. **30** (1983) 2669.
70) D. I. Okunbor and R. D. Skeel: J. Comput. Chem. **15** (1994) 72.
71) J. Candy and W. Rozmus: J. Comput. Phys. **92** (1991) 230.
72) É. Forest and R. D. Ruth: Physica D **43** (1990) 105.
73) M. Suzuki and K. Umeno: in *Computer Simulation Studies in Condensed-Matter Physics VI*, ed. D. P. Landau, K. K. Mon, and H.-B. Schüttler (Springer, Heidelberg, 1993) pp. 74–86.
74) N. Hatano and M. Suzuki: in *Quantum Annealing and Related Optimization Methods*, ed. A. Das and B. K. Chakrabarti (Springer, Heidelberg, 2005) pp. 37–68.
75) K. Tselios and T. E. Simos: J. Math. Chem. **35** (2004) 55.
76) H. Yoshida: Phys. Lett. A **150** (1990) 262.
77) H. Yoshida: in *Chaos, Resonance and Collective Dynamical Phenomena in the Solar System*, ed. S. Ferraz-Mello (Springer, Heidelberg, 1992) pp. 407–411.