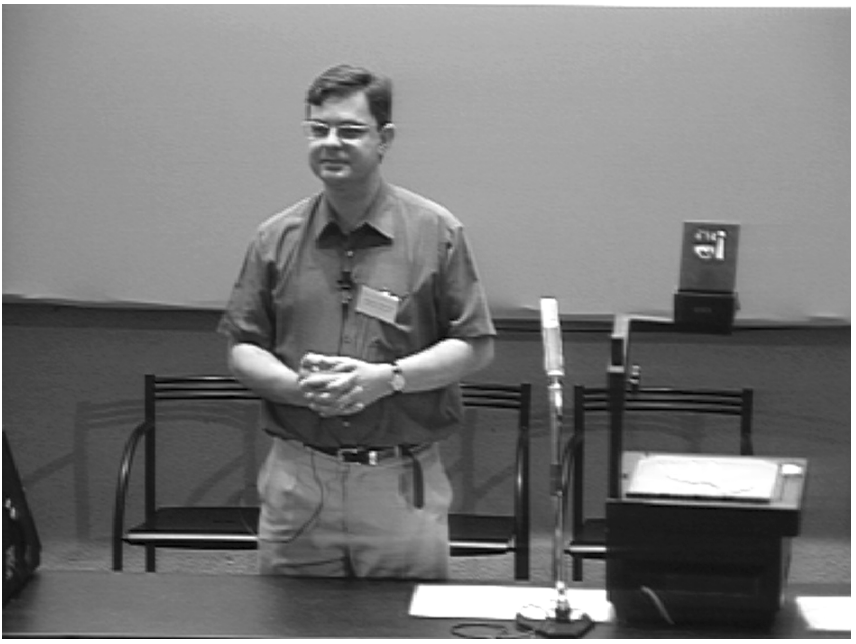
Ensemble Optimization Techniques for Classical and Quantum Systems

S. Trebst[1] and M. Troyer[2]

[1] Microsoft Research and Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA
trebst@comp-phys.org
[2] Theoretische Physik, ETH Zürich, 8093 Zürich, Switzerland
troyer@comp-phys.org



Matthias Troyer

S. Trebst and M. Troyer: *Ensemble Optimization Techniques for Classical and Quantum Systems*, Lect. Notes Phys. **703**, 591–640 (2006)
DOI 10.1007/3-540-35273-2_17

We present a review of extended ensemble methods and ensemble optimization techniques. Extended ensemble methods, such as multicanonical sampling, broad histograms, or parallel tempering aim to accelerate the simulation of systems with large energy barriers, as they occur in the vicinity of first order phase transitions or in complex systems with rough energy landscapes, such as spin glasses or proteins. We present a recently developed feedback algorithm to iteratively achieve an *optimal* ensemble, with the fastest equilibration and shortest autocorrelation times. In the second part we review time-discretization free world line representations for quantum systems, and show how any algorithm developed for classical systems, such as local updates, cluster updates or the extended and optimized ensemble methods can also be applied to quantum systems. An overview over the methods is followed by a selection of typical applications.

1 Introduction

In this chapter we will review recent developments in the simulation of lattice (and continuum) models by classical and quantum Monte Carlo simulations. Unbiased numerical methods are required to obtain reliable results for classical and quantum lattice model when interactions or fluctuations are strong, especially in the vicinity of phase transitions, in frustrated models and in systems where quantum effects are important. For classical systems, molecular dynamics or the Monte Carlo method are the methods of choice since they can treat large systems.

Both Monte Carlo and molecular dynamics simulations slow down in the vicinity of phase transitions or in disordered systems with rough energy landscapes, since the time scales to tunnel through energy barriers can become prohibitively long. Here Monte Carlo simulations have an advantage over molecular dynamics, since in Monte Carlo simulations both the dynamics and the ensemble can be changed to achieve faster tunneling through the energy barriers. Using modern sampling algorithms, such as cluster updates, extended ensemble methods or parallel tempering strategies most classical magnets can be efficiently simulated, with the computational effort scaling with a low power of the system size, and usually linear in system size. The notable exceptions are spin glasses, known to be nondeterministic-polynomially (NP) hard in more than two space dimensions [1] and where most likely no polynomial-time algorithm can exist [2].

In the first part of this chapter we will give a short overview of Monte Carlo simulations for classical lattice models in Sect. 2.1 and will then review the extended and optimized ensemble methods in Sect. 3. We will focus the discussion on a recently developed algorithm to iteratively achieve an *optimal* ensemble, with the fastest equilibration and shortest autocorrelation times.

For quantum magnets, quantum Monte Carlo (QMC) methods are also the method of choice whenever they are applicable. Over the last decade

efficient algorithms for classical Monte Carlo simulations have been generalized to quantum systems and systems with millions of quantum spins have been simulated [3]. In Sect. 4 we will present modern time-discretization free world line representations for quantum lattice models. They faithfully map the quantum system to an equivalent classical system with one more dimension. Efficient Monte Carlo algorithms developed for classical systems can also be applied to quantum systems, using these world line representations, as we will show in Sect. 5.

Unfortunately, in contrast to classical magnets, QMC methods are efficient only for non-frustrated magnets and for bosonic systems. Fermionic degrees of freedom or frustration in quantum systems usually lead to the "negative sign problem", when the weights of some configurations become negative [4]. These negative weights cannot be directly interpreted as probabilities in the Monte Carlo process and lead to cancellation effects in the sampling. As a consequence the statistical errors grow exponentially with inverse temperature and system size and the QMC methods are restricted to small systems and not too low temperatures.

2 The Monte Carlo Method for Classical Lattice Models

2.1 The Metropolis Algorithm

We start with a short review of the Monte Carlo method for calculating integrals of the form

$$\langle O \rangle = \frac{\int_\Omega \mathrm{d}x W(x) O(x)}{\int_\Omega \mathrm{d}x W(x)} \,, \tag{1}$$

where Ω is a discrete or continuous configuration space and $W(x)$ a not necessarily normalized weight function. We want to sample this integral in a Monte Carlo process by creating a sequence $\{x_i\}$ of N configurations, where each configuration is drawn according to the normalized probability distribution function

$$P(x) = \frac{W(x)}{\int_\Omega \mathrm{d}x W(x)} \,. \tag{2}$$

Under the assumption of uncorrelated samples x_i we can then estimate the expectation value (1) by the sample mean

$$\langle O \rangle \approx \frac{1}{N} \sum_{i=1}^{N} O(x_i) \,, \tag{3}$$

within a statistical error

$$\Delta = \sqrt{\frac{\mathrm{Var} O}{N}} = \sqrt{\frac{\langle O^2 \rangle - \langle O \rangle^2}{N}} \,. \tag{4}$$

Since we will, in general, not have a direct algorithm to create samples x_i according to the distribution $P(x_i)$ we will use a Markov process in which starting from an initial configuration x_0 a Markov chain of configuration is generated:

$$x_0 \to x_1 \to x_2 \to \ldots \to x_n \to x_{n+1} \to \ldots \ . \tag{5}$$

A transition matrix T_{xy} gives the transition probabilities of going from configuration x to configuration y in one step of the Markov process. As the sum of probabilities of going from configuration x to any other configuration is one, the columns of the matrix T are normalized:

$$\sum_y T_{xy} = 1 \ . \tag{6}$$

A consequence is that the Markov process conserves the total probability. Another consequence is that the largest eigenvalue of the transition matrix T is 1 and the corresponding eigenvector with only positive entries is the equilibrium distribution which is reached after a large number of Markov steps.

We want to determine the transition matrix T so that we asymptotically reach the desired probability $P(x)$ for a configuration i. A set of sufficient conditions is:

1. **Ergodicity:** It has to be possible to reach any configuration **x** from any other configuration **y** in a finite number of Markov steps. This means that for all **x** and **y** there exists a positive integer $n < \infty$ such that $(T^n)_{\mathbf{xy}} \neq 0$.
2. **Detailed balance:** The probability distribution $p_x^{(n)}$ changes at each step of the Markov process:

$$\sum_x p_x^{(n)} T_{xy} = p_{\mathbf{y}}^{(n+1)} \ , \tag{7}$$

but converges to the equilibrium distribution p_x. This equilibrium distribution p_x is an eigenvector with left eigenvalue 1 and the equilibrium condition

$$\sum_x p_x T_{xy} = p_y \tag{8}$$

must be fulfilled. It is easy to see that the detailed balance condition

$$\frac{W_{\mathbf{xy}}}{W_{\mathbf{yx}}} = \frac{p_{\mathbf{y}}}{p_x} \tag{9}$$

is sufficient.

The simplest Monte Carlo algorithm is the Metropolis algorithm [5] which can be outlined as follows:

- Starting with a configuration $x = x_i$ propose a new configuration y with an a-priori probability A_{xy}.

- Calculate the acceptance ratio

$$P_{xy} = \min\left(1, \frac{A_{yx}W(y)}{A_{xy}W(x)}\right) \tag{10}$$

  and accept the proposed configuration with probability P_{xy}. To do so we draw a uniform random number u in the interval $[0, 1[$ and choose $x_{i+1} = y$ if $u < P_{xy}$ and $x_{i+1} = x$ otherwise.
- Measure the quantity O for the new configuration x_{i+1} no matter whether the proposed configuration was accepted or not.

Since the samples created in this Markov chain are correlated (we only do small changes at each step), equation (4) for the statistical error needs to be modified to

$$\Delta = \sqrt{\frac{\mathrm{Var}\, O}{N}(1 + 2\tau_O)} \tag{11}$$

where τ_O is the integrated autocorrelation time of $O(x_i)$ in the Markov chain.

2.2 The Local Update Metropolis Algorithm for the Ising Model

We will next apply this Metropolis algorithm to simulations of the Ising ferromagnet with Hamilton function

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - g\mu_B h \sum_{i=1}^{N} \sigma_i \ , \tag{12}$$

where J is the exchange constant, h the magnetic field, g the Landé g-factor, μ_B the Bohr magneton, and N the total number of spins. The sum runs over all pairs of nearest neighbors i and j and $\sigma_i = \pm 1$ is the value of the Ising spin at site i.

To calculate the value of an observable, such as the mean magnetization at an inverse temperature $\beta = 1/k_B T$ with T being the temperature and k_B the Boltzmann constant, we need to evaluate

$$\langle m \rangle = \sum_c m(c) \exp(-\beta E(c))/Z, \tag{13}$$

where

$$m(c) = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \tag{14}$$

is the magnetization of the configuration c, $E(c)$ the energy of the configuration,

$$P(c) = \exp(-\beta E(c)) \tag{15}$$

the Boltzmann weight and

$$Z = \sum_c P(c) \tag{16}$$

the partition function, normalizing the weights.

As discussed above, Monte Carlo sampling can be performed on this sum using the Metropolis method. The simplest types of updates are local spin flips:

1. Pick a random site i. The a-priori probabilites A_{xy} are all just $1/N_{sites}$ for a system with N_{sites} spins.
2. Calculate the energy cost ΔE for flipping the spin at site i: $\sigma_i \rightarrow -\sigma_i$
3. Flip the spin with the Metropolis probability $\min[1, \exp(-\beta \Delta E)]$. If rejected, keep the original spin value.
4. Perform a measurement independent of whether the spin flip was accepted or rejected.

The same local update algorithm can be applied to systems with longer-range interactions and with coupling constants that vary from bond to bond. For more complex classical models, such as Heisenberg models, local updates will no longer consist of simple spin flips, but of arbitrary rotations of the local spin vectors.

2.3 Critical Slowing Down and Cluster Update Algorithms

Local update algorithms are easy to implement and work well away from phase transitions. Problems arise in the vicinity of continuous (second order) phase transitions, where these algorithms suffer from "critical slowing down" [6] and at first order phase transitions where there is a tunneling problem through free energy barriers.

At second order phase transitions the correlation length ξ diverges upon approaching the phase transition, and this causes the autocorrelation times τ_O to also diverge as

$$\tau_O \propto \min(L, \xi)^z \tag{17}$$

with a dynamical critical exponent of $z \approx 2$. L is the linear extent of the system. The origin of critical slowing down is the fact that close to the critical temperature large ordered domains of linear extent ξ are formed and the single spin updates are not effective in changing these large domains. The value $z \approx 2$ can be understood considering that the time for a domain wall to move a distance ξ by a random walk scales as ξ^2. The solution to critical slowing down are cluster updates, flipping carefully selected clusters of spins instead of single spins. Cluster update algorithms were originally invented by Swendsen and Wang for the Ising model [6] and soon generalized to $O(N)$ models, such as the Heisenberg model [7]. These cluster update algorithms are discussed in text books on classical Monte Carlo simulations and in computational physics text books. While most cluster algorithms require spin-inversion invariance and thus do not allow for external magnetic fields, extensions to spin models

in magnetic fields have been proposed [8,9]. An open source implementation of local and cluster updates for Ising, Potts, XY and Heisenberg models is available through the ALPS (Applications and Libraries for Physics Simulations) project [10] at the web page `http://alps.comp-phys.org/`.

The tunneling problem at first order phase transitions and for disordered systems, where tunneling times often diverge exponentially can be overcome using extended ensemble methods, which are the main topic of the next chapter.

3 Extended Ensemble Methods

3.1 First Order Phase Transitions and the Multicanonical Ensemble

While cluster updates can solve critical slowing down at second order phase transitions they are usually inefficient at first order phase transitions and in frustrated systems. Let us consider a first order phase transition, such as in a two-dimensional q-state Potts model with Hamilton function

$$H = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j} \ , \tag{18}$$

where the spins σ_i can now take the integer values $1, \ldots, q$. For $q > 4$ this model exhibits a first order phase transition, accompanied by exponential slowing down of conventional local update algorithms. The exponential slowdown is caused by the free energy barrier between the two coexisting metastable states at the first order phase transition.

This barrier can be quantified by considering the energy histogram

$$H_{\text{canonical}}(E) \propto g(E) P_{\text{Boltzmann}}(E) = g(E) \exp(-\beta E) \ , \tag{19}$$

which is the probability of encountering a configuration with energy E during the Monte Carlo simulation. Here

$$g(E) = \sum_c \delta_{E,E(c)} \tag{20}$$

is the density of states. Away from first order phase transitions, $H_{\text{canonical}}(E)$ has approximately Gaussian shape, centered around the mean energy. At first order phase transitions, where the energy jumps discontinuously the histogram $H_{\text{canonical}}(E)$ develops a double-peak structure. The minimum of $H_{\text{canonical}}(E)$ between these two peaks, which the simulation has to cross in order to go from one phase to the other, becomes exponentially small upon increasing the system size. This leads to exponentially large autocorrelation times.

This tunneling problem at first-order phase transitions can be relieved by extended ensemble techniques which aim at broadening the sampled energy space. Instead of weighting a configuration c with energy $E = E(c)$ using the Boltzmann weight $P_{\text{Boltzmann}}(E) = \exp(-\beta E)$ more general weights $P_{\text{extended}}(E)$ are introduced which define the extended ensemble. The configuration space is explored by generating a Markov chain of configurations

$$c_1 \to c_2 \to \ldots \to c_i \to c_{i+1} \to \ldots \,, \tag{21}$$

where a move from configuration c_1 to c_2 is accepted with probability

$$P_{\text{acc}}(c_1 \to c_2) = \min\left(1, \frac{P(c_2)}{P(c_1)}\right) = \min\left(1, \frac{W_{\text{extended}}(E_2)}{W_{\text{extended}}(E_1)}\right) \,. \tag{22}$$

In general, the extended weights are defined in a single coordinate, such as the energy, thereby projecting the random walk in configuration space to a random walk in energy space

$$E_1 = E(c_1) \to E_2 \to \ldots \to E_i \to E_{i+1} \to \ldots \,. \tag{23}$$

For this random walk in energy space a histogram can be recorded which has the characteristic form

$$H_{\text{extended}}(E) \propto g(E) W_{\text{extended}}(E) \,, \tag{24}$$

where the density of states $g(E)$ is fixed for the simulated system.

One choice of generalized weights is the multicanonical ensemble [11, 12] where the weight of a configuration c is defined as $W_{\text{multicanonical}}(c) \propto 1/g(E(c))$. The multicanonical ensemble then leads to a flat histogram in energy space

$$H_{\text{multicanonical}}(E) \propto g(E) W_{\text{multicanonical}}(E) = g(E)\frac{1}{g(E)} = \text{const.}, \tag{25}$$

removing the exponentially small minimum in the canonical distribution. After performing a simulation, measurements in the multicanonical ensemble are reweighted by a factor $W_{\text{Boltzmann}}(E)/W_{\text{multicanonical}}(E)$ to obtain averages in the canonical ensemble.

3.2 The Wang-Landau Algorithm

Since the density of states and thus the multicanonical weights are not known initially, a scalable algorithm to estimate these quantities is needed. The Wang-Landau algorithm [13, 14] is a simple but efficient iterative method to obtain good approximations of the density of states $g(E)$ and the multicanonical weights $W_{\text{multicanonical}}(E) \propto 1/g(E)$.

The algorithm starts with a (very bad) estimate of the density of states $g(E) = 1$ for all energies which is iteratively improved by a modification factor f in the following loop:

- Start with $g(E) = 1$ and a modification factor $f \approx \exp(1)$.
- Repeat
  - Reset a histogram of energies $H(E) = 0$.
  - Perform simulations until the histogram of energies $H(E)$ is "flat":
    - Pick a random site and propose a local update, e.g. by flipping the spin at the site, which changes the current configuration c to a new configuration c', and the energy from $E = E(c)$ to $E' = E(c')$.
    - Approximating multicanoncal weights with the current estimate of the density of states the update is accepted with probability $\min[1, g(E)/g(E')]$.
    - Increase the histogram at the current value of E: $H(E) \leftarrow H(E)+1$
    - Increase the estimate $g(E)$ at the current value of E: $g(E) \leftarrow f g(E)$.
  - Once $H(E)$ is "flat" and has "sufficient statistics", reduce $f \leftarrow \sqrt{f}$.
- Stop once f is sufficiently small, e.g. $f \approx \exp(10^{-6})$.

Only a few lines of code need to be changed in the local update algorithm for the Ising model, but a few remarks are helpful:

1. The initial value for f needs to be carefully chosen, $f = \exp(1)$ is only a rough guide. A good choice is picking the initial f such that $f^{N_{\text{sweeps}}}$ is approximately the total number of states $\sum_E g(E)$ (e.g. 2^N for an Ising model with N sites).
2. Checking for flatness of the histogram (e.g. the minimum is at least 80% of the mean) should be done only after a reasonable number of sweeps N_{sweeps}. One sweep is defined as one attempted update per site.
3. The flatness criterion is quite arbitrary. In order to ensure convergence of the estimated $g(E)$ it should be extended to enforce sufficient statistics, e.g. by requiring that each histogram entry is at least of the order of $1/\sqrt{\ln f}$ as pointed out in Refs. [15, 16].
4. The density of states $g(E)$ can become very large and easily exceed 10^{10000}. In order to obtain such large numbers the *multiplicative increase* $g(E) \leftarrow f g(E)$ is essential. A naive additive guess $g(E) \leftarrow g(E) + f$ would never be able to reach the large numbers needed.
5. Since $g(E)$ is so large, we only store its *logarithm*. The update step is thus $\ln g(E) \leftarrow \ln g(E) + \ln f$.

At the end, the density of states $g(E)$ needs to be normalized. Either a known ground state degeneracy (e.g. $g(E_{\text{GS}}) = 2$ in the Ising ferromagnet) or a known total number of states (e.g. $\sum_E g(E) = 2^N$ in the Ising model with N spins) or a combination of the two (e.g. $g(E_{\text{GS}}) \cdot \sum_E g(E) = 2^{N+1}$ for the Ising ferromagnet) can be used to normalize $g(E)$.

Besides overcoming the exponentially suppressed tunneling problem at first order phase transitions, the Wang-Landau algorithm calculates the generalized density of states $g(E)$ in an iterative procedure. The knowledge of the density of states $g(E)$ then allows the direct calculation of the free energy from the partition function (16). The internal energy, entropy, specific heat

and other thermal properties are easily obtained as well, by differentiating the free energy. By additionally measuring the averages $A(E)$ of other observables A as a function of the energy E, thermal expectation values can be obtained at arbitrary inverse temperatures β by performing just a single simulation:

$$\langle A(\beta)\rangle = \frac{\sum_E A(E)g(E)e^{-\beta E}}{\sum_E g(E)e^{-\beta E}} \; . \tag{26}$$

3.3 Markov Chains and Random Walks in Energy Space

The multicanonical ensemble and Wang-Landau algorithm both project a random walk in high-dimensional configuration space onto a one-dimensional random walk in energy space where all energy levels are sampled equally often. It is important to note that the random walk in configuration space, (21), is a biased Markovian random walk, while the projected random walk in energy space, (23), is non-Markovian, as memory is stored in the configuration. This becomes evident as the system approaches a phase transition in the random walk: While the energy no longer reflects from which side the phase transition is approached, the current configuration may still reflect the actual phase the system has visited most recently. In the case of the two-dimensional ferromagnetic Ising model, the order parameter for a given configuration at the critical energy $E_c \sim -1.41N$ will reveal whether the system is approaching the transition from the magnetically ordered (lower energies) or disordered side (higher energies).

This loss of information in the projection of the random walk in configuration space has important consequences for the random walk in energy space. Most strikingly, the local diffusivity of a random walker in energy space, which for a diffusion time t_D can be defined as

$$D(E, t_D) = \langle (E(t) - E(t + t_D))^2 \rangle / t_D \tag{27}$$

is *not* independent of the location in energy space. This is illustrated in Fig. 1 for the Ising ferromagnet. Below the phase transition around $E \sim -1.41\,N$ a clear minimum evolves in the local diffusivity. In this region large ordered domains are formed and by moving the domain boundaries through local spin flips only small energy changes are induced resulting in a suppressed local diffusivity in energy space.

Because of the strong energy dependence of the local diffusivity the simulation of a multicanonical ensemble sampling all energy levels equally often turns out to be suboptimal [17]. The performance of flat-histogram algorithms can be quantified for classical spin models such as the ferromagnet where the number of energy levels is given by $[-2N, +2N]$ and thereby scales with the number of spins N in the system. When measuring the typical round-trip time between the two extremal energies for multicanonical simulations, these round-trip times τ are found to scale like

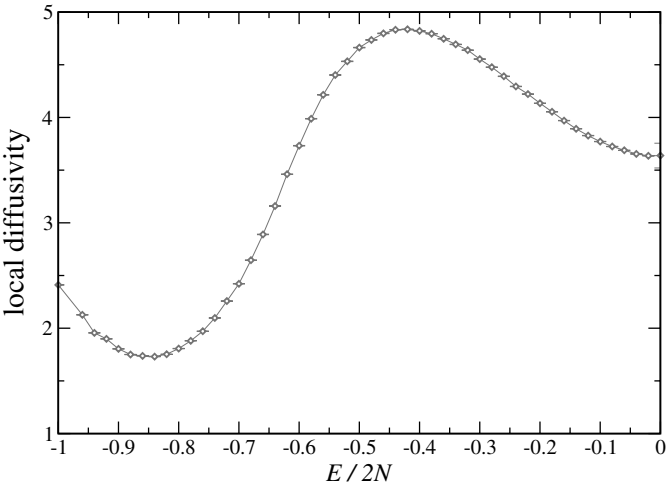


Fig. 1. Local diffusivity $D(E, t_D) = \langle (E(t) - E(t + t_D))^2 \rangle / t_D$ of a random walk sampling a flat histogram in energy space for the two-dimensional ferromagnetic Ising model. The local diffusivity strongly depends on the energy with a strong suppression below the critical energy $E_c \approx -1.41\, N$

$$\tau \sim N^2 L^z \ , \tag{28}$$

showing a power-law deviation from the N^2-scaling behavior of a completely unbiased random walk. Here z is a critical exponent describing the slowdown of a multicanonical simulation in the proximity of a phase transition [17,18]. The value of z strongly depends on the simulated model and the dimensionality of the problem. In two dimensions the exponent increases from $z = 0.74$ for the ferromagnet as one introduces competing interactions leading to frustration and disorder. The exponent becomes $z = 1.73$ for the fully frustrated Ising model which is defined by a Hamilton

$$H = \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j \ , \tag{29}$$

where the spins around any given plaquette of four spins are frustrated, e.g. by choosing the couplings along three bonds to be $J_{ij} = -1$ (ferromagnetic) and $J_{ij} = +1$ (antiferromagnetic) for the remaining bond. For the spin glass where the couplings J_{ij} are randomly chosen to be $+1$ or -1 exponential scaling ($z = \infty$) is found [17, 19]. Increasing the spatial dimension for the ferromagnet the exponent is found to decrease as $z \approx 1.81, 0.74$ and 0.44 for dimension $d = 1, 2$ and 3 and z vanishes for the mean-field model in the limit of infinite dimensions [18].

3.4 Optimized Ensembles

The observed polynomial slowdown for the multicanonical ensemble poses the question whether for a given model there is an optimal choice of sampling energies, $H_{\text{optimal}}(E)$ and corresponding weights $W_{\text{optimal}}(E)$, which eliminates the slowdown. To address this question an adaptive feedback algorithm has recently been introduced that iteratively improves the weights in an extended ensemble simulationt leading to further improvements in the efficiency of the algorithm by several orders of magnitude [20]. The scaling for the optimized ensemble is found to scale like $O([N \ln N]^2)$ thereby reproducing the behavior of an unbiased Markovian random walk up to a logarithmic correction.

At the heart of the algorithm is the idea to maximize a current j of walkers that move from the lowest energy level, E_-, to the highest energy level, E_+, or vice versa, in an extended ensemble simulation by varying the weights $W_{\text{extended}}(E)$. To measure the current a label is added to the walker that indicates which of the two extremal energies the walker has visited most recently. The two extrema act as "reflecting" and "absorbing" boundaries for the labeled walker: e.g., if the label is plus, a visit to E_+ does not change the label, so this is a "reflecting" boundary. However, a visit to E_- does change the label, so the plus walker is absorbed at that boundary. The behavior of the labeled walker is *not* affected by its label except when it visits one of the extrema and the label changes.

For the random walk in energy space, two histograms are recorded, $H_+(E)$ and $H_-(E)$, which for sufficiently long simulations converge to steady-state distributions which satisfy $H_+(E) + H_-(E) = H(E) = W(E)g(E)$. For each energy level the fraction of random walkers which have label "plus" is then given by $f(E) = H_+(E)/H(E)$. The above-discussed boundary conditions dictate $f(E_-) = 0$ and $f(E_+) = 1$.

The steady-state current to first order in the derivative is

$$j = D(E)H(E)\frac{df}{dE} \ , \tag{30}$$

where $D(E)$ is the walker's diffusivity at energy E. There is no current if $f(E)$ is constant, since this is equilibrium. Therefore the current is to leading order proportional to df/dE. Rearranging the above equation and integrating on both sides, noting that j is a constant and f runs from 0 to 1, one obtains

$$\frac{1}{j} = \int_{E_-}^{E_+} \frac{dE}{D(E)H(E)} \ . \tag{31}$$

To maximize the current and thus the round-trip rate, this integral must be minimized. However, there is a constraint: $H(E)$ is a probability distribution and must remain normalized which can be enforced with a Lagrange multiplier:

$$\int_{E_-}^{E_+} dE \left(\frac{1}{D(E)H(E)} + \lambda H(E) \right) \ . \tag{32}$$

To minimize this integrand, the ensemble, that is the weights $W(E)$ and thus the histogram $H(E)$ are varied. At this point it is assumed that the dependence of $D(E)$ on the weights can be neglected.

The optimal histogram, $H_{\text{optimal}}(E)$, which minimizes the above integrand and thereby maximizes the current j is then found to be

$$H_{\text{optimal}}(E) \propto \frac{1}{\sqrt{D(E)}} \,. \tag{33}$$

Thus for the optimal ensemble, the probability distribution of sampled energy levels is simply inversely proportional to the square root of the local diffusivity.

The optimal histogram can be approximated in a feedback loop of the form

- Start with some trial weights $W(E)$, e.g. $W(E) = 1/g(E)$.
- Repeat
  - Reset the histograms $H(E) = H_+(E) = H_-(E) = 0$.
  - Simulate the system with the current weights for N sweeps:
    - Updates are accepted with probablity $\min[1, W(E')/W(E)]$.
    - Record the histograms $H_+(E)$ and $H_-(E)$.
  - From the recorded histogram an estimate of the local diffusivity is obtained as

$$D(E) \propto \frac{1}{H(E)\frac{df}{dE}} \,, \quad f(E) = \frac{H_+(E)}{H(E)} \,, \qquad H(E) = H_+(E) + H_-(E) \,.$$

  - Define new weights as

$$W_{\text{optimized}}(E) = W(E)\sqrt{\frac{1}{H(E)} \cdot \frac{df}{dE}} \,.$$

  - Increase the number of sweeps for the next iteration $N \leftarrow 2N$.
- Stop once the histogram $H(E)$ has converged.

Again the implementation of this feedback algorithm requires to change only a few lines of code in the original local update algorithm for the Ising model. Some additional remarks are useful:

1. In contrast to the Wang-Landau algorithm the weights $W(E)$ are modified only after a batch of N sweeps, thereby ensuring detailed balance between successive moves at all times.
2. The initial value of sweeps N should be chosen large enough that a couple of round trips are recorded, thereby ensuring that steady state data for $H_+(E)$ and $H_-(E)$ are measured.
3. The derivative df/dE can be determined by a linear regression, where the number of regression points is flexible. Initial batches with the limited statistics of only a few round trips may require a larger number of regression points than subsequent batches with smaller round-trip times and better statistics.

4. Similar to the multicanonical ensemble the weights $W(E)$ can become very large, and storing the logarithms may be advantageous. The reweighting then becomes $\ln W_{\text{optimized}}(E) = \ln W(E) + (\ln \frac{df}{dE} - \ln H(E))/2$.

At the end of the simulation, the density of states can be estimated from the recorded histogram as $g(E) = H_{\text{optimized}}(E)/W_{\text{optimized}}(E)$ and normalized as described above.

Figure 2 shows the optimized histogram for the two-dimensional ferromagnetic Ising model. The optimized histogram is no longer flat, but a peak evolves at the critical region around $E_c \approx -1.41\ N$ of the transition. The feedback of the local diffusivity reallocates resources towards the bottlenecks of the simulation which have been identified by a suppressed local diffusivity.

The scaling of round-trip times is shown in Fig. 3 for the two-dimensional fully frustrated Ising model. The power-law slowdown of round-trip times for the flat-histogram ensemble $O(N^2 L^{1.73})$ is reduced to a logarithmic correction $O([N \ln N]^2)$ for the optimized ensemble in comparison to a completely unbiased random walk with $O(N^2)$-scaling. This scaling improvement results in a speedup by a nearly two orders of magnitude already for a system with some 128×128 spins.

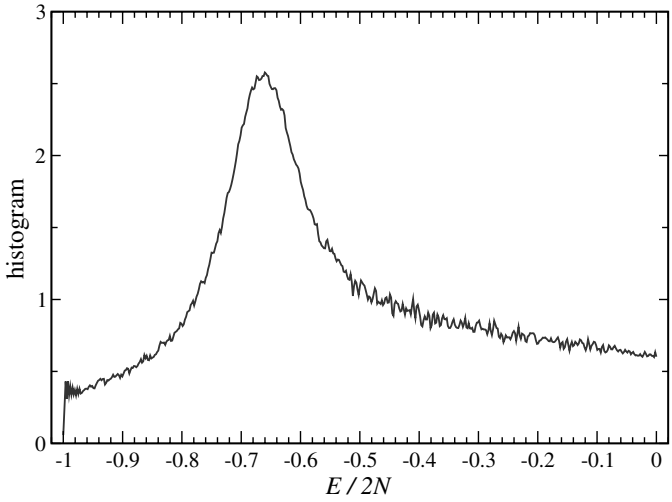


Fig. 2. Optimized histograms for the two-dimensional ferromagnetic Ising model. After the feedback of the local diffusivity a peak evolves near the critical energy of the transition $E_c \approx -1.41\ N$. The feedback thereby shifts additional resources towards the bottleneck of the simulation which were identified by a suppressed local diffusivity

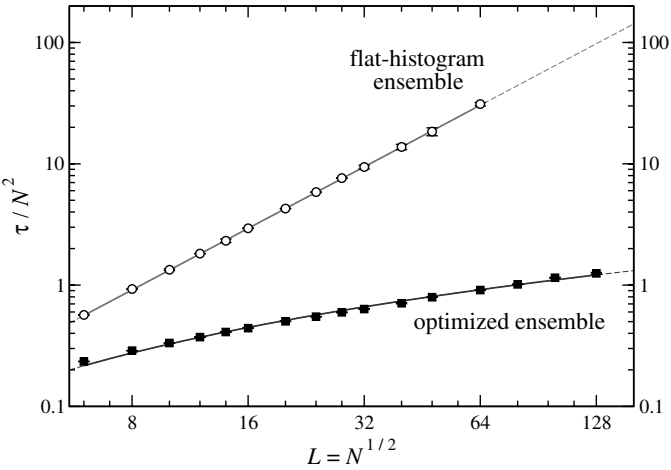


Fig. 3. Scaling of round-trip times for a random walk in energy space sampling a flat histogram (*open squares*) and the optimized histogram (*solid circles*) for the two-dimensional fully frustrated Ising model. While for the multicanonical simulation a power-law slowdown of the round-trip times $O(N^2 L^z)$ is observed, the round-trip times for the optimized ensemble scale like $O([N \ln N]^2)$ thereby approaching the ideal $O(N^2)$-scaling of an unbiased Markovian random walk up to a logarithmic correction

3.5 Simulation of Dense Fluids

Extended ensembles cannot only be defined as a function of energy, but in arbitrary reaction coordinates $\boldsymbol{R}$ onto which a random walk in configuration space can be projected. The generalized weights in these reaction coordinates $W_{\text{extended}}(\boldsymbol{R})$ are then used to bias the random walk along the reaction coordinate by accepting moves from a configuration c_1 with reaction coordinate $\boldsymbol{R}_1$ to a configuration c_2 with reaction coordinate $\boldsymbol{R}_2$ with probability

$$p_{\text{acc}}(c_1 \to c_2) = p_{\text{acc}}(\boldsymbol{R}_1 \to \boldsymbol{R}_2) = \min\left(1, \frac{W_{\text{extended}}(\boldsymbol{R}_2)}{W_{\text{extended}}(\boldsymbol{R}_1)}\right) . \tag{34}$$

The generalized weights $W_{\text{extended}}(\boldsymbol{R})$ can again be chosen in such a way that similar to a multicanonical simulation a flat histogram is sampled along the reaction coordinate by setting the weights to be inversely proportional to the density of states defined in the reaction coordinates, that is $W_{\text{extended}}(\boldsymbol{R}) \propto 1/g(\boldsymbol{R})$.

An optimal choice of weights can be found by measuring the local diffusivity of a random walk along the reaction coordinates and applying the feedback method to shift weight towards the bottlenecks in the simulation. This generalized ensemble optimization approach has recently been illustrated for the simulation of dense Lennard-Jones fluids close to the vapor-liquid equilibrium [21]. The interaction between particles in the fluid is described by a

pairwise Lennard-Jones potential of the form

$$\Phi_{\mathrm{LJ}}(R) = 4\epsilon \left[\left(\frac{\sigma}{R}\right)^{12} - \left(\frac{\sigma}{R}\right)^{6} \right] , \tag{35}$$

where ϵ is the interaction strength, σ a length parameter, and R the distance between two particles. It is this distance R between two arbitrarily chosen particles in the fluid that one can use as a new reaction coordinate for a projected random walk. Defining an extended ensemble with weights $W_{\mathrm{extended}}(R)$ and recording a histogram $H(R)$ during a simulation will then allow to calculate the pair distribution function $g(R) = H(R)/W_{\mathrm{extended}}(R)$. The pair distribution function $g(R)$ is closely related to the potential of mean force (PMF)

$$\Phi_{\mathrm{PMF}}(R) = -\frac{1}{\beta} \ln g(R) , \tag{36}$$

which describes the average interaction between two particles in the fluid in the presence of many surrounding particles.

For high particle densities and low enough temperatures shell structures will form in the fluid which are reminiscent of the hexagonal lattice of the solid structure at very low temperatures. These shell structures are revealed by a sinusoidal modulation in the PMF as illustrated in the lower panel of Fig. 4 for the case of a two-dimensional fluid. Thermal equilibration between the shells is suppressed by entropic barriers which form between the shells. Again, one can ask what probability distribution, or histogram, should be sampled along the reaction coordinate, in this case the radial distance R, so that equilibration between the shells is improved. Measuring the local diffusivity for a random walk along the radial distance R in an interval $[R_{\mathrm{min}}, R_{\mathrm{max}}]$ and subsequently applying the feedback algorithm described above optimized histograms $H(R)$ are found which are plotted in Fig. 4 for varying temperatures [21]. The feedback algorithm again shifts additional weight in the histogram towards the bottleneck of the simulation, in this case towards the barriers between the shells. Interestingly, additional peaks emerge in the optimized histogram as the temperature is lowered towards the vapor-liquid equilibrium. The minima between these peaks points to additional meta-stable configurations which occur at these low temperatures, namely interstitial states which occur as the shells around two particles merge as detailed in [21].

This example illustrates that for some simulations the local diffusivity and optimized histogram *itself* are very sensitive measures that can reveal interesting phenomena which are otherwise hard to detect in a numerical simulation. In general, a strong modulation of the local diffusivity for the random walk along a given reaction coordinate is a good indicator that the reaction coordinate itself is a good choice that captures some interesting physics of the problem.

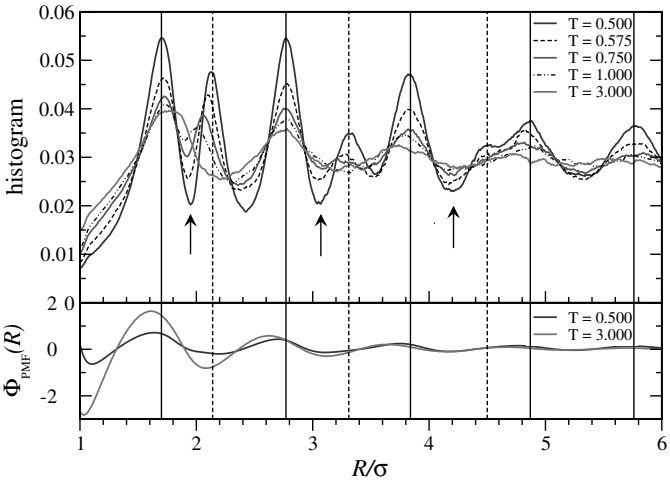


Fig. 4. Optimized histograms for the simulation of dense Lennard-Jones fluids

3.6 Parallel Tempering

The simulation of frustrated systems suffers from a similar tunneling problem as the simulation of first order phase transitions: local minima in energy space are separated by barriers that grow with system size. While the multicanonical or optimized ensembles do not help with the NP-hard problems faced by spin glasses, they are efficient in speeding up simulations of frustrated magnets without disorder.

An alternative to these extended ensembles for the simulation of frustrated magnets is the "parallel tempering" or "replica exchange" Monte Carlo method [22–25]. Instead of performing a single simulation at a fixed temperature, simulations are performed for M replicas at a set of temperatures $T_1, T_2, \ldots, T_M$. In addition to standard Monte Carlo updates at a fixed temperature, exchange moves are proposed to swap two replicas between adjacent temperatures. These swaps are accepted with a probability

$$\min[1, \exp(\Delta\beta\Delta E)], \tag{37}$$

where $\Delta\beta$ is the difference in inverse temperatures and ΔE the difference in energy between the two replicas.

The effect of these exchange moves is that a replica can drift from a local free energy minimum at low temperatures to higher temperatures, where it is easier to cross energy barriers and equilibration is fast. Upon cooling (by another sequence of exchanges) it can end up in a different local minimum on time scales that are much shorter compared to a single simulation at a fixed low temperature. This random walk of a single replica in temperature space is the analog of the random walk in energy space discussed for the extended

ensemble techniques. The complement of the statistical ensemble, defined by the weights $W_{\text{extended}}(E)$, is the particular choice of temperature points in the temperature set $\{T_1, T_2, \ldots, T_M\}$ for the parallel tempering simulation. The probability of sampling any given temperature T in an interval $T_i < T < T_{i+1}$ can then be approximated by $H(T) \propto 1/\Delta T$, where $\Delta T = T_{i+1} - T_i$ is the length of the temperature interval around the temperature T. This probability distribution $H(T)$ is the equivalent to the histogram $H(E)$ in the extended ensemble simulations. The ensemble optimization technique introduced above can thus be reformulated to optimize the temperature set in a parallel tempering simulation in such a way that the rate of round-trips between the two extremal temperatures, T_1 and T_M respectively, is maximized [26, 27].

Starting with an initial temperature set $\{T_1, T_2, \ldots, T_M\}$ a parallel tempering simulation is performed where each replica is labeled either "plus" or "minus" indicating which of the two extremal temperatures the respective replica has visited most recently. This allows to measure a current of replicas diffusing from the highest to the lowest temperature by recording two histograms, $h_+(T)$ and $h_-(T)$ for each temperature point. The current j is then given by

$$j = D(T)H(T)\frac{df}{dT} \,, \tag{38}$$

where $D(T)$ is the local diffusivity for the random walk in temperature space, and $f(T) = h_+(T)/(h_+(T) + h_-(T))$ is the fraction of random walkers which have visited the highest temperature T_M most recently. The probability distribution $H(T)$ is normalized, that is

$$\int_{T_1}^{T_M} H(T)\; dT = C \int_{T_1}^{T_M} \frac{dT}{\Delta T} = 1 \,, \tag{39}$$

where C is a normalization constant. Rearranging (38) the local diffusivity $D(T)$ of the random walk in temperature space can be estimated as

$$D(T) \propto \frac{\Delta T}{df/dT} \,. \tag{40}$$

Analog to the argument for the extended ensemble in energy space the current j is maximized by choosing a probability distribution

$$H_{\text{optimal}}(T) \propto \frac{1}{\sqrt{D(T)}} \propto \sqrt{\frac{1}{\Delta T}\,\frac{df}{dT}} \tag{41}$$

which is inversely proportional to the square root of the local diffusivity. The optimized temperature set $\{T'_1, T'_2, \ldots, T'_M\}$ is then found by choosing the n-th temperature point T'_n such that

$$\int_{T'_1}^{T'_n} H_{\text{optimal}}(T)\; dT = \frac{n}{M} \,, \tag{42}$$

where M is the number of temperature points in the original temperature set, and the two extremal temperatures $T_1' = T_1$ and $T_M' = T_M$ remain unchanged. Similarly to the algorithm for the ensemble optimization this feedback of the local diffusivity should be iterated until the temperature set is converged.

Figure 5 illustrates the optimized temperature sets for the Ising ferromagnet obtained by several iterations of the above feedback loop. After feedback of the local diffusivity temperature points accumulate near the critical temperature $T_c = 2.269$ of the transition. This is in analogy to the optimized histograms for the extended ensemble simulations where resources where shifted towards the critical energy of the transition, see Fig. 2.

It is interesting to note that for the optimized temperature set the acceptance rates for swap moves are not independent of the temperature. Around the critical temperature, where temperature points are accumulated by the feedback algorithm, the acceptance rates are higher as at higher/lower temperatures, where the density of temperature points becomes considerably smaller after feedback. The almost Markovian scaling behavior for the optimized random walks in either energy or temperature space is thus generated by a problem-specific statistical ensemble which is characterized neither by a flat histogram nor flat acceptance rates for exchange moves, but by a characteristic

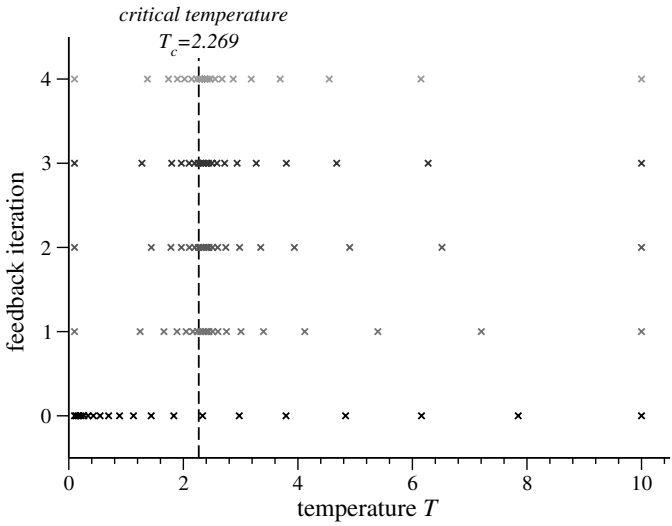


Fig. 5. Optimized temperature sets for the two-dimensional Ising ferromagnet. The initial temperature set with 20 temperature points is determined by a geometric progression for the temperature interval $[0.1, 10]$. After feedback of the local diffusivity the temperature points accumulate near the critical temperature $T_c = 2.269$ of the phase transition (*dashed line*). Similar to the ensemble optimization in energy space the feedback of the local diffusivity relocates resources towards the bottleneck of the simulation

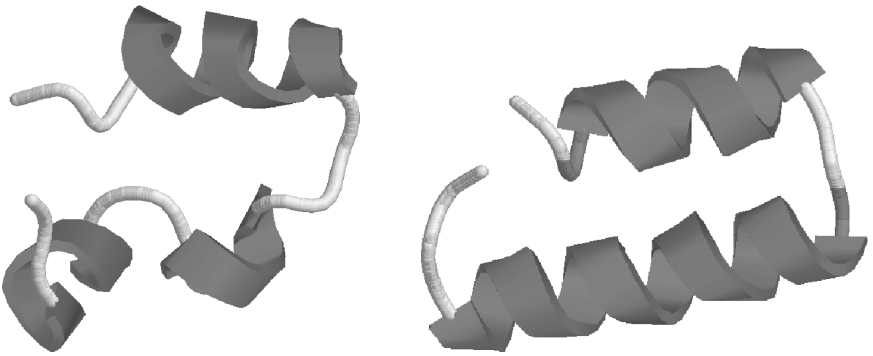


Fig. 6. Low-energy structure of the 36-residue chicken villin headpiece sub-domain HP-36. On the left the structure determined in NMR experiments is shown. The right panel shows the lowest-energy configuration found in a feedback-optimized all-atom parallel tempering simulation using the ECEPP/2 force field and an implicit solvent model. The root-mean square deviation of this structure to the structure on the left is $R_{\mathrm{RMSD}} = 3.8$ Å

probability distribution which concentrates resources at the minima of the measured local diffusivity.

3.7 Optimized Parallel Tempering Simulations of Proteins

The feedback-optimized parallel tempering technique [26] outlined in the previous section has recently been applied to study the folding of the 36-residue chicken villin headpiece sub-domain HP-36 [27]. Since HP-36 is one of the smallest proteins with well-defined secondary and tertiary structure [28] and at the same time with 596 atoms still accessible to numerical simulations, it has recently attracted considerable interest as an example to test novel numerical techniques, including molecular dynamics [29,30] and Monte Carlo [31,32] methods. The experimentally determined structure [28] which is deposited in the Protein Data Bank (PDB code 1vii) is illustrated in the left panel of Fig. 6.

Applying an all-atom parallel tempering simulation of the protein HP-36 in the ECEPP/2 force field [33] using an implicit solvent model [34] the authors of [27] have measured the diffusion of labeled replicas in temperature space. The simulated temperature interval is chosen such that at the lowest temperature $T_{\mathrm{min}} = 250$ K the protein is in a folded state and the highest temperature $T_{\mathrm{max}} = 1000$ K ensures that the protein can fully unfold for the simulated force field. The measured local diffusivity for the random walk between these two extremal temperatures is shown in Fig. 7. A very strong modulation of the local diffusivity is found along the temperature. Note the logarithmic scale of the ordinate. The pronounced minimum of the local diffusivity around $T \approx 500$ K points to a severe bottleneck in the simulation which by measurements of the specific heat has been identified as the helix-coil

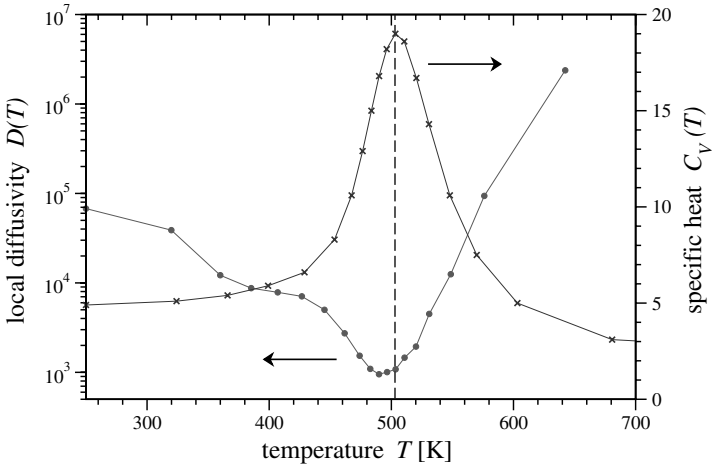


Fig. 7. Local diffusivity (*solid circles*) of the random walk in temperature space for a parallel tempering simulation of the 36-residue villin headpiece sub-domain HP-36. The diffusivity shows a strong modulation along the temperature, note the logarithmic scale of the ordinate. Slightly below the helix-coil transition around $T \approx 500$ K which is identified by a maximum in the specific heat (crosses, right ordinate) there is a strong suppression of the diffusivity

transition [31]. Above this transition the protein is in an extended unordered configuration, while below the helix-coil transition the protein is characterized by high helical content [31]. The shoulder in the local diffusivity in the temperature region 350 K $\leq T \leq$ 490 K points to a second bottleneck in the simulation, possibly caused by competing low-energy configurations with high helical content.

An optimized temperature set for the parallel tempering simulation of HP-36 in the ECEPP/2 force field can then be found by feeding back the local diffusivity applying the algorithm outlined above. Results for a temperature set with 20 temperature points are illustrated in Fig. 8 for an initial temperature set which similar to a geometric progression concentrates temperature points at low temperatures [27]. After the feedback temperature points concentrate around the bottleneck of the simulation, primarily around the helix-coil transition at $T \approx 500$ K and in the temperature regime 350 K $\leq T \leq$ 490 K below the transition where a shoulder in the local diffusivity was found.

In [27] it was demonstrated that by using the optimized temperature set in the simulations the low-energy configurations equilibrated considerably faster than in previous parallel tempering simulations [31]. As a consequence, the low energy structures are more compact and the configuration with lowest energy illustrated in Fig. 6 shows a root-mean square deviation to the experimentally determined structure of $R_{\mathrm{RMSD}} = 3.7$ Å. This deviation from the native structure is similar to results found by large-scale molecular dynamics

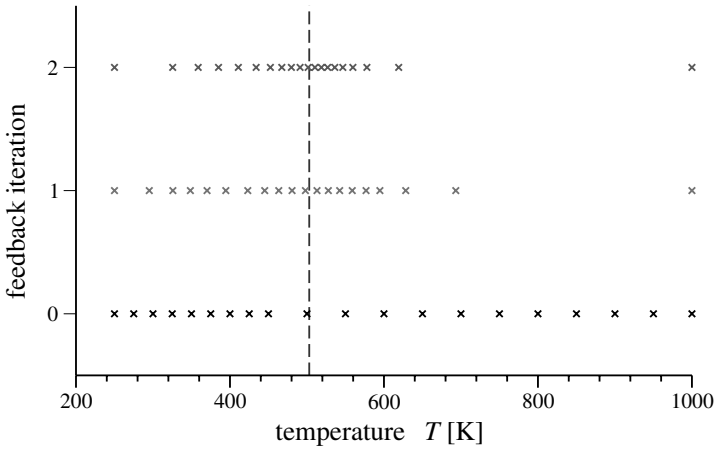


Fig. 8. Optimized temperature sets with 20 temperature points for the parallel tempering simulation of the 36-residue protein HP-36. The initial temperature set covers a temperature range 250 K $\leq T \leq$ 1000 K and concentrates temperature points at low temperatures similar to a geometric progression. After the feedback of the local diffusivity temperature points accumulate around the helix-coil transition at $T \approx 500$ K where the strong suppression of the local diffusivity points to a severe bottleneck

simulations [30] with a different force field. However, employing the optimized temperature set the Monte Carlo simulations consumed only a fraction of one percent of the computing time used for the molecular dynamics simulations. A detailed discussion of the application of the feedback method in the study of proteins is given in [27].

3.8 Simulation of Quantum Systems

Extended ensemble methods, such as the multicanonical ensemble, Wang-Landau sampling or parallel tempering can also be generalized to quantum systems [35, 36], as we will show in the next two sections.

4 Quantum Monte Carlo World Line Algorithms

4.1 The $S = 1/2$ Quantum XXZ Model

In this section we will generalize the Monte Carlo methods described in section 2 for classical spin systems to quantum spin systems. As an example we will use the spin-1/2 quantum Heisenberg or XXZ models with Hamiltonian

$$H = \sum_{\langle i,j \rangle} \left[J_z S_i^z S_j^z + J_{xy} \left(S_i^x S_j^x + S_i^y S_j^y \right) \right] - h \sum_{i=1}^{N} S_i^z \tag{43}$$

$$= \sum_{\langle i,j \rangle} \left[J_z S_i^z S_j^z + \frac{J_{xy}}{2} \left(S_i^+ S_j^- + S_i^- S_j^+ \right) \right] - h \sum_{i=1}^{N} S_i^z$$

where S_i^α are spin $S = 1/2$ operators fulfilling the standard commutation relations and in the second line we have replaced S_i^x and S_i^y by the spin raising and lowering operators S_i^+ and S_i^-.

The case $J_{xy} = 0$ corresponds to the classical Ising model (12) up to a change in sign: while in classical Monte Carlo simulations (where there is no difference in the thermodynamics of the ferromagnet and the antiferromagnet) a positive exchange constant J denotes the ferromagnet, the convention for quantum systems is usually opposite with a positive exchange constant denoting the antiferromagnet. The other limit $J_z = 0$ corresponds to the quantum XY-model, while $J_z = J_{xy}$ is the Heisenberg model.

4.2 Representations

The basic problem for Monte Carlo simulations of quantum systems is that the partition function is no longer a simple sum over classical configurations as in (16) but an operator expression

$$Z = \mathrm{Tr} \exp(-\beta H) \ , \tag{44}$$

where H is the Hamilton operator and the trace Tr goes over all states in the Hilbert space. Similarly the expression for an observable like the magnetization is an operator expression:

$$\langle m \rangle = \frac{1}{Z} \mathrm{Tr} \left[m \exp(-\beta H) \right] \ , \tag{45}$$

and the Monte Carlo method cannot directly be applied except in the classical case where the Hamilton operator H is diagonal and the trace reduces to a sum over all basis states. The first step of any QMC algorithm is thus the mapping of the quantum system to an equivalent classical system

$$\langle m \rangle = \frac{1}{Z} \mathrm{Tr} \left[m \exp(-\beta H) \right] = \sum_c m(c) W(C) \ , \tag{46}$$

where the sum goes over configurations c in an artificial classical system (e.g. a system of world lines), $m(c)$ will be the value of the magnetization or another observable as measured in this classical system and $W(C)$ the weight of the classical configuration. We will now present two different but related methods for this mapping, namely continuous time path integrals and the stochastic series expansion.

The Path-Integral Representation

The path-integral formulation of a quantum systems goes back to [37], and forms the basis of most QMC algorithms. Instead of following the historical route and discussing the Trotter-Suzuki (checkerboard) decomposition [38,39] for path integrals with discrete time steps $\Delta\tau$ we will directly describe the continuous-time formulation used in modern codes.

The starting point is a time-dependent perturbation expansion in imaginary time to evaluate the density matrix operator $\exp(-\beta H)$. Using a basis in which the S^z operators are diagonal we follow [40] and split the Hamiltonian $H = H_0 + V$ into a diagonal term H_0, containing the S^z term and an offdiagonal perturbation V, containing the exchange terms $(J_{xy}/2)(S_i^+ S_j^- + S_i^- S_j^+)$. In the interaction representation the time-dependent perturbation is $V(\tau) = \exp(\tau H_0) V \exp(-\tau H_0)$ and the partition function can be represented as:

$$\begin{aligned} Z &= \mathrm{Tr}\exp(-\beta H) = \mathrm{Tr}\left[\exp(-\beta H_0)\mathcal{T}\exp\int_0^\beta d\tau V(\tau)\right], \\ &= \mathrm{Tr}\left[\exp(-\beta H_0)\left(1 - \int_0^\beta d\tau_1 V(\tau_1) + \frac{1}{2}\int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 V(\tau_1)V(\tau_2) + \ldots\right)\right] \\ &= \sum_i \langle i|\left[\exp(-\beta H_0)\left(1 - \int_0^\beta d\tau_1 V(\tau_1) + \frac{1}{2}\int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 V(\tau_1)V(\tau_2) + \ldots\right)\right]|i\rangle, \end{aligned} \tag{47}$$

where the symbol $\mathcal{T}$ denotes time-ordering of the exponential and in the last line we have replaced the trace by a sum over a complete set of basis states $|i\rangle$, that are eigenstates of the local S^z operators. Note that, in contrast to a real time path integral, the imaginary time path integral always converges on finite systems of N spins at finite temperatures β, and the expansion can be truncated at orders $n \gg \beta J_{xy} N$.

Equation (47) is now just a classical sum of integrals and can be evaluated by Monte Carlo sampling in the intitial states $|i\rangle$, the order of the perturbation n and the times τ_i $(i = 1, \ldots, n)$. This is best done by considering a graphical world line representation of the partition function (47) shown in Fig. 9. The zero-th order terms in the sum $\sum_i \langle i| \exp(-\beta H_0)|i\rangle$ are given by straight world lines shown in Fig. 9a. First order terms do not appear since the matrix elements $\langle i|V|i\rangle$ are zero for the XXZ model. The first non-trivial terms appear in second order with two exchanges, as shown in Fig. 9b. A general configuration of higher order is depicted in Fig. 10a.

Since the XXZ Hamiltonian commutes with the z-component of total spin $\sum_i S_i^z$, the total magnetization is conserved and all valid configurations are represented by closed world lines as shown in Figs. 9 and 10. Models that break this conservation of magnetization, such as general XYZ models with different couplings in all directions, models with transverse fields coupling to S_i^x or higher spin models with single ion anisotropies $(S_i^x)^2$ or $(S_i^y)^2$ will in addition contain configurations with broken world line segments.

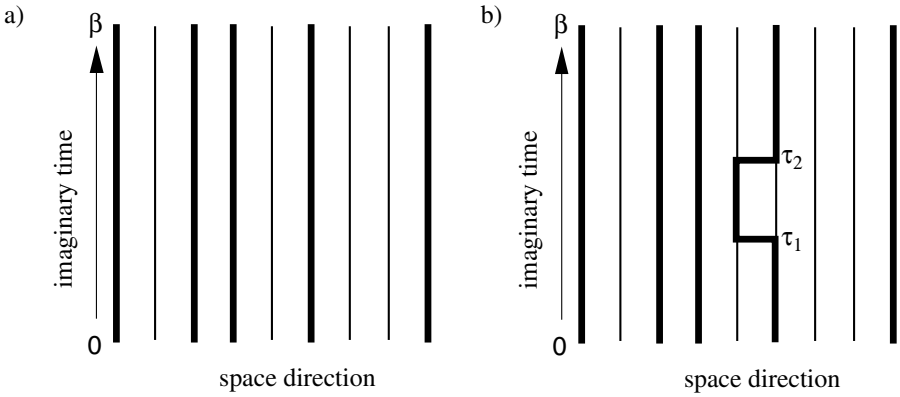


Fig. 9. Examples of simple world line configurations in imaginary time for a quantum spin model. Up-spins are shown by bold lines and down spins by thin lines. (**a**) a configuration in 0-th order perturbation theory where the spins evolve according to the diagonal term $\exp(-\beta H_0)$ and the weight is given by the classical Boltzmann weight of H_0. (**b**) a configuration in second order perturbation theory with two exchanges at times τ_1 and τ_2. Its weight is given by the matrix elements of the exchange processes and the classical Boltzmann weight of H_0 of the spins

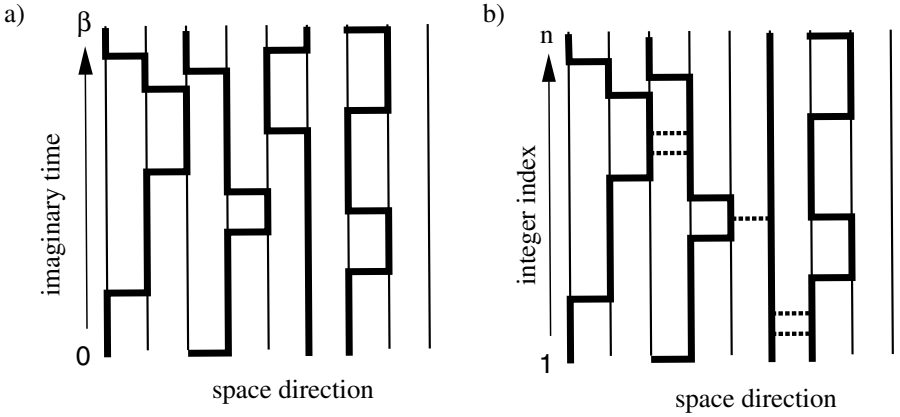


Fig. 10. Examples of world line configurations in (**a**) a path-integral representation where the time direction is continuous and (**b**) the stochastic series expansion (SSE) representation where the "time" direction is discrete. Since the SSE representation perturbs not only in offdiagonal terms but also in diagonal terms, additional diagonal terms are present in the representation, indicated by *dashed lines*

The Stochastic Series Expansion Representation

An alternative representation is the stochastic series expansion (SSE) [41], a generalization of Handscomb's algorithm [42] for the Heisenberg model. It starts from a Taylor expansion of the partition function in orders of β:

$$Z = \mathrm{Tr}\exp(-\beta H) = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \mathrm{Tr}(-H)^n \tag{48}$$

$$= \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{\{i_1,\ldots i_n\}} \sum_{\{b_1,\ldots b_n\}} \langle i_1| - H_{b_1}|i_2\rangle\langle i_2| - H_{b_2}|i_3\rangle \cdots \langle i_n| - H_{b_n}|i_1\rangle \ ,$$

where in the second line we decomposed the Hamiltonian H into a sum of single-bond terms $H = \sum_b H_b$, and again inserted complete sets of basis states. We end up with a similar representation as (47) and a related world-line picture shown in Fig. 10b.

The key difference is that the SSE representation is a perturbation expansion in *all* terms of the Hamiltonian, while the path-integral representation perturbs only in the off-diagonal terms. Although the SSE method thus needs higher expansion orders for a given system, this disadvantage is compensated by a simplification in the algorithms: only integer indices of the operators need to be stored instead of continuous time variables τ_i. Except in strong magnetic fields or for dissipative quantum spin systems [43, 44] the SSE representation is thus the preferred representation for the simulation of quantum magnets.

The Negative Sign Problem

While the mapping from the quantum average to a classical average in (46) can be performed for any quantum system, it can happen in frustrated quantum magnets, that some of the weights $W(C)$ in the quantum system are negative, as is shown in Fig. 11.

Since Monte Carlo sampling requires positive weights $W(C) > 0$ the standard way of dealing with the negative weights of the frustrated quantum magnets is to sample with respect to the unfrustrated system by using the absolute values of the weights $|W(C)|$ and to assign the sign, $s(c) \equiv \mathrm{sign}\, W(C)$ to the quantity being sampled:

$$\langle m\rangle = \frac{\sum_c m(c)W(C)}{\sum_c W(C)} = \frac{\sum_c m(c)s(c)|W(C)| \,/\sum_c |W(C)|}{\sum_c s(c)|W(C)| \,/\sum_c |W(C)|} \equiv \frac{\langle ms\rangle'}{\langle s\rangle'}. \tag{49}$$

While this allows Monte Carlo simulations to be performed, the errors increase exponentially with the particle number N and the inverse temperature β. To see this, consider the mean value of the sign $\langle s\rangle = Z/Z'$, which is just the ratio of the partition functions of the frustrated system $Z = \sum_c W(C)$ with weights $W(C)$ and the unfrustrated system used for sampling with

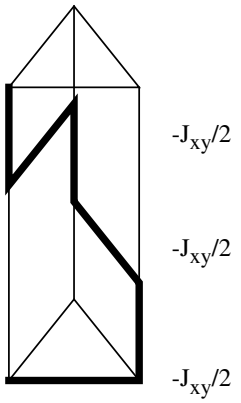


Fig. 11. Example of a frustrated world line configuration in a Heisenberg quantum antiferromagnet on a triangle. The closed world line configuration contains three exchange processes, each contributing a weight proportional to $-J_{xy}/2$. The overall is proportional to $(-J_{xy}/2)^3$ and is negative, causing a negative sign problem for the antiferromagnet with $J_{xy} > 0$

$Z' = \sum_c |W(C)|$. As the partition functions are exponentials of the corresponding free energies, this ratio is an exponential of the differences Δf in the free energy densities:$\langle s \rangle = Z/Z' = \exp(-\beta N \Delta f)$. As a consequence, the relative error $\Delta s/\langle s \rangle$ increases exponentially with increasing particle number and inverse temperature:

$$\frac{\Delta s}{\langle s \rangle} = \frac{\sqrt{\left(\langle s^2 \rangle - \langle s \rangle^2\right)/M}}{\langle s \rangle} = \frac{\sqrt{1 - \langle s \rangle^2}}{\sqrt{M}\langle s \rangle} \sim \frac{\exp(\beta N \Delta f)}{\sqrt{M}}. \tag{50}$$

Here M is the number of uncorrelated Monte Carlo samples. Similarly the error for the numerator increases exponentially and the time needed to achieve a given relative error scales exponentially in N and β.

It was recently shown that the negative sign problem is NP-hard, implying that almost certainly no solution for this exponential scaling problem exists [4]. Given this exponential scaling of quantum Monte Carlo simulations for frustrated quantum magnets, the QMC method is best suited for nonfrustrated magnets and we will restrict ourselves to these sign problem free cases in the following.

Measurements

Physical observables that can be measured in both the path-integral representation and the SSE representation include, next to the energy and the specific heat, any expectation value or correlation function that is diagonal in the basis set $\{|i\rangle\}$. This includes the uniform or staggered magnetization in the z direction, the equal time correlation functions and structure factor of the

z-spin components and the z-component uniform and momentum-dependent susceptibilites.

Offdiagonal operators, such as the magnetization in the x- or y-direction, or the corresponding correlation functions, structure factors and susceptibilities require an extension of the sampling to include configurations with broken world line segments. These are hard to measure in local update schemes (described in Sect. 4.3) unless open world line segments are already present when the Hamiltonian does not conserve magnetization, but are easily measured when non-local updates are used (see Sects. 4.4 and 4.5).

The spin stiffness ρ_s can be obtained from fluctuations of the winding numbers of the world lines [45], a measurement which obviously requires non-local moves that can change these winding numbers.

Dynamical quantities are harder to obtain, since the QMC representations only give access to imaginary-time correlation function. With the exception of measurements of spin gaps, which can be obtained from an exponential decay of the spin-spin correlation function in imaginary time, the measurement of real-time or real-frequency correlation functions requires an ill-posed analytical continuation of noisy Monte Carlo data, for example using the Maximum Entropy Method [46–48].

Thermodynamic quantities that cannot be expressed as the expectation value of an operator, such as the free energy or entropy cannot be directly measured but require an extended ensemble simulation, discussed in Sect. 5.

4.3 Local Updates

To perform a quantum Monte Carlo simulation on the world line representation, update moves that are ergodic and fulfill detailed balance are required. The simplest types of moves are again local updates. Since magnetization conservation prohibits the breaking of world lines, the local updates need to move world lines instead of just changing local states as in a classical model.

A set of local moves for a one-dimensional spin-1/2 model is shown in Fig. 12 [49, 50]. The two required moves are the insertion and removal of a pair of exchange processes (Fig. 12(a)) and the shift in time of an exchange process (Fig. 12(b)). Slightly more complicated local moves are needed for higher-dimensional models, for example to allow world lines to wind around elementary squares in a square lattice [51]. Since these local updates cannot

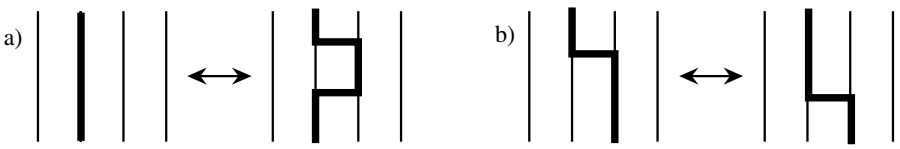


Fig. 12. Examples of local updates of world lines: (**a**) a pair of exchange processes can be inserted or removed; (**b**) an exchange process is moved in imaginary time

change global properties, such as the number of world lines (the magnetization) or their spatial winding, they need to be complemented with global updates [51].

While the local update world line and SSE algorithms enable the simulation of quantum systems they suffer from critical slowing down at second order phase transitions. Even worse, changing the spatial and temporal winding numbers usually has an exponentially small acceptance rate. While the restriction to zero spatial winding can be viewed as a boundary effect, changing the temporal winding number and thus the magnetization is essential for simulations at fixed magnetic fields.

4.4 Cluster Updates and the Loop Algorithm

The ergodicity problems of purely local updates and the critical slowing down observed also in quantum systems require the use of cluster updates. The loop algorithm [52] and its continuous time version [53], are generalizations of the classical cluster algorithms [6, 7] to quantum systems. They not only solve the problem of critical slowing down, but can also change the magnetization and winding numbers efficiently, avoiding the ergodicity problem of local updates. While the loop algorithm was initially developed for the path-integral representation it can also be applied to simulations in the SSE representation.

Since there exist extensive recent reviews of the loop algorithm [54,55], we will only outline the loop algorithm here. It constructs clusters of spins, similar to the Swendsen-Wang [6] clusters of the classical Ising model (Sect. 2.3). Upon applying the cluster algorithms to world lines in QMC we have to take into account that – in systems with conserved magnetization – the world lines may not be broken. This implies that a single spin cannot be flipped by itself, but, as shown in Fig. 13, connected world line segments of spins must be flipped together. These world line segments form a closed loop, hence the name "loop algorithm".

While the loop algorithm was originally developed only for spin-1/2 models it has been generalized to higher spin models [56–59] and anisotropic spin

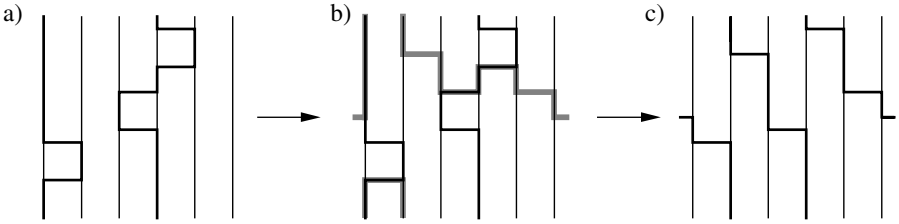


Fig. 13. A loop cluster update: (**a**) world line configuration before the update, where the world line of an up-spin is drawn as a *thick line* and that of a down-spin as a *thin line*; (**b**) world line configuration and a loop cluster (*grey line*); (**c**) the world line configurations after all spins along the loop have been flipped

models [60]. Since an efficient open-source implementation of the loop algorithm is available (see Sect. 4.6) we will not discuss further algorithmic details but refer interested readers to the reviews [54, 55].

4.5 Worm and Directed Loop Updates

The Loop Algorithm in a Magnetic Field

As successful as the loop algorithm is, it is restricted – as most classical cluster algorithms – to models with spin inversion symmetry. Terms in the Hamiltonian which break this spin-inversion symmetry, such as a magnetic field, are not taken into account during loop construction. Instead they enter through the acceptance rate of the loop flip, which can be exponentially small at low temperatures.

As an example consider two $S = 1/2$ quantum spins in a magnetic field:

$$H = J\mathbf{S}_1\mathbf{S}_2 - g\mu_B h(S_1^z + S_2^z) \tag{51}$$

In a field $g\mu_B h = J$ the singlet state $1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ with energy $-3/4J$ is degenerate with the triplet state $|\uparrow\uparrow\rangle$ with energy $1/4J - h = -3/4J$. As illustrated in Fig. 14a), we start from the triplet state $|\uparrow\uparrow\rangle$ and propose a loop shown in Fig. 14b). The loop construction rules, which ignore the magnetic field, propose to flip one of the spins and go to the intermediate configuration $|\uparrow\downarrow\rangle$ with energy $-1/4J$ shown in Fig. 14c). This move costs potential energy $J/2$ and thus has an *exponentially small acceptance rate* $\exp(-\beta J/2)$. Once we accept this move, immediately many small loops are built, exchanging the spins on the two sites, and gaining exchange energy $J/2$ by going to the spin singlet state. A typical world line configuration for the singlet is shown in Fig. 14d). The reverse move has the same exponentially small probability, since the probability to reach a world line configuration without any exchange term [Fig. 14c)] from a spin singlet configuration [Fig. 14d)] is exponentially small.

This example clearly illustrates the reason for the exponential slowdown: in a first step we *lose all potential energy*, before *gaining it back in exchange energy*. A faster algorithm could thus be built if, instead of doing the trade in one big step, we could trade potential with exchange energy in small pieces, which is exactly what the worm algorithm does.

The Worm Algorithm

The worm algorithm [40] works in an extended configuration space, where in addition to closed world line configurations one open world line fragment (the "worm") is allowed. Formally this is done by adding a source term to the Hamiltonian which for a spin model is

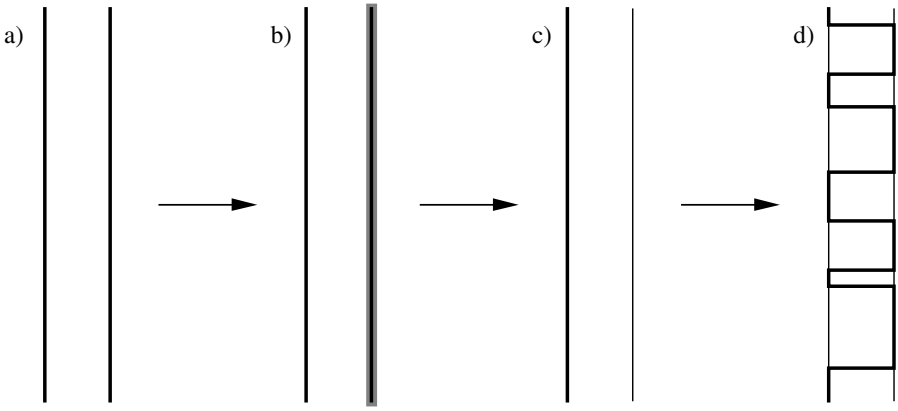


Fig. 14. A loop update for two antiferromagnetically coupled spins in a magnetic field with $J = g\mu_B h$. (**a**) Starting from the triplet configuration $|\uparrow\uparrow\rangle$, (**b**) a loop is constructed, proposing to go to (**c**), the intermediate configuration $|\uparrow\downarrow\rangle$, which has an exponentially small acceptance rate, and finally into configurations like (**d**) which represent the singlet state $1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. As in the previous figure a *thick line* denotes an up-spin and a *thin line* a down-spin

$$H_{\text{worm}} = H - \eta \sum_i (S_i^+ + S_i^-) \; . \tag{52}$$

This source term allows world lines to be broken with a matrix element proportional to η. The worm algorithm now proceeds as follows: a worm (i.e. a world line fragment) is created by inserting a pair (S_i^+, S_i^-) of operators at nearby times, as shown in Fig. 15a,b). The ends of this worm are then moved randomly in space and time [Fig. 15c)], using local Metropolis or heat bath updates until the two ends of the worm meet again as in Fig. 15d). Then an update which removes the worm is proposed, and if accepted we are back in a configuration with closed world lines only, as shown in Fig. 15e). This algorithm is straightforward, consisting just of local updates of the worm ends in the extended configuration space but it can perform nonlocal changes. A worm end can wind around the lattice in the temporal or spatial direction and that way change the magnetization and winding number.

In contrast to the loop algorithm in a magnetic field, where the trade between potential and exchange energy is done by first losing all of the potential energy, before gaining back the exchange energy, the worm algorithm performs this trade in small pieces, never suffering from an exponentially small acceptance probability. While it is not as efficient as the loop algorithm in zero magnetic field (the worm movement follows a random walk while the loop algorithm can be interpreted as a self-avoiding random walk), the big advantage of the worm algorithm is that it remains efficient in the presence of a magnetic field.

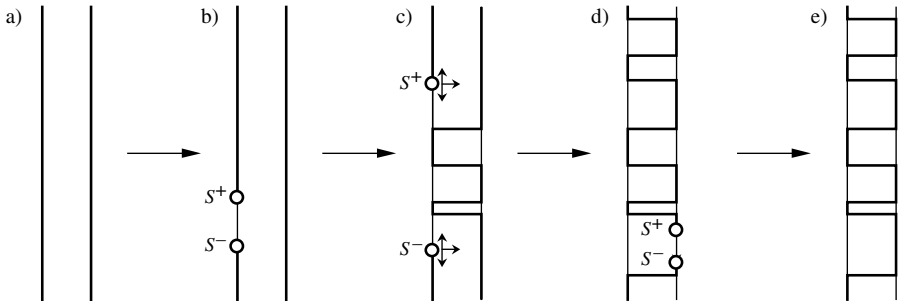


Fig. 15. A worm update for two antiferromagnetically coupled spins in a magnetic field with $J = g\mu_B h$. (**a**) starting from the triplet configuration $|\uparrow\uparrow\rangle$ a worm is constructed in (**b**) by inserting a pair of S^+ and S^- operators. (**c**) these "worm end" operators are then moved by local updates until (**d**) they meet again, when a move to remove them is proposed, which leads to the closed world line configuration (**e**). As in the two previous figures a *thick line* denotes an up-spin and a *thin line* a down-spin

The Directed Loop Algorithm

Algorithms with a similar basic idea as the worm algorithm in the path-integral representations are the operator-loop update [61,62] and the directed-loop algorithms [63] which can be formulated in both an SSE and a world-line representation. Like the worm algorithm, these algorithms create two world line discontinuities, and move them around by local updates. The main difference to the worm algorithm is that here these movements do not follow an unbiased random walk but have a preferred direction, always trying to move away from the last change, which further speeds up the simulations.

4.6 Open Source Implementations: the ALPS Project

The loop, worm and directed loop algorithms can be used for the simulation of a wide class of quantum magnets. They are of interest not only to theoretical physicists, but also to experimentalists who want to fit experimental measurements to theoretical models. The wide applicability of these methods has led to the publication of open-source versions of these algorithms as part of the ALPS project (Algorithms and Libraries for Physics Simulations) [10] on the web page `http://alps.comp-phys.org/`.

4.7 Applications

We will finally present typical applications of the above algorithms by reviewing a small and necessarily biased selection.

The loop algorithm has been applied to a wide range of problems, ranging from purely theoretical questions to experimental data fitting. Below we list a

selection of applications that provide an overview over the possibilities of the loop algorithm. The first simulation using the loop algorithm was an accurate determination of the ground state properties (staggered magnetization, spin stiffness and spin wave velocity) of the square-lattice spin-1/2 quantum Heisenberg antiferromagnet [64]. In a similar spirit the uniform susceptibility, correlation length and spin gap of spin ladder models [65,66] and integer spin chains [59] was calculated, confirming the presence of a spin gapped ground state in even-leg spin ladders and integer spin chains.

As the loop algorithm is efficient also at critical points, it has been used in the first high accuracy simulations of the critical properties of quantum phase transitions by studying the Néel to quantum paramagnet transition in two-dimensional quantum spin systems [67], for a determination of the low-temperature asymptotic scaling of two-dimensional quantum Heisenberg antiferromagnets [3,58,68], and for accurate calculations of the Néel temperature of anisotropic quasi-one and quasi-two dimensional antiferromagnets [69].

The loop algorithm is not only restricted to toy models, but can be applied to realistic models of quantum magnets. Comparisons to experimental measurements are done by fitting simulation data to experimental measurements, as for alternating chain compounds [70], spin ladder materials [71] or frustrated square lattice antiferromagnets [72]. In the latter material the sign problem due to frustration limits the accuracy. As an example we show in Fig. 16 the good quality of a fit of QMC data to experimental measurements on the spin ladder compound $SrCu_2O_3$.

Another interesting application is to simulate realistic models for quantum magnets, using exchange constants calculated by ab-initio methods. Comparing these ab-initio QMC data to experimental measurements, as done for a series of vanadates [73] and for ladder compounds [71] allows to quantitatively check the ab-initio calculations.

The worm and directed loop algorithms are applied when magnetic fields are present. Typical examples include the calculation of magnetization curves of quantum magnets [74], the determination of the first order nature of the spin flop transition in two dimensions [75] and the calculation of phase diagrams of dimerized quantum magnets in a magnetic field [76].

5 Extended Ensemble Methods for Quantum Systems

In this section we will present generalizations of extended ensemble simulations to world line quantum Monte Carlo simulations, in particular:

- histogram reweighting,
- parallel tempering,
- extended ensemble methods.

Histogram reweighting allows to extract information at a temperature different than (but close to) the temperature at which the simulation is performed. This is especially useful when studying critical phenomena, where a

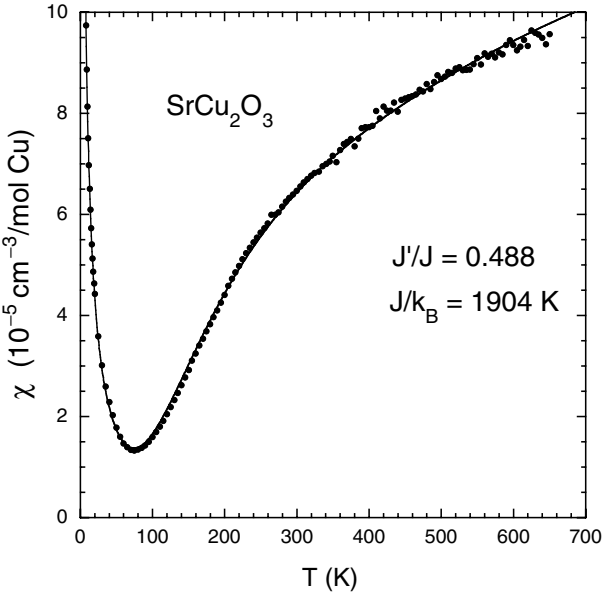


Fig. 16. Fits of experimental measurements of the uniform susceptibility of $SrCu_2O_3$ to the results of QMC simulations, determining a coupling $J \approx 1904\,\mathrm{K}$ along the chains of the ladder and a ratio $J'/J \approx 0.488$ for the inter-chain to intra-chain coupling

single simulation can provide information for the whole critical region around the phase transition.

Parallel tempering and the extended ensemble methods (such as multicanonical simulations and Wang-Landau sampling) speed up simulations at and below phase transitions and are especially useful at first order phase transitions or for frustrated systems. Note, however, that frustrated quantum spin systems generally suffer from the negative sign problem. Since the negative sign problem arises as a property of the representation and does not depend on the ensemble or the updates, the scaling will remain exponential even when using improved sampling algorithms, in contrast to classical simulations where extended ensemble algorithms and parallel tempering can dramatically speed up the simulations.

Another advantage of extended ensemble simulations is the ability to directly calculate the density of states and from it thermodynamic properties such as the entropy or the free energy that are not directly accessible in canonical simulations. In the following we will again use quantum magnets as concrete examples. A generalization to bosonic and fermionic models will always be straightforward.

5.1 Generalizing Extended Ensembles to Quantum Systems

Since simulations of quantum systems suffer from the same problems as classical simulations, the extension of these generalized sampling schemes to quantum systems is highly desired. The extension is not immediately obvious since the partition function of a quantum system cannot be cast in the classical form

$$Z = \sum_E g(E) e^{-\beta E} \tag{53}$$

unless the complete spectrum of the Hamilton operator H is known.

Instead of a representation like (53) we will aim for a generalized representation of the form

$$Z = \sum_c W(c) = \sum_{\boldsymbol{\lambda}} g(\boldsymbol{\lambda}) p(\boldsymbol{\mu}, \boldsymbol{\lambda}) \,, \tag{54}$$

where $\boldsymbol{\lambda}$ describes values of properties $\boldsymbol{\Lambda}$ of the configuration that are sampled, $\boldsymbol{\mu}$ are external parameters such as temperature or coupling constants, and $p(\boldsymbol{\mu}, \boldsymbol{\lambda})$ is the weight of that configuration. The generalized density of states $g(\boldsymbol{\lambda})$ is a sum over all configurations c with the property $\boldsymbol{\Lambda}(c) = \boldsymbol{\lambda}$

$$g(\boldsymbol{\lambda}) = \sum_c \delta_{\Lambda(c),\lambda} \tilde{W}(c) \,, \tag{55}$$

where the reduced weight $\tilde{W}(c) = W(c)/p(\boldsymbol{\mu}, \boldsymbol{\lambda})$ of a configuration c shall not depend on the parameters $\boldsymbol{\mu}$. By defining as

$$A(\boldsymbol{\lambda}) = \frac{1}{g(\boldsymbol{\lambda})} \sum_c \delta_{\boldsymbol{\Lambda}(c),\lambda} A(c) \tilde{W}(c) \tag{56}$$

the "microcanonical" average of A for configurations with $\boldsymbol{\Lambda}(c) = \lambda$ we can obtain thermal averages $\langle A(\boldsymbol{\mu}) \rangle$ at arbitrary parameters $\boldsymbol{\mu}$:

$$\langle A(\boldsymbol{\mu}) \rangle = \frac{\sum_{\boldsymbol{\lambda}} A(\boldsymbol{\lambda}) g(\boldsymbol{\lambda}) p(\boldsymbol{\mu}, \boldsymbol{\lambda})}{\sum_{\boldsymbol{\lambda}} g(\boldsymbol{\lambda}) p(\boldsymbol{\mu}, \boldsymbol{\lambda})} \,. \tag{57}$$

In a classical simulation we might choose the desired property as the energy: $\boldsymbol{\lambda} = E$, the external parameter the inverse temperature: $\boldsymbol{\mu} = \beta$, the weight the Boltzmann weight $p(\beta, E) = \exp(-\beta E)$. The reduced weight $\tilde{W}(c) = 1$ and, hence, $g(E)$ the standard density of states.

The generalized notation makes sense even for a classical system. Consider, for example, an Ising antiferromagnet in a magnetic field:

$$H = J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \,. \tag{58}$$

If we are interested in properties at a fixed inverse temperature β as a function of the magnetization we can choose as external parameter the magnetic field

$\boldsymbol{\mu} = h$ and as property of the system the magnetization $\boldsymbol{\lambda} = M = \sum_i \sigma_i$, giving a representation:

$$Z = \sum_M g(M) e^{\beta h M}. \tag{59}$$

If we are interested in properties as a function of both temperature and magnetization we might pick a two-dimensional representation. As external parameters we choose the inverse temperature and magnetic field $\boldsymbol{\mu} = (\beta, h)$. The corresponding properties of the system are $\boldsymbol{\lambda} = (E_J, M)$, where the magnetic energy E_J is defined as

$$E_J = J \sum_{\langle i,j \rangle} \sigma_i \sigma_j \ . \tag{60}$$

This gives a representation

$$Z = \sum_{E_J, M} g(E_J, M) e^{-\beta E_J + \beta h M} \ . \tag{61}$$

Continuous Time Path Integrals

To apply generalized sampling schemes to quantum systems in the path integral representation we cast (47) into the form

$$Z = \int dE_0 \sum_{n=0}^{\infty} g(E_0, n) \beta^n e^{-\beta E_0} \ , \tag{62}$$

where the diagonal energy contribution E_0 is the value of the diagonal part of the Hamiltonian H_0 in each configuration.

Comparing to (54) we have as control parameter the inverse temperature $\boldsymbol{\mu} = \beta$ and need two properties of the configuration $\boldsymbol{\lambda} = (E_0, n)$.

As in classical systems we might be interested in the dependence on a magnetic field h instead of the temperature, and rewrite (47) in a form very similar to the classical one as

$$Z = \int dM g(M) e^{\beta h M} \ , \tag{63}$$

where the magnetization M of a configuration is defined as

$$M = \frac{1}{\beta} \left[(\tau_1 + \beta - \tau_n) \left\langle i_1 | \sum_r S_r^z | i_1 \right\rangle + \sum_{i=2}^{n} (\tau_i - \tau_{i-1}) \left\langle i_i | \sum_r S_r^z | i_i \right\rangle \right]. \tag{64}$$

Like in the classical systems, similar expression, can be derived for the dependency on any parameters of interest.

Stochastic Series Expansion

In the stochastic series expansion a one-dimensional representation is sufficient to calculate properties as a function of the temperature:

$$Z = \sum_{n=0}^{\infty} g(n)\beta^n, \tag{65}$$

which is just the high temperature expansion of the partition function.

5.2 Histogram Reweighting

In a classsical system the thermal average of a quantity A at an inverse termperature β

$$\langle A(\beta)\rangle = \frac{1}{Z}\sum_c A_c e^{-\beta E_c}, \tag{66}$$

where A_c is the measurement of the observable A in the configuration c and E_c the energy of that configuration is usually estimated by the sample mean in a Monte Carlo simulation

$$\langle A(\beta)\rangle \approx \overline{A} = \frac{1}{M}\sum_i A_{c_i}. \tag{67}$$

This sampling scheme gives results only for the inverse temperature β, but actually there is much more information available than just the simple average (67). For example, in the search for a phase transition a range of temperatures needs to be explored and information at a nearby inverse temperature $\beta' \approx \beta$ can be obtained from a simulation performed at β. This is done by reweighting the configurations sampled with the Boltzmann weight $p_c = \exp(-\beta E_c)$ to obtain averages for the Boltzmann weight $p'_c = \exp(-\beta' E_c)$:

$$\langle A(\beta')\rangle \approx \frac{\sum_i A_{c_i} p'_{c_i}/p_{c_i}}{\sum_i p'_{c_i}/p_{c_i}} = \frac{\sum_i A_{c_i} e^{-\Delta\beta E_{c_i}}}{\sum_i e^{-\Delta\beta E_{c_i}}}\ , \tag{68}$$

where $\Delta\beta = \beta' - \beta$.

Instead of storing the full time series of measurements $\{A_{c_i}\}$ and energies $\{E_{c_i}\}$, it is sufficient to store a histogram $H(E)$, counting how often the energy level E occurs in the time series $\{E_{c_i}\}$, and the average $A(E)$ of all the measurements performed on configurations with energy E [77, 78]. Since the histogram $H(E)$ is a statistical estimator for the product $g(E)p(E)$, the average $\langle A(\beta')\rangle$ can be calculated from (57) as a sum over all energies

$$\langle A(\beta')\rangle \approx \frac{\sum_E H(E)A(E)e^{-\Delta\beta E}}{\sum_E H(E)e^{-\Delta\beta E}}\ . \tag{69}$$

In a model with continuous energy spectrum, such as the Heisenberg model, the energy range is divided into discrete bins of width ΔE and the histograms are constructed for these bins.

Histogram reweighting works well only if the configurations sampled at the inverse temperature β are also relevant at β', requiring that $\Delta\beta$ is small. Otherwise the errors become too large since there will not be sufficient entries in $H(E)$ for the energies E important at β'.

Multiple histograms obtained at different temperatures can be used to broaden the accessible temperature range [77, 78].

Generalized Histogram Reweighting

Histogram reweighting can not only be performed in the temperature, but also in any of the coupling constants. Using the generalized representation (54), we can calculate expectation values at coupling constant $\boldsymbol{\mu}' \approx \boldsymbol{\mu}$ from simulation performed for coupling constants $\boldsymbol{\mu}$ by reweighting as

$$\langle A(\boldsymbol{\mu}')\rangle \approx \frac{\sum_{\boldsymbol{\lambda}} H(\boldsymbol{\lambda}) A(\boldsymbol{\lambda}) p(\boldsymbol{\mu}', \boldsymbol{\lambda})/p(\boldsymbol{\mu}, \boldsymbol{\lambda})}{\sum_{\boldsymbol{\lambda}} H(\boldsymbol{\lambda}) p(\boldsymbol{\mu}', \boldsymbol{\lambda}/p(\boldsymbol{\mu}, \boldsymbol{\lambda})}, \tag{70}$$

since the recorded histogram $H(\boldsymbol{\lambda})$ is an estimator for $g(\boldsymbol{\lambda})p(\boldsymbol{\mu}, \boldsymbol{\lambda})$.

For example, to investigate a phase transition as a function of the magnetic field h in an Ising antiferromagnet with Hamilton function (58) one would construct a histogram $H(M)$ of the magnetization $M = \sum_i \sigma_i$, and store the averages of the energy $E(M)$ and any observable $A(M)$ as a function of magnetization. Reweighting to a new field strength $h' = h + \Delta h$ is then easily done:

$$\langle A(h')\rangle \approx \frac{\sum_M H(M) A(M) e^{\beta \Delta h M}}{\sum_M H(M) e^{\beta \Delta h M}}. \tag{71}$$

Similar expressions are readily derived for quantum systems. Here we only give the expressions for temperature reweighting in the path integral representation:

$$\langle A(\beta')\rangle \approx \frac{\int dE_0 \sum_{n=0}^{\infty} H(n, E_0) A(n, E_0) e^{-\Delta\beta E_0} (\beta'/\beta)^n}{\int dE_0 \sum_{n=0}^{\infty} H(n, E_0) e^{-\Delta\beta E_0} (\beta'/\beta)^n}\ , \tag{72}$$

and the SSE representation:

$$\langle A(\beta')\rangle \approx \frac{\sum_{n=0}^{\Lambda} H(n) A(n) (\beta'/\beta)^n}{\sum_{n=0}^{\Lambda} H(n) (\beta'/\beta)^n}\ . \tag{73}$$

The integral in the path integral equation is again replaced by a sum over entries in a binned energy histogram.

Since the histogram $H(\boldsymbol{\lambda})$ is strongly peaked around the thermal expectation values for the observable $\boldsymbol{\Lambda}$ at a given set of parameters $\boldsymbol{\mu}$, the density of states $g(\boldsymbol{\lambda}) = H(\boldsymbol{\lambda})/p(\boldsymbol{\mu}, \boldsymbol{\lambda})$ can be accurately estimated only in a small

region of phase space. Consequently, histogram reweighting can only be used to estimate averages at nearby parameters $\boldsymbol{\mu}' \approx \boldsymbol{\mu}$, where the same states are relevant. In order to explore larger parameter regions, parallel tempering or generalized ensembles with "flat" histograms can be used.

5.3 Parallel Tempering

Parallel tempering, introduced in Sect. 3.6 can be generalized in the same way. Using the generalized representation (54) we can write the combined weights of two configurations c_i and c_{i+1} with properties $\boldsymbol{\Lambda}(c_i) = \boldsymbol{\lambda}_i$ and $\boldsymbol{\Lambda}(c_{i+1}) = \boldsymbol{\lambda}_{i+1}$ simulated at parameters $\boldsymbol{\mu_i}$ and $\boldsymbol{\mu_{i+1}}$ as

$$\tilde{W}(c_i)\tilde{W}(c_{i+1})p(\boldsymbol{\mu}_i, \boldsymbol{\lambda}_i)p(\boldsymbol{\mu}_{i+1}, \boldsymbol{\lambda}_{i+1}) \tag{74}$$

before the swap and

$$\tilde{W}(c_{i+1})\tilde{W}(c_i)p(\boldsymbol{\mu}_i, \boldsymbol{\lambda}_{i+1})p(\boldsymbol{\mu}_{i+1}, \boldsymbol{\lambda}_i) \tag{75}$$

after the swap. Since the reduced weights $\tilde{W}(c)$ do not depend on the parameters $\boldsymbol{\mu}$, the Metropolis acceptance probability for the swap is

$$\min\left[1, \frac{p(\boldsymbol{\mu}_i, \boldsymbol{\lambda}_{i+1})p(\boldsymbol{\mu}_{i+1}, \boldsymbol{\lambda}_i)}{p(\boldsymbol{\mu}_i, \boldsymbol{\lambda}_i)p(\boldsymbol{\mu}_{i+1}, \boldsymbol{\lambda}_{i+1})}\right], \tag{76}$$

which reduces to (37) for the usual parallel tempering in temperature.

Applying parallel tempering to the magnetic field h in a classical Monte Carlo simulation and choosing a set of magnetic field strengths $\{h_i\}$, a swap between configurations at neighboring field strengths is then accepted with probability

$$\min\left[1, e^{-\beta(h_{i+1}-h_i)(M(c_{i+1})-M(c_i))}\right], \tag{77}$$

where $M(c)$ is the magnetization of the configuration c.

For a quantum system a parallel tempering swap in temperature is accepted with a probability

$$\min\left[1, e^{(\beta_{i+1}-\beta_i)(E_0(c_{i+1})-E_0(c_i))}\left(\frac{\beta_{i+1}}{\beta_i}\right)^{n_i-n_{i+1}}\right] \tag{78}$$

for continuous time path integrals and

$$\min\left[1, \left(\frac{\beta_{i+1}}{\beta_i}\right)^{n_i-n_{i+1}}\right] \tag{79}$$

in the SSE representation, where n_i and n_{i+1} refer to the order of the respective configuration.

The expressions for parallel tempering in coupling constants instead of temperatures can be derived in a similar fashion.

Optimal Temperature Sets

The algorithm to determine optimized temperature sets presented for classical systems in Sect. 3.6 can now be applied without modifications to the quantum case.

5.4 Wang-Landau Sampling and Optimized Ensembles

Just like histogram reweighting or parallel tempering, the multicanonical ensemble, Wang-Landau sampling and the optimized ensemble algorithms can not only be applied to the energy but to arbitrary observables $\boldsymbol{\Lambda}$, by choosing the generalized multicanonical weight of a configuration c with $\boldsymbol{\Lambda}(c) = \boldsymbol{\lambda}$:

$$p(c) = \frac{\tilde{W}(c)}{g(\boldsymbol{\lambda})} = \frac{W(c)}{p(\boldsymbol{\mu}, \boldsymbol{\lambda}) g(\boldsymbol{\lambda})} \; . \tag{80}$$

The Wang-Landau algorithm can again be used to iteratively determine the generalized density of states $g(\boldsymbol{\lambda})$, and a flat histogram $H(\boldsymbol{\lambda})$ will be obtained. After the simulation, thermal averages $\langle A(\boldsymbol{\mu}) \rangle$ at arbitrary parameters $\boldsymbol{\mu}$ can be obtained using (57).

For example, to perform a multicanonical ensemble simulation in the magnetization M instead of the energy E, in an Ising model (58) we consider the density of states for the magnetization $g(M)$ of the magnetization $\boldsymbol{\lambda} = M = \sum_i \sigma_i$ and use a mixed weight

$$p(E_J, M) = e^{-\beta E_J} \frac{1}{g(M)} \tag{81}$$

where the exchange energy is defined as

$$E_J = J \sum_{\langle i,j \rangle} \sigma_i \sigma_j \; . \tag{82}$$

After the simulation, magnetic field dependent expectation values can be obtained at arbitrary values of the magnetic field $\boldsymbol{\mu} = h$:

$$\langle A(h) \rangle = \frac{\sum_M A(M) g(M) e^{\beta h M}}{\sum_M g(M) e^{\beta h M}} \tag{83}$$

Wang-Landau Sampling for Quantum Systems – High Temperature Expansion

It should now be obvious that Wang-Landau sampling and similar algorithms can be applied to quantum systems, by using the generalized density of states for quantum systems introduced in Sect. 5.1.

For thermal representations as a function of inverse temperature β, SSE offers a generalized density of states $g(n)$ in just the expansion order n, while

the path integral representation requires a two-dimensional density of states $g(E_0, n)$ as a function of the non-interacting energy E_0 and the expansion order n. Since it is our experience that one-dimensional histograms perform, in general, better than higher dimensional histograms, we will focus only on the stochastic series expansion representation, for which the quantum version of Wang-Landau sampling was first introduced [35, 36].

In the following we discuss Wang-Landau sampling in the order n of a configuration, which is the equivalent of Wang-Landau sampling in energy space for classical systems. Following the generalized Wang-Landau algorithm, we replace the weight $W(c)$ of a configuration by the new weight (80) which here is

$$\frac{W(c)}{\beta^n g(n)} \; . \tag{84}$$

Normalization of $g(n)$ is simple, since $g(0)$ is just the total number of basis states (e.g. $(2+1)^N$ in a quantum spin model with N spins of size S.).

After the simulation, the partition function can easily be calculated as

$$Z = \sum_{n=0}^{\infty} g(n)\beta^n \; , \tag{85}$$

and observables at arbitrary temperatures are calculated using (57), which here reads:

$$\langle A(\beta) \rangle = \frac{1}{Z} \sum_{n=0}^{\infty} A(n) g(n) \beta^n \; . \tag{86}$$

In any simulation the sums $\sum_{n=0}^{\infty}$ have to be truncated at some order K. What is the effect of this truncation? For canonical simulations we chose $K > O(N\beta)$, such that contributions from orders $n > K$ were negligibly small, and orders $n > K$ were never reached in the simulation. We could then ignore the cutoff. Using Wang-Landau sampling, the cutoff similarly restricts the validity of the results to temperatures where $g(K)\beta^K$, and hence contributions from terms $n > K$ are small. The cutoff K thus sets an upper bound for the accessible inverse temperatures β. In Fig. 17 results of calculations for the free energy F, entropy S and specific heat C of an $N = 10$ site antiferromagnetic Heisenberg chain, and compare to exact results. Using 10^8 sweeps, which can be performed in a few hours on a PC, the errors can be reduced down to the order of 10^{-4}. The cutoff was set to $K = 250$, restricting the accessible temperatures to $T \gtrsim 0.05J$. The sudden departure of the Monte Carlo data from the exact values below this temperature clearly shows this limit, which can be pushed lower by increasing K. The sudden deviation becomes even more pronounced in larger systems and provides a reliable indication for the range of validity of the results.

To illustrate the efficiency of the algorithm close to a thermal second order phase transition, we consider in our second example the Heisenberg antiferromagnet on a simple cubic lattice. From simulations of systems with L^3 sites,

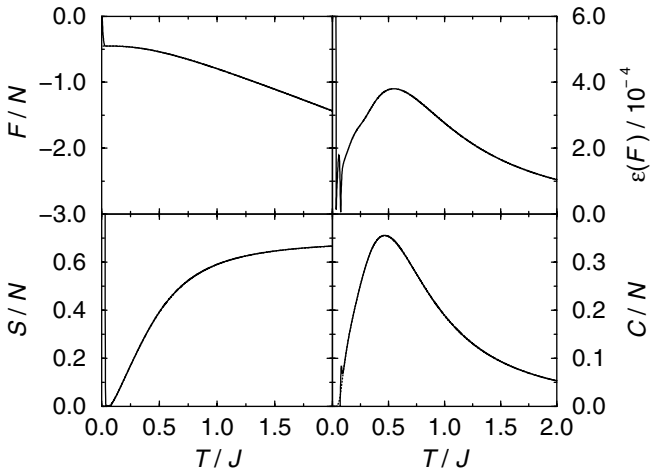


Fig. 17. Free energy (F), entropy (S) and specific heat (C) of an $N = 10$ site antiferromagnetic Heisenberg chain. Solid line correspond to the MC results, indistinguishable from the dotted lines for the exact results. Also shown is the relative error $\varepsilon(F)$ of F compared to the exact result

$L = 4, 6, 8, 12, 16$, we can calculate the staggered structure factor $S(\pi, \pi)$ for any value of the temperature using the measured histograms. Figure 18 shows the scaling plot of $S(\pi, \pi)/L^{2-\eta}$ with $\eta = 0.034$. The estimate for the critical temperature $T_c = 0.947J$, obtained in less than a day on a PC, compares well with earlier estimates [79].

Wang-Landau Sampling for Quantum Systems – Perturbation Expansion

Instead of performing a high temperature expansion, which is well suited to the investigation of finite temperature phase transitions, we can also apply Wang-Landau sampling to a perturbation expansion, better suited for quantum phase transitions. Instead of scanning a temperature range we vary one of the interactions at fixed temperature. Defining the Hamiltonian as $H = H_0 + \lambda V$ we can write the partition function equation as

$$Z = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \mathrm{Tr}(-H_0 - \lambda V)^n \equiv \sum_{n_\lambda=0}^{\infty} g(n_\lambda) \lambda^{n_\lambda} \ , \tag{87}$$

where n_λ counts the powers of λ in the weight of a configuration. In this formulation of the algorithm, the cutoff K restricts the value of coupling parameter λ up to which the perturbation expansion is reliable.

The simplest case is when the parameter λ multiplies all terms on a subset of the bonds. An example is the bilayer Heisenberg quantum antiferromagnet with Hamiltonian

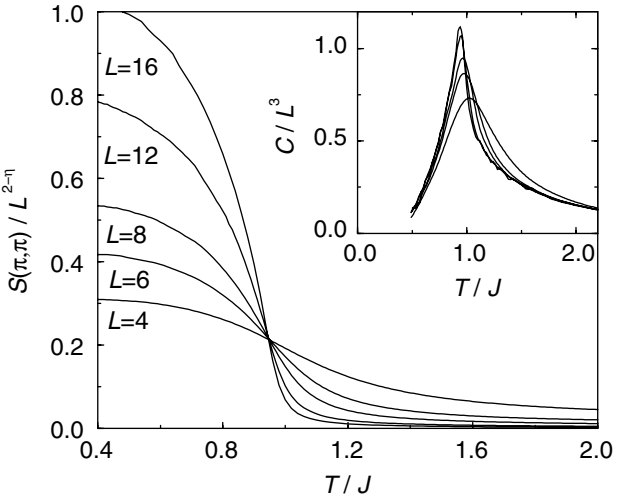


Fig. 18. Scaling plot of the staggered structure factor of a cubic antiferromagnet as a function of temperature, obtained from simulations at a fixed temperature for various lattice sizes. The inset shows the specific heat as a function of temperature. The cutoff $K = 500(L/4)^3$ restricts the accessible temperature range to $T \gtrsim 0.4J$

$$H_{\text{bilayer}} = J \sum_{l=1}^{2} \sum_{\langle i,j \rangle} \boldsymbol{S}_{i,l} \boldsymbol{S}_{j,l} + J' \sum_{\langle i,j \rangle} \boldsymbol{S}_{i,1} \boldsymbol{S}_{j,2} \ , \tag{88}$$

where $\boldsymbol{S}_{i,l}$ is the spin operator on site i in layer l. This model consists of two square lattices with coupling J between nearest neighbors inside each square lattice layer, and a coupling J' between adjacent spins in different layers, and we set $\lambda = J/J'$. For such models the algorithm remains very simple, and again only a few acceptance rates need to be changed in the code.

To normalize $g(n_\lambda)$ there are two options. If H_0 can be solved exactly, $g(0)$ can be determined directly. Otherwise, the normalization can be fixed using the high temperature expansion version of the algorithm to calculate $Z(\beta)$ at any fixed value of λ. Even without normalization we can still obtain entropy and energy differences.

We consider as an example the quantum phase transition in the bilayer Heisenberg antiferromagnet. Its ground state changes from quantum disordered to Néel ordered as the ratio $\lambda = J/J'$ of intra-plane (J) to inter-plane (J') coupling is increased [80]. From the histograms generated within *one* simulation we can calculate the staggered structure factor $S(\pi, \pi)$ of the system at *any value* of λ. In Fig. 19 we show a scaling plot of $S(\pi, \pi)/L^{2-z-\eta}$ as a function of λ. In short simulations, taking only a few days on a PC, we find the quantum critical point at $\lambda = 0.396$, which again compares well with earlier results.

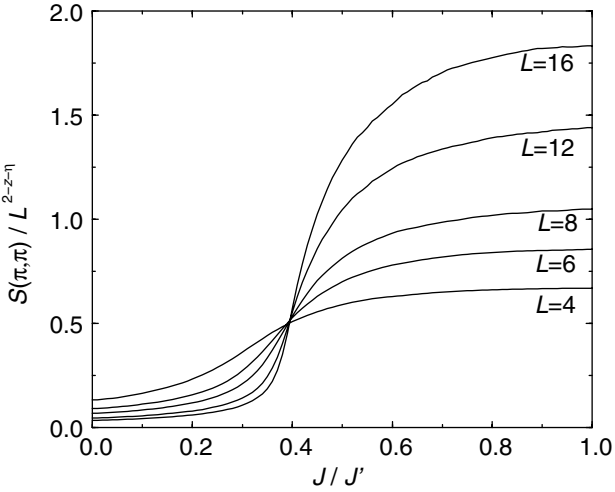


Fig. 19. Scaling plot of the staggered structure factor of a Heisenberg bilayer as a function of the coupling ratio $\lambda = J/J'$. Results are shown for various linear system sizes L. The temperature was chosen $\beta J' = 2L$, low enough to be in the scaling regime. The cutoff $K = 8L^3$ was chosen large enough to cover the coupling range $J/J' \lesssim 1$. The dynamical critical exponent of this model is $z = 1$ and $\eta = 0.034$

Optimized Ensembles for Quantum Systems

As in classical Monte Carlo simulations it turns out that a flat histogram $H(n)$ of the expansion orders n is not optimal, and again an optimized ensemble can be derived.

6 Summary

In this chapter we have reviewed two recent developments in the field of classical and quantum Monte Carlo simulations. In the first part we have presented a short review of extended ensemble techniques, including multicanonical and parallel tempering simulations. Counter widespread assumptions "flat-histogram" multicanonical ensembles, or parallel tempering with constant, "flat" acceptance rates are not optimal. We have reviewed a recently developed iterative feedback algorithm to obtain an optimal multicanonical ensemble or an optimal choice of temperature set respectively that for a given model will maximize equilibration. We have shown examples ranging from classical spin systems, over dense liquids to protein folding.

In the second part we have given a short introduction and review of modern world line quantum Monte Carlo algorithms for quantum systems, which are free of any time discretization errors. We have highlighted the fact that

the various world line representations map the quantum system to an effective classical system, and shown how efficient sampling schemes developed for classical simulations, such as histogram reweighting, parallel tempering, extended and optimized ensembles, can be applied also to quantum systems.

With examples ranging from quantum magnets to protein folding the ensemble optimization techniques have already been demonstrated to be useful for a wide range of models and scales, and will certainly help with the simulation of many other complex systems in the future.

References

1. F. Barahona (1982) On the computational complexity of Ising spin glass models. *J. Phys. A* **15**, p. 3241
2. S. Cook (1971) The complexity of theorem-proving procedures. *Conference Record of Third Annual ACM Symposium on Theory of Computing*, pp. 151–158
3. J. Kim and M. Troyer (1998) Low temperature behavior and crossovers of the square lattice quantum Heisenberg antiferromagnet. *Phys. Rev. Lett.* **80**, p. 2705
4. M. Troyer and U.-J. Wiese (2005) Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94**, p. 170201
5. N. Metropolis, A. R. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953) Equation of state calculations on fast computing machines. *J. of Chem. Phys.* **21**, p. 1087
6. R. Swendsen and J.-S. Wang (1987) Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58**, p. 86
7. U. Wolff (1989) Collective Monte Carlo updating for spin systems. *Phys. Rev. Lett.* **62**, p. 361
8. O. Redner, J. Machta, and L. F. Chayes (1998) Graphical representations and cluster algorithms for critical points with fields. *Phys. Rev. E* **58**, p. 2749
9. H. Evertz, H. Erkinger, and W. von der Linden (2002) New cluster method for the Ising mode. In: *Computer Simulations in Condensed Matter Physics*, eds. D. Landau, S. P. Lewis, H.-B. Schüttler, vol. **XIV**, Springer, Berlin, p. 123
10. F. Alet, P. Dayal, A. Grzesik, A. Honecker, M. Körner, A. Läuchli, S. Manmana, I. McCulloch, F. Michel, R. Noack, G. Schmid, U. Schollwöck, F. Stöckli, S. Todo, S. Trebst, M. Troyer, P. Werner, and S. Wessel (2005) The ALPS project: open source software for strongly correlated systems. *J. Phys. Soc. Jpn. Suppl.* **74**, p. 30
11. B. A. Berg and T. Neuhaus (1991) Multicanonical algorithms for first order phase transitions. *Phys. Lett. B* **267**, p. 249
12. B. A. Berg and T. Neuhaus (1992) Multicanonical ensemble: A new approach to simulate first-order phase transitions. *Phys. Rev. Lett.* **68**, p. 9
13. F. Wang and D. P. Landau (2001) Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **86**, p. 2050
14. F. Wang and D. P. Landau (2001) Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram. *Phys. Rev. E* **64**, p. 056101

15. C. Zhou and R. N. Bhatt (2005) *Phys. Rev. E* **72**, p. 025701(R)
16. H. K. Lee, Y. Okabe, and D. P. Landau (2006) Convergence and Refinement of the Wang-Landau Algorithm. *Comp. Phys. Comm.* **175**, p. 36
17. P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith (2004) Performance limitations of flat-histogram methods. *Phys. Rev. Lett.* **92**, p. 097201
18. Y. Wu, M. Körner, L. Colonna-Romano, S. Trebst, H. Gould, J. Machta, and M. Troyer (2005) Overcoming the critical slowing down of flat-histogram Monte Carlo simulations: Cluster updates and optimized broad-histogram ensembles. *Phys. Rev. E* **72**, p. 046704
19. S. Alder, S. Trebst, A. K. Hartmann, and M. Troyer (2004) Dynamics of the Wang-Landau algorithm and Complexity of rare events for the three-dimensional bimodal Ising spin glass. *J. Stat. Mech.* P07008
20. S. Trebst, D. A. Huse, and M. Troyer (2004) Optimizing the ensemble for equilibration in broad-histogram Monte Carlo simulations. *Phys. Rev. E* **70**, p. 046701
21. S. Trebst, E. Gull, and M. Troyer (2005) Optimized ensemble Monte Carlo simulations of dense Lennard-Jones fluids. *J. Chem. Phys.* **123**, p. 204501
22. R. H. Swendsen and J. Wang (1986) Replica Monte Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett.* **57**, p. 2607
23. E. Marinari and G. Parisi (1992) Simulated tempering: A new Monte Carlo scheme. *Europhys. Lett.* **19**, p. 451
24. A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov (1992) *J. Chem. Phys.* **96**, p. 1776
25. K. Hukushima and Y. Nemoto (1996) Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.* **65**, p. 1604
26. H. G. Katzgraber, S. Trebst, D. A. Huse, and M. Troyer (2006) *J. Stat. Mech* p. P03018
27. S. Trebst, M. Troyer, and U. H. E. Hansmann (2006) Optimized parallel tempering simulations of proteins. *J. Chem. Phys.* **124** p. 174903
28. J. C. McKnight, D. S. Doering, P. T. Matsudaira, and P. S. Kim (1996) A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* **260**, p. 126
29. Y. Duan and P. A. Kollman (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **282**, p. 740
30. B. Zagrovic, C. D. Snow, S. Khaliq, M. R. Shirts, and V. S. Pande (2002) Native-like mean structure in the unfolded ensemble of small proteins. *J. Mol. Biol.* **323**, p. 153
31. C.-Y. Liu, C.-K. Hu, and U. H. E. Hansmann (2003) Parallel tempering simulations of HP-36. *Proteins: Struct., Funct., Genet.* **52**, p. 436
32. U. H. E. Hansmann (2004) Simulations of a small protein in a specifically designed generalized ensemble. *Phys. Rev. E* **70**, p. 012902
33. M. J. Sippl, G. Némethy, and H. A. Sheraga (1984) Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O-H...O=C hydrogen bonds from packing configurations. *J. Phys. Chem.* **88**, p. 6231
34. T. Ooi, M. Oobatake, G. Nemethy, and H. A. Scheraga (1987) Accessible surface -areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc. Natl. Acad. Sci.* **84**, p. 3086
35. M. Troyer, S. Wessel, and F. Alet (2003) Flat histogram methods for quantum systems: algorithms to overcome tunneling problems and calculate the free energy. *Phys. Rev. Lett.* **90**, p. 120201

36. M. Troyer, F. Alet, and S. Wessel (2004) Histogram methods for quantum systems: from reweighting to Wang-Landau sampling. *Braz. J. of Physics* **34**, p. 377
37. R. Feynman (1953) Atomic theory of liquid helium near absolute zero. *Phys. Rev.* **91**, p. 1301
38. H. Trotter (1959) On the product of semi-groups of operators. *Proc. Am. Math. Soc.* **10**, p. 545
39. M. Suzuki (1976) Relationship between d-dimensional quantal spin systems and (d+1)-dimensional Ising systems – Equivalence, Critical Exponents and Systematic Approximants of the Partition Function and Spin Correlations. *Prog. Theor. Phys.* **56**, p. 1454
40. N. V. Prokofev, B. V. Svistunov, and I. S. Tupitsyn (1998) Exact, complete, and universal continuous-time worldline Monte Carlo approach to the statistics of discrete quantum systems. *JETP* **87**, p. 310
41. A. Sandvik and J. Kurkijärvi (1991) Quantum Monte Carlo simulation method for spin systems. *Phys. Rev. B* **43**, p. 5950
42. D. Handscomb (1962) The Monte Carlo method in quantum statistical mechanics. *Proc. Cambridge Philos. Soc.* **58**, p. 594
43. S. Sachdev, P. Werner, and M. Troyer (2004) Universal conductance of quantum wires near the superconductor-metal quantum transition. *Phys. Rev. Lett.* **92**, p. 237003
44. P. Werner, K. Völker, M. Troyer, and S. Chakravarty (2005) Phase diagram and critical exponents of a dissipative Ising spin chain in a transverse magnetic field. *Phys. Rev. Lett.* **94**, p. 047201
45. E. L. Pollock and D. M. Ceperley (1987) Path-integral computation of superfluid densities. *Phys. Rev. B* **36**, p. 8343
46. M. Jarrell and J. Gubernatis (1996) Bayesian inference and the analytic continuation of imaginary time Monte Carlo data. *Physics Reports* **269**, p. 133
47. W. von der Linden (1995) Maximum-entropy data analysis. *Applied Physics A* **60**, p. 155
48. K. S. D. Beach (2004) Identifying the maximum entropy method as a special limit of stochastic analytic continuation. *cond-mat/0403055*
49. M. Suzuki, S. Miyashita, and A. Kuroda (1977) Monte Carlo simulation of quantum spin systems. I. *Prog. Theor. Phys.* **58**, p. 1377
50. N. V. Prokofev, B. V. Svistunov, and I. S. Tupitsyn (1996) Exact quantum Monte Carlo process for the statistics of discrete systems. *JETP Lett.* **64**, p. 911
51. M. S. Makivić and H. Q. Ding (1991) Two-dimensional spin-1/2 Heisenberg antiferromagnet: A quantum Monte Carlo study. *Phys. Rev. B* **43**, p. 3562
52. H. G. Evertz, G. Lana, and M. Marcu (1993) Cluster algorithm for vertex models. *Phys. Rev. Lett.* **70**, p. 875
53. B. Beard and U. Wiese (1996) Simulations of discrete quantum systems in continuous Euclidean time. *Phys. Rev. Lett.* **77**, p. 5130
54. H. G. Evertz (2003) The loop algorithm. *Adv. in Physics* **52**, p. 1
55. N. Kawashima and K. Harada (2004) Recent developments of world-line Monte Carlo methods. *J. Phys. Soc. Jpn.* **73**, p. 1379
56. N. Kawashima and J. Gubernatis (1994) Loop algorithms for Monte Carlo simulations of quantum spin systems. *Phys. Rev. Lett.* **73**, p. 1295
57. N. Kawashima and J. Gubernatis (1995) Generalization of the Fortuin-Kasteleyn transformation and its application to quantum spin simulations. *J. Stat. Phys.* **80**, p. 169

58. K. Harada, M. Troyer and N. Kawashima (1998) The two-dimensional spin-1 quantum Heisenberg antiferromagnet at finite temperatures. *J. Phys. Soc. Jpn.* **67**, p. 1130
59. S. Todo and K. Kato (2001) Cluster algorithms for general-S quantum spin systems. *Phys. Rev. Lett.* **87**, p. 047203
60. N. Kawashima (1996) Cluster algorithms for anisotropic quantum spin models. *J. Stat. Phys.* **82**, p. 131
61. A. Sandvik (1999) Stochastic series expansion method with operator-loop update. *Phys. Rev. B* **59**, p. R14157
62. A. Dorneich and M. Troyer (2001) Accessing the dynamics of large many-particle systems using the stochastic series expansion. *Phys. Rev. E* **64**, p. 066701
63. O. Syljuasen and A. W. Sandvik (2002) Quantum Monte Carlo with directed loops. *Phys. Rev. E* **66**, p. 046701
64. U.-J. Wiese and H.-P. Ying (1992) Blockspin cluster algorithms for quantum spin systems. *Phys. Lett. A* **168**, p. 143
65. B. Frischmuth, B. Ammon, and M. Troyer (1996) Susceptibility and low-temperature thermodynamics of spin-1/2 Heisenberg ladders. *Phys. Rev. B* **54**, p. R3714
66. M. Greven, R. J. Birgeneau, and U. J. Wiese (1996) Monte Carlo study of correlations in quantum spin ladders. *Phys. Rev. Lett.* **77**, p. 1865
67. M. Troyer, M. Imada, and K. Ueda (1997) Critical exponents of the quantum phase transition in a planar antiferromagnet. *J. Phys. Soc. Jpn.* **66**, p. 2957
68. B. B. Beard, R. J. Birgeneau, M. Greven, and U.-J. Wiese (1998) Square-lattice Heisenberg antiferromagnet at very large correlation lengths. *Phys. Rev. Lett.* **80**, p. 1742
69. C. Yasuda, S. Todo, K. Hukushima, F. Alet, M. Keller, M. Troyer, and H. Takayama (2005) Néel temperature of quasi-low-dimensional Heisenberg antiferromagnets. *Phys. Rev. Lett.* **94**, p. 217201
70. D. C. Johnston, M. Troyer, S. Miyahara, D. Lidsky, K. Ueda, M. Azuma, Z. Hiroi, M. Takano, M. Isobe, Y. Ueda, M. A. Korotin, V. I. Anisimov, A. V. Mahajan, and L. L. Miller (2000) Magnetic susceptibilities of spin-1/2 antiferromagnetic Heisenberg ladders and applications to ladder oxide compounds. *cond-mat/0001147*
71. D. C. Johnston, R. K. Kremer, M. Troyer, X. Wang, A. Klümper, S. L. Budko, A. F. Panchula, and P. C. Canfield (2000) Thermodynamics of spin S=1/2 antiferromagnetic uniform and alternating-exchange Heisenberg chains. *Phys. Rev. B* **61**, p. 9558
72. R. Melzi, P. Carretta, A. Lascialfari, M. Mambrini, M. Troyer, P. Millet, and F. Mila (1999) $Li_2VO(Si,Ge)O_4$, a prototype of a two-dimensional frustrated quantum Heisenberg antiferromagnet. *Phys. Rev. Lett.* **85**, p. 1318
73. M. A. Korotin, I. S. Elfimov, V. I. Anisimov, M. Troyer, and D. I. Khomskii (1998) Exchange interactions and magnetic properties of the layered vanadates CaV_2O_5, MgV_2O_5, CaV_3O_7, and CaV_4O_9. *Phys. Rev. Lett.* **83**, p. 1387
74. F. Woodward, A. Albrecht, C. Wynn, C. P. Landee, and M. Turnbull (2002) Two-dimensional S= 1/2 Heisenberg antiferromagnets: Synthesis, structure, and magnetic properties. *Phys. Rev. B* **65**, p. 144412
75. G. Schmid, S. Todo, M. Troyer, and A. Dorneich (2002) Finite-temperature phase diagram of hard-core bosons in two dimensions. *Phys. Rev. Lett.* **88**, p. 167208

76. O. Nohadani, S. Wessel, B. Normand, and S. Haas (2004) Universal scaling at field-induced magnetic phase transitions. *Phys. Rev. B* **69**, p. 220402
77. A. Ferrenberg and R. Swendsen (1988) New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* **61**, p. 2635
78. A. Ferrenberg and R. Swendsen (1989) Optimized Monte Carlo data analysis. *Phys. Rev. Lett.* **63**, p. 1195
79. A. Sandvik (1998) Critical temperature and the transition from quantum to classical order parameter fluctuations in the three-dimensional Heisenberg antiferromagnet. *Phys. Rev. Lett.* **80**, p. 5196
80. A. Sandvik (1994) Order-disorder transition in a two-layer quantum antiferromagnet. *Phys. Rev. Lett.* **72**, p. 2777