National Research University Higher School of Economics

As a manuscript

Fadeeva Marina Alexandrovna

On the properties of Wang-Landau algorithm

Dissertation summary

for the purpose of obtaining academic degree Doctor of Philosophy in Applied Mathematics

Academic supervisor:
Doctor of Sciences in Physics and Mathematics
professor Shchur L.N.

Moscow – 2023

Problem statement

Relevance

High-performance computing has been gaining increasing importance in both scientific tasks and practical applications. These computations are most efficiently performed on hybrid supercomputer systems. The development of algorithms and methods to utilise the full power of hybrid high-performance computing systems is one of the important task. One of the promising approach for modeling tasks is the method of direct calculation of the density of states known as the Wang-Landau method. This method is applicable to a wide range of problems, including polymer physics, spin systems, and optimization tasks. Despite its widespread application (with over 3 thousand citations of the original paper [1] on Google Scholar), the method in its original form exhibits limited accuracy. The study of the accuracy and applicability of this method for solving problems on high-performance computational systems is undoubtedly a important issue. Thus, this project is cross-disciplinary due to involving the development of computation methods and the application of the approaches and algorithms to test theoretical hypotheses in statistical physics.

Aims and tasks of the study

The aim of the study is to develops new methods for studying of computational problems of statistical physics. The objectives of the study are:

- Develop a numerical method for studying problems of statistical physics with adjustable accuracy.
- Estimate the tunneling time and mixing time in the Wang-Landau algorithm.
- Develop a parallel algorithm with controlled precision for the Wang-Landau method.
- Apply the developed modification of the algorithm to study models of statistical mechanics.

Degree of problem development

In statistical mechanics, a partition function has an important role in the study of the properties of various systems. A partition function describes the system in the state of thermodynamic equilibrium. For a canonical ensemble partition function is defined as

$$Z = \sum_j e^{-E_j/k_B T} \tag{1}$$

where $E_i = \mathcal{H}(X_i)$ is Hamiltonian or energy, that describes the state of the system in the configuration X_i; k_B - Boltzmann's constant; T - temperature. Thermodynamic parameters of the system can be obtained from the partition function and its derivatives. For systems with a small number of interacting particles, the partition function can be computed analytically. However, the number of all possible configurations grows as 2^N even for a system composed of N elements. Consequently, the computational problem becomes extremely complex. Therefore methods for numerically approximating the partition function are used.

Such methods include Monte Carlo algorithms, which are based on various representations of the partition function. Each representation of the partition function can be associated with a numerical algorithm for its computation.

The canonical notation (1) of the partition function is used in the first method of the Monte Carlo family, the Metropolis algorithm [2]. This algorithm belongs to the class of Monte Carlo methods with Markov chains. A new state of the system X_m is generated from the previous one X_k with transition probability that depends on the energy difference between the initial and final states $\Delta E_{km} = E(X_m) - E(X_k)$. Here $E(X_m)$ is the energy corresponding to the configuration X_m and $E(X_k)$ is the energy corresponding to the configuration X_k. The probability of transition to a new state is determined by the Metropolis probability

$$P = min[1, e^{\Delta E_{km}/k_B T}] \tag{2}$$

Multicanonical Monte-Carlo (or MuCa) is based on multicanonical representation

$$Z = \sum_k^{N_E} g(E_k) W(E_k), \tag{3}$$

where $g(E_k)$ is density of states with energy E_k, $W(E_k)$ - weight function.

The cluster representation of the partition function [3]

$$Z = \sum_{\text{bonds}} p^b (1-p)^n q^{N_c} \tag{4}$$

enables the introduction of the Wolff algorithm [4] and the Swendsen-Wang algorithm [5]. Here, $p = 1 - \exp^{-J/k_B T}$, where J is the interaction constant between system elements; N_c is the number of clusters, and q is the number of components in the system.

Wang-Landau algorithm

The Wang-Landau algorithm was proposed in 2001 and is described in [1, 6]. It provides direct computation of the density of states of a system. The algorithm is applied to any discrete system that can be described by a finite set of configuration states $\mathbf{X} = X_1, X_2, \ldots, X_m$. Each configuration corresponds to a different energy level $E(X_i)$. For such systems it is possible to pass from the summary by configurations to the summary by energy levels in the notation of the partition function:

$$Z = \sum_j e^{-E_j/k_B T} = \sum_{E_n} g(E_n) e^{-E_n/k_B T}, \tag{5}$$

where k_B - Boltzmann's constant, T - temperature. Function $g(E_n)$ is a density of states (DoS). DoS describes the number of states with energy E_n. Since $g(E_n)$ is independent of temperature, this representation of the partition function makes it possible to calculate, for example, the free energy and heat capacity at any value of temperature

$$E(\beta) = \langle E \rangle = \frac{\sum_{i=0}^{N_E} E_i g(E_i) e^{-\beta E_i}}{\sum_{i=0}^{N_E} g(E_i) e^{-\beta E_i}}, \tag{6}$$

$$C(\beta) = \beta^2 (\langle E^2 \rangle - \langle E \rangle^2). \tag{7}$$

The DoS is calculated by random walk with Wang-Landau probability over the energy space of the system. During the walk, two histograms [1] are accumulated. The first is the current value of the logarithm of the density of states $S(E) = \log(g(E))$. The second is the auxiliary histogram $H(E)$, which contains information about the number of visits to each energy level. At the start of the algorithm, $H(E)$ is initialised with zeros and $S(E)$ with ones. The initial configuration of the system is set, and its energy E_k is calculated. The initial value of the parameter $f = \exp(1)$ is set. Further steps of the algorithm are arranged as follows: 1) the state of the system is changed and the energy of the new state E_m is calculated; 2) the transition from the state with energy E_k to the state with energy E_m occurs with the Wang-Landau probability:

$$P_{WL}(E_k, E_m) = \min\left(1, \frac{\tilde{S}(E_k)}{\tilde{S}(E_m)}\right), \tag{8}$$

where $\tilde{S}(E_k) = \log(\tilde{g}(E_k))$ is the current estimation of DoS. The next step is to update the auxiliary histogram $H(E_k) \to H(E_k) + 1$, and the current estimation $\tilde{S}(E_k) \to \tilde{S}(E_k) f$. Step 1 and 2 are repeated until the histogram $H(E_k)$ is sufficiently "flat" (e.g., at the 5% [1] level). The value of the parameter f is then updated as a function of the square root $f_i = \sqrt{f_{i-1}}$, and the histogram is reset $H(E_k) = 0$. Steps 1 and 2 are then repeated again. The algorithm terminates when the parameter f reaches some desired value of f_{end}.

Modifications of the Wang-Landau algorithm

The Wang-Landau algorithm has been widely used in various fields of science. The method is used for modeling polymers [7, 8] and protein chains [9, 10], for optimization [11]. However several questions regarding the performance and accuracy of the algorithm required an explanation, such as: how to choose the optimal "flatness" value of the histogram $H(E)$? How does the "flatness" requirement of the histogram $H(E)$ affect the final accuracy of the computation? Is the choice of the square root function optimal for f?

The last question was answered in two independent papers [12] and [11] suggesting a modification of the 1/t-WL algorithm and the Stochastic approximation Monte Carlo (SAMC) algorithm.

1/t-Wang-Landau modification was proposed [12, 13]. The modified algorithm consists of two steps. The first stage is similar to Wang-Landau steps [1] except that the test of the histogram $H(E)$ "flattness" was replaced with test $H(E) \neq 0$. The function $F = \ln f$ also remains the same $F_i = F_{i-1}/2$ as long as the condition $F_i \leq N_E/t$ is satisfied. When $F_i > N_E/t$ the algorithm proceeds to the second stage, the histogram $H(E)$ is no longer checked. The parameter F is updated according to the new rule $F_i = N_E/t$. Here, t is the number of elementary spin flips and N_E is the number of levels in the system.

The modification of SAMC [11, 14] is similar to the one described above [12, 13]. An additional parameter t_0 is introduced, representing the time at which the algorithm transitions to the second stage, measured in elementary spin flips. The parameter F is updated as a function of $F = \frac{t_0}{max(t,t_0)}$.

The convergence conditions for the function $F = 1/t$ are formulated based on the theory of stochastic approximation [11]

$$\sum_{i=0}^{\infty} F_i = \infty, \sum_{i=0}^{\infty} F_i^{\zeta} < \infty, \tag{9}$$

where $\zeta \in (1, 2)$. It was proven [15] that the most optimal choice of the function is $F(t) \propto 1/t^{\alpha}$ with $1 \leq \alpha < 2$. Moreover, the results of numerical experiments indicate that $\alpha = 1$ is the best option.

The 1/t-WL modification resolved the problem with the systematic error of the algorithm. As the number of algorithm steps increased, the deviation from the exact value reached a constant value. Some of the first descriptions of this problem were provided by authors in the articles [16, 17]. The presence of a statistical error proportional to the square root $\sqrt{f}$ was demonstrated. The article [12] also demonstrated the limited accuracy of calculations using the Ising model as an example. The deviation $\epsilon(E,t) = |1 - \frac{\ln g_n(E,t)}{\ln g_{ex}(E)}|$ of the estimated value of the density of states $g_n(E,t)$ from the exact $g_{ex}(E)$ was calculated at different number of algorithmic steps t. The numerical results showed the error rate reaches a constant after a certain value of the parameter $F = \ln(f)$, and does not decrease with increasing number of algorithm steps. A simi-

lar numerical experiment for the 1/t-WL method demonstrated that $\epsilon(E,t)$ decreases with the number of steps proportionally to $1/t$.

A review of other modifications also shows that the condition (9) imposed on F is critical in the convergence of the method. For example, the [16] presented suggestions regarding the importance of the "plane" of the auxiliary histogram $H(E)$ in the accuracy of the computation, and proposed the idea of fixing the parameter $F = 1$ after 14 iterations. Numerical experiments have demonstrated that this approach gives a meaningful improvement in accuracy only at the early steps of the algorithm compared to the original [1] method; however, as the number of steps increases, the algorithm does not give a significant improvement.

The frequency of updating the values of $g(E)$ is also not a critical factor in the accuracy of computations. In the article [18], the authors propose reducing the frequency of updating the values of $g(E)$ in the Wang-Landau algorithm from every spin flip to every Monte Carlo step, i.e., updating $g(E)$ after every L^2 elementary spin flips. The performance of this approach is compared with the original Wang-Landau algorithm with histogram flatness levels $H(E)$ of 80% and 90%, as well as with the 1/t-modification. Based on the comparison of the results of calculations of three characteristics of the system, specifically the temperature of the maximum heat capacity T_c and the critical indices β and γ, the authors conclude the accuracy of the Wang-Landau algorithm is improved when the density of states is updated more rarely compared to the original algorithm and the 1/t-WL modification. However, this approach requires several times more CPU time. For example, if we compare the results obtained with the 1/t-WL algorithm with the stopping criterion $\ln f = 5 \cdot 10^{-8}$ and the proposed modification ($T_c = 2.26916(12)$, $\beta = 0.1259(21)$ and $T_c = 2.26904(25)$, $\beta = 0.12494(68)$, respectively). then the latter takes about 7 times longer [18].

Two characteristic times of the Wang-Landau algorithm

There are two characteristic times in the Wang-Landau algorithm, which are tunneling time and mixing time.

Tunneling time is related to the initial stage of the 1/t-WL algorithm, and is a characteristic time for algorithms that require the histogram [19] criterion of "flatness". This time is also known as first-passage time [20]. In terms of the Wang-Landau method, this is the time for which the algorithm reaches the level with maximum energy E_{max}, starting from the minimum E_{min}, for the first time. It is known that for random walk, the tunneling time scales as $t_{tun} \approx N_E^2$ [20], i.e., for a two-dimensional Ising model with lattice size L, it is $t_{tun} \approx L^4$. An estimation of this time for the Wang-Landau algorithm for the Ising model is given in [19].

In the final stage of the algorithm, the estimation of $\tilde{g}(E)$ is in the vicinity of the exact value of $g(E)$, and the steps of the algorithm are a Markov random walk over the energy spectrum. The time in which the Markov process converges to an equilibrium state is called the mixing time t_{mix}. It is known that this time is inversely proportional to the spectral gap G, defined by the difference between the higher eigenvalues of the

transition matrix [35]

$$t_{mix} = \frac{1}{G} = \frac{1}{\lambda_1 - \lambda_2}. \tag{10}$$

Estimation of this time for the Wang-Landau algorithm has not been done before.

Parallel implementations of the Wang-Landau algorithm

Among the existing approaches to the realization of the parallel Wang-Landau algorithm, two groups can be distinguished:

- Algorithms where random walks are performed across the entire spectrum and shared or distributed memory is used [21–23].
- Algorithms where the energy spectrum is divided into segments, and independent random walks are conducted in each segment using distributed memory [24] - [33].

We will further examine the proposed implementations of both approaches.

Random walks across the entire energy spectrum.

Let us start with the case when several random walks are performed over the entire energy spectrum. There are implementations of this approach for systems with distributed memory using MPI [21] and for shared memory using OpenMP [22].

In case of MPI [21] distributed memory, a specific time interval is set for communication between independent random walks to update the collective DoS value. An aggregated histogram across all walks is used to assess the histogram's flatness. Determining the optimal time interval for information exchange between walks is a weak point of this approach, as it could lead to computational inefficiency.

The implementation of the algorithm using OpenMP [22] is organized as follows: multiple copies of the system are created, each corresponding to its own initial state of the random number generator. Within each copy, standard WL (Wang-Landau) steps are executed, involving the reading and modification of the state density $g(E)$ and the flatness histogram $H(E)$. These entities are shared and writable by each copy. The histogram $H(E)$ flatness checking step is also performed by all processes; however, the permission to modify the F parameter is granted only to the processor where the flatness condition is met with a certain accuracy. Afterward, the F parameter becomes readable to all processes.

This approach to parallelization of the algorithm is quite simple, as it doesn't require rewriting the entire algorithm's logic and instead involves adding directives defining the work with shared memory. Specifically, to establish the order of operations on $g(E)$ and $H(E)$ and avoid premature overwriting of their values, ATOMIC or CRITICAL directives are used. As the experiments of the authors of [22] have shown, the CRITICAL directive slows down the algorithm to a point where its performance lags behind a regular sequential implementation. On the other hand, using the ATOMIC

directive results in speed improvements for the parallel implementation. However, the assumption of premature overwriting of common elements gives the most time-efficient implementation. The lost values of $g(E)$ and $H(E)$ due to premature overwriting are compensated for by additional algorithm steps, which require less time than waiting in the queue for updates the DoS value and the $H(E)$ histogram. According to the authors, this approach doesn't lead to a loss of accuracy. However, as it is shown in [23] this may be true only for some models and doesn't guarantee the convergence of the density of states to the accurate solution. This is due to the fact that while all energy levels are equally probable, the likelihood of visiting those levels is not equal. Some of the parallel random walks "get stuck" at low energy levels, and contribute to the overall histogram $H(E)$ shared among all walks. However, they do not traverse the entire spectrum properly, thereby overestimating the low-energy levels. A solution could be a stricter "levelness" criterion, which in turn increases the running time of the algorithm and loses the advantage of a parallel implementation over a sequential one. Therefore, to solve this problem, the authors propose to use a non-uniform parameter f: the less likely a level is to be visited, the higher the value of parameter f is used to adjust the current DoS.

Random walks across segments of the energy spectrum

In 2013, T. Vogel and colleagues [24–26] introduced a parallel implementation of the Wang-Landau algorithm by dividing the energy spectrum into overlapping segments that can exchange configurations with each other.

The concept of exchanging configurations is borrowed from the Replica Exchange Monte Carlo algorithm [27], where multiple copies of a single system are simulated in parallel at different temperatures. Periodically, configurations are exchanged between replicas, so that replicas with low temperature receive the configurations of the system with high temperatures and vice versa. This approach helps overcome the issue of getting trapped in local minimum at low temperatures and increase the accuracy of calculating thermodynamic properties of the system [28]. The independent nature of simulating each replica allows for scalability on high-performance clusters, and it provides system characteristics across a specified temperature range.

In the proposed Replica Exchange Wang-Landau algorithm (REWL) by Vogel [24, 25], the energy spectrum is divided into h overlapping intervals. Within each of these h intervals, m independent random walks are conducted using the classical Wang-Landau algorithm. Consequently, the density of states $g(E)$ and the flatness histogram $H(E)$ are computed independently. After a certain number of steps, configurations are exchanged between two random walks from neighboring windows with probability satisfying the detailed equilibrium principle, denoted as P_{acc}:

$$P_{acc} = min[1, \frac{g_i(E(X))}{g_i(E(Y))} \frac{g_j(E(Y))}{g_j(E(X))}] \tag{11}$$

i and j are randomly selected random walks from neighboring intervals, X and Y are

their corresponding configurations at a given time, $E(X)$ and $E(Y)$ are the energy of the system with configuration X and Y.

It's important to pay attention to the choice of the overlap fraction o between the intervals h. The value of o should be chosen to avoid two extremes. On one hand, a high overlap fraction would be ineffective as it would result in conducting multiple walks almost across the entire energy spectrum. In such a case, parallelizing the algorithm would lose its purpose. On the other hand, an excessively small overlap fraction wouldn't provide a sufficient level of probability for exchanging configurations between windows. The authors of the algorithm suggest selecting the overlap fraction at the level of 75% [25, 26, 32, 33].

The scientific novelty of the study

1. A new mathematical object was introduced for the Wang-Landau method - the transition matrix across the discrete energy spectrum.

2. It has been analytically and numerically demonstrated that the transition matrix converges to a stochastic matrix when the required accuracy in calculating the density of states is achieved.

3. A criterion for the accuracy of calculation of density of states has been proposed as the deviation of the largest eigenvalue of the transition matrix from unity.

4. A method for estimating the scaling of density of states calculations is proposed. The concept of mixing time has been introduced as an estimate of the scalability of computation time.

5. A parallel modification with controlled accuracy for the Wang-Landau method was introduced.

6. The developed technique is applied to the study of the four-component Potts model on a hexagonal lattice, resolving the discussion about referring this model to the universality class of the four-component Potts model.

Summary

In the **first chapter**, the modification of the 1/t-Wang-Landau algorithm is discussed by introducing a transition matrix between energy levels with the Wang-Landau probability.

The auxiliary function $H(E)$ in the Wang-Landau algorithm accumulates information about the number of visits to each energy level, however, this function does not represent information about the energy levels between which the transition was made. An additional function $U(E_k, E_m)$ was introduced in the algorithm. It is a square matrix of size $N_E \times N_E$, where N_E is the number of energy levels in the system. At the start of the algorithm, the matrix is initialized with zeros. During the random walk, when we move from energy level E_k to energy level E_m, the corresponding matrix element increases by one $U(E_k, E_m) \to U(E_k, E_m) + 1$. We also compute the normalized matrix $T(E_k, E_m) = U(E_k, E_m)/H$, where $H = \sum_{k,m} U(E_k, E_m)/N_E$. The simulation results show that as the algorithm converges to the exact value of $g(E)$, the normalized matrix $T(E_k, E_m)$ tends to become stochastic, i.e., the sum of its elements on both columns and rows is close to unity. Correspondingly, its largest eigenvalue tends to $\lambda_1 \to 1$ [34].

The analytical solution of the matrix was obtained for the one-dimensional Ising model with periodic boundary conditions. It was shown that the matrix elements are probabilities of transition from level E_k to level E_m and the total probability is composed of two factors

$$T(E_k, E_m) = \min\left(1, \frac{g(E_k)}{g(E_m)}\right) P(E_k, E_m). \tag{12}$$

The first component is the Wang-Landau probability P_{WL} to move from one energy level to another. $P(E_k, E_m)$ is the probability of transition from the configuration of the system with energy E_k to the configuration of the system with energy E_m. The complete system of expressions for the non-diagonal elements of the matrix $T(E_k, E_m)$:

$$\begin{aligned} T(E_k, E_m) = \min\left(1, \frac{g(E_k)}{g(E_m)}\right) \sum_{i=0}^{2k} \frac{N_i(k, L) Q_i^{E_k \to E_m}}{g(E_k)}, \\ N_i(k, L) = \frac{L}{k} C_{2k-i}^{i} C_{L-2k-1}^{2k-i-1}, \\ N_2 k(k, L) = 2\delta_{L,2k}, \\ Q_i^{E_k \to E_{k-1}} = \frac{i}{L}, \\ Q_i^{E_k \to E_{k+1}} = \frac{L - 4k + i}{L}, \\ Q_i^{E_k \to E_k} = \frac{4k - 2i}{L}, \\ g(E_k) = 2C_L^{2k}. \end{aligned} \tag{13}$$

The analytical solution based on the exact value of DoS also confirms that the sum of elements in both rows and columns is equal to one. Based on this property, we propose a control of the accuracy of the algorithm as difference of the largest eigenvalue of the transition matrix in the energy space from unity

$$\delta = |1 - \lambda|. \tag{14}$$

It was demonstrated that with an exact value of $g(E)$, the Wang-Landau algorithm becomes a Markov process.

In **second chapter** the estimation of characteristic times of the modified algorithm is analyzed and the methodology of their calculation is given. The tunneling time has been estimated using a two-dimensional Ising model with periodic boundary conditions as an example. It was found that this time scales with increasing system size as $t_{tun} = L^{=4.743(7)}$.

Since the implementation of the transition matrix into the algorithm, it became possible to estimate the characteristic time for the second stage of the 1/t-WL algorithm. It is known that this time is inversely proportional to the difference between the largest eigenvalues of the transition matrix: $t_{mix} = \frac{1}{|\lambda_1 - \lambda_2|}$ [35]. For the one-dimensional Ising model, the elements of the matrix $T(E_k, E_m)$ are known (see equation 13). It was determined that $t_{mix} = L^{2.19}$. For the two-dimensional Ising model, the exact solution of the matrix is unknown. Therefore, Wang-Landau random walks were run with an initially given exact value of $g(E)$ (computed with [36]). $g(E)$ was kept unchanged while the transition matrix $U(E_k, E_m)$ had been accumulated during the algorithm. The normalized matrix $\tilde{T}(E_k, E_m)$ was used to compute the eigenvalues. It was found that the mixing time grows as $t_{mix} = L^{4.28(4)}$ for the two-dimensional Ising model.

Tunneling and mixing times were also found for the q-component Potts model, and the results are shown in Table . For the estimation of mixing time, first the $\tilde{g}(E)$ estimate was found using the 1/t-WL algorithm with transition matrices, then the algorithm steps were run with $\tilde{g}(E)$ as initial values.

q	t_{tun}	t_{mix}
3	5.(3)	3.(3)
4	4.9(4)	3.(3)
5	4.7(2)	1.(7)
8	4.(9)	1.7(6)
10	4.8(1)	1.(3)

Table 1: Tunneling and mixing times of the Wang-Landau algorithm for the Potts model on a square lattice with different numbers of q components.

In the **third chapter**, an efficient implementation approach for the parallel Wang-Landau algorithm is discussed. The modification is based on the concept of random walks across segments (windows) of the energy spectrum, which intersect with each other at a certain overlap fraction $o\%$. Within each window, in addition to functions

$g(E)$ and $H(E)$, a transition matrix $\tilde{T}(E_k, E_m)$ is accumulated. Periodically, the values of the DoS for the entire spectrum are updated using the formula: $G_i = \frac{1}{k}\sum_{j=0}^{k} g^j(E_i)$, where k is the number of windows where the energy level E_i appears, and $g_j(E_i)$ is the accumulated DoS value of the energy level E_i in window j. Simulation results demonstrated that both for the entire energy spectrum and for the windows, the condition of convergence of the matrix to a stochastic one holds. Within each window, as the number of steps t increases, the largest eigenvalue tends to approach unity. Within each window, tunneling time was calculated in two ways: based on a uniform distribution ($g(E_i) = 1$) and based on the exact value ($g(E_i) = g_{exact}(E_i)$) [36]. In both cases we estimated tunneling time starting from left edge of the window and ending on the right, and vice versa. The results shed light on the characteristics of tunneling time distribution within spectrum windows at the initial stages of the algorithm and on its final stage. By using the example of calculating the heat capacity $C(\beta)$ for the Ising model and its relative error, we demonstrated that the proposed approach for computing the density of states is simpler to implement compared to the one suggested earlier in [24], while still providing sufficient computational accuracy.

The **fourth chapter** presents the results of applying the modified method to solve problems in statistical physics. It's known that the Potts model on a square lattice with a number of components $q > 4$ exhibits a first-order phase transition, whereas for $q \leq 4$, it's a second-order transition. The motivation for this study arises from several publications (e.g., [37]) that claim the existence of a first-order phase transition in the four-component Potts model on a hexagonal lattice. This contradicts the universality theory [38], which states that the type of phase transition is unaffected by local configurations, but is influenced by the dimensionality of the system and its Hamiltonian. By employing the modified 1/t - Wang-Landau method with criterion accuracy as $\delta = |1 - \lambda_1|$, we calculated the density of states $g(E)$ for the four-component Potts model on a hexagonal lattice. Using $g(E)$, functions of temperature β such as energy $E(\beta)$, heat capacity $C(\beta)$, and Binder cumulant $B_E(\beta)$ were calculated using the formulas

$$E(\beta) = \langle E \rangle = \frac{\sum_{i=0}^{N_E} E_i g(E_i) e^{-\beta E_i}}{\sum_{i=0}^{N_E} g(E_i) e^{-\beta E_i}}, \tag{15}$$

$$\langle E^2 \rangle = \frac{\sum_{i=0}^{N_E} E_i^2 g(E_i) e^{-\beta E_i}}{\sum_{i=0}^{N_E} g(E_i) e^{-\beta E_i}}, \tag{16}$$

$$C(\beta) = \beta^2(\langle E^2 \rangle - \langle E \rangle^2), \tag{17}$$

$$B_E(\beta) = 1 - \frac{\langle E^4 \rangle}{3\langle E^2 \rangle^2} \tag{18}$$

It is a widely recognized fact that models belonging to this universality class display specific heat with multiplicative logarithmic corrections. The the result of approximation of the numerical data of the maximum specific heat is in good agreement with the analytical value [39]. The analysis results demonstrate that the Potts model with $q = 4$ components on a hexagonal lattice are consistent with the universality theory.

Thesis statements for defending

- A new element of the Wang-Landau method is introduced - the transition matrix over a discrete energy space. It has been demonstrated that the matrix tends to become stochastic as density of states calculations approach the exact value.
- An approach has been developed to control the accuracy of calculating the density of states in the Wang-Landau algorithm.
- For the first time, the characteristic time in the Wang-Landau algorithm, the mixing time, has been estimated.
- A parallel implementation of the Wang-Landau algorithm with adjustable accuracy is proposed. In addition, a simpler algorithmic method for combining the density of states over the entire spectrum is proposed.
- The developed method is used to determine the nature of the phase transition in the four-component Potts model on a hexagonal lattice.

Personal contribution of the author

The ideas and hypotheses discussed in the dissertation were developed jointly with the supervisor. The dissertation author personally developed the software code, conducted computational experiments, and writing the text of the articles.

General findings of the study

- The transition matrix of random walk with the Wang-Landau probability was introduced. The frequency of transitions from states with one energy to a state with another energy is accumulated. Expressions for the transition matrix for the one-dimensional Ising model with periodic boundary conditions are obtained analytically. It is shown that the matrix elements are transition probabilities from the level E_m to E_k, and the transition probability consists of two components: 1) the Wang-Landau transition probability $P_{wl} = \min\left(1, \frac{g(E_m)}{g(E_k)}\right)$; 2) the probabilities of being at a given energy level. It was shown that the normalized matrix tends to be stochastic as the number of simulation steps increases, respectively, the largest eigenvalue a matrix tends to unity $\lambda_1 \to 1$. Based on this property, we propose an accuracy criterion as the difference of the largest eigenvalue from unity: $\delta = |\lambda_1 - 1|$.

- Mixing time and tunneling time of the Wang-Landau algorithm were estimated. A dependence between the system size and the tunneling time was found. An estimation of the mixing time for the Ising model is done using the spectral gap $G = |\lambda_1 - \lambda_2|$ of the transition matrix with Wang-Landau probability. The mixing time $t_{mix} = 1/G$, is an estimation of the characteristic time in the final stage of the 1/t Wang-Landau algorithm.

- A parallel implementation of the Wang-Landau algorithm is proposed. It's based on the transition matrix and a more simplified approach to combine the density of states of the entire spectrum. It is shown that the transition matrices within each window also tends to become stochastic like the transition matrix in a sequential realization. A new approach to computing the density of states of the entire spectrum from parallel windows is proposed. It is demonstrated that the new approach allows to compute thermodynamic functions of the system without loss of accuracy while reducing the computation time.

Approbation of the results

The work underwent approbation at the following conferences:

- The study of phase transitions in the Potts model by the Wang-Landau method, Russian Supercomputing Days, 27-28 of September, 2021
- On phase transition in four-component hexagonal lattice Potts model, CSP2020, October 12-16 , 2020
- Analytical structure of transition matrix in Wang-Landau algorithm, International Conference on Computer Simulation in Physics and beyond, October 9-12,2017, Moscow, Russia

Certificate of state registration of software

- Certificate of state registration of software №2019613934 "Implementation of the Wang-Landau algorithm with energy spectrum discretization", 2019г.
- Certificate of state registration of software №2017663374 "Construction of transition matrix for the one-dimensional Ising model in the Wang-Landau algorithm", 2017г.

The list of articles published in the journals included in international citation system WoS or Scopus

- Barash L. Y., Fadeeva M. A., Shchur L. N. Control of accuracy in the Wang-Landau algorithm //Physical Review E. – 2017. – T. 96. – №. 4. – C. 043307. (Scopus/WoS Q1)
- Fadeeva M., Shchur L. On the mixing time in the Wang-Landau algorithm //Journal of Physics: Conference Series. – IOP Publishing, 2018. – T. 955. – №. 1. – C. 012028. (Scopus/WoS Q4)
- Fadeeva M. A., Shchur L. N. Simulation of a Four-Component Potts Model on a Hexagonal Lattice by the Wang–Landau Method with Controlled Accuracy //Journal of Experimental and Theoretical Physics. – 2022. – T. 135. – №. 6. – C. 869-875. (Scopus/WoS Q3)

The list of articles published in the journals included in Russian Science Citation Index

- M.A.Fadeeva, The study of phase transitions in the Potts model by the Wang-Landau method, Russian Supercomputing Days: Proceedings of the International Conference. September 27–28, 2021, Moscow, Russia / Ed. by Vl. Voevodin. – Moscow : MAKS Press, 2021. – 192 p., p.87-96

References

[1] Wang F., Landau D. P. Efficient, multiple-range random walk algorithm to calculate the density of states //Physical review letters. – 2001. – Т. 86. – №. 10. – С. 2050.

[2] Metropolis N., Ulam S. The monte carlo method //Journal of the American statistical association. – 1949. – Т. 44. – №. 247. – С. 335-341.

[3] Fortuin C. M., Kasteleyn P. W. On the random-cluster model: I. Introduction and relation to other models //Physica. – 1972. – Т. 57. – №. 4. – С. 536-564.

[4] Wolff U. Lattice field theory as a percolation process //Physical review letters. – 1988. – Т. 60. – №. 15. – С. 1461.

[5] Swendsen R. H., Wang J. S. Nonuniversal critical dynamics in Monte Carlo simulations //Physical review letters. – 1987. – Т. 58. – №. 2. – С. 86.

[6] Wang F., Landau D. P. Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram //Physical Review E. – 2001. – Т. 64. – №. 5. – С. 056101.

[7] Zablotskiy S. V., Ivanov V. A., Paul W. Multidimensional stochastic approximation Monte Carlo //Physical Review E. – 2016. – Т. 93. – №. 6. – С. 063303.

[8] Mark P. Taylor, Wolfgang Paul, Kurt Binder Phase transitions of a single polymer chain: A Wang–Landau simulation study // J. Chem. Phys. 131, 114907 (2009)

[9] Rastogi C. et al. Accurate and sensitive quantification of protein-DNA binding affinity //Proceedings of the National Academy of Sciences. – 2018. – Т. 115. – №. 16. – С. E3692-E3701.

[10] Wüst T., Reith D., Virnau P. Sequence determines degree of knottedness in a coarse-grained protein model //Physical review letters. – 2015. – Т. 114. – №. 2. – С. 028102.

[11] Liang F., Liu C., Carroll R. J. Stochastic approximation in Monte Carlo computation //Journal of the American Statistical Association. – 2007. – Т. 102. – №. 477. – С. 305-320.

[12] Belardinelli R. E., Pereyra V. D. Fast algorithm to calculate density of states //Physical Review E. – 2007. – Т. 75. – №. 4. – С. 046701.

[13] Belardinelli R. E., Pereyra V. D. Wang-Landau algorithm: A theoretical analysis of the saturation of the error //The Journal of chemical physics. – 2007. – Т. 127. – №. 18. – С. 184105.

[14] Liang F., Liu C., Carroll R. J. Stochastic approximation in Monte Carlo computation //Journal of the American Statistical Association. – 2007. – Т. 102. – №. 477. – С. 305-320.

[15] Zhou C. et al. Optimal modification factor and convergence of the Wang-Landau algorithm //Physical Review E. – 2008. – Т. 78. – №. 4. – С. 046705.

[16] Lee H. K., Okabe Y., Landau D. P. Convergence and refinement of the Wang–Landau algorithm //Computer physics communications. – 2006. – Т. 175. – №. 1. – С. 36-40.

[17] Zhou C., Bhatt R. N. Understanding and improving the Wang-Landau algorithm //Physical Review E. – 2005. – Т. 72. – №. 2. – С. 025701.

[18] Caparica Á. A., Cunha-Netto A. G. Wang-landau sampling: Improving accuracy //Physical Review E. – 2012. – Т. 85. – №. 4. – С. 046702.

[19] Dayal P. et al. Performance limitations of flat-histogram methods //Physical Review Letters. – 2004. – Т. 92. – №. 9. – С. 097201.

[20] Redner S. A guide to first-passage processes. – Cambridge university press, 2001.

[21] Khan M. O., Kennedy G., Chan D. Y. C. A scalable parallel Monte Carlo method for free energy simulations of molecular systems //Journal of computational chemistry. – 2005. – Т. 26. – №. 1. – С. 72-77.

[22] Zhan L. A parallel implementation of the Wang–Landau algorithm //Computer Physics Communications. – 2008. – Т. 179. – №. 5. – С. 339-344.

[23] Yin J., Landau D. P. Massively parallel Wang–Landau sampling on multiple GPUs //Computer Physics Communications. – 2012. – Т. 183. – №. 8. – С. 1568-1573.

[24] Vogel T. et al. Generic, hierarchical framework for massively parallel Wang-Landau sampling //Physical review letters. – 2013. – Т. 110. – №. 21. – С. 210603.

[25] Vogel T. et al. Scalable replica-exchange framework for Wang-Landau sampling //Physical Review E. – 2014. – Т. 90. – №. 2. – С. 023302.

[26] Vogel T. et al. Exploring new frontiers in statistical physics with a new, parallel Wang-Landau framework //Journal of Physics: Conference Series. – IOP Publishing, 2014. – Т. 487. – №. 1. – С. 012001.

[27] Hukushima K., Nemoto K. Exchange Monte Carlo method and application to spin glass simulations //Journal of the Physical Society of Japan. – 1996. – Т. 65. – №. 6. – С. 1604-1608.

[28] Earl D. J., Deem M. W. Parallel tempering: Theory, applications, and new perspectives //Physical Chemistry Chemical Physics. – 2005. – Т. 7. – №. 23. – С. 3910-3916.

[29] Hansmann U. H. E. Parallel tempering algorithm for conformational studies of biological molecules //Chemical Physics Letters. – 1997. – Т. 281. – №. 1-3. – С. 140-150.

[30] Y. Sugita and Y. Okamoto, Chem. Phys. Lett., 1999, 314, 141.

[31] Sugita Y. et al. Replica-exchange methods for biomolecular simulations //Biomolecular Simulations. – Humana, New York, NY, 2019. – С. 155-177.

[32] Li Y. W. et al. A new paradigm for petascale Monte Carlo simulation: Replica exchange Wang-Landau sampling //Journal of Physics: Conference Series. – IOP Publishing, 2014. – Т. 510. – №. 1. – С. 012012.

[33] Perera D. et al. Replica-Exchange Wang—Landau Sampling: Pushing the Limits of Monte Carlo Simulations in Materials Sciences //TMS 2015 144th Annual Meeting Exhibition. – Springer, Cham, 2015. – С. 811-818.

[34] Ф.Р. Гантмахер, Теория матриц. Наука, М., 1967г. с.354-362.

[35] S. Boyd, P. Diaconis, L. Xiao, Fastest mixing Markov chain on a graph, SIAM Review

[36] Beale P.D. Exact Distribution of Energies in the Two-Dimensional Ising Model //Physical review letters.1996.Т.76.№. 1.С. 78.

[37] Муртазаев А. К. и др. ФАЗОВЫЕ ПЕРЕХОДЫ И ТЕРМОДИНАМИЧЕСКИЕ СВОЙСТВА МОДЕЛИ ПОТТСА С ЧИСЛОМ СОСТОЯНИЙ СПИНА q= 4 НА ГЕКСАГОНАЛЬНОЙ РЕШЕТКЕ //ЖЭТФ. – 2019. – Т. 156. – С. 502-506.

[38] M. Fisher, The renormalization group in the theory of critical behavior, Rev. Mod. Phys. **46** 597 (1974).

[39] Salas J., Sokal A. D. Logarithmic corrections and finite-size scaling in the two-dimensional 4-state Potts model //Journal of statistical physics. – 1997. – Т. 88. – №. 3. – С. 567-615.

[40] Barash L. Y., Fadeeva M. A., Shchur L. N. Control of accuracy in the Wang-Landau algorithm //Physical Review E. – 2017. – Т. 96. – №. 4. – С. 043307.