**Noname manuscript No.**
(will be inserted by the editor)

# Matrix-power energy-landscape transformation for finding spin-glass ground states

**Markus Manssen** · **Alexander K. Hartmann**

Received: date / Accepted: date

**Abstract** A method for solving binary optimization problems was proposed by Karandashev and Kryzhanovsky that can be used for finding ground states of spin glass models. By taking a power of the bond matrix the energy landscape of the system is transformed in such a way, that the global minimum should become easier to find. In this paper we test the combination of the new approach with various algorithms, namely simple random search, a cluster algorithm by Houdayer and Martin, and the common approach of parallel tempering. We apply these approaches to find ground states of the three-dimensional Edwards-Anderson model, which is an NP-hard problem, hence computationally challenging. To investigate whether the power-matrix approach is useful for such hard problems, we use previously computed ground states of this model for systems of size $10^3$ spins. In particular we try to estimate the difference in needed computation time compared to plain parallel tempering.

**Keywords** spin glass model · binary minimization · energy landscape transformation · Monte Carlo method · NP-hardness

## 1 Introduction

Spin glasses in physics are magnetic systems characterized by random competing interactions [1–4]. Models of them prominently feature disorder and frustration resulting in a very complex energy landscape. However these models also represent optimization problems, that can be applied in many other scientific fields [2,4]. Unfortunately the actual problem, which is to find the systems' ground states, turns out to be NP-hard [5]. Consequently many elaborate algorithms were adopted over the years, specifically to find ground states or to investigate spin glasses at finite temperature and performing the limit $T \to 0$ [6–13]. Also special computers [14] were build with this challenge in mind. For this kind of optimization problems Karandashev et al. [15] more recently proposed a complementary approach, in which one tries to simplify the problem with a simple transformation of the energy landscape. Approaches based on the same basic idea but using other transformations [16–19]

M. Manssen · A. K. Hartmann
Institute of Physics, Carl von Ossietzky University of Oldenburg, Carl-von-Ossietzky-Strasse 11, 26111 Oldenburg, Germany
E-mail: markus.manssen@uni-oldenburg.de

have been applied to few optimization problems. Nevertheless, though the idea is tantalizing and Ref. [15] and its follow-up [20] have shown it to basically work, it has yet to prove, whether it can actually be useful and compete with established methods when applying it to really hard optimization problems. The goal of this work is to test this approach with a few different algorithms on the 3D Edwards-Anderson model[21] and try to estimate the amount of time, which can be saved by using it.

## 2 Model

The Edwards-Anderson model [21] of spin glasses consists of a $D$-dimensional lattice of side length $L$ filled with $N = L^D$ Ising spins $s_i = \pm 1$. The spins are coupled by nearest-neighbor bonds $J_{ij} = \pm 1$ with periodic boundary conditions. Given a spin configuration $\mathbf{S}$ and the bond matrix $\mathbf{J}$ with standard deviation $\sigma_J$ the state's energy can be calculated as

$$E(\mathbf{S}) = -\frac{1}{N^2 \sigma_J} \sum_{i,j} J_{ij} s_i s_j. \tag{1}$$

This definition follows [15] with the sum prefactor acting as a normalization to make different matrices comparable. There is two major differences between our setup and the one in [15]. Firstly we are looking at $D = 3$ systems instead of $D = 2$, which is important, because the problem of finding ground states for spin glass models was found to be NP-hard[5] in three or higher dimensions. Secondly the mentioned paper uses gaussian couplings $J_{ij}$. In that case spin flips resulting in no change of energy $\Delta E = 0$ are extremely unlikely to appear, while for our binary case they are very common, which allows for a better convergence of the algorithms.

## 3 Algorithms

### 3.1 Simple random search

The first and simplest algorithm that comes to mind when trying to find a ground state for this model is probably the simple random search (SRS). This was the focus of [15], inspired by the dynamics of the Hopfield model of neural networks [22], which can be seen an equivalent to spin-glass models. The algorithm acts as a simple descent in the energy landscape from a random starting point:

1. Iterate over all spins. Flip every spin with $\Delta E < 0$.
2. Stop, if no spin was flipped in the last sweep. Otherwise return to step 1.

This method can be expected to get stuck in local minima of the energy landscape most of the time. Hence the search would have to be repeated many times from different random starting configurations for a realistic chance of finding the global minimum. However in contrast to the paper [15] we were not able to actually find ground states using this algorithm verbatim, because we study here a harder problem. We can however make a simple but effective amendment by also allowing spin flips with $\Delta E = 0$, which makes the algorithm more similar to a Monte Carlo simulation at zero temperature. Of course the stop criterion has to be changed as well: the search terminates, if no spin flip with $\Delta E < 0$ could be found in the last 100 sweeps. The improvement is quite noticeable and should not be underestimated. However it is only helpful, because spin flips with $\Delta E = 0$ are quite common, as mentioned

before. For easy distinction we will call this the altered SRS and the former the unaltered SRS.

### 3.2 Houdayer-Martin algorithm

In a follow-up paper [20] Karandashev et al. also utilized a cluster algorithm introduced by Houdayer and Martin [12] (HM). In the original paper [12] it was used as a local search in a more sophisticated genetic renormalization algorithm. The basics idea of HM is to flip clusters of spins instead of single spins. In detail, it works as follows:

1. Pick any spin with $\Delta E < 0$ or, if none exists, simply a random spin with uniform probability. Virtually flip spin and use it as seed for new cluster. Set total cluster energy change $\Delta E_c = \Delta E$.
2. Pick neighbor spin of current cluster with smallest $\Delta E$. Virtually flip spin and add it to cluster, and update $\Delta E_c = \Delta E_c + \Delta E$.
3. Continue at step 4, if no spins are left or if 20 spins were flipped since the cluster achieved the smallest total energy change $\Delta E_c$. Otherwise return to step 2.
4. Flip the cluster, which achieved a minimal $\Delta E_c$, but only if $\Delta E_c < 0$.
5. Stop, as soon as three times the resulting clusters have not been flipped. Otherwise return to step 1.

This algorithm also considers flips which seem bad at first, and can therefore escape local minima. But though it is much better than the unaltered SRS, surprisingly it was not able to find ground states either. Consequently we alter it in basically the same way as with the SRS. Spins with $\Delta E = 0$ are preferred over random ones as seeds and clusters with no change in energy are flipped as well. And a much more lenient stop criterion is applied, which waits for 1000 consecutive failures to find a cluster, that lowers the energy. Again this improves the algorithm noticeably, considerably slowing it down at the same time.

### 3.3 Monte Carlo and parallel tempering

A common physical approach to the problem is to evolve the system at a certain temperature $T$ using a Monte Carlo simulation(MC)[23]. For Metropolis dynamics used hereafter one simply picks for each step a random spin (with uniform probability) and flips it with probability $\min(1, \exp(-\Delta E/T))$. Theoretically one can reach equilibrium at low temperature, hence the ground state should be encountered. Nevertheless, for such systems it takes a very long time to actually reach equilibrium, since the dynamics get stuck in metastable states due the existence of many low-lying local minima. This issue can be tackled by using the Metropolis-coupled Markov-chain Monte Carlo ($MC^3$) algorithm [24], now known in Physics as the replica-exchange or parallel-tempering (PT) approach [25,8].

1. Create $i = 1, \ldots, M$ replicas of the given samples of the disorder each assigned to a different temperature $T_i$ with $T_1 < T_2 < \ldots < T_M$.
2. Perform a Monte Carlo sweep of $N$ updates for each replica.
3. $M-1$ times pick random pair $(i, i+1)$ of replicas at neighboring temperatures $T_i, T_{i+1}$ and exchange them with probability

$$p_{\text{exch}} = \min\left\{1, \exp\left[-\left(\frac{1}{T_i} - \frac{1}{T_{i+1}}\right) \cdot (E_i - E_{i+1})\right]\right\}.$$

4. Stop after a specified number of iterations. Otherwise return to step 2.

Here we simply chose a set of 21 equally spaced temperatures $T_1 < T_2 \ldots < T_{21}$ in $[0.3, 1.35]$. A well-considered choice of tailored temperatures can improve the efficiency of the algorithm but such would be overdue for the tests we will perform. But we perform a cheap overrelaxation sweep, which simply iterates over all spins flipping those with $\Delta E = 0$, after each Monte Carlo sweep for a bit of improvement. At the end we pick the lowest temperature replica and apply the unaltered SRS once, just to make it more likely that we do not strand just a few steps away from the global minimum.

### 3.4 Double-descend algorithm

Karandashev et al. proposed a different approach for finding ground states[15]. The basic idea is to consider the $k$-th power of the bond matrix $\mathbf{J}^k$ resulting in a change of the energy landscape. Particularly the global minimum is expected to deepen while expanding its attraction area. But at the same time the global minimum shifts a little in configuration space. Explicitly Karandashev et al. [15] state a deepening by a factor of 1.35 and a shift by a Hamming distance of $0.026N$ for the case $k = 2$. Based on this the double-descent algorithm for value $k$ (DD$k$) involves two main steps:

1. Start from random configuration and search for the ground state of the system with bond matrix $\mathbf{J}^k$, $k > 1$.
2. Continue from the configuration obtained in this way to search for the ground state of the original system with bond matrix $\mathbf{J}$.

In the first step the global minimum of the transformed system should be easier to find and, because of its small distance, the global minimum of the original system should be easily reached from there in the second step. The actual search in both steps can be carried out with any algorithm of choice. In Ref. [15] the simple random search was used and the best performance was observed for the choice $k = 3$. The follow-up paper [20] also used the Houdayer-Martin algorithm giving better results. Further adjustment has to be made if one wants to use Monte Carlo or parallel tempering on a transformed system, because the typically occurring energies will differ in magnitude for different values of $k$. Hence, one would have to use different sets of temperatures. By simply multiplying the exponentiated matrix with the ratio of original and new standard deviation $\sigma_J$ this is avoided. However these combinations will not be discussed here in detail, as they turned out to have no notable advantage over temperature-less algorithms for this purpose. Normal Monte Carlo simulations fare quite similarly to the SRS, while taking much longer. Parallel tempering suffers extreme slowdown for $k > 1$ due to the reasons outlined in Sect. 4.2. Though tests showed an increase in success rates, it was not large enough to vindicate a closer look under these circumstances.

## 4 Test results

### 4.1 Success rates

We implemented the double-descend method using the sparse matrix package included in the Eigen library [26]. First we will take a look at the distribution of energies found by basic SRS and the DD3 incorporating it in Fig. 1 together with the distribution of ground-state

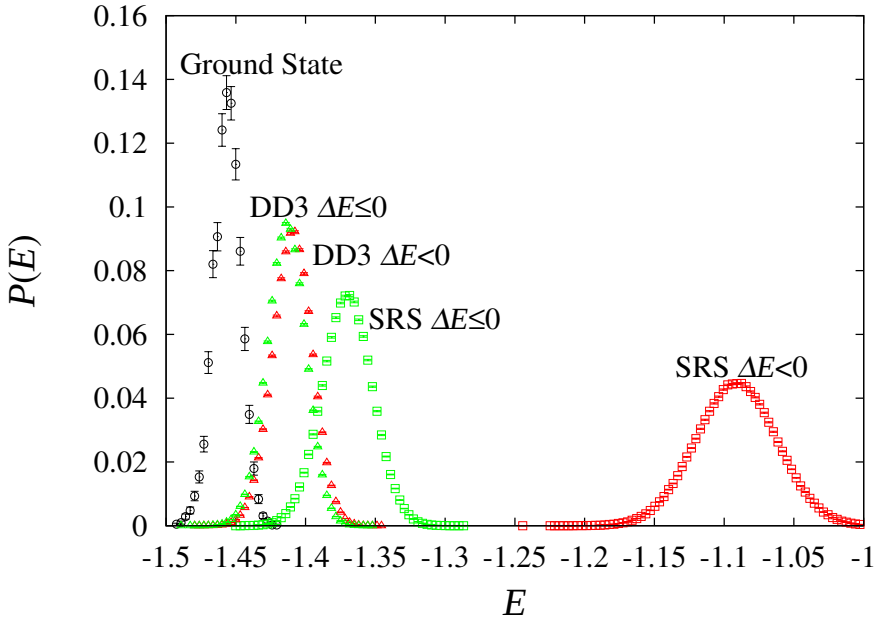


**Fig. 1** Energy distribution $P(E)$ of ground states of the samples and of the energies of the configurations found by the unaltered ($\Delta E < 0$) and altered ($\Delta E \leq 0$) simple random search. On top of that we differentiate between the simple use of the SRS ($k = 1$) and the variant using a double-descend approach DD3 with $k = 3$.

energies. We considered 1000 samples of the disorder for $L = 10$. The ground states had been calculated previously using the algorithm detailed in [9,10]. Note that only the availability of presumably true ground states of the NP-hard minimum-energy spin-glass problem renders the evaluation of the DD*k* algorithms reliable. For testing the different algorithms considered in this work, for each sample (and type of algorithm) $10^3$ independent runs with random initial spin configurations were performed, and the lowest-energy configuration was kept as potential ground state each time. Like in [15] the DD3 version of the unaltered SRS is much closer to the ground states, suggesting a better chance of finding them. However the adjusted SRS (allowing for spin flips with $\Delta E \leq 0$) is already better and yet combined with the DD3 the improvement turns out to be much smaller compared to the adjusted SRS approach. This is just one example of what we found holds for all algorithms we tested with this approach: The better an algorithm is on its own, i.e., for $k = 1$, the less it will benefit from the double-descend approach.

Next we take a look at the measured probability $P_0$ of finding a ground state as a function of the matrix power $k$ in Fig. 2. It can be seen, that even powers $k$ do worse than odd powers most of the time and can thus be mostly neglected. Contrary to findings in [15] $k = 3$ does not seem to be optimal with $P_0 < 10^{-5}$. Hence, when using only 1000 independent runs, one will not find a true ground states for most samples. Nevertheless, the accuracy seems to steadily increase only becoming worse for $k > 20$. However this benefit does not come without a cost, as we will see in the following section.

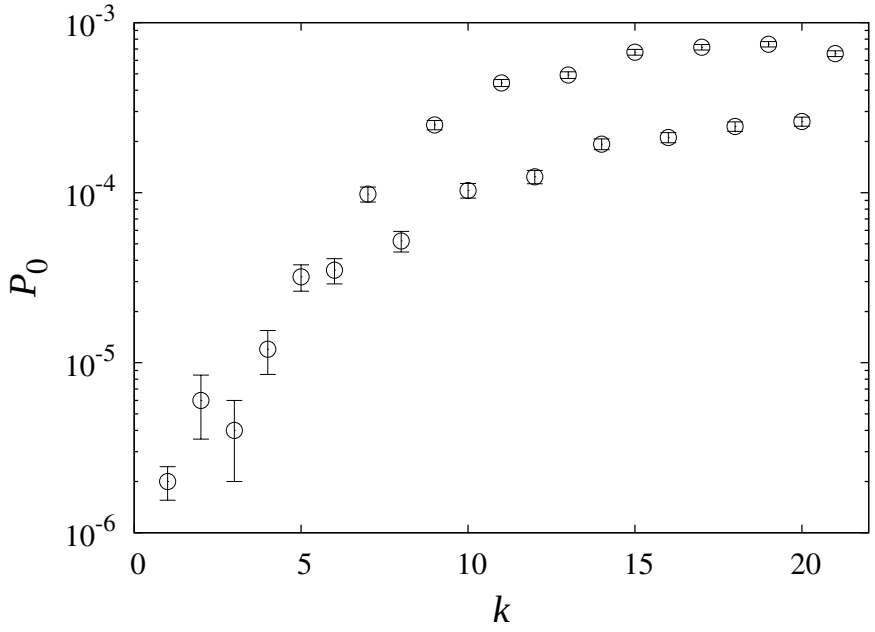


**Fig. 2** Probability $p_0$ of finding a ground state using the simple random search in a double-descend approach with different matrix powers $k$.

### 4.2 Timing

To evaluate the efficiency of the algorithms, we timed all algorithms on the same computer, an Intel Core 2 Duo E6550, differentiating between the setup time $t_{\text{Setup}}$ and the run time $t_{\text{Run}}$. The setup time $t_{\text{Setup}}$ is the time needed for setting up an algorithm including the calculation of the matrix powers $\mathbf{J}^k$ and thus increases immensely with $k$ in the same way for all algorithms. As we have to repeat the algorithm many times, this is usually eclipsed by the run time $t_{\text{Run}}$ of an algorithm spent after the setup, which is shown in Fig. 3. As evident the run time increases as well. The first increase happens from $k = 1 \longrightarrow k = 2$ because it executes the incorporated algorithm twice for all $k > 1$. The further growth happens because the number of non-zero elements in the bond matrix increases with growing power $k$. This second effect then saturates, in this case for $k > 10$, because almost half of all matrix elements become non-zero. For complex algorithms this slowdown for growing value of $k$ becomes quite notable and at the same time the benefit of considering powers of the bond matrix becomes smaller, as seen before. Notably the inclusion of $\Delta E = 0$ flips for the SRS and HM algorithms become useless, because for $k > 3$ spin flips with $\Delta E = 0$ rarely occur anymore. As a consequence, we used the original variant for the first ($k > 1$) part, i.e., the descent was stopped if no $\Delta E < 0$ flip was encountered (unaltered variant). This saved quite a bit of running time, while hardly altering the probability of finding ground states.

Now we include the effect of multiple restarts with random initial configurations in our considerations. Assuming $P_0$ is again the probability of finding a ground state in a single run, the probability of finding at least one ground state after $M$ runs is $q = 1 - (1 - P_0)^M$. Using this we are able to compare the efficiency of different algorithms by defining a weighted

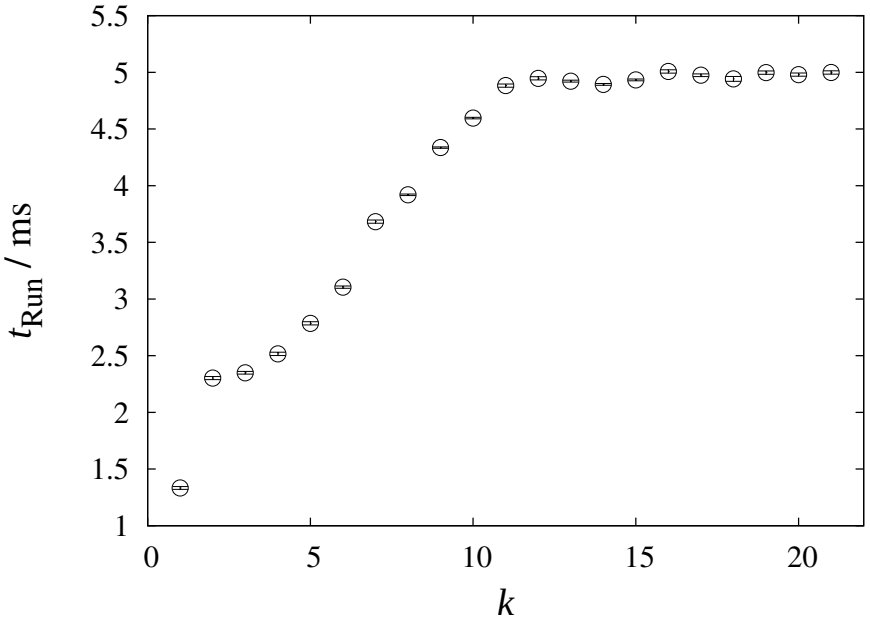


**Fig. 3** Run time $t_{\mathrm{Run}}$ of the simple random search in a double-descend approach with different matrix powers $k$.

time

$$t_q = t_{\mathrm{Setup}} + t_{\mathrm{Run}} \frac{\log(1-q)}{\log(1-P_0)} \tag{2}$$

needed for finding ground states, assuming a target probability $q$. We choose $q = 0.99$ here, but any other choice would just differ by a prefactor. Fig. 4 shows this weighted time for the combinations of the simple random search, Houdayer-Martin algorithm and parallel tempering, which are the most efficient of their kind (the underlying data of $P_0$, $t_{\mathrm{Setup}}$ and $t_{\mathrm{Run}}$ is not shown here for briefness). For the simple random search we used a combination of unaltered SRS in the first step and altered SRS in the second one. The same unaltered SRS also proved most practical when combined as the first step together with the altered Houdayer-Martin algorithm in the second. While this turned out to be useful for parallel tempering of short durations, for $10^4$ iterations any combination tried only served to lower the probability $P_0$, Thus, we show here as an example (marked by PT in Fig.4) the combination with unaltered HM in the first step where the matrix powers $k > 1$ are considered.

So far, when looking at average running times, the setup building on the Houdayer-Martin algorithm seems most successful with plain parallel tempering being slower by more than one order of magnitude, despite having about 70% success rate on a single run. However our comparison so far has disregarded that for different samples of the disorder for some samples the ground states are much harder to find than for others. Thus, we actually find a distribution of vastly different probabilities $P_0$ for different samples. For Fig. 5 we calculated the weighted times $t_q$ from the individual $P_0$ values which were obtained by averaging over $10^5$ independent runs per sample for HM ($k = 21$) and plain PT with $10^4$ iterations averaged over $10^2$ independent runs per sample. Using these histograms of $P_0$ we determined the distributions $P(t_q)$ using logarithmic binning. Now the difference in timing

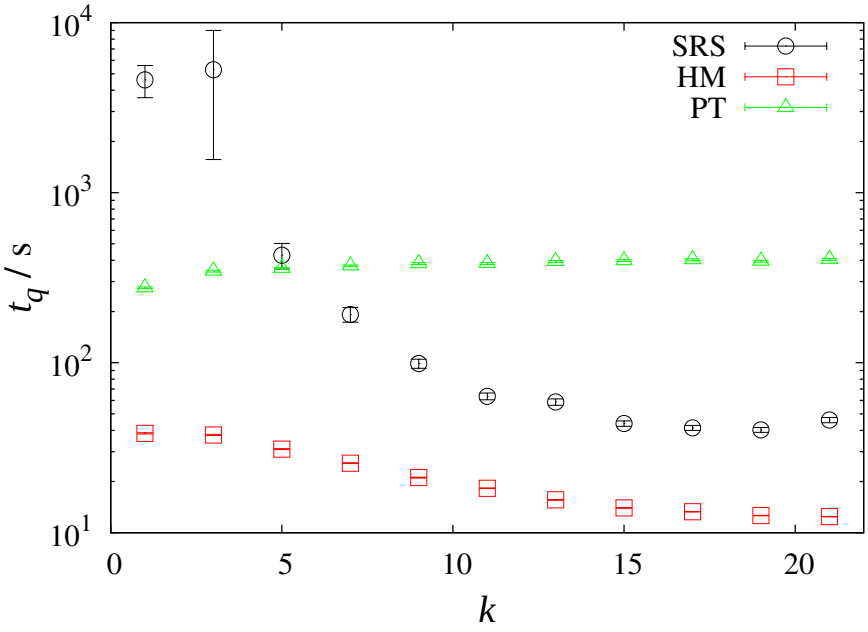


**Fig. 4** Weighted times $t_q$ for finding ground states with the simple random search(SRS), the Houdayer-Martin algorithm(HM) and parallel tempering(PT) with $10^4$ iterations in a double-descend approach with different matrix powers $k$.

between both algorithms looks much smaller. In both cases the distribution has a long tail complying with a power law of exponent $\gamma = -1.5$. This can be seen as a sign of the universal hardness of the problem. Note furthermore that the expectation value of the distribution is not defined, which means that there are rare instances with very long running time, which dominate the average. Hence, the spin glass ground-state problem is very hard.

Note also that for the result shown in Fig. 4 behavior was skewed, because the HM combination can for some samples find ground states with a few short runs, while PT always needs at least one slow run. To dependably succeed for all samples would take both a more similar amount of time.

## 5 Conclusion

As we have seen, the double-descend approach, based on considering powers of the interaction matrix in the first descent, can enhance the probability of finding ground states for almost all algorithms used with it, with the big exception being long-time parallel tempering. But at the same time the first descend in the transformed energy landscape also greatly slows down compared to the second one. Furthermore we found that the better an algorithm is the less of an improvement it can receive from this. Because of this we opt for using simpler algorithms for the descend in the transformed landscape.

The definition of a weighted time $t_q$ in (2) allowed us to compare different algorithm combinations under the assumption, that we can do multiple runs, i.e., restarts from different initial random conditions. Although a double descend approach incorporating the Houdayer-

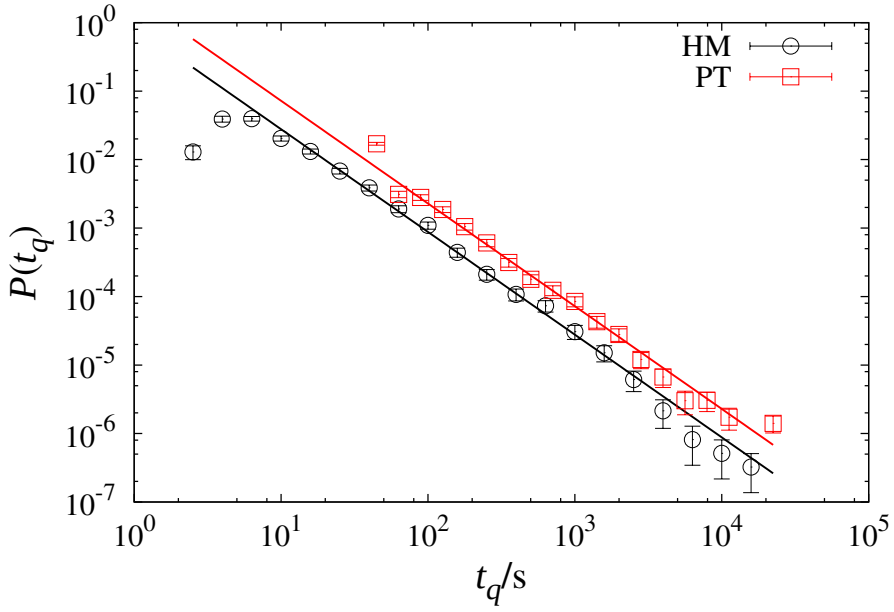


**Fig. 5** Distribution of the weighted time $t_q$ for finding ground states with the Houdayer-Martin algorithm(HM) in a double-descend approach with $k = 21$ on one hand and plain parallel tempering(PT) with $10^4$ iterations on the other hand. The added lines are power laws with exponent $\gamma = -1.5$.

Martin algorithm seemed to come out on top of the established PT, a look at the distribution of $t_q$ showed only little advantage. Considering we have not even made any effort to tune the PT to the problems at hand, PT, as a very simple to implement method, appears still to be a meaningful choice when computing spin-glass ground states. Also so far we have only compared to already known ground states, but for use in practice we would first have to establish a method for assessing whether the ground state has actually been reached, while many such checks are already available for PT. Overall the double descend approach, while an interesting concept, does not seem to have, in the moment, practical use above other approaches.

**Acknowledgements** The simulations were performed at the C. v. O. Universität Oldenburg on the HERO cluster funded by the DFG (INST 184/108-1 FUGG) and the ministry of Science and Culture (MWK) of the Lower Saxony State.

## References

1. Binder, K., Young, A.: Spin-glasses: Experimental facts, theoretical concepts and open questions. Rev. Mod. Phys. **58**, 801 (1986)
2. Mézard, M., Parisi, G., Virasoro, M.A.: Spin glass theory and beyond. World Scientific, Singapore (1987)
3. Young, A.P. (ed.): Spin glasses and random fields. World Scientific, Singapore (1998)
4. Nishimori, H.: Statistical Physics of Spin Glasses and Information Processing: An Introduction. Oxford University Press (2001)
5. Barahona, F.: On the computational complexity of Ising spin glass models. J. Phys. A: Math. Gen. **15**(10), 3241 (1982)

6. Swendsen, R.H., Wang, J.S.: Replica Monte Carlo Simulation of Spin-Glasses. Phys. Rev. Lett. **57**(21), 2607–2609 (1986)
7. Hartmann, A.K.: Cluster-exact approximation of spin glass ground states. Physica A **224**, 480–488 (1999)
8. Hukushima, K., Nemoto, K.: Exchange Monte Carlo Method and Application to Spin Glass Simulations. J. Phys. Soc. Jpn. **65**(6), 1604–1608 (1996)
9. Hartmann, A.K.: Scaling of stiffness energy for three-dimensional $\pm$J Ising spin glasses. Phys. Rev. E **59**, 84 (1999)
10. Hartmann, A.K.: Calculation of ground states of four-dimensional $\pm$J Ising spin glasses. Phys. Rev. E **60**(5), 5135 (1999)
11. Wang, F., Landau, D.P.: Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram. Phys. Rev. E **64**(5), 056,101 (2001)
12. Houdayer, J., Martin, O.C.: Hierarchical approach for computing spin glass ground states. Phys. Rev. E **64**(5), 056,704 (2001)
13. Hartmann, A.K., Rieger, H.: Optimization Algorithms in Physics. Wiley-VCH, Weinheim (2001)
14. Belletti, F., Cotallo, M., Cruz, A., Fernandez, L.A., Gordillo-Guerrero, A., Guidetti, M., Maiorano, A., Mantovani, F., Marinari, E., Martin-Mayor, V., Muñoz-Sudupe, A., Navarro, D., Parisi, G., Perez-Gaviro, S., Rossi, M., Ruiz-Lorenzo, J.J., Schifano, S.F., Sciretti, D., Tarancon, A., Tripiccione, R., Velasco, J.L., Yllanes, D., Zanier, G.: Janus: An FPGA-Based System for High-Performance Scientific Computing. Comput. Sci. Eng. **11**(1), 48–58 (2009)
15. Karandashev, Y.M., Kryzhanovsky, B.V.: Transformation of Energy Landscape in the Problem of Binary Minimization. Doklady Mathematics **80**(3), 927–931 (2009)
16. Gu, J., Huang, X.: Efficient local search space smoothing: a case study of the traveling salesman problem (tsp). IEEE Trans. Systems Man Cybernet. **24**, 728–735 (1994)
17. Schneider, J.J., Dankesreiter, M., Fettes, W., Morgenstern, I., Schmid, M., Singer, J.M.: Search-space smoothing for combinatorial optimization problems. Physica A **243**, 77–112 (1997)
18. Zhang, Y., Kihara, D., J., S.: Local energy landscape flattening: parallel hyperbolic monte carlo sampling of protein folding. Proteins **48**, 192–201 (2002)
19. Pritchard-Bell, A., Shell, M.S.: Smoothing Protein Energy Landscapes by Integrating Folding Models with Structure Prediction. Biophys. J. **101**(9), 2251–2259 (2011)
20. Karandashev, Y.M., Kryzhanovsky, B.V.: Increasing the attraction area of the global minimum in the binary optimization problem. J. Glob. Optim. pp. 1–19 (2012)
21. Edwards, S.F., Anderson, P.W.: Theory of spin glasses. J. Phys. F: Met. Phys. **5**(5), 965–974 (1975)
22. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. PNAS **79**(8), 2554–2558 (1982)
23. Newman, M.E.J., Barkema, G.T.: Monte Carlo methods in statistical physics. Oxford University Press (1999)
24. Geyer, C.: Markov Chain Monte Carlo Maximum Likelihood. In: Computing Science and Statistics, Proceedings of the 23rd Symposium on the Interface, pp. 156–163. Interface Foundation of North America (1991). URL http://purl.umn.edu/58440
25. Marinari, E., Parisi, G.: Simulated Tempering: A New Monte Carlo Scheme. Europhys. Lett. **19**(6), 451 (1992)
26. Guennebaud, G., Jacob, B., et al.: Eigen. http://eigen.tuxfamily.org (2012)