Hierarchical BOA, Cluster Exact Approximation, and Ising Spin Glasses

Martin Pelikan[1], Alexander K. Hartmann[2], and Kumara Sastry

[1] Dept. of Math and Computer Science, 320 CCB
University of Missouri at St. Louis
One University Blvd.
St. Louis, MO 63121
pelikan@cs.umsl.edu
[2] Institut für Theoretische Physik
Universität Göttingen
Friedrich-Hund-Platz 1
37077 Göttingen, Germany
hartmann@physik.uni-goettingen.de
[3] Illinois Genetic Algorithms Laboratory (IlliGAL)
University of Illinois at Urbana-Champaign
104 S. Mathews Ave.
Urbana, IL 61801, USA
kumara@illigal.ge.uiuc.edu

Abstract. This paper analyzes the hierarchical Bayesian optimization algorithm (hBOA) on the problem of finding ground states of Ising spin glasses with $\pm J$ couplings in two and three dimensions. The performance of hBOA is compared to that of the simple genetic algorithm (GA) and the univariate marginal distribution algorithm (UMDA). The performance of all tested algorithms is improved by incorporating a deterministic hill climber based on single-bit flips and cluster exact approximation (CEA). The results show that hBOA significantly outperforms GA and UMDA on a broad spectrum of spin-glass instances and that CEA enables all tested algorithms to solve larger spin-glass instances. Using advanced hybrid methods created by combining competent genetic and evolutionary algorithms with advanced local searchers thus proves advantageous in this challenging class of problems.

1 Introduction

Ising spin glasses are prototypical models for disordered systems and have played a central role in statistical physics during the last three decades [1–4]. Examples of experimental realizations of spin glasses are metals with magnetic impurities, e.g. gold with a small fraction of iron added. Spin glasses represent also a large class of challenging problems for optimization algorithms [5–7] where the task is to minimize energy of a given spin-glass instance [8–13]. States with the lowest energy are called *ground states* and thus the problem of minimizing the

energy of spin-glass instances can be formulated as the problem of finding ground states of these instances. There are two main challenges that must be tackled to find ground states of spin glasses efficiently and reliably: (1) There are many local optima in the energy landscape (the number of local optima may grow exponentially with problem size). (2) The local minima are often surrounded by high-energy configurations, which make it difficult for local operators to escape the local optimum once they get trapped in it. That why even state-of-the-art Markov chain Monte Carlo (MCMC) methods require exponential time to locate ground states [14].

This paper analyzes the hierarchical Bayesian optimization algorithm (hBOA) [15, 16] on a broad spectrum of instances of the problem of finding ground states of Ising spin glasses. The performance of hBOA is compared to that of the simple genetic algorithm (GA) and the univariate marginal distribution algorithm (UMDA). We also describe cluster exact approximation (CEA) [17], which provides an efficient method to perform large updates of spin glass configurations to decrease their energy. CEA is then incorporated into hBOA and GA to improve their performance, and the resulting hybrids are tested on a number of spin-glass problem instances.

The paper is organized as follows. Section 2 describes the problem of finding ground states of Ising spin glasses. Section 3 briefly describes algorithms hBOA, GA and UMDA; additionally, the section describes the deterministic hill climber and cluster exact approximation, which are incorporated into all tested algorithms to improve their performance. Section 4 presents and discusses experiments. Finally, Section 5 summarizes and concludes the paper.

2 Ising Spin Glass

A very simple model to describe a finite-dimensional Ising spin glass is typically arranged on a regular 2D or 3D grid where each node i corresponds to a spin s_i and each edge $\langle i, j\rangle$ corresponds to a coupling between two spins s_i and s_j. Each edge has a real value associated with it that defines the relationship between the two connected spins. To approximate the behavior of the large-scale system, periodic boundary conditions are often used that introduce a coupling between the first and the last element along each dimension.

For the classical Ising model, each spin s_i can be in one of two states: $s_i = +1$ or $s_i = -1$. Note that this simplification corresponds to highly anisotropic systems, which do indeed exist in some experimental situations. Nevertheless, the two-state Ising model comprises all basic effects also found in models with more degrees of freedom. A specific set of coupling constants define a spin-glass instance. Each possible setting of all spins is called a spin configuration.

Given a set of coupling constants $J_{i,j}$, and a configuration of spins $C = \{s_i\}$ $(i = 1, \ldots, n)$, the energy can be computed as

$$E(C) = \sum_{\langle i,j\rangle} s_i J_{i,j} s_j \ , \tag{1}$$

where the sum runs over all couplings $\langle i, j\rangle$.

Given a set of coupling constants, the usual task in statistical physics is to integrate a known function over all possible configurations of spins, assuming the Boltzmann distribution of spin configurations; that means, the probability of each configuration C is proportional to $\exp(-E(C)/T)$ where $E(C)$ is energy of C and T is the temperature. From the physics point of view, it is also interesting to know the ground states (configurations associated with the minimum possible energy). Finding extremal energies then corresponds to sampling the Boltzmann distribution with temperature approaching 0 and thus the problem of finding ground states is simpler *a priori* than integration over a wide range of temperatures. However, most of the conventional methods based on sampling the above Boltzmann distribution fail to find the ground states because they get often trapped in a local minimum [14].

In order to obtain a quantitative understanding of the disorder in a spin glass system introduced by the random spin-spin couplings, one generally analyzes a large set of random spin-glass instances for a given distribution of the spin-spin couplings. For each spin glass instance, the optimization algorithm is applied and the results are analyzed to obtain a measure of computational complexity. Here we consider the $\pm J$ spin glass, where each spin-spin coupling constant is set randomly to either $+1$ or -1 with equal probability.

3 Compared Algorithms

This section outlines the algorithms compared in this paper: (1) The hierarchical Bayesian optimization algorithm (hBOA), (2) the genetic algorithm (GA), and (3) the univariate marginal distribution algorithm (UMDA). hBOA and UMDA are estimation of distribution algorithms [18–20], where standard recombination and mutation operators are replaced by building and sampling probabilistic models. Additionally, the section describes the deterministic hill climber and cluster exact approximation, which are used to improve performance of compared algorithms.

In all algorithms, candidate solutions are represented by n-bit binary strings where each bit specifies the value of one of the n spins (0 represents state -1 and 1 represents state $+1$).

3.1 Hierarchical Bayesian optimization algorithm (hBOA)

The hierarchical Bayesian optimization algorithm [15, 16] evolves a population of candidate solutions. The population is initially generated at random according to a uniform distribution over all n-bit strings. Each iteration starts by selecting a population of promising solutions using any common selection method of genetic and evolutionary algorithms, such as tournament and truncation selection. We use binary tournament selection. New solutions are generated by building a Bayesian network with decision trees [21, 22] for the selected solutions and sampling the built Bayesian network. To ensure useful diversity maintenance, the new candidate solutions are incorporated into the original population using restricted tournament replacement (RTR) [23]. The run is terminated when termination criteria are met.

3.2 Genetic algorithm (GA)

The genetic algorithm (GA) [24, 25] also evolves a population of candidate solutions represented by fixed-length binary strings. The first population is generated at random. Each iteration starts by selecting promising solutions from the current population. New solutions are created by applying variation operators to the population of selected solutions. Specifically, crossover is used to exchange bits and pieces between pairs of candidate solutions and mutation is used to perturb the resulting solutions. Here we use one-point crossover and bit-flip mutation. The new candidate solutions are incorporated into the original population using RTR. The run is terminated when termination criteria are met.

3.3 Univariate marginal distribution algorithm (UMDA)

The univariate marginal distribution algorithm (UMDA) [18] also evolves a population of candidate solutions represented by binary strings, starting with a random population. Each iteration starts by selection. Then, the probability vector is learned that stores the proportion of 1s in each position of the selected population. Each bit of a new candidate solution is then set to 1 with the probability equal to the proportion of 1s in this position; otherwise, the bit is set to 0. Consequently, the variation operator of UMDA preserves the proportions of 1s in each position while decorrelating different string positions. The new candidate solutions are incorporated into the original population using RTR. The run is terminated when termination criteria are met.

The only difference between hBOA and the UMDA variant discussed in this paper is the type of the probabilistic model used to model promising candidate solutions and generate the new ones. The comparison between hBOA and UMDA should therefore indicate whether in this problem domain effective exploration necessitates complex probabilistic models that can efficiently encode large-order interactions between spins, as it is the case for hBOA. For analogical reasons, the comparison between hBOA and GA will indicate whether it is important to use advanced variation operators that adapt to the problem at hand like in hBOA.

3.4 Deterministic hill climber

We incorporate a deterministic hill climber (DHC) is to hBOA, GA and UMDA to improve their performance. DHC takes a candidate solution represented by an n-bit binary string on input. Then, it performs one-bit changes on the solution that lead to the maximum improvement of solution quality (maximum decrease in energy). DHC is terminated when no single-bit flip improves solution quality and the solution is thus locally optimal. Here, DHC is used to improve every solution in the population before the evaluation is performed. The hybrids created by incorporating DHC into hBOA, GA and UMDA are referred to as hBOA+DHC, GA+DHC and UMDA+DHC, respectively.

3.5 Cluster Exact Approximation (CEA)

Due to the complex structure of the energy landscape of spin glasses, many local minima exist, which have energies very close to the ground-state energy. Usually these minima differ from the true ground states by flips of large domains. Hence, as already mentioned, the minima are surrounded by high energy barriers from the viewpoint of single-spin-flip dynamics. This leads to a very small efficiency of algorithms which apply single-bit (spin) changes as DHC. For this reason, we also consider *cluster exact approximation* [17], which provides an efficient method that can change many spins at the same time optimally (assuming that the remaining spins remain fixed).

CEA starts by constructing a non-frustrated cluster of spins. The selected cluster is first transformed so that all interactions become ferromagnetic (negative coupling). All spins outside the cluster are fixed and treated as local magnetic fields. All cluster spins are computed leading to an optimal spin configuration with respect to the non-cluster spins. The computation can be performed in polynomial time based on graph-theoretical methods [26, 27]: an equivalent network is constructed [28], the maximum flow is calculated [29, 30] and the spins of the cluster are set to orientations leading to a minimum in energy.

The CEA update step ensures that the spins in the cluster minimize energy of the considered spin glass assuming that the remaining (non-cluster) spins remain fixed to their current values. Each CEA iteration either decreases the energy or the energy remains the same, which is the case when all cluster spins were already set to their optimal values.

In this work, we use CEA to improve all obtained candidate solutions and we repeat the CEA update step until the update fails to decrease the energy for a predefined number of iterations; specifically, the bound on the number of failures is $\sqrt{n}$ for 2D spin glasses and it is $\sqrt[3]{n}$ for 3D spin glasses.

4 Experiments

This section presents and discusses experimental results. The section starts by describing problem instances used in the tests. Next, experiments are described. Finally, the experimental results are presented and discussed.

4.1 Tested Spin-Glass Instances

Both 2D and 3D Ising spin-glass instances with $\pm J$ couplings and periodic boundary conditions were considered. To analyze scalability, for two-dimensional spin glasses, instances of size 6×6 ($n = 36$ spins) to 50×50 ($n = 2500$ spins) have been considered; 1000 random instances have been generated for each problem size. For three-dimensional spin glasses, instances of size $4 \times 4 \times 4$ ($n = 84$ spins) to $10 \times 10 \times 10$ ($n = 1000$ spins) have been considered. In experiments on 3D spin glasses without CEA, only 8 random instances have been generated for each problem size because of the increased computational resources required to solve the 3D instances.

4.2 Description of Experiments

All compared algorithms use binary tournament selection to select promising solutions. As a replacement strategy, RTR is used where the window size w is set to the number of bits in solution strings but it is always ensured to be at most 5% of the population size, $w = \min(n, N/20)$. GA+DHC and GA+CEA use one-point crossover with the probability of crossover $p_c = 0.6$ and bit-flip mutation with the probability of flipping each bit $p_m = 1/n$.

For each problem instance, bisection is run to determine the minimum population size to ensure convergence in 5 independent runs (out of 5 runs total). Each run is terminated either when the algorithm has found the optimum or when the algorithm has failed to find the optimum for a large number of iterations. The optimum for most 2D instances was verified with the branch-and-cut algorithm provided at the Spin-Glass Ground State Server at the University of Köln [31]. The remaining 2D instances with ground states were obtained from S. Sabhapandit and S. N. Coppersmith from the University of Wisconsin who identified the ground states using flat-histogram Markov chain Monte Carlo simulations [14]. All 3D instances with their ground states were obtained from previous simulations of one of the authors [10].

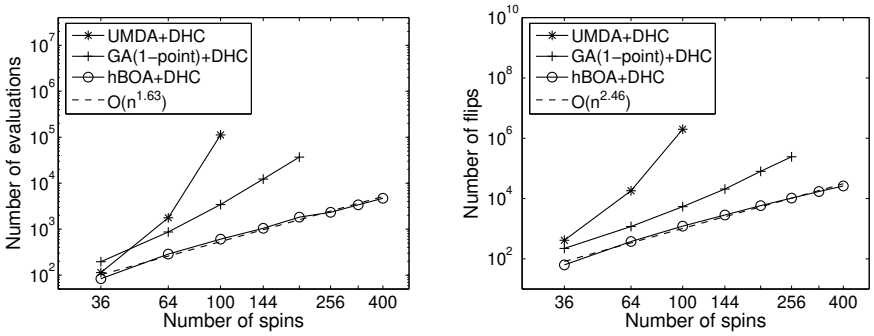
The upper bound on the number of iterations (generations) is determined empirically so that the number of iterations is sufficiently large for all tests. In general, the bound on the number of iterations for GA+DHC is larger than that for hBOA+DHC and UMDA+DHC because of the slower mixing with one-point crossover [32].

The performance of hBOA+DHC, GA+DHC and UMDA+DHC is measured by the number of evaluated spin glass configurations until the optimum has been found. Additionally, we show the number of DHC updates (steps of local search). Since one update step of CEA is usually more computationally expensive than the entire evaluation of a spin configuration [33], the time complexity of hBOA+CEA and GA+CEA is measured by the number of iterations of CEA as opposed to the number of evaluations.

4.3 Results

Figure 1 compares the performance of hBOA+DHC, UMDA+DHC and GA+DHC on two-dimensional spin glasses with $\pm J$ couplings and periodic boundary conditions. The results indicate that the number of evaluations and the number of DHC flips for hBOA+DHC grows with a low-order polynomial of problem size, i.e., with $O(n^{1.63})$ and with $O(n^{2.46})$, respectively. Furthermore, the results show that hBOA significantly outperforms GA+DHC and UMDA+DHC in both the number of evaluations and the number of DHC flips. The worst performance is achieved by UMDA+DHC, the time complexity of which grows faster than polynomially.

Recall that for spin glasses, one-point crossover performs relatively well because one-point crossover rarely breaks important interactions between spins due to the used representation. Nonetheless, this behavior cannot be generalized to other similar slowly equilibrating problems that exhibit different energy landscapes, such as protein folding or polymer dynamics.



(a) Number of evaluations. (b) Number of flips.

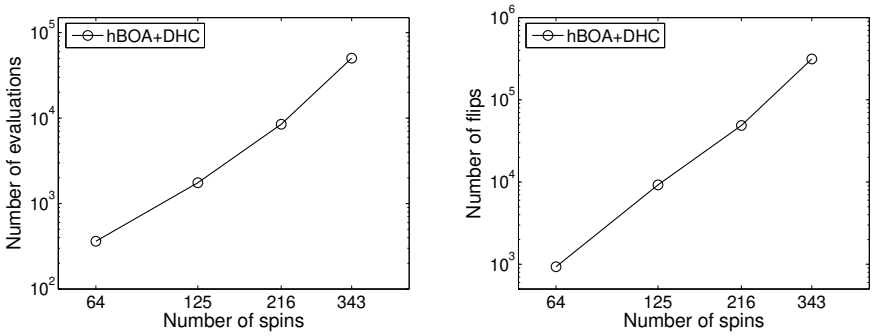
Fig. 1. Performance of hBOA+DHC, UMDA+DHC, and GA+DHC on random 2D $\pm J$ Ising spin glasses.

For two-dimensional Ising spin glasses, a polynomial algorithm [34, 35] with complexity $O(n^{3.5})$ exists that computes the number of states at each energy level, including the ground state. It was shown [36] that on two-dimensional $\pm J$ Ising spin glasses, hBOA+DHC achieves asymptotic performance of the polynomial algorithm without any prior knowledge about spin glasses.

Figure 2 shows the performance of hBOA+DHC on three-dimensional $\pm J$ spin glasses. Since both GA+DHC and UMDA+DHC have not been capable of solving most three-dimensional instances even with enormous computational resources, we only include the results for hBOA+DHC. The results show that the performance of hBOA+DHC appears to grow exponentially fast. This behavior is expected because the problem of finding ground states of three-dimensional spin glasses is NP-complete [37] and it is thus unlikely that there exists an algorithm that can solve this class of problems in polynomial time. However, we see that hBOA+DHC is still capable of solving instances of several hundreds spins, which are intractable with most standard optimization algorithms, such as genetic algorithms and simulated annealing.

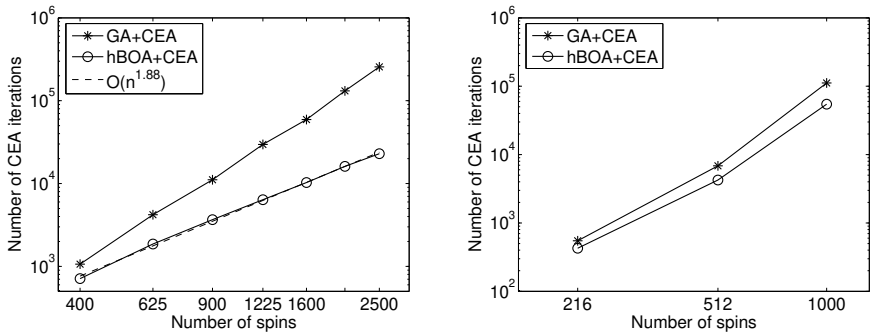
Figure 3a shows the performance of hBOA+CEA and GA+CEA on two-dimensional Ising spin glasses with $\pm J$ couplings. The results indicate that hBOA+CEA significantly outperforms GA+CEA and thus hBOA retains superior performance even with CEA. The results also show that incorporating CEA leads to a somewhat faster asymptotic growth of time complexity with problem size; on the other hand, the use of CEA provides a significant decrease of running time for the tested range of problems and, consequently, much larger problem sizes can be treated currently as compared to hBOA+DHC. Nonetheless, based on these results, it can be hypothesized that hBOA+DHC will become faster than hBOA+CEA for much larger spin-glass instances.

Figure 3b shows the performance of hBOA+CEA and GA+CEA on three-dimensional Ising spin glasses with $\pm J$ couplings. The results indicate that the performance of both algorithms grows faster than polynomially even with the



(a) Number of evaluations. (b) Number of flips.

Fig. 2. Performance of hBOA+DHC on random 3D $\pm J$ Ising spin glasses.



(a) 2D $\pm J$ Ising spin glass. (b) 3D $\pm J$ Ising spin glass.

Fig. 3. Performance of hBOA+CEA and GA+CEA on $\pm J$ Ising spin glasses.

use of CEA as is expected from the NP-completeness of this problem. However, CEA improves the performance of GA significantly and makes the difficult three-dimensional instances tractable even with GA. Nonetheless, hBOA+CEA still retains superior performance.

5 Summary and Conclusions

This paper tested the hierarchical Bayesian optimization algorithm (hBOA) on a large number of instances of the problem of finding ground states of Ising spin glasses with random couplings in two and three dimensions. The performance of hBOA was compared to that of the simple genetic algorithm (GA) and the univariate marginal distribution algorithm (UMDA). All algorithms were hybridized by using either a simple deterministic hill climber (DHC) or the cluster exact approximation (CEA). The results showed that hBOA significantly outperforms all other compared methods in all cases and that in some classes of Ising spin glasses, it achieves the performance of best polynomial-time analytical methods.

The results presented in this paper thus confirm that using hierarchical decomposition for solving difficult optimization problems with little problem-specific knowledge holds a big promise and that advanced estimation of distribution algorithms offer a robust and scalable class of optimization algorithms applicable to important classes of difficult problems.

Acknowledgments

The authors would like to thank Fabien Alet, Martin V. Butz, David E. Goldberg, Jiri Ocenasek, Kumara Sastry, Simon Trebst and Matthias Troyer.

This work was supported by the NSF CARREER grant ECS-0547013, by the Research Award and the Research Board at the University of Missouri, and by the *VolkswagenStiftung* (Germany) within the program "Nachwuchsgruppen an Universitäten". Most experiments were done using the hBOA software developed by Martin Pelikan and David E. Goldberg at the University of Illinois at Urbana-Champaign.

References

1. Binder, K., Young, A.: Spin-glasses: Experimental facts, theoretical concepts and open questions. Rev. Mod. Phys. **58** (1986) 801
2. Mezard, M., Parisi, G., Virasoro, M.: Spin glass theory and beyond. World Scientific, Singapore (1987)
3. Fischer, K., Hertz, J.: Spin Glasses. Cambridge University Press, Cambridge (1991)
4. Young, A., ed.: Spin glasses and random fields. World Scientific, Singapore (1998)
5. Hartmann, A.K., Rieger, H.: Optimization Algorithms in Physics. Wiley-VCH, Weinheim (2001)
6. Hartmann, A.K., Rieger, H., eds.: New Optimization Algorithms in Physics. Wiley-VCH, Weinheim (2004)
7. Hartmann, A.K., Weigt, M.: Phase Transitions in Combinatorial Optimization Problems. Wiley-VCH, Weinheim (2005)
8. Mühlenbein, H., Mahnig, T.: Convergence theory and applications of the factorized distribution algorithm. Journal of Computing and Information Technology **7**(1) (1999) 19–32
9. Naudts, B., Naudts, J.: The effect of spin-flip symmetry on the performance of the simple GA. Parallel Problem Solving from Nature (1998) 67–76
10. Hartmann, A.K.: Ground-state clusters of two, three and four-dimensional +/-J Ising spin glasses. Phys. Rev. E **63** (2001) 016106
11. Van Hoyweghen, C.: Detecting spin-flip symmetry in optimization problems. In Kallel, L., et al., eds.: Theoretical Aspects of Evolutionary Computing. Springer, Berlin (2001) 423–437
12. Pelikan, M., Goldberg, D.E.: Hierarchical BOA solves Ising spin glasses and MAXSAT. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2003) **II** (2003) 1275–1286
13. Fischer, S., Wegener, I.: The Ising model on the ring: Mutation versus recombination. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004) (2004) 1113–1124

14. Dayal, P., Trebst, S., Wessel, S., Würtz, D., Troyer, M., Sabhapandit, S., Coppersmith, S.: Performance limitations of flat histogram methods and optimality of Wang-Langdau sampling. Physical Review Letters **92**(9) (2004) 097201
15. Pelikan, M., Goldberg, D.E.: Escaping hierarchical traps with competent genetic algorithms. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001) (2001) 511–518
16. Pelikan, M.: Hierarchical Bayesian optimization algorithm: Toward a new generation of evolutionary algorithms. Springer-Verlag (2005)
17. Hartmann, A.K.: Cluster-exact approximation of spin glass ground states. Physica A **224** (1996) 480
18. Mühlenbein, H., Paaß, G.: From recombination of genes to the estimation of distributions I. Binary parameters. Parallel Problem Solving from Nature (1996) 178–187
19. Larrañaga, P., Lozano, J.A., eds.: Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation. Kluwer, Boston, MA (2002)
20. Pelikan, M., Goldberg, D.E., Lobo, F.: A survey of optimization by building and using probabilistic models. Computational Optimization and Applications **21**(1) (2002) 5–20
21. Chickering, D.M., Heckerman, D., Meek, C.: A Bayesian approach to learning Bayesian networks with local structure. Technical Report MSR-TR-97-07, Microsoft Research, Redmond, WA (1997)
22. Friedman, N., Goldszmidt, M.: Learning Bayesian networks with local structure. In Jordan, M.I., ed.: Graphical models. MIT Press, Cambridge, MA (1999) 421–459
23. Harik, G.R.: Finding multimodal solutions using restricted tournament selection. Proceedings of the International Conference on Genetic Algorithms (ICGA-95) (1995) 24–31
24. Holland, J.H.: Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, MI (1975)
25. Goldberg, D.E.: Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading, MA (1989)
26. Claiborne, J.: Mathematical Preliminaries for Computer Networking. Wiley, New York (1990)
27. Swamy, M., Thulasiraman, K.: Graphs, Networks and Algorithms. Wiley, New York (1991)
28. Picard, J.C., Ratliff, H.: Minimum cuts and related problems. Networks **5** (1975) 357
29. Träff, J.: A heuristic for blocking flow algorithms. Eur. J. Oper. Res. **89** (1996) 564
30. Tarjan, R.: Data Structures and Network Algorithms. Society for industrial and applied mathematics, Philadelphia (1983)
31. Spin Glass Ground State Server. `http://www.informatik.uni-koeln.de/ls_juenger/research/sgs/sgs.html` (2004) University of Köln, Germany.
32. Sastry, K., Goldberg, D.E.: Analysis of mixing in genetic algorithms: A survey. IlliGAL Report No. 2002012, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory, Urbana, IL (2002)
33. Middleton, A., Fisher, D.S.: The three-dimensional random field Ising magnet: Interfaces, scaling, and the nature of states. Phys. Rev. B **65** (2002) 134411
34. Galluccio, A., Loebl, M.: A theory of Pfaffian orientations. I. Perfect matchings and permanents. Electronic Journal of Combinatorics **6**(1) (1999) Research Paper 6.

35. Galluccio, A., Loebl, M.: A theory of Pfaffian orientations. II. T-joins, k-cuts, and duality of enumeration. Electronic Journal of Combinatorics **6**(1) (1999) Research Paper 7.
36. Pelikan, M., Ocenasek, J., Trebst, S., Troyer, M., Alet, F.: Computational complexity and simulation of rare events of Ising spin glasses. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004) **2** (2004) 36–47
37. Barahona, F.: On the computational complexity of Ising spin glass models. Journal of Physics A: Mathematical, Nuclear and General **15**(10) (1982) 3241–3253