Extremal optimization of graph partitioning at the percolation threshold



1999 J. Phys. A: Math. Gen. 32 5201

(http://iopscience.iop.org/0305-4470/32/28/302)

J. Phys. A: Math. Gen. **32** (1999) 5201–5211. Printed in the UK PII: S0305-4470(99)00928-2

# Extremal optimization of graph partitioning at the percolation threshold

Stefan Boettcher

Physics Department, Emory University, Atlanta, GA 30322, USA
and
Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Received 12 January 1999

**Abstract.** The benefits of a recently proposed method to approximate hard optimization problems are demonstrated on the graph partitioning problem. The performance of this new method, called extremal optimization (EO), is compared with simulated annealing (SA) in extensive numerical simulations. While generally a complex (NP-hard) problem, the optimization of the graph partitions is particularly difficult for sparse graphs with average connectivities near the percolation threshold. At this threshold, the relative error of SA for large graphs is found to diverge relative to EO at equalized runtime. On the other hand, EO, based on the extremal dynamics of self-organized critical systems, reproduces known results about optimal partitions at this critical point quite well.

## 1. Introduction

The optimization of systems with many degrees of freedom with respect to some cost function is a frequently encountered task in physics and beyond [1]. In cases where the relation between individual components of the system is frustrated [2], such a cost function often exhibits a complex 'landscape' [3] over the space of all configurations. For growing system size, the cost function may exhibit an exponentially increasing number of unrelated local extrema separated by sizable barriers which makes the search for the exact, optimal solution usually unreasonably costly. Thus, it is of great importance to develop fast and reliable methods to find near-optimal solutions for such problems.

The observation of certain physical processes, in particular the annealing of disordered materials, have lead to general-purpose optimization methods such as 'simulated annealing' (SA) [5, 6]. SA applies the formalism of equilibrium statistical mechanics and in principle only requires the cost function as input. Thus, it is applicable to a variety of problems. But the performance of SA is hard to assess in general, even when limited to the standard combinatorial optimization problems. Aside from a multitude of adjustable parameters that crucially determine the quality of SA's performance in a particular context, typical combinatorial optimization problems themselves possess various parameters that may change the landscape and SA's behaviour drastically [7].

In this paper we will explore the properties of a new general-purpose method, called extremal optimization (EO) [4], in comparison with SA. In contrast to SA, EO is based on ideas from non-equilibrium physics. As the basis for comparison we will use the graph partitioning problem (GPP), a standard NP-hard combinatorial optimization problem [8] with similarities to disordered spin systems. We find that the GPP has a critical point as a function of the

0305-4470/99/285201+11$30.00  

connectivity of graphs, with a less complex phase at lower connectivities. This critical point is related to the percolation transition of the graphs. Near this critical point, the performance of SA markedly deteriorates while EO produces only small errors.

This paper is organized as follows: in the section 2 we describe the philosophy behind the EO method, in section 3 we introduce the graph partitioning problem, and section 4 we present the algorithms and the results obtained in our numerical comparison of SA and EO, followed by conclusions in section 5.

## 2. Extremal optimization

EO provides an entirely new approach to optimization [4], based on the non-equilibrium dynamics of systems exhibiting self-organized criticality (SOC) [9]. SOC often emerges when a system is dominated by the evolution of extremely atypical degrees of freedom [10].

A simple example of such a dynamical system which inspired the development of EO is the Bak–Sneppen model [11]. There, species are represented by a number between 0 and 1 that indicates their 'fitness,' located on the sites of a lattice. The smallest number (representing the worst adapted species) at each update is discarded and replaced with a new number drawn from a uniform distribution on [0, 1]. Without any interactions, all the numbers in the system would eventually become 1. But obvious interdependences between species provide constraints for balancing the systems' fitness with that of each species: the change of fitness in one species impacts the fitness of an interrelated species. In the Bak–Sneppen model, the fitness values on all sites neighbouring the smallest number at that time step are simply replaced with new random numbers as well†. After a certain number of such updates, the system organizes itself into a highly correlated state known as self-organized criticality (SOC) [9].

In the SOC state, almost all species have reached a fitness above a certain threshold. But these species merely possess what is referred to as punctuated equilibrium [11, 12], because the co-evolutionary activity is bound to return in a chain reaction where a weakened neighbour can undermine one's own fitness. Fluctuations that rearrange the fitness of many species abound and can rise to the size of the system itself, making any possible configuration accessible. Hence, such non-equilibrium systems provide a high degree of adaptation for most entities in the system without limiting the scale of change towards even better states.

EO attempts to utilize this phenomenology to obtain near-optimal solutions for optimization problems [13]. For instance, in a spin glass system [1] we may consider as fitness for each spin its contribution to the total energy of the system. EO would search for ground state configurations by perturbing preferentially spins with large contributions. As in the Bak–Sneppen model, such perturbations would be local, random rearrangements of those poorly adapted spins, allowing for better as well as for worse outcomes at each update. In the same way as systems exhibiting SOC get driven recurrently towards a small subset of attractor states through a sequence of 'avalanches' [9, 14], EO can fluctuate widely to escape local optima while the extremal selection process ensures recurrent approaches to many near-optimal configurations. Especially in exploring low-temperature properties of disordered spin systems, those qualities may help to avoid the extremely slow relaxation behaviour faced by heat-bath-based approaches [15]. In that, EO provides an approach alternative—and apparently equally capable [4]—to genetic algorithms, which are often the only means to illuminate those important properties [16]. The partitioning of sparse graphs as discussed here is particularly pertinent in preparation for similar studies on actual spin glasses.

† In the Bak–Sneppen model, the relation between species is not specified in detail. In reality, some fixed ('quenched') structure may exist between any two species that determines how the adaptive change in one effects the other.

It has been observed that many optimization problems exhibit critical points that separate off phases with simple cases of a generally hard problem [17]. Near such a critical point, finding solutions becomes particularly difficult for local search methods which proceed by exploring for an existing solution some neighbourhood in configuration space. There, near-optimal solutions become widely separated with diverging barrier heights between them. It is not surprising that search methods based on heat-bath techniques like SA are not particularly successful in this highly correlated state [15]. In contrast, the driven dynamics of EO does not possess any temperature-control parameters to successively limit the scale of its fluctuations. Our numerical results in section 4 show that EO's performance does not diminish near such a critical point. A non-equilibrium approach like EO may thus provide a general-purpose optimization method that is complementary to SA: while SA has the advantage far from this critical point, EO appears to work well 'where the *really* hard problems are' [17].

## 3. Graph partitioning

To illustrate the properties of EO and its differences with SA, we focus in this paper on the well-studied graph partitioning problem (GPP). In particular, we will consider the GPP near a phase transition where the optimization problem becomes especially difficult and possesses many similarities with physical systems.

### *3.1. Formulation of the problem*

The graph (bi-)partitioning problem is easy to formulate: take $N$ points where $N$ is an even number, let any pair of two points be connected by an edge with a certain probability, divide the points into two sets of equal size $N/2$ such that the number of edges connecting both sets, the 'cutsize' $m$, is minimal: $m = m_{\mathrm{opt}}$. The global constraint of an equal division of the points between the sets places this problem generally among the hardest problems in combinatorial optimization, requiring a computational effort that would grow faster than any power of $N$ to determine the *exact* solution with certainty [8]. The two physically motivated optimization methods, SA and EO, which we focus on here, usually obtain *approximate* solutions in polynomial time.

For random graphs, the GPP depends on the probability $p$ with which any two points in the system are connected. Thus, $p$ determines the total number of edges in an instance, $L = pN(N-1)/2$ on average, and its mean connectivity per point, $\alpha = p(N-1)$ on average. Alternatively, we can formulate a 'geometric' GPP by specifying $N$ randomly distributed points in the two-dimensional unit square which are connected with each other if they are located within a distance $d$ of one another. Then, the average expected connectivity $\alpha$ of such a graph is given by $\alpha = N\pi d^2$. This form of the GPP has the advantage of a simple graphical representation, as shown in figure 1.

It is known that geometric graphs are harder to optimize than random graphs [18]. The characteristics of the GPP for random and geometric graphs at low connectivity appear to be very different due to the dominance of long loops and short loops, respectively, and we present results for both types of graphs here. In fact, in the case of random graphs the structure is locally tree-like which allows for a mean-field treatment that yields exact results [19–21]. In turn, the geometric case corresponds to continuum percolation of 'soft' (overlapping) circles for which precise numerical results exist [22]. Finally, we also try to determine the average ground state energy of a dilute ferromagnetic system on a cubic lattice at fixed (zero) magnetization, which amounts to the equal partitioning of 'up' and 'down' spins while minimizing the interface between both types [20]. Here, each vertex of the lattice holds a $\pm$-spin, and any two nearest-

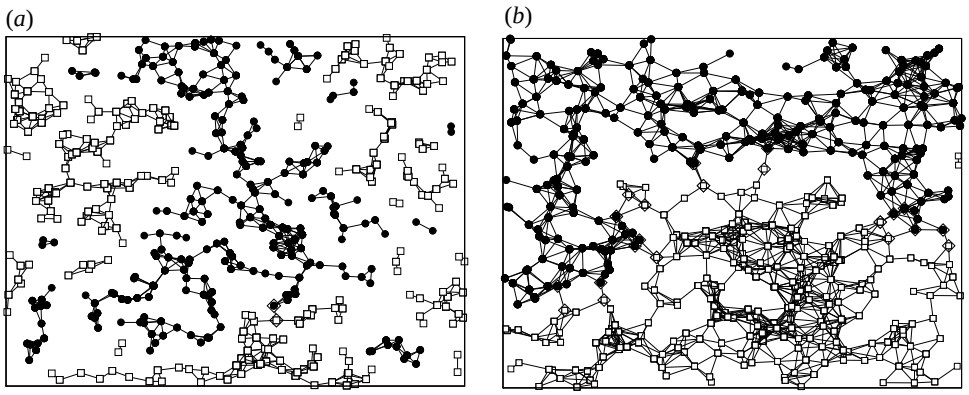


**Figure 1.** Two random geometric graphs, $N = 500$, with $\alpha = 4$ (*a*) and $\alpha = 8$ (*b*) in an optimized configuration found by EO. At $\alpha = 4$ the graph barely percolates, with merely one 'bad' edge (between points of opposite sets, masked by diamonds) connecting a set of 250 round points with a set of 250 square points, thus $m_{\rm opt} = 1$. For the denser graph in (*b*), EO reduced the cutsize to $m_{\rm opt} = 13$.

neighbour spins either possess a ferromagnetic coupling of unit strength or are unconnected. The probability that a coupling exists is fixed such that the average connectivity of the system is $\alpha$.

### *3.2. Graph partitioning and percolation*

Like many other optimization problems, the GPP exhibits a critical point as a function of its parameters [17]. In case of the GPP we observe this critical point as a function of the connectivity $\alpha$ of graphs, with the cutsize $m_{\rm opt}$ as the order parameter. In fact, the critical point of partitioning is closely linked to the percolation threshold of graphs. In our numerical simulations we proceed by averaging over many instances of a class of graphs and try to reproduce well known results from the corresponding percolation problem. Of course, using stochastic optimization methods (instead of cluster enumeration) is neither an efficient nor a precise means to determine percolation thresholds. But in turn we also obtain some valuable information about the scaling behaviour of the average cost $\langle m_{\rm opt}\rangle$ for optimal partitions near the threshold that goes beyond the percolating properties of these graphs.

We note, in accordance with [23], that the critical point separates between hard cases and easy-to-solve cases of the GPP. The transition is related to the corresponding percolation problem for the graphs in the following manner: if the mean connectivity $\alpha$ is very small, the graph of $N$ points consists mainly of disconnected, small clusters or isolated points which can be enumerated and sorted into two equal partitions in polynomial time with no edges between them ($m_{\rm opt} = 0$). If $\alpha$ is large and the probability that any two points are connected is $p = {\rm O}(1)$, almost all points are connected into one giant cluster with $m_{\rm opt} = {\rm O}(N^2)$, and almost any partition leads to an acceptable solution. But when $p = {\rm O}(1/N)$, i.e. $\alpha = {\rm O}(1)$, the distribution of cluster sizes is broad, and the partitioning problem becomes nontrivial. Obviously, as soon as a cluster of size $> N/2$ appears, $m_{\rm opt}$ must be positive. In this sense, we observe for $N \to \infty$ a sharp, percolation-like transition at an $\alpha_{\rm crit}$ with the cutsize $m_{\rm opt}$ as the order parameter.

For random graphs it is known that a cluster of size $N$ exists for $\alpha > 1$ [19], but only for $\alpha > \alpha_{\rm c} = 2\ln 2 \approx 1.386$ do we find a cluster of size $> N/2$ [20]. Geometric graphs in

$D = 2$ are known to percolate at about $\alpha = 4.5$ [22], and we would expect $\alpha_c$ for the GPP to be slightly larger than that. Also, the dilute ferromagnet should exhibit a non-trivial energy when the fraction of occupied bonds reaches slightly beyond the critical point $p_c \approx 0.2488$ for bond percolation on a cubic ($D = 3$) lattice [24], i.e. for connectivities $\alpha > 2Dp_c \approx 1.493$.

## 4. Numerical experiments

### *4.1. SA algorithm*

In SA [5], we try to minimize a global cost function given by $f = m + \mu(P_1 - P_2)^2$, where $P_1$ and $P_2$ are the number of points in the respective sets. Allowing the size of the sets to fluctuate is required to improve SA's performance in outcome and computational time at the cost of an arbitrary parameter $\mu$ to be determined. Then, starting at a 'temperature' $T_0$, the annealing schedule proceeds with $lN$ trial Monte Carlo steps on $f$ by tentatively moving a randomly chosen point from one set to the other (which changes $m$) to equilibrate the system. This move is accepted, if $f$ improves or if the Boltzmann factor $\exp[(f_{\rm old} - f_{\rm new})/T]$ is larger than a randomly drawn number between 0 and 1. Otherwise the move is rejected and the process continues with another randomly chosen point. After that, we set $T_i = T_{i-1}(1-\epsilon)$, equilibrate again for $lN$ trials, and so on, until the MC acceptance rate drops below $A_{\rm stop}$ for $K$ consecutive temperature levels. At this point the optimization process can be considered 'frozen' and the configuration should be near-optimal, $m \approx m_{\rm opt}$ (and balanced, $P1 = P2$). While SA is intuitive, controlled, and of very general applicability, its performance in practice is strongly dependent on the multitude of parameters which have to be arduously tuned. For us it is thus expedient (and most unbiased!) to rely on an extensive study of SA for graph partitioning [18] which determined $\mu = 0.05$, $T_0 = 2.5$, $\epsilon = 0.04$, $A_{\rm stop} = 2\%$, and $K = 5$. [18] set $l = 16$, but performance improved noticeably for our choice, $l = 64$.

### *4.2. EO algorithm*

In EO [4], each point $i$ obtains a 'fitness' $\lambda_i = g_i/(g_i + b_i)$ where $g_i$ and $b_i$ are the number of 'good' and 'bad' edges that connect that point within its set and across the partition, respectively. (We fix $\lambda_i = 1$ for isolated points.) Of course, point $i$ has an individual connectivity of $\alpha_i = g_i + b_i$ while the overall mean connectivity of a graph is given by $\alpha = \sum_i \alpha_i/N$. The current cutsize is given by $m = \sum_i b_i/2$. At all times, an ordered list $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_N$ is maintained where $\lambda_n$ is the fitness of the point with the $n$th rank in the list.

At each update we draw two numbers, $1 \leqslant n_1, n_2 \leqslant N$, from a probability distribution

$$P(n) \sim n^{-\tau}. \tag{1}$$

Then we pick the points which are elements $n_1$ and $n_2$ of the rank-ordered list of fitnesses. (We repeat a drawing of $n_2$ until we obtain a point that is from the opposite set than $n_1$.) These two points swap sets *no matter what* the resulting new cutsize $m$ may be, in notable distinction to the (temperature-) scale-dependent Monte Carlo update in SA. Then, these two points, and all points they are connected to ($2\alpha$ on average), re-evaluate their fitness $\lambda$. Finally, the ranked list of $\lambda$ is reordered using a 'heap' at a computational cost $\propto \alpha \ln N$, and the process is started again. We repeat this process for a number of update steps per run that rises linearly with system size, and we store the best result generated along the way. Note that no scales are introduced into the process, since the selection follows a scale-free power-law distribution $P(n)$ and since—unlike in a heat bath—all moves are accepted, allowing for fluctuations on all scales. Instead of a global cost function, the rank-ordered list of fitnesses provides the information about optimal configurations. This information emerges in a self-organized

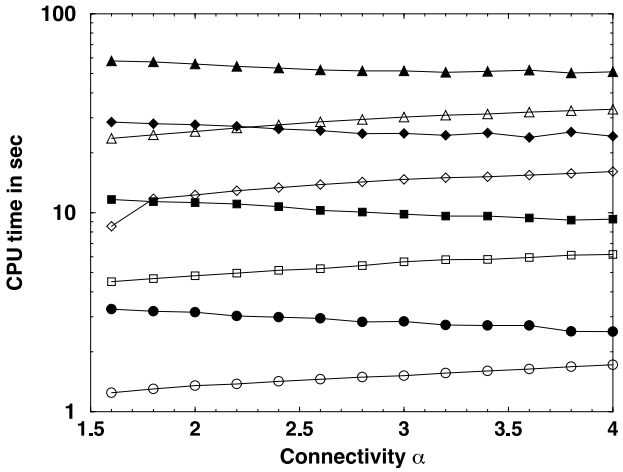


**Figure 2.** Typical runtime comparison of SA and EO for the ferromagnet on a 200 MHz Pentium. Runtimes generally rise $\propto N$, fall with the connectivity $\alpha$ for SA (filled symbols), but rise $\propto \alpha$ for EO (unfilled symbols). Circles refer to an average runtime of graphs with $N = 512 = 8^3$ points, squares to $N = 1728 = 12^3$, diamonds to $N = 4096 = 16^3$, and triangles to $N = 8000 = 20^3$.

manner merely by selecting with a bias *against* badly adapted points, instead of 'breeding' better ones [11].

There is merely one parameter, the exponent $\tau$ in the probability distribution in equation (1), that controls the selection process and optimizes the performance of EO. In initial studies, we determined $\tau = 1.4$ as the optimal value for all graphs. It is intuitive that such an optimal value of $\tau$ should exist: if $\tau$ is too small, points would be picked purely at random with no gradient towards a good partition, while if $\tau$ is too large, only a small number of points with particularly bad fitness would be chosen over and over again, confining the system to a poor local optimum. It is a surprising numerical result that this value of $\tau$ appears to be rather universal, independent of $N$, $\alpha$, and the type of graph considered.

### *4.3. Testbed of graphs*

In our numerical simulations we have generated random and $2D$ geometric graphs of varying connectivity by choosing $p$ or $d$, respectively. For any instance of a graph labelled by a 'connectivity $\alpha$', the actual connectivity not only varies from point to point, but also the mean connectivity of such graphs follows a normal distribution. (In particular for geometric graphs it is shifted to lower values due to the loss of connectivity at the boundaries.) For $N = 500$, 1000, 2000, 4000, 8000, and 16 000, we varied the connectivity between $\alpha = 1.25$ and $\alpha = 5$ for random graphs, and $\alpha = 4$ and $\alpha = 10$ for geometric graphs. Then, for each $\alpha$ we generated 16 different instances of graphs, identical for SA and EO. On each instance, we performed eight (32) optimization runs for random (geometric) graphs, both for EO and SA. For each run, we used a new random seed to establish an initial partition of the points. SA's runs terminate when the system freezes. We terminated EO-runs after $200N$ updates, leading to a comparable runtime for both methods.

For the dilute ferromagnet, we fixed the number of couplings to obtain a specific average connectivity $\alpha$. Those couplings were then placed on random links between nearest-neighbour spins to generate an instance. We used 16 instances, and 16 runs for each, at connectivities $1.6 \leqslant \alpha \leqslant 4$. Here, we only used $100N$ updates for EO, and the temperature length of $16N$

recommended in [18] but with a higher starting temperature for SA to optimize performance at a comparable runtime for both methods, as shown in figure 2.

### *4.4. Evaluation of results*

*4.4.1. Comparison of EO and SA.* We evaluate the performance of SA and EO separately. For each method, we only take its best result for each instance and average those best results at any given connectivity $\alpha$ to obtain the mean cutsize for that method as a function of $\alpha$ and $N$. To compare EO and SA, we determine the relative error of SA with respect to the best result found by either method for $\alpha \geqslant \alpha_c$. Figures 3(*a*)–(*c*) show how the error of SA diverges with increasing $N$ near to $\alpha_c$ for each class of graphs.

Depending on the type of graph under consideration, the quality of the SA results may vary. The data for random graphs in figure 3(*a*) only shows a relatively weak deficit in SA's performance relative to EO. Near $\alpha_c = 2 \ln 2 = 1.386$, SA's relative error remains modest, and only grows very weakly with increasing $N$. For large connectivities $\alpha$, SA quickly becomes the superior method for random graphs, which may be due to their increasingly homogeneous structure (i.e. low barriers between optima) that does not favour EO's large fluctuations. On the other hand, the averages obtained by EO appear to be very smooth (see the scaling in figure 4(*a*)) whereas the apparent noise in figure 3(*a*) indicates large variations between instances for the SA results.

The very rugged structure of geometric graphs near the percolation threshold, $\alpha_c \approx 4.5$ (see figure 1(*a*)), is most problematic for SA, leading to huge errors which appear to increase linearly with $N$. Barriers between optima are high within each graph, now favouring EO's propensity for large fluctuations. On the scale of figure 3(*b*), error bars attached to the data (which we have generally omitted) would hardly be significant. But experience shows that both methods exhibit large variations in results between instances which is in large part due to actual variations in the structure between geometric graphs.

The results for the dilute ferromagnet exhibit a mix of the two previous cases. Since the points are arranged on a $D = 3$-lattice, the structure of these graphs is definitely geometrical, but local connectivities are limited to the $2D = 6$ nearest neighbours that each point possesses. Again, SA's error is huge and appears to diverge about linearly near the threshold, $\alpha_c \approx 1.5$. But due to the limited rage of connectivities, graphs soon become rather homogeneous for increasing $\alpha$ which in turn appears to favour SA away from the transition, especially for larger graphs. (For larger $N$, any local structure gets quickly averaged out due to the local limits on the connectivity, whereas an unlimited range of local structures can emerge in the geometric graphs above.)

*4.4.2. Scaling of EO-data near the transition.* For the data obtained with EO, we make an ansatz

$$\langle m_{\mathrm{opt}} \rangle \sim N^{\nu} \, (\alpha - \alpha_c)^{\beta} \tag{2}$$

to scale the data for all $N$ onto a single curve, as shown in figures 4(*a*)–(*c*). From the scaling ansatz we can extract an estimate for $\alpha_c$ to compare with percolation results as a measure of the accuracy of the data obtained with EO. Furthermore, we also obtain a numerical estimates for the exponents $\nu$ and $\beta$ which characterize the transition. The exponent $\nu$, describing the finite-size scaling behaviour, could be infered from general, global properties of a class of graphs. For instance, $\nu = 1$ for random graphs because any global property of these graphs is extensive [20]. On the other hand, the exponent $\beta$, describing the scaling of the order parameter near the transition, is related to the intricate structure of the interface needed to separate points

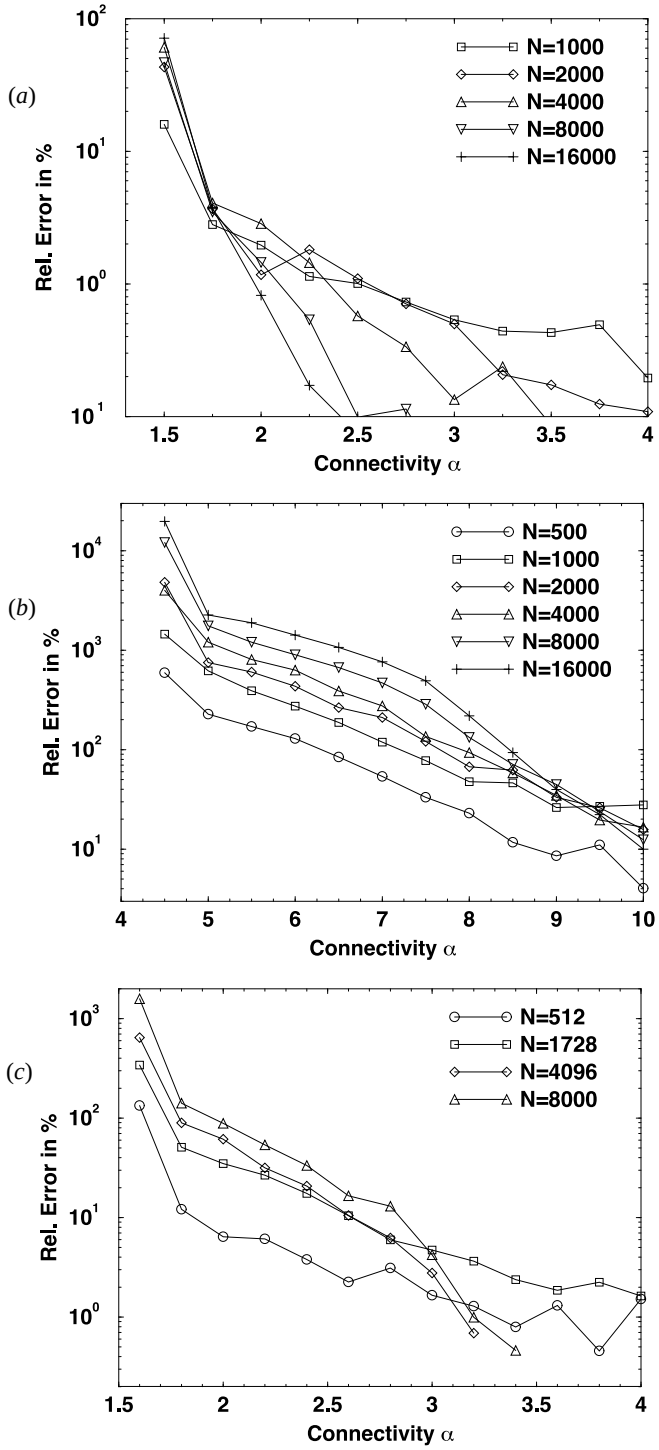


**Figure 3.** Plot of the error of SA relative to the best result found on (*a*) random graphs, (*b*) geometric graphs, and (*c*) the dilute ferromagnet as a function of the mean connectivity $\alpha$. While the error near $\alpha_c$ only increases slowly for random graphs, it appears to increase linearly with $N$ for the ferromagnet and for geometric graphs. At fixed but large connectivities, SA increasingly gains on EO for rising $N$.

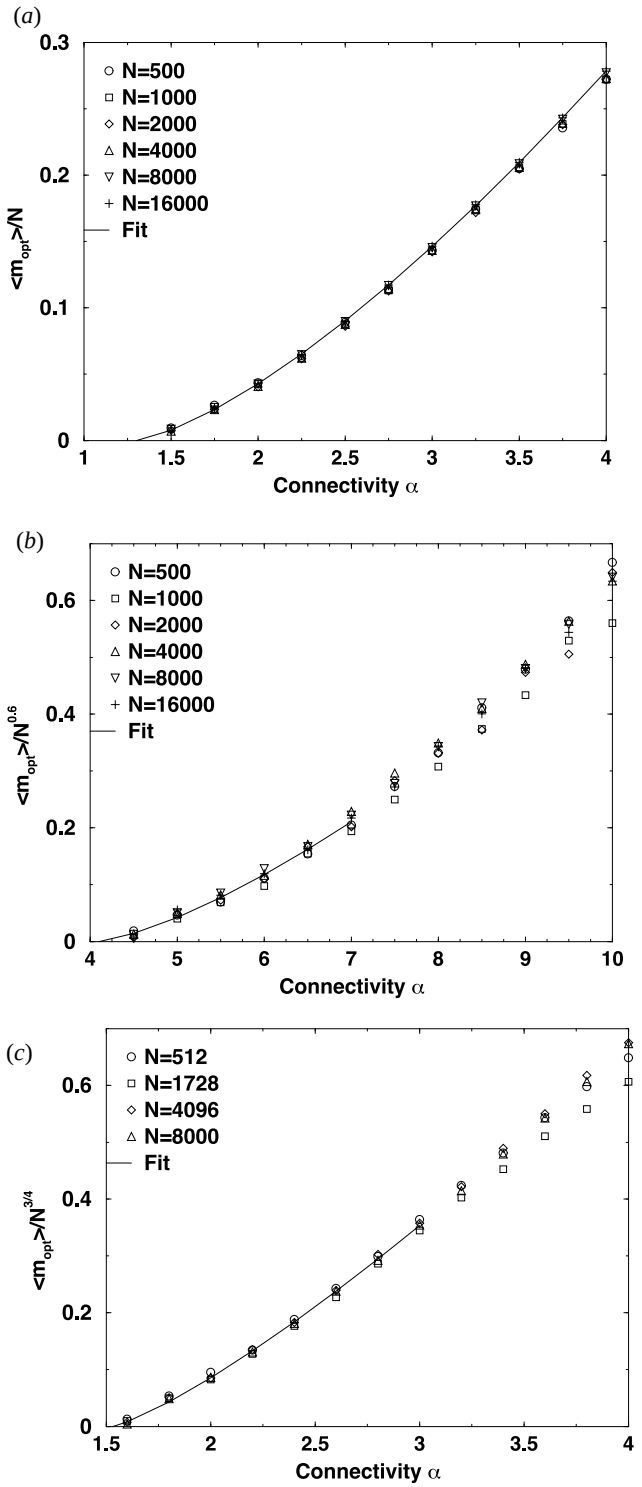


**Figure 4.** Scaling plot of the data from EO according to equation (2) for (*a*) random graphs, (*b*) geometric graphs, and (*c*) the dilute ferromagnet as function of the mean connectivity $\alpha$. The scaling parameters and fits are discussed in the text.

into equal-sized partitions. Thus, we would expect $\beta$ to be non-trivial even for random graphs. (To our knowledge, no previous predictions for these exponents exist.)

For random graphs in figure 4(*a*), the scaling ansatz in equation (2) is particularly convincing. We verify that $\nu = 1$ and obtain $\beta = 1.2$. From the fit we also obtain $\alpha_c \approx 1.30$, just slightly below the exact value of 1.38 [20]. The fit produces an error of about $\pm 0.1$ in the determination of $\beta$, which would ignore any error received through the limited number of instances averaged over, or any bias due to the shortcomings of EO to approach the exact optima. A satisfactory fit in turn would indicate that such errors should be negligible.

For geometric graphs in figure 4(*b*), we found the best scaling for $\nu = 0.6$. Since we used only 16 different instances to average over at each $N$ and $\alpha$, the data gets very noisy for larger connectivities due to large fluctuations in the optimal cutsizes between those instances and/or EO's inability to find good approximations. We chose to fit only points up to $\alpha = 7$ and obtained $\beta = 1.4$ and $\alpha_c \approx 4.1$, even smaller than the critical value for percolation, 4.5. Obviously, the obtained values are very poor, but at least indicate EO's ability to approximate the optimal cutsizes with bounded error near the transition.

The data for the dilute ferromagnet in figure 4(*c*) appears to scale well for $\nu = 0.75$. Since EO's performance is falling behind that of SA for $\alpha > 3$ we only fit to smaller values of $\alpha$ and obtain $\beta = 1.15$ and $\alpha_c = 1.55$, as desired just slightly larger than the value for percolation, 1.49. We estimate the error from the fit for each of these values to be about $\pm 0.05$.

*4.4.3. Fixed-valence graphs.* Finally, we have also performed a study on graphs where points are linked at random, but where the connectivity $\alpha$ *at each point* is fixed. These graphs have been investigated previously theoretically [20, 25] and numerically using SA [26]. While $\alpha$ is now fixed to be an integer, we cannot tune ourselves arbitrarily close to a critical point. Furthermore, the problem is non-trivial only when $\alpha \geqslant 3$. These graphs have the property that at a given $\alpha$ and $N$ the optimal cutsizes between instances vary little, and only few instances are needed to determine $\langle m_{\rm opt} \rangle$ with good accuracy.

In our simulations we found that for larger values of $\alpha$, SA and EO both confirm the results in [26] quite well. But for $\alpha = 3$, the lowest non-trivial connectivity, we did observe significant differences between EO and the study in [26]. Banavar *et al* found by averaging five instances each at various values of $N$ ($450 \leqslant N \leqslant 4000$), a normalized average energy

$$E = -1 + \frac{4\langle m_{\rm opt} \rangle}{\alpha N} \tag{3}$$

of $-0.840$, presumably correct to the digits given. We found by averaging over 32 instances, using eight EO runs on each, for $N = 1024$, 2048, and 4096 that $E = -0.844 \pm 0.001$. But this result is still significantly higher than some theoretical predictions [20, 25], and we will investigate whether longer runtimes may further reduce the cutsizes for these graphs [27].

## 5. Conclusions

In this paper we have demonstrated that EO, a new optimization method derived from non-equilibrium physics, may provide excellent results exactly where SA fails. While further studies will be necessary to understand (and possibly, predict) the behaviour of EO, we have used it here to analyse the phase transition in the NP-hard graph partitioning problem. The results illustrate convincingly the advantages of EO and produce a new set of scaling exponents for this transition for a variety of different graphs.

## Acknowledgments

I thank A Percus, P Cheeseman, D S Johnson, D Sherrington and K Y M Wong for very helpful discussions.

## References

[1] Mezard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[2] Toulouse G 1977 *Commun. Phys.* **2** 115
[3] H Frauenfelder *et al* (ed) 1997 *Landscape Paradigms in Physics and Biology* (Amsterdam: Elsevier)
[4] Boettcher S and Percus A G 1999 Artificial intelligence *Preprint* cond-mat/9901351 to appear
[5] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671
[6] Černy V 1985 *J. Optim. Theor. Appl.* **45** 41
[7] Sorkin G B 1991 *Algorithmica* **6** 367–418
[8] Garey M R and Johnson D S 1979 *Computers and Intractability, A Guide to the Theory of NP-Completeness* (New York: Freeman)
[9] Bak P, Tang C and Wiesenfeld K 1987 *Phys. Rev. Lett.* **59** 381
[10] Paczuski M, Maslov S and Bak P 1996 *Phys. Rev.* E **53** 414
[11] Bak P and Sneppen K 1993 *Phys. Rev. Lett.* **71** 4083
[12] Gould S J and Eldridge N 1977 *Paleobiology* **3** 115–51
[13] Boettcher S and Percus A G 1999 *GECCO-99: Proc. Genetic and Evolutionary Computation Conf.* (San Francisco, CA: Morgan Kaufmann) to appear
[14] Dhar D 1990 *Phys. Rev. Lett.* **64** 1613
[15] For those issues see, for instance, Binder K (ed) 1987 *Applications of the Monte Carlo Method in Statistical Physics* (Berlin: Springer) especially section 1.1.5 and related references
[16] Houdayer J and Martin O C 1998 *Ising Spin Glasses in a Magnetic Field Preprint* cond-mat/9811419
[17] Cheeseman P, Kanefsky B and Taylor W M 1991 *Proc. IJCAI-91* ed J Mylopoulos and R Rediter (San Mateo, CA: Morgan Kaufmann) pp 331–7
[18] Johnson D S, Aragon C R, McGeoch L A and Schevon C 1989 *Oper. Res.* **37** 865
[19] Erdös P and Rényi A 1973 *The Art of Counting* ed J Spencer (Cambridge, MA: MIT Press)
[20] Mezard M and Parisi G 1987 *Europhys. Lett.* **3** 1067
Wong K Y M and Sherrington D 1987 *J. Phys. A: Math. Gen.* **20** L793
[21] Janson S, Knuth D E, Luczak T and Pittel B 1993 *Random Struct. Algorithms* **4** 233–358
[22] Balberg I 1985 *Phys. Rev.* B **31** R4053
[23] Fu Y T and Anderson P W 1986 *J. Phys. A: Math. Gen.* **19** 1605
[24] Stauffer D and Aharony A 1992 *Percolation Theory* (London: Taylor and Francis)
[25] Wong K Y M and Sherrington D 1987 *J. Phys. A: Math. Gen.* **20** L793
Wong K Y M, Sherrington D, Mottishaw P, Dewar R and De Dominicis C 1988 *J. Phys. A: Math. Gen.* **21** L99
[26] Banavar J R, Sherrington D and Sourlas N 1987 *J. Phys. A: Math. Gen.* **20** L1
[27] Boettcher S and Percus A G in preparation