A New Automatic Simulated Annealing–type Global Optimization Scheme *

M. Shore † P. D. Coddington ‡ G. C. Fox § E. Marinari ¶ C. H. Wu ||

Abstract

A study of global optimization schemes is presented. Simulated annealing, a general and proven technique for minimizing functions with many coexisting states, is used as a foundation for the development of a new, more automatic approach, called simulated tempering. This novel method upholds the eminent attribute of simulated annealing – the probabilistic guarantee of convergence upon a global minimum. It is unique, however, in that system equilibrium is never disturbed. Although simulated tempering is in its infancy, it is promising, indeed, as the preliminary results suggest. The two–dimensional Ising Spin Glass Model, an NP–complete optimization problem, is used as the test case. Its theoretical formulation is briefly addressed.

1 Introduction

Global optimization, also known as multivariate and combinatorial optimization, is a technique for minimizing complex functions dependent on many variables. Such problems are prevalent in science and, with the advent of high–performance computing, have attracted considerable attention [8] [14]. Intuitively, the formulation consists of a function to be minimized, otherwise known as the *cost function*, and a number of constraints, which specify the feasibility of options. Consider, for example, the task of finding the minimum value of a function which defines a surface is 3–space. The value(s) of the function at the critical points suggest possible optima of which one is convinced after applying the second derivative test (luckily, there were no saddle points!) Armed with two or three discrete results, it is clear which one is smallest (or largest). This method may be characterized as *efficient*, i.e. with it, one is able to solve all *instances*, or *dimensions*, of the problem in a computation time bounded by a polynomial the size of the instance. (In the example above, the instance would equal 3.)

In theory, analytical techniques could be generalized to N–space. In practice, however, there exits a class of optimization problems for which analytical schemes, like the above,

*Project conducted as part of the 1993 Research Experiences for Undergraduates (REU) Program in High-Performance Computing conducted by the Northeast Parallel Architectures Center (NPAC) at Syracuse University. The National Science Foundation provides primary funding for the 1993 NPAC REU Program, through NSF Grant CDA-9200577. Additional funding is provided by the GE Foundation Faculty for the Future Program; the Vice President for Research and Computing, the Dean of the College of Engineering and Computer Science, the Dean of the College of Arts and Sciences, and NPAC, Syracuse University.

†Research Apprentice, 1993 NPAC REU Program; Majors: Computer Science, Music, Florida State University.

‡Project Leader, NPAC, Syracuse, NY.

§Director, NPAC, Syracuse, NY; Professor of Physics, Syracuse University, Syracuse, NY.

¶Associate Director, NPAC, Syracuse, NY; Professor of Physics, Syracuse University, Syracuse, NY; Professor of Physics, Università di Roma *Tor Vergata*, Via E. Cernevale, 00173 Roma, Italy.

||Graduate Assistant, CIS, Syracuse University, Syracuse, NY.

are not realistic. Take, for instance, a macroscopic crystal. N in this case would be on the order of 10^{23} – even contemplation of a numeric calculation is absurd! Indeed, there are no known efficient algorithms for optimizing many hard problems. (The Travelling Salesman Problem and variations of the Ising spin glass model are classic examples.) This class of problems is called *NP–complete* (non–polynomial time complete) and is the motivation for this work.

Simulated annealing is a popular and mature approach to combinatorial optimization and forms the foundation for this project. It is roughly analogous to the physical process of annealing, where a melted substance is gradually cooled until a crystalline state is achieved. In a simulation, the "frozen" substance corresponds to the minimal *cost*, or *ground state*, of the system. (The term *system* is used loosely, as this methodology is, in fact, a general one. For example, it can refer to a collection of cities in the Travelling Salesman Problem or to a group of units in complex circuit design.) Simulated annealing is an iterative method which works due to a simple premise: If the cost obtained by evaluating a configuration of a system at temperature t is less than that of temperature $t{+}1$, then the system is updated to that "less expensive" state. If, on the other hand, the cost of the proposed configuration is higher, it is accepted probabilistically, i.e. it may or may not be accepted. This process has a spectacular implication. By accepting "upward" moves some of the time, this algorithm is able to "jump" out of local minima in its search! In fact, simulated annealing guarantees an optimal solution [5], if truly gradual temperature decrements are observed.

As one may imagine, several ingredients are required to develop a general–purpose global optimization technique. At present, methods that employ simulated annealing need to address a number difficult questions. What is an optimal rate for decreasing the temperature? How long must the system remain at the current temperature before a move to a lower one can be made? What is a suitable criteria for convergence? Clearly, a working implementation would envelope the simple simulated annealing algorithm in an array of complex, computationally–intensive routines. Indeed, this presents the main obstacle for using this technique.

Recently, *simulated tempering*, a novel simulated annealing based approach was introduced which possesses promising attributes [10], [17]. Similar to annealing, the tempering scheme draws on an analogy with a physical process. Instead of gradual cooling however, the desired texture of some substance is achieved by repeated cooling and heating. Theoretically, this technique performs several crucial operations automatically, that require sophisticated implementation for simulated annealing , all the while preserving the salient feature of this class of optimization methods – the probabilistic guarantee of convergence upon the global minimum.

This work may be thought of as a preliminary step towards the development of a general multivariate optimization scheme. The study is organized naturally in two phases: an exploration of simulated annealing followed by the investigation of simulated tempering. The structure of this paper is dictated by the same format. After a brief presentation of the theoretical background for the experimental setting, the two–dimensional Ising spin glass, simulated annealing and simulated tempering are covered. Results and observations are discussed last.

2 The Ising Spin Glass Model

The Ising model, named after, Ernst Ising [1], is certainly a monumental contribution to physics. It aims to describe *phase transitions*, phenomena which occur when a small change in a parameter such as temperature causes a large–scale change in the state of the system. Although a thorough discussion of the Ising model is well outside the scope of this work, it is nevertheless important to describe the experimental setting for this study. In addition, the formulation of the Ising model conveniently addresses several issues of statistical mechanics relevant to simulated annealing–type optimization schemes.

In brief, the model consists of an arrangement of binary values, or *spins*, on a lattice. Every lattice site may correspond to an electron, for instance, with a tiny magnetic field. An independent variable, ς_i, is assigned to each lattice site, $i = 1, \ldots, N$, which collectively represent the configuration space, X, of the system. An essential ingredient in the Ising model is the sum over all possible configurations. Since there are 2^N possible configurations, this sum is clearly enormous if N is at all large.

The total energy of the system is described by the *Hamiltonian* (a natural choice for the cost function). Only nearest–neighbor interactions and interactions of lattice sites with an external field contribute to this quantity. A formal definition is in order. For each configuration, $\varsigma = (\varsigma_1, \ldots, \varsigma_N)$

$$H = H(\varsigma) = -\sum_{<i,j>} E\varsigma_i\varsigma_j - \sum_i J\varsigma_i, \tag{1}$$

where E and J are the energies associated with interactions with nearest–neighbors and with the external field, respectively.

The *partition function*, the central object in statistical mechanics, also plays a fundamental role for the Ising model. From it, one may derive all of the important thermodynamic features of the modeled physical system.

$$Z = Z(\beta, E, J, N) = \sum_{\varsigma=\pm 1} \exp(-\beta H(\varsigma)). \tag{2}$$

Here any dimensions the Hamiltonian may have are cancelled due to the β parameter. Typically, $\beta = 1/kT$, where k is Boltzmann's constant and T is temperature. Since the temperature, T, is incorporated into β, they will be used interchangeably throughout the remainder of this paper. The reader may wish to consult [4], an excellent introduction to the Ising model.

3 Simulated Annealing

As was alluded to in the introduction, a general simulated annealing implementation requires several finely tuned parameters. Perhaps two are most crucial: the *decorrelation* time, which refers to the amount of time (i.e. the number of iterations) the system is kept at a given temperature for sufficient "thermalization," and the *annealing schedule*, or the temperature scale, T, which corresponds to the rate at which the system is "cooled." Though the latter has attracted considerable attention, the former is seldom rigorously addressed. The sections following the formulation of the simulated annealing method describe an adaptive annealing schedule. Due to the constraints of time, the decorrelation criterion was not investigated.

[1] This model was actually originated by Ising's supervisor, Lenz. However, it was Ising who fist published results on the model and hence his name was adopted.

3.1 First formulation

One of the early successful applications of the computer was the simulation of systems with many degrees of freedom. This was motivated by experimental condensed matter and high energy physicists who wished to study the interactions of many particles. Unfortunately, to find the equations of state of models consisting of perhaps millions of elements is not feasible on even today's high–performance supercomputers. The science of statistical mechanics addresses this problem by providing heuristics for accurate estimation of the behavior of *microsystems*, systems consisting of a few elements, based on *macrosystems*, systems consisting of many.

In a typical scenario, a model may be comprised of a number of spheres in a box which interact by forces that decrease with distance. The energy content of the system could be calculated from the distances between the particles. A simulation may proceed as follows. A new configuration is constructed from the current one by imposing a random displacement to a randomly selected particle at each step of the simulation. If the energy of this new state is lower than that of the current state the displacement is accepted unconditionally and the system is updated to the new configuration. If the energy of the new state is higher by $\Delta\varepsilon$ it is accepted with the probability

$$\exp\left(\frac{-\Delta\varepsilon}{kT}\right). \tag{3}$$

This is the fundamental procedure of simulated annealing and is known as the Metropolis step. It was shown [11] [14] that many iterations of the Metropolis method of state generation lead to a distribution of states in which the probability of finding the system in microstate i with energy ε_i is given by

$$P(\varepsilon_i) = \frac{\exp(-\varepsilon_i/kT)}{\Sigma_j \exp(-\varepsilon_j/kT)}. \tag{4}$$

(Notice that the denominator is the partition function.)

The probability function described by equation (4) is known as the Boltzmann density function and has several interesting implications. At very high temperatures, for instance, each state has almost equal chances of being the current state. Consider an experiment which operates on a sample of size N from the space of configurations, X. It is clear that as $T \to \infty$, the probability of finding the system in any of N states approaches $\frac{1}{N}$. For low temperatures, however, only those configurations which result in low energies have a high probability of being selected, as should be clear by similar reasoning.

3.2 The annealing schedule

Since the formulation of simulated annealing as a general–purpose global optimization scheme by Kirkpatrick *et al.* [9] and Cerný [3], the annealing schedule has been the subject of much contemplation [1] [6] [7] [13] [15] [16]. This is not surprising since the rate at which a system is cooled is not merely a performance consideration, but also one of correctness of the algorithm. It is illustrative to draw further on the physical process of annealing. Consider, for instance, a melted substance which is gradually cooled and thus driven toward a crystalline state. If the temperature is decreased too quickly, the resulting crystal may have many defects or even lack all crystalline order.

Most approaches to dynamically adjusting the annealing schedule utilize a concept known as *constant thermodynamic speed.* Intuitively, a condition is imposed that forces the level of statistical uncertainty between two contiguous configurations to be constant. In our experimental implementation statistical uncertainty was obtained by calculating standard deviations of an ensemble of energies associated with configurations of a "thermalized" system (i.e. one which is ready to move to a lower temperature). Nulton and Salamon [13] offer an excellent discussion of this idea.

Following is a summary of the way the said adaptive annealing schedule works.

1. Thermalize the system at temperature, $T = t$.
2. Calculate the standard deviation of an ensemble of energies resulting at $T = t$, $sd(E(t))$.
3. Calculate the subsequent temperature, $T = t + 1$, by

$$T(t+1) = T(t)\exp\left(\frac{-\lambda T(t)}{sd(E(t))}\right). \tag{5}$$

FIG. 1. *The algorithm for the adaptive annealing schedule.*

The λ and the number of energies in the ensemble are constants currently determined *a priori*. (A feasible method for obtaining these values dynamically is proposed in section 5.3.)

4 Simulated Tempering

It should be clear that the fundamental problem with the simulated annealing method is that the system is driven out of equilibrium with every temperature decrement. Though sophisticated techniques for generating temperature schedules and decorrelation times have been investigated, the algorithm suffers in two ways. First, its implementation becomes significantly more complex, and second, additional computational time is incurred. Simulated tempering, a new method developed by Marinari and Parisi [10], overcomes these difficulties more elegantly and proficiently.

4.1 Formulation

The development of the simulated tempering algorithm was motivated by models with phase transitions of order one or greater. Such systems are characterized by coexisting states at T_c (i.e. "freezing"). The original experimental setting, for instance, involved the Random Field Ising Model which has a first order phase transition, i.e. two simultaneous states at T_c: a high and low energy state [17]. Such problems present some difficulty for the annealing approach due to a strictly decreasing temperature scale (the algorithm will tend to get stuck in *one* of these states). The sampling scheme of simulated tempering, however, is not constrained by a decreasing temperature schedule. A simulation of tempering, in fact, treats the temperature parameter, T, stochastically. In actuality, the temperature selection procedure is quite similar to the Metropolis method of phase space evolution.

The tempering procedure results in important advantages over annealing. Namely, the system *always* remains in equilibrium and a broader exploration of local minima is possible. It is realized in the following way. A discrete variable, m, $(m = 1, \ldots, M)$, is added to the configuration space, X. Intuitively, one may regard this step as an introduction of another stochastic parameter which represents a thermal quantity. This allows for a Monte

Carlo procedure to be performed for temperature selection, and thus represents a deviation from one fundamental standard simulated annealing premise – a continual decrease in temperature. The probability distribution is given by

$$P(X, m) \propto \exp(-H_t(X, m)), \tag{6}$$

where β is incorporated into the variable, H_t (H_t is not actually the Hamiltonian because of the m parameter. In the context of this work, H_t will stand for a new term, the tempering Hamiltoninan), resulting from the new configuration space,

$$H_t(X, m) \equiv \beta_m H(X) - g_m. \tag{7}$$

The g_m parameter will be considered in more detail in section 5.2. For the present, it is sufficient to state that it is a normalizing energy value that is required to make equation (7) true due to the introduction of m into the new configuration space.

It is clear that this formulation of H_t results in the Boltzmann distribution for $\beta_m = \beta$ (i.e. in the case of simple phase space, X). In the relevant instance of discrete β's, however, the probability of selecting a particular one is described by

$$P(m) \equiv \exp(-(\beta_m f_m - g_m)), \tag{8}$$

where f_m represents the corresponding free energy.

Indeed, equation (8) is proportional to

$$Z(\beta_m) \exp(g_m). \tag{9}$$

Since the aim of the tempering method is to freely sample configurations resulting from various temperatures, a careful estimation of g_m will prove to be fruitful. In fact, all of the *p(m)'s* will equal $\frac{1}{M}$ if $g_m = -\ln(Z_m)$, an equality obtained by trivial manipulation of equation (9). This process incorporates temperature into the stochastic procedure of phase space evolution.

1. Thermalize system at β_1.
2. Let β_m be the current temperature. Change to neighboring β's in the following way.
   - Generate random number, ρ, $0 < \rho < 1$.
   - If $\rho < \frac{1}{2}$ attempt to move to β_{m-1}; otherwise attempt to move to β_{m+1}.
   - *Accept* the new temperature with probability $\exp(-\Delta H_t)$.
3. Repeat step 2.

FIG. 2. *Temperature sampling by simulated tempering.*

In the algorithm described in Figure 2, β acceptance simply means that the system is ready to move to a new temperature. This is a subtle point that merits comment, as the tempering scheme requires that the last action of the second step of the algorithm be taken some of the time. At present, the aim is to develop an implementation such that acceptance be approximately equal for increasing and decreasing temperatures. In addition, the *acceptance rate*, or a percentage of accepted temperatures over time, for increasing and decreasing temperatures is believed to be optimal when it is around 50 %. (The reader may wish to consult [10] for the initial formulation of the simulated tempering method.)

5 Results and Discussion

As is so often that case, time has presented a hindrance. Indeed, a major ingredient to the annealing–type approaches to multivariate optimization, the decorrelation criterion, has yet to be investigated. For instance, it would be rather gratuitous to consider performance (e.g. timings and iterations) when the current annealing code uses a conservative estimate for the decorrelation (thermalization) time and the tempering implementation ignores the issue all together and has not yet been tested with the dynamic temperature schedule. In addition, since the primary effort revolved around the method and not the problem, details of the results concerning the discovered ground states of the spin glass models will not be considered in this text. Suffice it to say that both the simulated annealing and simulated tempering experiments consistently converged on minimum states with energies of around -362.0 and -1430.0 for the 16×16 and 32×32 size models, respectively. At any rate, this study presents several truly enlightening discoveries.

Problem size	Thermalization sweeps	Total sweeps	Metropolis sweeps	Initial β	Final β
16 x 16	500	700	10	0.30	1.0180
32 x 32	500	1000	50	0.20	1.5498

Table 1. *Runtime parameters for the annealing and tempering simulations* [2].

λ	Ensemble size
10.0	50.0

Table 2. *Runtime parameters for the adaptive annealing schedule described in section 3.2*

5.1 A fast optimal annealing schedule?

In order to converge upon a global minimum with probabilistic certainty a logarithmic cooling schedule must be observed [8]. In practice, however, researchers seldom adhere to this fundamental requirement because of the incurred computational cost and commonly use exponential schedules, e.g.,

$$T_{k+1} = \alpha T_k, \quad 0 < \alpha < 1, \tag{10}$$

justified only by expediency. For many applications the true global optimum is not actually necessary – oftentimes one *close* to the global optimum is sufficient. Despite this observation, since annealing is a relatively delicate process, one can imagine how this implementation aspect may result in inappropriate results.

This brings about the first major assertion of this work: the dynamic annealing schedule outlined in section 3.2 is, in the context of the two–dimensional Ising spin glass, exponential in nature, as Figure 3 depicts. This finding clearly merits rigorous investigation, as its possible ramifications are dramatic. Indeed, it is likely that one will be able to apply the *fast* simulated annealing algorithm, i.e. one utilizing an exponential temperature schedule, *and* be guaranteed an optimal result.

[2]The final β values fluctuated slightly due to variations in the temperature schedules.

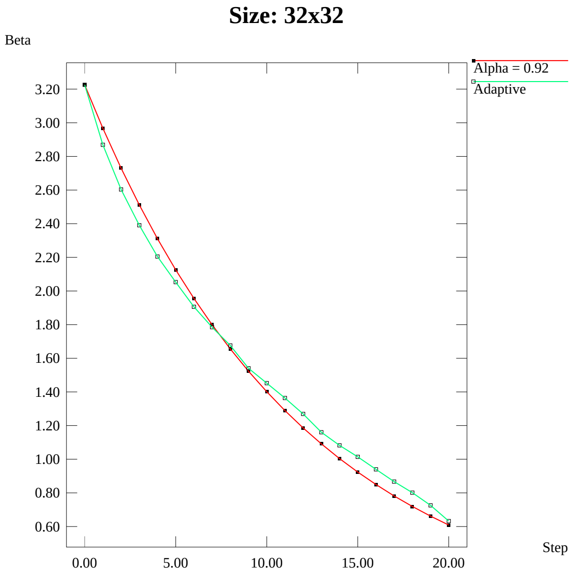

FIG. 3. *Comparison of dynamic and exponential annealing schedules*

It is important to note, however, that by no means is the dynamic cooling schedule superfluous. There are two primary reasons for this fact. First, annealing schedules vary with problem size, and second, the temperatures selected by the constant thermodynamic speed annealing schedule may not resemble an exponential curve for other hard optimization problems. For the present intents and purposes and in light of the fact that generating temperatures dynamically is a CPU intensive process, an efficient implementation may utilize the dynamic scheme only for the initial pass. The α parameter in equation (10) for generating the exponential schedule can then be extrapolated and used for generating an exponential cooling scale in subsequent runs.

5.2 Implementation considerations for simulated tempering

The primary question for the tempering scheme is how to achieve the kind of temperature acceptance described in section 4.1. Since it is the change in the Hamiltonian, ΔH, that governs temperature acceptance, a reasonably accurate method is needed to "guess" where a suitable next temperature will lie. This is truly a sensitive matter, as if the guess is erroneous enough, temperature acceptance will be greatly affected. Consider an example where a move is attempted from β_m to β_{m+1}. The change in the H_t is calculated in the following way.

$$\Delta H_t = E\Delta\beta_m - \Delta g_m \tag{11}$$

where E is the instantaneous energy value, *H(X)*, $\Delta\beta_m = \beta_{m+1} - \beta_m$ and $\Delta g_m = g_{m+1} - g_m$. The reader may recall from section 4.1 that new β's are accepted with probability $e^{-\Delta H_t}$. Due to the fact the the quantity ΔH_t in (11) is directly dependent on the value of Δg_m, careful approximations are a must.

Theoretically, the criterion for new β acceptance is quite simple. In fact, all that is required is that Δg_m be an intermediate energy value [10], say $E(g_m)$, such that

$$E(\beta_m) < E(g_m) < E(\beta_{m+1}). \tag{12}$$

The condition described in equation (12) is realized in a straightforward fashion: $E(g_m)$ is determined from the average of the energies at β_m and β_{m+1} (or β_m and β_{m-1}, as the case may be). From this formulation one is able to increase the accuracy of the estimation by subdividing the interval between β's [17]. A general equation is given by

$$\Delta g_m = E(g_m) = \frac{\Delta\beta_m}{n}\left\{\sum_{i=1}^{n}(E_{i,m}) - \frac{1}{2}(E_{1,m} + E_{n,m})\right\}, \tag{13}$$

where n is the number of subdivisions.

The transition probabilities for visiting five β values are illustrated in Figure 4. The reader may recall from the formulation of the simulated tempering algorithm that two aspects are desirable: similar acceptance values for increasing and decreasing temperatures, and an overall acceptance rate of 50 %. In that light, considerable improvement is seen in the $n = 5$ case.

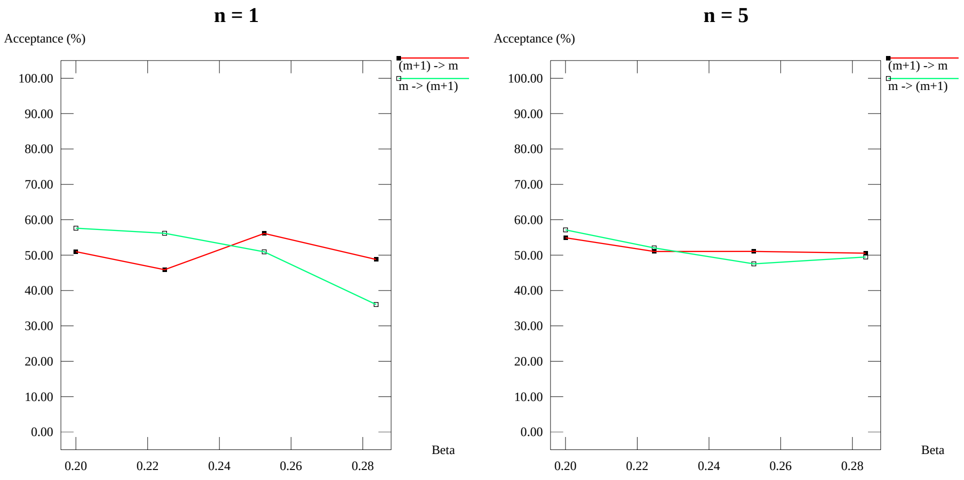

FIG. 4. *Transition probabilities for a 32 × 32 Ising spin glass model.*

The current tempering implementation further improves the Δg_m estimate. Since temperatures are revisited, previous Δg's are averaged in with the current one. This procedure presents an additional subtle advantage. Since the Metropolis method for energy calculation is probabilistic, the possibility of bogus results is real and quite inevitable. The new method of weight assignment to the Δg parameter reduces the ill–effects of the procedure over the course of the simulation. It is important to add that these results were obtained from a new implementation of the tempering algorithm. In addition to systematically updating the Δg parameter, this experimental version generated the temperature schedule (and thus the Δg's) internally. Prior implementations required a temperature list and a corresponding Δg list produced by simulated annealing runs.

As Figure 5 points out, the actual simulation of tempering only occurs on pre–existing temperatures. In fact, new temperatures are only added when the algorithm attempts to move to a *lower* temperature not previously visited. Furthermore, the current method assumes that $\beta_0 = 0$ (i.e. $t_0 = \infty$) and thus requires a downward move in the beginning of the simulation. This also implies, of course, that a move *below* β_0 would never take place.

1. Thermalize the system at β_0.
2. Generate a new β.
3. Perform the tempering step as in 2 of Figure 2.
    - If attempt to move to β_{m+1} not previously visited, generate a new β.
4. Repeat step 3.

FIG. 5. *The new simulated tempering implementation.*

5.3 Simulated tempering and the temperature schedule

To this point, the discussion considered theoretical and implementation details. It is consequential to recall the underlying motivation, however: the development of an *efficient* and *effective* global optimization method. How can simulated tempering contribute to this end?

Consider the graphs in Figure 6. The trend depicted by the transition probabilities at different α values is self–evident. Indeed, this heuristic provides feasible automation. Currently, the constants in the formulation of both the dynamic and the exponential temperature schedules are not only named *a priori*, but are initially determined quite crudely (i.e. guessed). The β acceptance rates obtained through the transition probabilities from tempering simulations can be easily utilized to adjust the scheduling parameters at runtime. In the case of an exponential temperature schedule, for example, this would entail the manipulation of the α variable until the desired performance is achieved.

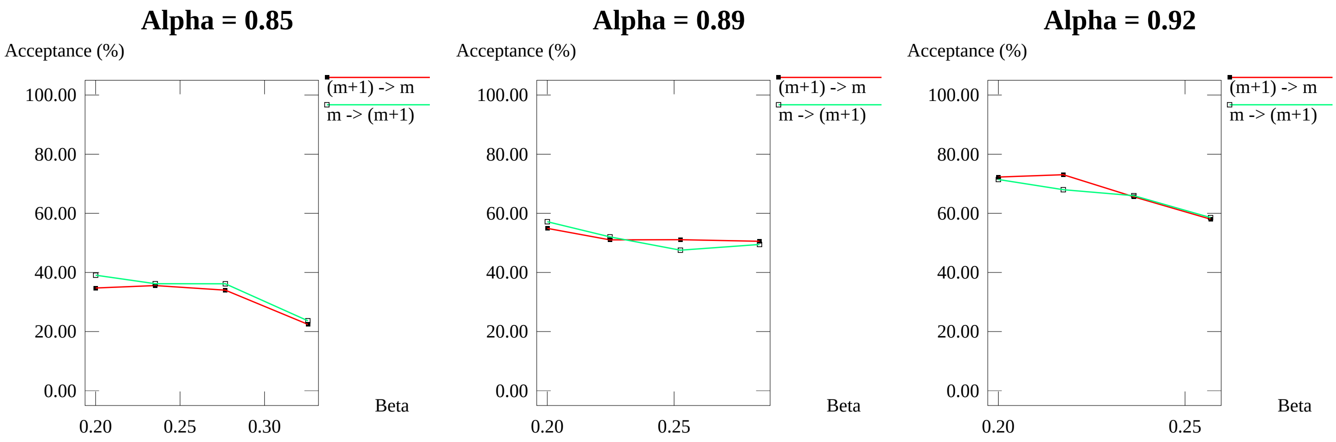

FIG. 6. *Acceptance comparison for different α values for a 32 $\times$ 32 model. n, the subdivision value, is 5. The $\alpha = 0.89$ case is an almost ideal realization of the requirements described in section 4.1.*

5.4 Overlap

In sampling ground states of systems with perhaps millions of local minima, how can one be certain that the algorithm is not visiting the *same* (or only a few) local minima throughout the simulation? Indeed, this is a legitimate and interesting question for all applications of combinatorial optimization. Fortunately, there is an elegant and straightforward solution.

The aim here is to derive a quantitative way to observe how a system evolves. Intuitively, this is achieved by superimposing two configurations and noting their similarities and differences. Consider two local state configurations, ω^α and ω^β, occurring at times t and $t+1$, respectively. Let V be the total volume of the system (with the Ising spin glass model, the volume is simply the number of lattice sites). The *overlap* [2], ω, is then given by

$$Q_{\alpha\beta} = \frac{1}{V}\sum_{i=1}^{V}\omega_i^\alpha\omega_i^\beta, \tag{14}$$

It is clear that if ω^α and ω^β are exactly the same, $Q_{\alpha\beta}$ would equal 1. Conversely, if the configurations were exactly opposite, -1 would be the result.

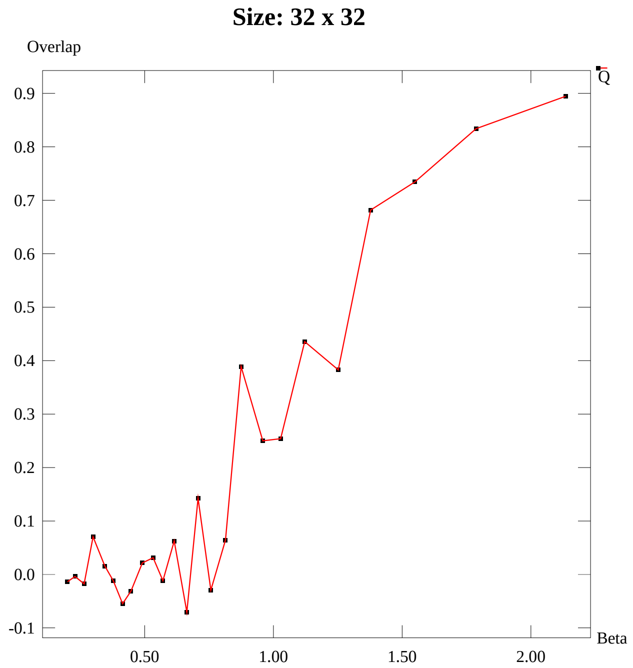

FIG. 7. *A plot of overlap as a function of β. As a point of interest, this data was obtained from a simulation of annealing. In a simulation of tempering one would expect a more undulating graph.*

A module for calculating this quantity has been implemented and tested in the context of simulated annealing and has truly provided an abundance of information. At high temperatures, for example, one might expect Q to fluctuate about 0. As the system freezes, however, the overlap calculation would approach 1. Figure 7 illustrates this phenomenon. (To see why this is so, recall from section 3.1 that the probability of incurring changes to the configuration space at low temperatures is small.) In addition, the notion of overlap can serve as a robust dynamic criterion for the termination of the algorithm. Given a proper cooling schedule and sufficient decorrelation times, one can conclude with confidence that the system is in a ground state if the overlap calculation from contiguous configurations approaches 1 to within some small value, ϵ. (ϵ would, of course, have to be determined experimentally.)

6 Conclusions

This work has brought to light several important contributions to the study of multivariate optimization. First, it has shown experimentally that the exponential annealing schedule is optimal when applied to the two–dimensional Ising spin glass model. Next, results from an improved implementation of the simulated tempering algorithm were presented. Satisfactory temperature transition probabilities clearly indicate that the algorithm is functioning properly. A captious point that has not yet been overtly stated is the following. Previous attempts of simulating tempering on larger Ising spin glass models have been problematic (mainly due to poor estimation of the Δg parameter) [17]. The current implementation, however, has performed equally well on the 32×32 size lattice as on the 16×16 and has shown trends to that suggest comparable performance for even larger problems. (Larger system sizes were not investigated because the current tempering implementation is lacking the autocorrelation function.) Finally, the usefulness of simulated tempering was illustrated with regard to the determination of an optimal cooling schedule.

Although these findings further confirm the conjecture that simulated tempering is a truly viable method for combinatorial optimization, there is, as was previously mentioned, crucial aspects that have still to be investigated. The realm of the Ising spin glass model, for instance, calls for additional research: the simulation of even larger models and the determination of an *actual* ground state are two avenues which are already under consideration. For the latter, parallelization and the decorrelation gauge are obvious requisites. Lastly, before a statement about the generality of the said scheme can be made, the method needs to be applied to other hard optimization problems, such as the Travelling Salesman Problem.

Acknowledgements

The primary author wishes to express his gratitude to two especially influential people: Professor Edward Bogucz and T. J. Willis. Dr. Bogucz is the embodiment of the most noble ideals of academia (a realm plagued with bureaucracy and politics). T.J., in addition to many intriguing discussions with regard to this work, has been my late–night compadre and, perhaps more importantly, has shown me the that there is only one way to drink coffee – *black*.

References

[1] Andresen, B., Hoffmann, K. H., Mosegaard, K., Hulton, J., Pedersen, J., M. and Salamon, P., *On lumped models for thermodynamic properties of simulated annealing problems*, J. Phys., vol. 49, 1485–1492, 1988.

[2] Bachas, C. P., *Spin–glass model of crystal surfaces*, Phys. Rev. Lett., vol. 54, no. 1, 53-55, 1985.

[3] Cerný, V., *A thermodynamical approach to the travelling salesman problem: An efficient simulation algorithm*, Report, Comenius University, Bratislava, Czechoslovakia.

[4] Cipra, B. A., *An introduction ot the Ising Model*, Math. Oper. Res., 937-959, 1987.

[5] Geman, S. and Geman, D., *Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration in images*, IEEE Trans. Patt. Anal. Mac. Int., vol. 6, 721–741, 1984.

[6] Hajek, B. *Cooling schedules for optimal annealing*, Mathematics of Operations Research, vol. 13, no. 2, 311–329, 1988.

[7] Hoffmann, K. H. and Salamon, P., *The optimal simulated annealing schedule for a simple model*, J. Phys. A: Math. Gen., vol. 23, 3511–3523, 1990.

[8] Ingber, L. *Simulated annealing: Practice versus theory*, Statistics and Computing, 1993, to appear.

[9] Kirkpatrick, S., Gellat, C. D., Jr. and Vecchi, M. P., *Optimization by simulated annealing*, Science, vol. 220, 671–680, 1983.

[10] Marinari, E. and Parisi, G. *Simulated tempering: A new Monte Carlo scheme*, Europhys. Lett., vol. 19, 451, 1992.

[11] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, E. *Equation of state calculation be fast computing machines*, J. Chem. Phys., vol. 21, 1087–1092, 1953.

[12] Mirkin, G., Vasudevan, K. and Cook, F. A., *A comparison of several cooling schedules for simulated annealing implemented on a residual statistics problem*, Geophys. Res. Lett., vol. 20, no. 1, 77–80, 1993.

[13] Nulton, J. D. and Salamon, P., *Statistical mechanics of combinatorial optimization*, Phys. Rev. A, vol. 37, no. 4, 1351–1356, 1988.

[14] Otten, R. H. J. M., van Ginneken, L. P. P. P., *The Annealing Algorithm*, Kluwer Academic Publishers, Dordrecht, 1989.

[15] Rees, S. and Ball, R. C., *Criteria for an optimum simulated annealing schedule for problems of the travelling salesman type*, J. Phys. A: Math. Gen., vol. 20, 1239–1249, 1987.

[16] Salamon, P., Nulton, J. D., Harland, J. R., Pedersen, J., Ruppeiner, G. and Liao, L., *Simulated annealing with constant thermodynamic speed*, Comp. Phys. Comm., vol. 49, 423–428, 1988.

[17] Wu, C. H., *On a study of simulated tempering as an optimization method*, Technical Report, NPAC, Syracuse University, 1993.