



Verification and rectification of the physical analogy of simulated annealing for the solution of the traveling salesman problem

著者	Hasegawa M.
journal or publication title	Physical review E
volume	83
number	3
page range	036708
year	2011-03
権利	(C)2011 American Physical Society
URL	http://hdl.handle.net/2241/113241

doi: 10.1103/PhysRevE.83.036708

Verification and rectification of the physical analogy of simulated annealing for the solution of the traveling salesman problem

M. Hasegawa

Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba 305-8573, Japan

(Received 25 December 2010; revised manuscript received 2 February 2011; published 31 March 2011)

The aim of the present study is to elucidate how simulated annealing (SA) works in its finite-time implementation by starting from the verification of its conventional optimization scenario based on equilibrium statistical mechanics. Two and one supplementary experiments, the design of which is inspired by concepts and methods developed for studies on liquid and glass, are performed on two types of random traveling salesman problems. In the first experiment, a newly parameterized temperature schedule is introduced to simulate a quasistatic process along the scenario and a parametric study is conducted to investigate the optimization characteristics of this adaptive cooling. In the second experiment, the search trajectory of the Metropolis algorithm (constant-temperature SA) is analyzed in the landscape paradigm in the hope of drawing a precise physical analogy by comparison with the corresponding dynamics of glass-forming molecular systems. These two experiments indicate that the effectiveness of finite-time SA comes not from equilibrium sampling at low temperature but from downward interbasin dynamics occurring before equilibrium. These dynamics work most effectively at an intermediate temperature varying with the total search time and thus this effective temperature is identified using the Deborah number. To test directly the role of these relaxation dynamics in the process of cooling, a supplementary experiment is performed using another parameterized temperature schedule with a piecewise variable cooling rate and the effect of this biased cooling is examined systematically. The results show that the optimization performance is not only dependent on but also sensitive to cooling in the vicinity of the above effective temperature and that this feature is interpreted as a consequence of the presence or absence of the workable interbasin dynamics. It is confirmed for the present instances that the effectiveness of finite-time SA derives from the glassy relaxation dynamics occurring in the “landscape-influenced” temperature regime and that its naive optimization scenario should be rectified by considering the analogy with vitrification phenomena. A comprehensive guideline for the design of finite-time SA and SA-related algorithms is discussed on the basis of this rectified analogy.

DOI: [10.1103/PhysRevE.83.036708](https://doi.org/10.1103/PhysRevE.83.036708)

PACS number(s): 02.60.Pn, 05.10.-a, 89.20.-a, 64.70.P-

I. INTRODUCTION

In recent decades, various research topics in computer science have been studied using concepts and methods developed in the physics community. A profound characterization of computational complexity using the concept of phase transition [1–4] is an intriguing approach in this line of work. Computational complexity theory [5] deals with the inherent difficulty of computational tasks and is applicable to their feasibility assessment. The concept of phase transition plays a meaningful role in describing the nature of computationally difficult-to-solve problems: The hardest problem instances cluster in the transition range and the complexity class is associated with the characteristics of this transition. Optimization is also an attractive subject in this interdisciplinary research field. Some physically relevant properties of combinatorial optimization problems have been examined by focusing on their mathematical form shared with frustrated disordered systems [6–8]. In addition, one often employs heuristic algorithms to solve hard optimization problems. Also on this point, several algorithms have been developed in the physics community and applied to the solution of problems arising in physics and other areas of science and engineering [9,10].

In the present paper, we look at one of those heuristic optimization methods, simulated annealing (SA) [11,12], which was proposed in the early days of this growing interdisciplinary field. An optimization problem is usually formulated as a task of finding a solution minimizing a given cost function, and the SA algorithm is designed to find such an optimal (or

suboptimal) solution by emulating the sampling of low energy states of a physical system using the Monte Carlo method. There is now an abundant literature on algorithmic properties of SA, including its optimization performance [10,13,14]. Furthermore, several essential improvements in sampling quality have been achieved to overcome the slow-convergence problem [15]. For these reasons, one may not be able to appreciate further efforts to comprehend the optimization characteristics of this historic approach. We see, however, many attempts to eliminate barriers among various complex systems including hard optimization problems [16–19]. From the present interdisciplinary point of view, it would appear that a correct understanding of the function of SA has historically been hampered by a misguided analogy to physical phenomena. Because of this situation, we consider the optimization mechanism of finite-time SA by restarting our investigation from the verification of its conventional naive optimization scenario. In the remainder of this introduction, the outline of the study is presented after a brief review of the relevant background.

Originally, the approach of SA for optimization is based on the idea of equilibrium statistical mechanics [7,11,12]: By replacing the feasible solution and its cost with the microscopic state and its energy, respectively, the Metropolis algorithm (MA) [20] for equilibrium sampling is employed with decreasing temperature. In this analogy, a globally optimal solution of the optimization problem corresponds to a ground state of the physical system. It follows that one can expect to find an optimal or suboptimal solution with high

probability at low temperatures. Under such cold conditions, however, a search process, that is, a sequential sampling process, tends to be stuck in some subset of the solution space. This is because a neighborhood function, which determines the connectivity of the solution space, generally creates a complex multimodal landscape of the cost function and the process is apt to be confined in one or a few valley structures at low temperatures. As a result, convergence to the equilibrium cost distribution is very slow and the algorithm fails to work in its finite-time implementation.

To extract the most out of this optimizer, one has elaborated the design of the temperature schedule. It has been generally recommended that the system should be cooled from high temperature as slow as possible to keep it in its equilibrium state at each temperature. Furthermore, a reduction of the cooling rate near the temperature analogous to the freezing point has been considered to be an effective strategy [7,11]. This idea is based on the analogy between the optimal solution and the ideal crystal, and therefore is implemented to avoid yielding unfavorable results like amorphous states. In practice, the freezing temperature of an optimization problem is manifested as a specific-heat peak; the specific heat is evaluated from the cost fluctuation at each temperature [11].

Although this recipe for optimization has been generally mentioned in literature, one should recall that this is merely an envisioned scenario. In fact, physical annealing as a heat treatment of metals [21,22] and glasses [23] is not generally the solidification process of liquid. This treatment is normally intended for property modification of materials in the solid state and is performed often in the following sequence: First the material is heated up to a prescribed temperature (below the freezing temperature); then it is kept at this temperature for a certain period of time and/or carefully cooled over a narrow temperature range; and last it is cooled down to the initial temperature. Although one may expect energetic stabilization also in this sequence, such an expected effect has not been explicitly considered in the standardized framework of the algorithm. Meanwhile, a new experiment has shown that in the solution of some typical optimization problems, a constant-temperature search after rapid quenching outperforms a search with an elaborated cooling schedule in the same amount of time [24]. Even before that, it was shown that there exists an effective temperature for finite-time optimization and that in some instances, this temperature is different from the specific-heat-peak (or the cost-fluctuation-peak) temperature and is found on its low temperature side [25,26]. Strangely enough, this crude but workable search strategy taking advantage of such an effective temperature is more analogous to the actual heat treatment by physical annealing than is the standardized SA approach with slow cooling.

Given the historical success of SA on both the theoretical and the practical sides, one might say that this review makes the situation unnecessarily complicated by taking the word out of context. However, precise knowledge of similarities and differences between complex physical systems and hard optimization problems should be necessary if one attempts to create a mutually beneficial relationship between the physics and optimization communities [18] or to discuss their research topics on an equal footing [19,27,28]. In the analogy between the two fields, complex multimodal cost functions of hard

optimization problems correspond to rugged energy surface landscapes of complex physical systems such as glass. The characteristics of the geometry of and the system's behavior in such landscapes have been mainly investigated for physical systems. Hence, complementary studies for optimization problems and search trajectories are expected to be useful in fertilizing the integrated research field. In addition, a good understanding of the optimization mechanism of SA should be essential for fair competition with other algorithms. After the proposal of SA, several improvements have been achieved through the use of other concepts and methods developed in the physics community. These include the adoption of constant thermodynamic speed [29], generalization inspired by Tsallis statistics [30,31], application of extended ensemble methods [32,33], and utilization of quantum fluctuation [34]. The optimization qualities of these algorithms, and also of some others (e.g., Refs. [35,36]), were often compared with those of SA to demonstrate their relative superiority. Identifying the origin of the effectiveness of SA, the other methods as well, should be needed for the appropriate design of such competition experiments and for a better understanding of each algorithm.

Taking this situation into account, we reconsider the optimization characteristics of finite-time SA on one of the hard optimization problems, the traveling salesman problem (TSP) [37,38]. Our goal is to obtain a reasonable interpretation of its optimization mechanism by precise analogy with physical phenomena. To this end, the following two and one supplementary experiments were designed with the help of concepts and methods developed for studies on complex physical systems. The first experiment was designed to investigate the optimization characteristics of SA in line with the conventional naive optimization scenario. For this purpose, a new quasiequilibrium condition was introduced by applying the concept of effective ergodic convergence [39,40] and used as an adaptive stopping criterion employed at each temperature. This framework of the temperature schedule enabled us to simulate successive quasiequilibrations in a stepwise cooling and to test the validity of the scenario. The second experiment was designed to analyze the search trajectory of the MA through the observation of its transition process among the basins of attraction in the solution space. The analysis was performed in a manner similar to the mapping-onto-minima approach [41–43] in the hope of drawing an explicit analogy to the physical process by reference to the corresponding interbasin dynamics of molecular model systems. These two experiments showed that there exists an effective temperature for finite-time optimization by the Monte Carlo dynamics and suggested that its physical counterpart is a temperature relevant to the glass transition. To examine the search function of the glassy behavior in the approach of SA directly, a supplementary experiment was designed and performed using temperature schedules with a piecewise variable cooling rate. The results showed that the optimization performance is not only dependent on but also sensitive to the cooling rate in a restricted temperature range including the above effective temperature and that in this range the glassy relaxation dynamics naturally play a primary role as an optimizer.

This work is a significant extension of some preliminary studies [44–46]; the present paper describes the systematized

new results and also discusses comprehensively the optimization characteristics of finite-time SA and SA-related algorithms. All the experiments were designed for a proper understanding of the optimization mechanism without a direct attempt to propose new improvements of the algorithm; for this reason, the experiments were sometimes performed under conditions unsuited for practical use. For the sake of completeness, the experimental details are redescribed in the relevant sections below. The remainder of this paper is organized as follows: In Sec. II, we define the TSPs that will serve as test beds and introduce the standard framework of the SA algorithm. In Sec. III, the methods of the first two experiments are described in detail. After presenting their results, we discuss the origin of the effectiveness of finite-time SA, including the physical implication of the effective temperature. In Sec. IV, the method of the supplementary experiment is described in detail and after presenting the results, we summarize the optimization characteristics of finite-time SA and discuss some SA-related topics. In Sec. V, we conclude our study.

II. PRELIMINARIES

A. Random TSPs

The TSP [37,38] is a combinatorial optimization problem to find the shortest tour that passes through each of the given cities once and returns to the starting city. We consider the following two types of random problems: (i) random Euclidean TSP (RE-TSP), where the locations of N cities are sampled uniformly in the unit square region and the intercity distance is computed under the Euclidean metric; (ii) random distance matrix TSP (RD-TSP), where the intercity distances between N cities are sampled uniformly in the unit interval. Even in the latter type, we consider only the symmetric case, namely, that the intercity distance does not depend on the order of two cities.

B. Standard SA

Johnson *et al.* have performed detailed experimental analyses of SA for these random TSPs [38] and other representative problems [47,48]. They introduced a parameterized framework of the cooling schedule without regard to the physical analogy and investigated systematically the optimization performance of their basic algorithm. We adopt this standardized framework as a reference; however, we proceed with our study always keeping physical counterparts in mind.

First, we introduce some notation and terminology used throughout this paper. Let x be a tour (a feasible solution), $f(x)$ be the cost function defined by the tour length, and $\mathcal{N}(x)$ be the neighborhood function. Throughout the study, we use a 2-opt neighborhood [38], which is defined as a set of tours constructed by any change of two intercity paths from the tour x . We use the notation x_n ($n = 0, 1, 2, \dots$) to represent the (actual) search history and x^* to represent the incumbent solution (the best-so-far solution). A locally optimal solution is a solution that does not have any improved solution in its neighborhood. A basin is defined as a set of solutions attracted to the same locally optimal solution by a simple local search, that is, a repetition of the move to the best solution in the neighborhood. Let $y(x)$ be the locally optimal solution inside the basin including the solution x and we refer to its cost

$f(y(x))$ as the cost of the basin. [Obviously, $y(x)$ can be found by a simple local search starting from the solution x .] We denote the temperature by T ; however, we often use the logarithmic temperature $\Theta (= \log_{10} T)$.

The standard framework of the SA algorithm is described as follows.

[SA1] Select an initial temperature T_s and a cooling ratio γ .

[SA2] Set $n := 0$ and generate an initial solution x_0 randomly. Set $T := T_s$ and $x^* := x_0$.

[SA3] Implement a single step of the MA and increment n by one. (At this point, the current solution is x_n and the current incumbent solution is x^* .)

[SA4] If the stopping criterion for the MA search (the constant-temperature search) is satisfied, go to [SA5]. Otherwise, return to [SA3].

[SA5] If the termination condition of the algorithm is satisfied, output x^* and stop. Otherwise, set $T := \gamma T$, $x_0 := x_n$, and $n := 0$; and return to [SA3].

In a single step of the MA in [SA3], the following [MA1] is implemented.

[MA1] Choose a trial solution $x'_n \in \mathcal{N}(x_n)$ randomly and set $\Delta := f(x'_n) - f(x_n)$. If $\Delta < 0$, accept the trial solution and set $x_{n+1} := x'_n$; furthermore, if $f(x_{n+1}) < f(x^*)$, renew the incumbent solution, $x^* := x_{n+1}$. If $\Delta \geq 0$, accept the trial solution with the probability $\exp(-\Delta/T)$ and set $x_{n+1} := x'_n$; with the complementary probability, reject the trial solution and set $x_{n+1} := x_n$.

We adopt the following stopping criterion in [SA4] and termination condition in [SA5] after Johnson and McGeoch [38]. The stopping criterion in [SA4] is that the counter n reaches a predetermined constant value L ; periodic reduction of temperature by multiplying by a constant factor is called geometric cooling. The termination condition in [SA5] is that five consecutive reductions of temperature proceed without any renewal of the incumbent solution and without the acceptance ratio going above 2%. Here, the acceptance ratio is the ratio of the case that a trial solution is accepted as the next solution in the MA search. We refer to this algorithm simply as the standard SA.

Before closing this section, we need to complement the initial setting of the algorithm. In the beginning of our series of experiments, an initial acceptance probability p_s is selected instead of an initial temperature T_s . The reason is that the interesting temperature range is unknown in advance and therefore the use of temperature is unsuited for our systematic investigation. In practice, the conversion was made according to the functional relationship between acceptance probability and temperature; this relationship was determined in the preliminary runs of the standard SA by averaging the measured acceptance ratios at each temperature.

III. EXPERIMENTAL ANALYSIS I

A. Adaptive SA

In the standard SA algorithm, the number of steps allocated in the MA search is constant regardless of the temperature. If we want to simulate successive quasiequilibria in this stepwise cooling, we have to introduce some sort of relaxation or equilibration time into the stopping criterion for each

MA search. For this purpose, in the first experiment, we consider many independent search processes simultaneously and evaluate the degree of equilibration by measuring the difference among the cost distributions each of which is accumulated in each process.

To describe the method more fully, let I be the number of search processes and $x^{(i)}$ be the solution of the i th process. We introduce the measure Ω_n defined by the following equation:

$$\Omega_n = \frac{1}{I} \sum_{i=1}^I (F_n^{(i)} - \bar{F}_n)^2, \quad (1)$$

where

$$F_n^{(i)} = \frac{1}{n+1} \sum_{k=0}^n f(x_k^{(i)}) \quad (2a)$$

and

$$\bar{F}_n = \frac{1}{I} \sum_{i=1}^I F_n^{(i)}. \quad (2b)$$

That is, Ω_n is the variance of the average costs obtained in individual search processes and is expected to approach zero in equilibration ($n \rightarrow \infty$). Therefore, this measure is calculated in every time step and if the condition defined by the inequality $\Omega_n/\Omega_0 < \varepsilon$ is satisfied for the first time, all the MA searches are stopped and the temperature is reduced, where ε is a preset parameter ($0 < \varepsilon < 1$) that quantifies the intended degree of equilibration. In what follows, this condition is called the quasiequilibrium condition, and SA with this adaptive stopping criterion is referred to as the adaptive SA. Its algorithm is organized as follows.

[ASA1] Select an initial temperature T_s , a quasiequilibrium constant ε , a cooling ratio γ , and the number of search processes, I .

[ASA2] Set $n := 0$ and generate I initial solutions $x_0^{(i)}$ ($i = 1, 2, \dots, I$) randomly. (The note on the I independent runs with respect to the superscript i is omitted below without confusion.) Set $T := T_s$ and $x^{(i)*} := x_0^{(i)}$.

[ASA3] Calculate Ω_0 .

[ASA4] Implement a single step of the MA. (Namely, implement [MA1] for each of the $x_n^{(i)}$'s.) Increment n by one. (At this point, the current solutions are $x_n^{(i)}$'s and the current incumbent solutions are $x^{(i)*}$'s.)

[ASA5] If the termination condition of the algorithm is satisfied, output $x^{(i)*}$ and stop. Otherwise, calculate Ω_n .

[ASA6] If the quasiequilibrium condition is satisfied, set $T := \gamma T$, $x_0^{(i)} := x_n^{(i)}$, and $n := 0$; and return to [ASA3]. Otherwise, return to [ASA4].

As described in the previous section, the initial temperature T_s is determined by the initial acceptance probability p_s in [ASA1]. To specify the termination condition in [ASA5], we use the global counter to count the accumulated number of search steps in the individual cooling process. The algorithm is terminated if this global counter reaches a predetermined value of the total number of search steps, n_c .

The use of the measure Ω_n is inspired by the concept of effective ergodic convergence proposed by Mountain and Thirumalai [39,40]. They considered several glass-forming molecular model systems and probed the approach to ergodic behavior in each system by monitoring the difference, for

example, among the time-averaged energies of individual molecules.

A parametric study was conducted to investigate how our adaptive cooling influences the optimization performance evaluated by the cost of the basin including the incumbent solution. In this experiment, the framework of the cooling schedule is described by the five parameters, p_s (or T_s), ε , γ , I , and n_c , and their values were selected in the following way. All possible combinations of a value of p_s approximately from $\{0.9, 0.5, 0.1\}$, a value of ε from $\{0.1, 0.3, 0.5, 0.7\}$, and a value of γ from $\{0.95, 0.75, 0.55\}$ were selected for the first three parameters. The values of I and n_c were chosen by taking account of the system size N and the computation time. For some system sizes, the dependence of the performance on the number of processes was examined by varying I .

B. Hidden search dynamics of MA

In the second experiment, we analyze the search dynamics of the MA through the observation of successive interbasin transitions occurring in the landscape of the cost function. The experiment is performed in the following manner. First, a tentative target temperature is selected for operational convenience, and a standard SA is employed from a sufficiently high temperature until the temperature falls below the tentative value for the first time. Then the temperature is fixed at the final value and after resetting $n(=0)$, an MA search is performed for a long time. Throughout this constant-temperature search, we observe the change of the basin's cost along the search trajectory. In what follows, we regard the final temperature as an actual target temperature, which is denoted as T_c ($\Theta_c = \log_{10} T_c$), and refer to the history described by $f(y(x_n))$ ($n = 0, 1, 2, \dots$) as the hidden search dynamics.

This method of analysis is similar to the mapping-onto-minima approach introduced by Stillinger and Weber [41–43], and this approach has been successfully used for studies on liquid and glass. In the discussion of the present experimental results, we make reference to two computer simulations of glass formation in Lennard-Jones (LJ) systems [49,50]. This is because both provide data useful for bridging between optimization dynamics and physical processes. On the basis of similarities found in the successive interbasin dynamics, we discuss the physical implication of the search characteristics of the MA. This will be described in Sec. III D.

The experiment was conducted at various levels of temperature over the interesting temperature range found in the first experiment. The initial temperature T_s of the preceding SA was taken to be 10 ($\Theta_s = 1$), which is high enough for the acceptance probability to exceed 90% for any instance considered in the present study. The cooling ratio γ was fixed at 0.95. The hidden search dynamics occurring after a rapid quenching ($L = 1$) were observed at all target temperatures used in the experiment; the influence of the precooling rate on the following MA dynamics was also examined by varying L at several temperatures.

C. Results

The experiments were performed for $N = 32(\approx 10^{1.5})$, 100, 316($\approx 10^{2.5}$), 1000, and 3162($\approx 10^{3.5}$) on both types of

TSPs. Although the results are shown only for the case of RE-TSP with $N = 316$, the characteristic features described below were similarly observed in the other cases. The initial temperatures selected for this illustrative case in the first experiment were $T_s \approx 1.661, 0.306$, and 0.089 ($\Theta_s \approx 0.22, -0.51$, and -1.05) in the descending order of the three values of p_s . In practice, the same set of values was used for the other RE-TSP instances because these values were little affected by the system size. [For the record, the corresponding three temperature values for RD-TSP were $T_s \approx 2.378, 0.438$, and 0.135 ($\Theta_s \approx 0.38, -0.36$, and -0.87), irrespective of the system size.] It should be mentioned that for low values of p_s , the average cost level of the solutions likely to be found at the initial temperature deviates considerably from that of the initial random solutions. This causes an unavoidable rise of the measured acceptance ratio from the predetermined probability in the beginning of the run; this deviation was ignored in the present experiment.

We begin with the results of the first experiment. Under the present experimental conditions, as shown just below, the adaptive SA outperforms the standard one if we select an appropriate set of parameter values. What is more important is that although its cooling schedule is determined by the five parameters, its optimization performance evaluated by $f(y(x^*))$ appears to be directly influenced by the final temperature T_e ($\Theta_e = \log_{10} T_e$) that is secondarily determined at the end of the run. To illustrate these features, Fig. 1 shows the average performance of both the standard and the adaptive SAs against the (average) final logarithmic temperature for $p_s \approx 0.9$. The results displayed in Fig. 1(a) were obtained by the following procedure. First, a standard SA with $L = N^2$ was implemented $I(=2^3)$ times independently for each of the three values of γ . Then, by using the same values of T_s and I as in this standard SA, an adaptive SA was performed for each of the four values of ε for each γ . The total number of search steps in the latter experiment was equalized to that spent in the former one; more

specifically, the number of search steps spent in the standard SA was averaged over the I runs and truncated to an integer for each γ . From this sequence, a pair of results, average $f(y(x^*))$ and (average) Θ_e , was obtained for each of the three cases of the standard SA and the 12 cases of the adaptive SA; here, each average was taken over the I processes. In addition, the average behavior over several different instances (different city configurations for RE-TSP) was observed throughout the first experiment. Thus, the results averaged again over the same ten instances are plotted in the present figure. We find that, as mentioned above, the adaptive SA can surpass the standard one in the average performance for each γ . More important, its final temperatures attained under the most workable conditions appear to be close to each other regardless of the value of γ . Figure 1(b) shows the results with the same adaptive SA as in Fig. 1(a) except that the total number of search steps was extended and equalized to the longest case (i.e., $\gamma = 0.95$). Surprisingly, the results lie roughly on a single line regardless of the combination of the parameter values. Quasistatic cooling, which has been recommended in the conventional naive optimization scenario, is ideally realized by letting $\varepsilon \ll 1$ and $\gamma \approx 1$. The present results, however, indicate that the uncontrollable final temperature is a major factor in the extraction of optimizing ability in the finite-time implementation of the algorithm. Figure 2 depicts four examples of the cooling schedule constructed in the case $\varepsilon = 0.3$ and $\gamma = 0.75$ shown in Fig. 1(b). As is expected from the physical analogy, the length of the MA search increases with decreasing temperature; as a result, most of the search time is allocated near and at the final temperature in this adaptive cooling.

The results for all the combinations of the values of p_s , ε , and γ are shown in Fig. 3 for $I = 2^3$ and 2^5 ; the average performance is plotted against the final logarithmic temperature as in Fig. 1(b). For each value of n_e , a total of 72 results are shown using the same symbol. (Eight different instances are considered here.) As in Fig. 1(b), these 72 results lie roughly on a single curve and it changes its position in the lower left direction with increasing n_e . Thus, we find that the optimization performance is directly influenced by the total number of search steps and by the final temperature. The other parameters, for example the quasiequilibration constant

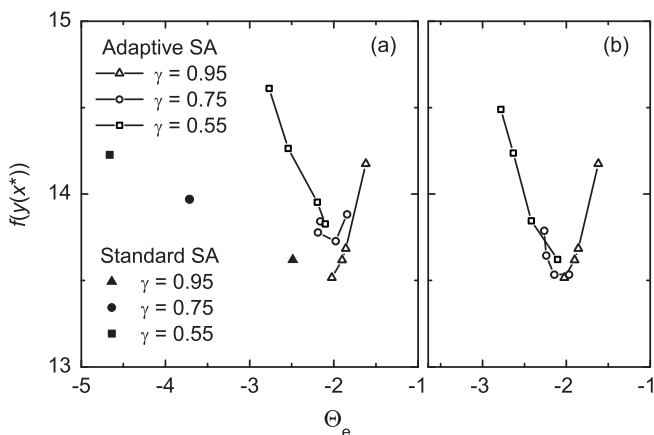


FIG. 1. Optimization performance (RE-TSP with $N = 316$; $p_s \approx 0.9$). The average value of $f(y(x^*))$ is plotted against the (average) final logarithmic temperature. See text for experimental details. (a) Results with the standard SA (solid symbols) and the adaptive SA (open symbols). The latter results for the four ε 's are connected by line segments for each γ (from right to left, $\varepsilon = 0.1, 0.3, 0.5$, and 0.7). (b) Results with the adaptive SA. The total number of search steps is equalized to that in the case $\gamma = 0.95$ in (a).

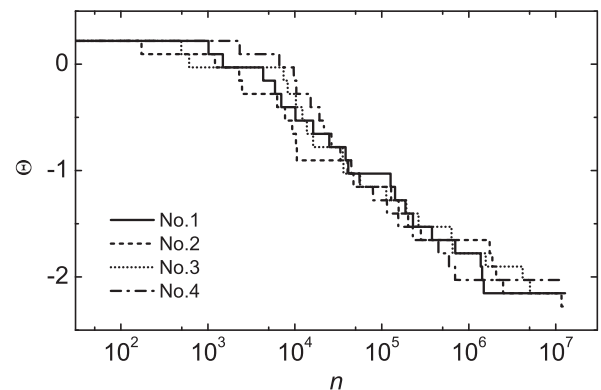


FIG. 2. Examples of the adaptively constructed cooling schedule (RE-TSP with $N = 316$; $p_s \approx 0.9$, $\varepsilon = 0.3$, $\gamma = 0.75$, and $I = 2^3$). Depicted are the results for the first four of ten instances.

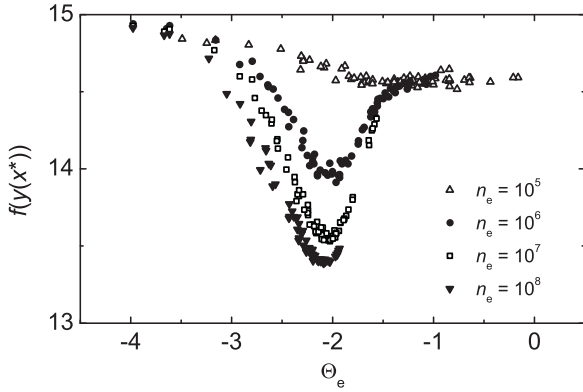


FIG. 3. Optimization performance $f(y(x^*))$ (RE-TSP with $N = 316$). For each n_e , a total of 72 results [36 results (see text for details) for $I = 2^3$ and those for $I = 2^5$] are plotted against the final logarithmic temperature using the same symbol. After being averaged over I search processes, the results are averaged again over the same eight instances.

and the cooling ratio, appear to influence the performance indirectly, only through the first two factors. Note that the optimum final temperature slightly decreases with the progress of the search.

During this adaptive cooling, the specific heat was measured at each temperature. The specific heat $C(T)$ is defined, as usual, as

$$C(T) = \frac{\langle f(x)^2 \rangle - \langle f(x) \rangle^2}{T^2}, \quad (3)$$

where $\langle \cdot \rangle$ denotes the time average, which was computed in each MA search, though the formula holds only in the equilibrium state. The results averaged over the $I(=2^5)$ processes are shown in Fig. 4 for the case $p_s \approx 0.9$, $\varepsilon = 0.1$, and $\gamma = 0.95$; plotted are the results for the first four of the eight instances. We easily notice that a single peak appears near $\Theta = -1.4$ regardless of the problem instance. Note that the above optimum final temperature ($\Theta_c \approx -2.0$) is clearly different from this maximizing temperature and is on its lower side. This feature is consistent with the known results [25,26]

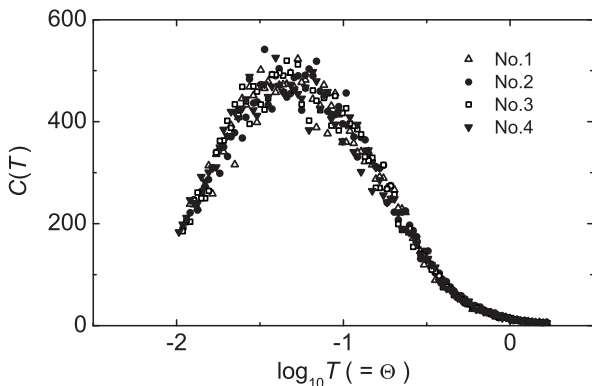


FIG. 4. Specific heat $C(T)$ (RE-TSP with $N = 316$; $p_s \approx 0.9$, $\varepsilon = 0.1$, $\gamma = 0.95$, $I = 2^5$, and $n_e = 10^8$). The results averaged over I search processes are plotted for the first four of eight instances.

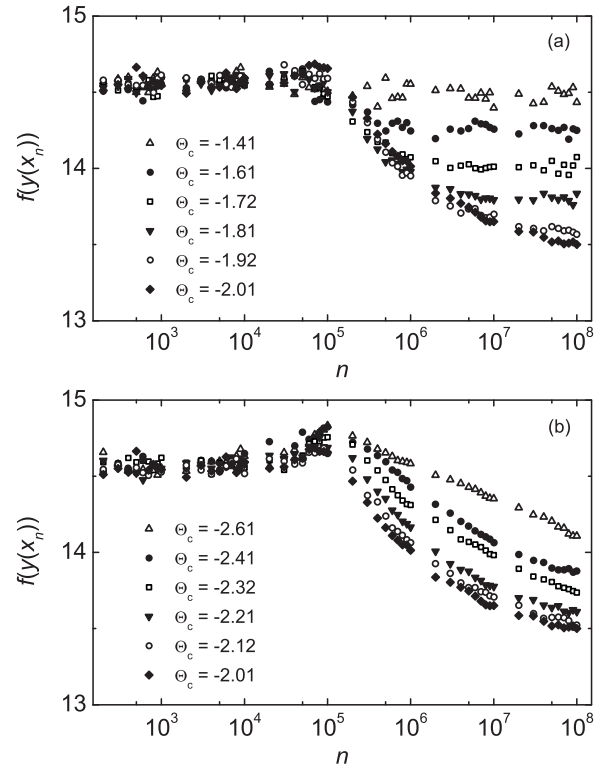


FIG. 5. Hidden search dynamics $f(y(x_n))$ (RE-TSP with $N = 316$; $I = 2^5$). Plotted are the results averaged over I search processes. (a) $\Theta_c \geq -2.01$; (b) $\Theta_c \leq -2.01$. The preceding standard SA is implemented under the conditions, $\Theta_s = 1$, $\gamma = 0.95$, and $L = 1$.

mentioned earlier. For reference convenience, we denote the specific-heat-peak temperature by \hat{T} ($\hat{\Theta} = \log_{10} \hat{T}$).

Next we turn to the results of the second experiment. The hidden search dynamics $f(y(x_n))$ after a rapid quenching ($L = 1$) are shown in Fig. 5. The results were averaged over $I(=2^5)$ independent runs on a single instance identical to the first instance in the first experiment. The average dynamics can be classified into three categories according to the target temperature; the characteristic features are summarized as follows. In the high temperature regime [$\Theta_c \gtrsim \hat{\Theta}$; the result for $\Theta_c = -1.41$ in Fig. 5(a) is one illustration], the system explores many basins whose costs distribute around the average value; the average cost of the basin is almost equal regardless of the temperature. In contrast, when the target temperature falls below $\hat{\Theta}$, transitions to lower basins occur after the initial $\sim 10^5$ search steps. At relatively high temperatures, these downward dynamics continue for a while and then change to roughly horizontal transitions [the intermediate temperature regime, Fig. 5(a)]. As the temperature decreases, the lifetime of these downward dynamics becomes longer and the system settles into more low-lying basins. Once the target temperature falls below some lower temperature, $\Theta_c \approx -2.0$ for the present instance, the unidirectional downward dynamics appear to last beyond the observation time [the low temperature regime, Fig. 5(b)]. It was confirmed for the present instance that in the successful case of the adaptive SA, most of the search time was spent near this characteristic temperature; more specifically, for each of the best five combinations of the parameter

values for $n_e = 10^8$ in the first experiment, the last more than 80%–90% of the total number of search steps were spent in a logarithmic temperature range included between -1.89 and -2.10 . Note that the downward interbasin dynamics, when they occur, become sluggish with decreasing temperature. As a result, if the performance is evaluated by the final basin's cost $f(y(x_{n_e}))$, the optimum temperature decreases with the progress of the search. [As can be seen from Fig. 5(a), the best target temperatures are $\Theta_c = -1.81$ for $n_e = 5 \times 10^5$, $\Theta_c = -1.92$ for $n_e = 5 \times 10^6$, and $\Theta_c = -2.01$ for $n_e = 5 \times 10^7$.] In other words, the optimum temperature is determined by the trade-off between the quickness of the descending motion and the fitness of the accessible basin.

Of course, these observations depend on the precooling rate; however, as shown just below, it does not affect the time required for the system to settle down to the horizontal transition state. In Fig. 6, the results for four different precooling rates are shown for three target temperatures. For $L \leq 10^3$, no qualitative change from the result for $L = 1$ (the most rapid quenching here) is observed for all the three temperatures. This is probably because the time for the cooling is shorter than the initial transient period. In contrast, for $L \geq 10^4$, the downward dynamics, when they occur ($\Theta_c < \hat{\Theta}$), begin during the cooling step and the following MA dynamics start from a low-lying basin. As can be seen from Figs. 6(b) and 6(c), however, this preprocessing does not affect the time needed for the settlement.

D. Discussion

These results show the existence of an effective temperature that depends on the total number of search steps. Contrary to the traditional recommendation, this temperature clearly deviates from the specific-heat-peak temperature. To discuss the physical implication of the present results, first we review two relevant molecular-dynamics studies on different glass-forming LJ systems. In both studies, the system's behavior was characterized by the time variation of the inherent-structure energy (the minimum energy), which dynamics are equivalent to the hidden search dynamics observed in our second experiment.

One of these studies was made by Sastry *et al.* on the equilibrium dynamics of a binary LJ mixture [49]. This model was originally proposed for a two-component alloy and yet it has been widely used as a model glass former. In this study, a stepwise cooling was done from an equilibrated high temperature state and equilibration at each of the successive temperatures was carried out by periodic rescaling of the molecular velocities in a certain period of time. The results showed that the interbasin dynamics can be classified into three categories according to the temperature and that this classification is strongly affected by the underlying energy landscape. The relevant characteristics are summarized as follows: In the high temperature regime ($T > T_1$), the system explores a wide range of minimum energies and the average value does not vary significantly with the temperature. In contrast, in the intermediate temperature regime ($T_1 > T > T_2$), the sampled minimum energy becomes progressively biased toward lower energies with decreasing temperature. By correlating these results with the motion of molecules, it

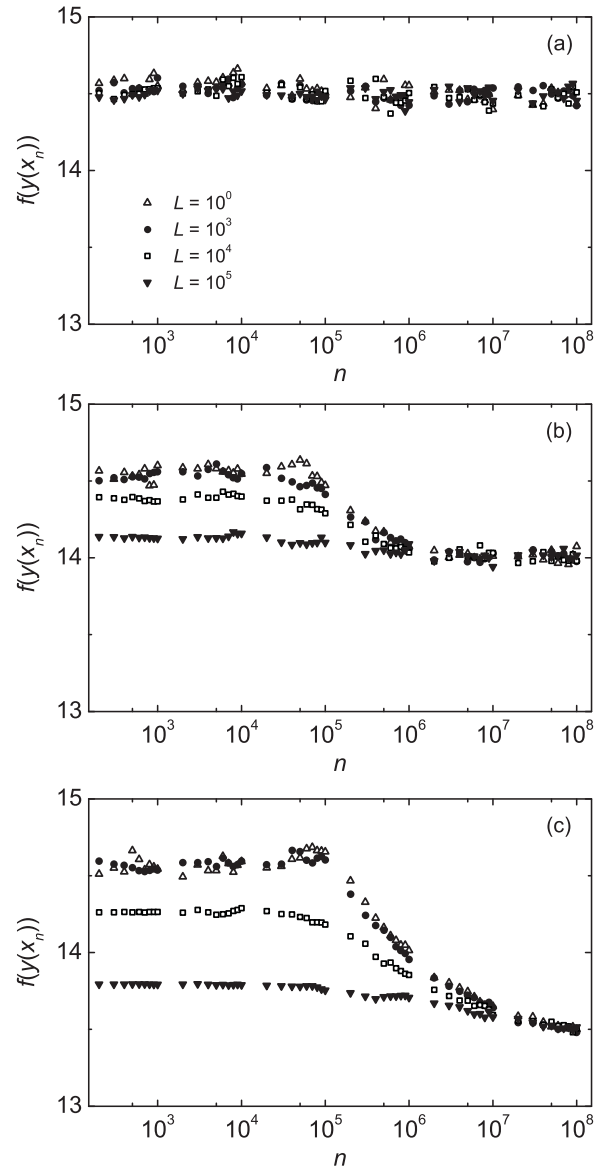


FIG. 6. Hidden search dynamics $f(y(x_n))$ (RE-TSP with $N = 316$; $I = 2^5$). The results averaged over I search processes are plotted for four values of L . (a) $\Theta_c = -1.41$; (b) $\Theta_c = -1.72$; (c) $\Theta_c = -2.01$. The values of the other parameters in the preceding standard SA are $\Theta_s = 1$ and $\gamma = 0.95$.

was found that the crossing T_1 corresponds to the onset of “slow dynamics” and the system becomes influenced by the underlying energy landscape. Last, in the low temperature regime ($T < T_2$), the system is frozen and confined in a single or a few basins on the observation time scale. The temperature T_2 depends on the cooling rate in a manner analogous to the experimental glass transition, namely, slower cooling lower this temperature.

The other study was made by Shinjo on the relaxation dynamics of a single-component LJ system [50], which was used as a model of argon. There, unconstrained dynamics after rapid quenching were observed at several temperatures: The experimental procedure is similar to that in our second experiment with a small value of L . This study showed that the interbasin dynamics can be classified into four categories

according to the temperature. The relevant characteristics, including those not described in the text but seen from the figures, are summarized as follows: In the high temperature regime ($T > T_1$), the system wanders over a wide range of minimum energies throughout the run; the system was considered to be in the liquid state. Contrastingly, in the intermediate temperature regime ($T_1 > T > T_2$), first the system wanders over a wide range of minimum energies and then it settles into a single basin or a few basins with almost the same minimum energy; this change was regarded as a transition from the supercooled liquid to the solid state. The sampled minimum energy gradually decreases with large fluctuation in the wandering phase. (This can be seen from Fig. 2 in Ref. [50].) Next in the low-intermediate temperature regime ($T_2 > T > T_3$), the wandering phase disappears and a stepwise decrease in minimum energy is observed during the run. It should be noted that over the last two temperature regimes, the reachable minimum energy appears to go down as the temperature decreases (Figs. 1–3 in Ref. [50]). Last, in the low temperature regime ($T < T_3$), the system stays at a single basin or a few basins with almost the same minimum energy throughout the run, and this energy is conversely higher than the lowest minimum energy found in the third regime (Fig. 4 in Ref. [50]). By correlating these results with the number of accessible local minima, it was predicted that the temperature T_2 would be the glass transition temperature.

The connection between their molecular and our optimization systems is a little unclear because we have no good grounds for believing that not only the two dynamics—the Hamiltonian and the Monte Carlo dynamics—but also the two landscapes—the landscape of the LJ systems and that of the TSPs with the 2-opt neighborhood—are equivalent in the present context. We can find, however, that there is a striking resemblance between these two systems in the temperature dependence of the successive interbasin dynamics. That is, the horizontal transition state of the Monte Carlo dynamics, the level of which state decreases with decreasing temperature [Fig. 5(a)], is analogous to the equilibrium state of the Hamiltonian dynamics in the “landscape-influenced” temperature regime [49]. Hence, the preceding downward transitions can be regarded as a relaxation step (as observed in Ref. [50]) to the equilibrium state. These physical implications suggest that the conventional physical analogy is a misguided one and if and when it is modified, the following two features should be taken into account.

The first is that the effectiveness of finite-time SA comes from the downward interbasin dynamics occurring before equilibrium, not from the equilibrium sampling after quasistatic cooling. In the first experiment, we found that the search performed in a limited temperature range plays a direct role in the effective use of SA. As seen from the second experiment, this range includes the temperature at which the relaxation continues just until the end of the observation and the equilibration time is not affected by the rate of the preceding cooling. These findings strongly suggest that the optimizing ability of the Monte Carlo dynamics derives from the relaxation dynamics near the effective intermediate temperature.

The second is that this effective temperature is specified by the condition that the equilibration time is comparable to the

observation time. This means that the effective temperature can be identified using the Deborah number [51,52]. The Deborah number is defined as the ratio of relaxation time to observation time; thus, if we use the above equilibration time as a measure of the relaxation time, the effective temperature is identified as the temperature at which the number is unity. This view is consistent with the finding that the effective temperature decreases with the extension of the search time, which fact was observed in both of our experiments, as well as in the previous study [24]. The Deborah number was originally introduced to discuss the rheological properties of materials, and its applicability to the characterization of the glass transition is discussed in Ref. [52]. Whether this application is appropriate or not, the present findings strongly suggest that the effective temperature is related not to the crystallization phenomenon but to a vitrificationlike process because this temperature is not intrinsic to each problem instance but is variable with the observation time like the glass transition temperature. It is worthy to note that although the concept of time has no intrinsic meaning in the Monte Carlo dynamics, this characterization of the effective temperature serves its purpose, irrespective of the dynamics, because the Deborah number is a dimensionless number.

From this discussion, the downward dynamics analogous to relaxation phenomena are expected to play a key role also in the optimization process with slow cooling. To confirm this feature directly, another experiment for SA was designed to examine systematically the effect of cooling in various temperature ranges. The details of this supplementary analysis are described in the next section.

IV. EXPERIMENTAL ANALYSIS II

A. Rate-cycling experiment

We introduce here another design of the temperature schedule for the supplementary experiment. This experiment is similar to what is called the temperature-cycling experiment [53] except that the derivative of the temperature instead of the temperature itself is changed cyclically. This variant experiment is referred to as the rate-cycling experiment for short below. The temperature-cycling experiment has also been used in studies on liquid and glass, and in particular on aging phenomena [53–56].

The framework of the cooling schedule is reconstructed in the following manner. First, an initial logarithmic temperature Θ_s , a final logarithmic temperature Θ_e , and a total search time t_e (or a total number of search steps, n_e) are selected. Then, the cooling process is divided into three stages so that the cooling in the intended logarithmic temperature range, $\Theta_c + \frac{1}{2}\Delta\Theta > \Theta > \Theta_c - \frac{1}{2}\Delta\Theta$, is performed as its second stage. We refer to this intermediate temperature range simply as the target range. The cooling rate in logarithmic temperature scale is selected and fixed in each stage so that the rates in the first and third stages are equal and are λ times the rate in the second stage. In this framework, the six parameters, Θ_s , Θ_e , Θ_c , $\Delta\Theta$ (the logarithmic temperature width of the target range), λ (the cooling rate ratio), and t_e ($=n_e$), uniquely determine the cooling schedule.

For experimental convenience, we use the following expression of the logarithmic temperature schedule in the k th

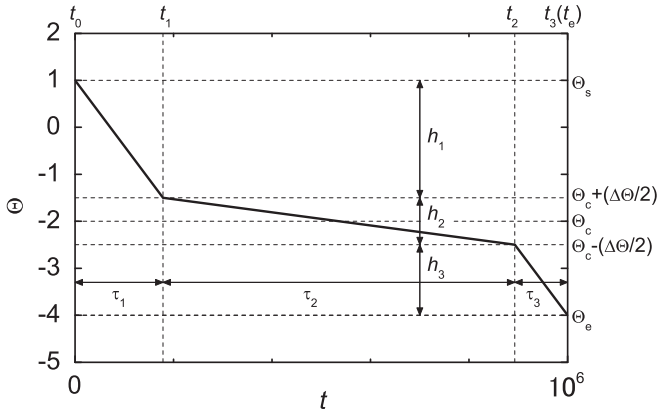


FIG. 7. Example of the cooling schedule $\Theta(t)$ ($\Theta_s = 1$, $\Theta_e = -4$, $\Theta_c = -2.0$, $\Delta\Theta = 1.0$, $\lambda = 10$, and $t_e = 10^6$). See text for symbols.

stage:

$$\Theta(t) = \Theta(t_{k-1}) - \frac{t - t_{k-1}}{\tau_k} h_k \quad (t_{k-1} \leq t \leq t_k), \quad (4)$$

where t_{k-1} and t_k are the beginning and the ending time of this stage, respectively, and h_k and τ_k are the logarithmic temperature width and the allocated time in this stage, respectively; hence, t_k equals the accumulated time of τ_j 's up to $j = k$ ($t_0 = 0$). An example of the cooling schedule is depicted in Fig. 7. The values of h_k 's are determined by Θ_s , Θ_e , Θ_c , and $\Delta\Theta$, and those of τ_k 's are determined by h_k 's, λ , and t_e by the following equations: $\tau_1 = (h_1/H)t_e$, $\tau_2 = (\lambda h_2/H)t_e$, and $\tau_3 = (h_3/H)t_e$, where $H = h_1 + \lambda h_2 + h_3$. For $\lambda > 1$, the cooling in the second stage is slower than that in the other stages and thus much time is allocated in the target range. Conversely, for $\lambda < 1$, the cooling is faster and little time is allocated in the target range. Note that if $\lambda = 1$, then the schedule is equivalent to the geometric cooling schedule.

The algorithm for this rate-cycling experiment is described as follows.

[RCSA1] Determine the logarithmic temperature schedule $\Theta(t)$; that is, select the values of the six parameters, Θ_s , Θ_e , Θ_c , $\Delta\Theta$, λ , and t_e ($=n_e$).

[RCSA2] Set $n := 0$ and generate an initial solution x_0 randomly. Set $\Theta := \Theta(0)$ and $x^* := x_0$.

[RCSA3] Implement a single step of the MA, namely, [MA1] (the logarithmic temperature Θ has to be converted to the temperature T in advance). Increment n by one. (The current solution is x_n and the current incumbent solution is x^* .)

[RCSA4] If the termination condition, $n = n_e$, is satisfied, output x^* and stop. Otherwise, set $\Theta := \Theta(n)$, and return to [RCSA3].

The optimization performance and the hidden search dynamics were investigated for the same instances as in the second experiment. The parameter values were selected in the following way. The values of Θ_s and Θ_e were fixed at 1 and -4 , respectively; these values as well as the values of Θ_c and also t_e (or n_e) were selected to cover the whole interesting temperature range found in the first two experiments. The value of λ was selected from $\{100, 10, 1, 0.1, 0.01\}$. The value of $\Delta\Theta$ was taken to be 1.0 and for some system sizes,

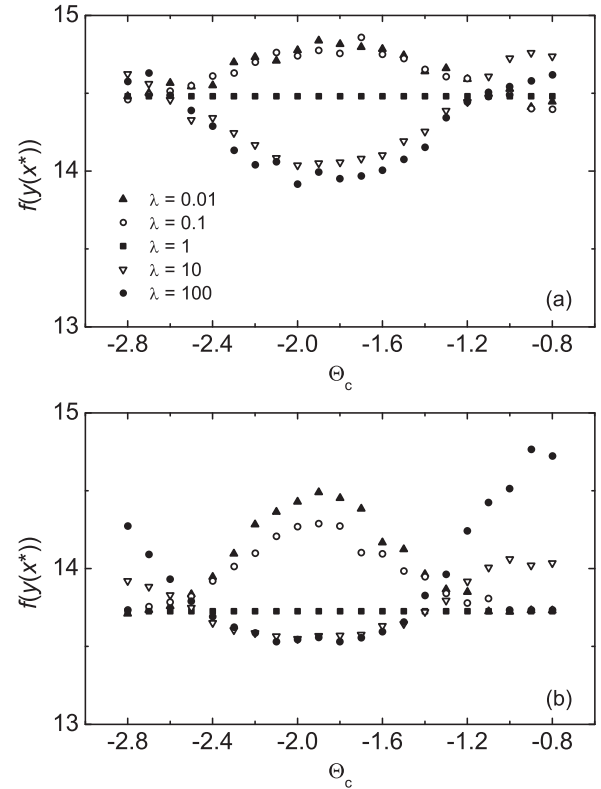


FIG. 8. Optimization performance $f(y(x^*))$ (RE-TSP with $N = 316$; $\Theta_s = 1$, $\Theta_e = -4$, $\Delta\Theta = 1.0$, and $I = 2^5$). The results are averaged over I search processes and plotted against the target logarithmic temperature Θ_c . (a) $n_e = 10^6$; (b) $n_e = 10^7$.

the experiment was repeated with varying $\Delta\Theta$, the values of which were selected from the range from 0.2 to 1.8. The results will be presented only for $\Delta\Theta = 1.0$ as typical illustrations because in the characteristic features described below, no qualitative differences were observed in other cases.

B. Results

The optimization performance evaluated by $f(y(x^*))$ is shown in Fig. 8 for $n_e = 10^6$ and 10^7 ; the results were averaged over $I (=2^5)$ independent runs. We find that for both values of n_e , the average performance is improved by allocating much time to the search performed in some intermediate temperature range ($\lambda > 1$). What is more important here is that the best target range extends on both sides of the effective temperature found in the first two experiments. Conversely, in the case that little time is allocated ($\lambda < 1$), the average performance is deteriorated. In other words, the performance is sensitive to cooling crossing the effective temperature.

Figure 9 shows the hidden search dynamics $f(y(x_n))$ for the case $\lambda = 100$ in Fig. 8(b); the results were averaged over the I processes. In this case, the search time allocated in the second stage is more than 96% of the total search time. We find that, if the system is not in the initial transient period, the downward interbasin dynamics begin at approximately the time that the logarithmic temperature falls below $\hat{\Theta}$, regardless of the target temperature. (The time steps passing through $\hat{\Theta}$ are $n \approx 6.9 \times 10^6$ for $\Theta_c = -1.2$, $n \approx 3.1 \times 10^6$ for $\Theta_c = -1.6$, and $n \approx 2.3 \times 10^5$ for $\Theta_c = -2.0$.) For $\Theta_c \gtrsim -1.2$, the system freezes

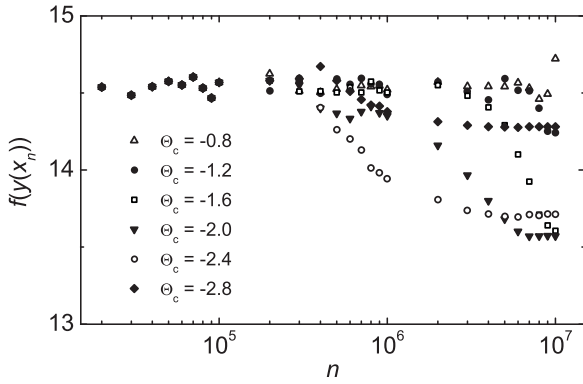


FIG. 9. Hidden search dynamics $f(y(x_n))$ (RE-TSP with $N = 316$; $\Theta_s = 1$, $\Theta_c = -4$, $\Delta\Theta = 1.0$, $\lambda = 100$, $n_c = 10^7$, and $I = 2^5$). Plotted are the results averaged over I search processes.

before the downward dynamics work effectively. In contrast, for $\Theta_c \lesssim -2.4$, the system appears to be confined in a valley structure in the landscape some time during the second stage. Consequently, the optimization performance is maximized in the intermediate case that the descending motion continues most successfully on the experimental time scale.

C. Discussion

As was expected, the optimization performance was maximized in the case that the system was cooled slowly in some intermediate temperature range including the effective temperature. The performance was not only dependent on but also sensitive to the cooling rate in this limited temperature range. The results of the hidden search dynamics were also unsurprising ones. Overall, the downward interbasin dynamics worked most successfully in the above influential temperature range. Thus, we conclude that the process of relaxation to equilibrium plays a primary role in finite-time optimization by SA at least for the present TSPs. Conventionally, the geometric cooling schedule has been used and standardized probably because of its easiness and cost performance. There is, however, room for essential improvement within the framework of monotonic cooling under time constraint; it can be achieved by taking the function of the relaxation dynamics appropriately into account.

The applicability of the present findings to other optimization problems should be tested because the present results obviously depend on the landscape structure determined by the cost and neighborhood functions. We leave this issue for future work and in the remainder of this section, several comments on SA and SA-related topics are made on the basis of the meaningful relationship with vitrification phenomena.

We confirmed that slow cooling crossing the temperature analogous to the glass transition temperature was recommendable for the effective use of finite-time SA. From the standpoint of the present work, it is appropriate to point out that a similar strategy has already been proposed for actual annealing. It was made by Narayanaswamy [57], at approximately the same time as the dawn of SA, as a result of a phenomenological analysis of the optimum schedule for annealing flat glass. There, the analysis was designed to find a

schedule that minimizes either residual stress or annealing time. The resulting schedule includes a slow cooling step (see Fig. 6 in Ref. [57]) in which the temperature decreases crossing the glass transition temperature and the annealing point of the considered material, soda-lime glass [23,58,59]. This schedule is qualitatively similar to that used in the present rate-cycling experiment with $\lambda > 1$, though the former was unknown to the author at the time of the experiment. After the proposal of SA, its modeling using the Markov chain approach enabled us to describe a cooling schedule that ensures asymptotic convergence to the globally optimal solution(s) with probability one. A successful description of such a schedule [13,14] might have led one to believe that convergence is the principal mechanism for optimization by SA. In a practical sense, however, the assumption of ergodicity is not relevant for this kind of complex system because it has the nature of effectively broken ergodicity [60,61]. With this feature in mind, we confirmed that the glassy relaxation dynamics naturally function as a crucial ingredient in the present optimization. It may be said that the rectified analogy restores the normal relationship between simulated and actual annealing.

The next comment concerns the use of nonmonotonic temperature schedules. This line of improvement can be found, for example, in the thermal cycling [62] and bouncing [63] approaches. If we reflect the rectified analogy in the design of the reheating schedule, heating up to or slightly beyond the range corresponding to the glass transition range may be a rational strategy. This idea comes from the stabilization phenomenon observed in glass-forming materials [23]: When a cooled glass is reheated, the glass does not return its history experienced in the cooling process but relaxes in the glass transition range into another state possibly reached by a slower cooling. Note that in monotonic cooling from high temperature, more time is dissipated for the initial transient period with the increase of the system size. For this reason, nonmonotonic cooling is a more realistic approach to deal with large-size instances. It should be emphasized that, whether monotonic or not, the rectified analogy may provide a comprehensive guideline for the effective design of the temperature schedule.

Finally, a remark is made on threshold algorithms [14]. The name is intended to cover any kind of algorithm having a threshold by which one determines whether or not to accept a deteriorated move to the neighborhood. The SA algorithm belongs to this group and in fact a parameterized exponential function is used as a threshold (see [MA1] in Sec. II B). The threshold accepting method [64] and the generalized SA algorithm [30,31] are other known examples. In the conventional finite-time SA, as confirmed in the present work, what is important is not its intended statistics but the effective use of the glassy relaxation dynamics observed as downward interbasin transitions in the solution space. Therefore, it is quite natural to consider that the other threshold algorithms work by the same mechanism. If this is correct, for fair competition among threshold algorithms and also for a comprehensive understanding of their optimization characteristics, each algorithm has to be designed and implemented in such a way that the glassy relaxation dynamics function successfully in the landscape of the cost function. The positive role of these

dynamics in the search process of several threshold algorithms has been affirmatively confirmed in a recent experimental analysis conducted by the author [65]. This confirmation was made by applying the present series of experiments to some of the same TSP instances.

V. CONCLUSION

The optimization mechanism of finite-time SA was reconsidered by starting from the verification of its conventional naive optimization scenario and by aiming at the elucidation of the precise connection with physical phenomena. A series of three experiments was designed by applying concepts and methods developed for studies on liquid and glass and was conducted on two types of random TSPs up to a system size of practical interest. We confirmed that there exists an effective temperature in finite-time optimization by SA and found that in the cooling process near this effective temperature, downward interbasin dynamics successfully function as an

optimizer. By comparison with the corresponding results of computer experiments on LJ systems, these dynamics were related to a slow approach to equilibrium and thus the effective temperature can be identified as the temperature at which the Deborah number is unity. It is concluded that the effectiveness of finite-time SA derives from the glassy relaxation dynamics occurring in the landscape-influenced temperature regime and that the naive optimization scenario should be rectified on the basis of the analogy with vitrification phenomena, at least for the present TSPs. This rectification is expected to restore the normal relationship between simulated and actual annealing and to develop a comprehensive understanding of the optimization characteristics of the SA and related algorithms.

ACKNOWLEDGMENT

This work was supported in part by JSPS KAKENHI (Grant No. 20656017).

-
- [1] P. Cheeseman, B. Kanefsky, and W. M. Taylor, in *Proceedings of the 12th International Conference on Artificial Intelligence (IJCAI-91)*, edited by J. Mylopoulos and R. Reiter, Vol. 1 (Morgan Kaufmann, San Mateo, 1991), p. 331.
- [2] S. Kirkpatrick and B. Selman, *Science* **264**, 1297 (1994).
- [3] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, *Nature (London)* **400**, 133 (1999).
- [4] R. Monasson, in *Complex Systems*, edited by J.-P. Bouchaud, M. Mézard, and J. Dalibard (Elsevier, Amsterdam, 2007), p. 1.
- [5] M. R. Garey and D. S. Johnson, *Computers and Intractability, A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).
- [6] S. Kirkpatrick and G. Toulouse, *J. Phys. (Paris)* **46**, 1277 (1985).
- [7] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond*, Part 2 (World Scientific, Singapore, 1987).
- [8] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, Chap. 9 (Oxford University Press, Oxford, 2001).
- [9] A. K. Hartmann and H. Rieger, eds., *New Optimization Algorithms in Physics* (Wiley-VCH, Berlin, 2004).
- [10] J. J. Schneider and S. Kirkpatrick, *Stochastic Optimization* (Springer, Berlin, 2006).
- [11] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, *Science* **220**, 671 (1983).
- [12] V. Černý, *J. Optim. Theory Appl.* **45**, 41 (1985).
- [13] E. H. L. Aarts and J. H. M. Korst, in *Essays and Surveys in Metaheuristics*, edited by C. C. Ribeiro and P. Hansen (Kluwer Academic Publishers, Boston, 2002), p. 1, and references therein.
- [14] E. H. L. Aarts, J. H. M. Korst, and P. J. M. van Laarhoven, in *Local Search in Combinatorial Optimization*, edited by E. H. L. Aarts and J. K. Lenstra (Wiley, Chichester, 1997), p. 91.
- [15] Y. Iba, *Int. J. Mod. Phys. C* **12**, 623 (2001).
- [16] J. -P. Bouchaud, M. Mézard, and J. Dalibard, eds., *Complex Systems* (Elsevier, Amsterdam, 2007).
- [17] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, 2003).
- [18] A. K. Hartmann and H. Rieger, *Optimization Algorithms in Physics* (Wiley-VCH, Berlin, 2002).
- [19] P. Salamon, P. Sibani, and R. Frost, *Facts, Conjectures, and Improvements for Simulated Annealing* (Society for Industrial and Applied Mathematics, Philadelphia, 2002).
- [20] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [21] R. K. Singh, D. Kumar, and B. Chen, in *Encyclopedia of Applied Physics*, edited by G. L. Trigg, Vol. 21 (Wiley-VCH, Weinheim, 1997), p. 67.
- [22] H. Ohtani, in *Materials Science and Technology*, edited by R. W. Cahn, P. Haasen, and E. J. Kramer, Vol. 7 (VCH, Weinheim, 1992), p. 147.
- [23] G. W. Scherer, in *Materials Science and Technology*, edited by R. W. Cahn, P. Haasen, and E. J. Kramer, Vol. 9 (VCH, Weinheim, 1991), p. 119.
- [24] M. Fielding, *SIAM J. Optim.* **11**, 289 (2000).
- [25] I. Morgenstern, in *Heidelberg Colloquium on Glassy Dynamics*, edited by J. L. van Hemmen and I. Morgenstern (Springer-Verlag, Heidelberg, 1987), p. 399.
- [26] K. H. Hoffmann, D. Würtz, C. de Groot, and M. Hanf, in *Parallel and Distributed Optimization*, edited by M. Grauer and D. B. Pressmar (Springer-Verlag, Heidelberg, 1991), p. 155.
- [27] K. H. Hoffmann and J. C. Schön, *Found. Phys. Lett.* **18**, 171 (2005).
- [28] J. C. Schön, *J. Phys. A* **30**, 2367 (1997).
- [29] P. Salamon, J. D. Nulton, J. R. Harland, J. Pedersen, G. Ruppeiner, and L. Liao, *Comput. Phys. Commun.* **49**, 423 (1988).
- [30] C. Tsallis and D. A. Stariolo, *Physica A* **233**, 395 (1996).
- [31] T. J. P. Penna, *Phys. Rev. E* **51**, R1 (1995).
- [32] K. Kimura and K. Taki, in *Proceedings of the 13th IMACS World Congress on Computation and Applied Mathematics (IMACS'91)*, edited by R. Vichnevetsky and J. J. H. Miller, Vol. 2 (Criterion Press, Dublin, 1991), p. 827.
- [33] U. H. E. Hansmann and Y. Okamoto, *J. Comput. Chem.* **14**, 1333 (1993).

- [34] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, *Chem. Phys. Lett.* **219**, 343 (1994).
- [35] S. Boettcher and A. G. Percus, *Phys. Rev. Lett.* **86**, 5211 (2001).
- [36] G. Zaránd, F. Pázmándi, K. F. Pál, and G. T. Zimányi, *Phys. Rev. Lett.* **89**, 150201 (2002).
- [37] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys, eds., *The Traveling Salesman Problem, A Guided Tour of Combinatorial Optimization* (Wiley, Chichester, 1985).
- [38] D. S. Johnson and L. A. McGeoch, in *Local Search in Combinatorial Optimization*, edited by E. H. L. Aarts and J. K. Lenstra (Wiley, Chichester, 1997), p. 215.
- [39] R. D. Mountain and D. Thirumalai, *J. Phys. Chem.* **93**, 6975 (1989).
- [40] D. Thirumalai, R. D. Mountain, and T. R. Kirkpatrick, *Phys. Rev. A* **39**, 3563 (1989).
- [41] F. H. Stillinger and T. A. Weber, *Phys. Rev. A* **25**, 978 (1982).
- [42] F. H. Stillinger and T. A. Weber, *Phys. Rev. A* **28**, 2408 (1983).
- [43] F. H. Stillinger and T. A. Weber, *Science* **225**, 983 (1984).
- [44] M. Hasegawa, *AIP Conf. Proc.* **708**, 747 (2004).
- [45] M. Hasegawa, *J. Phys. Soc. Jpn.* **74**, 2872 (2005).
- [46] M. Hasegawa, *AIP Conf. Proc.* **982**, 796 (2008).
- [47] D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon, *Oper. Res.* **37**, 865 (1989).
- [48] D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon, *Oper. Res.* **39**, 378 (1991).
- [49] S. Sastry, P. G. Debenedetti, and F. H. Stillinger, *Nature (London)* **393**, 554 (1998).
- [50] K. Shinjo, *Phys. Rev. B* **40**, 9167 (1989).
- [51] M. Reiner, *Phys. Today* **17**, 62 (1964).
- [52] J. M. Stevels, *J. Non-Cryst. Solids* **6**, 307 (1971).
- [53] E. Vincent, J. P. Bouchaud, J. Hammann, and F. Lefloch, *Philos. Mag. B* **71**, 489 (1995).
- [54] L. Bellon, S. Ciliberto, and C. Laroche, e-print [arXiv:cond-mat/9905160v2](https://arxiv.org/abs/cond-mat/9905160v2).
- [55] R. L. Leheny and S. R. Nagel, *Phys. Rev. B* **57**, 5154 (1998).
- [56] F. Alberici, P. Doussineau, and A. Levelut, *Europhys. Lett.* **39**, 329 (1997).
- [57] O. S. Narayanaswamy, *J. Am. Ceram. Soc.* **64**, 109 (1981).
- [58] D. G. Holloway, *The Physical Properties of Glass* (Wykeham, London, 1973).
- [59] R. H. Doremus, in *Encyclopedia of Applied Physics*, edited by G. L. Trigg, Vol. 7 (VCH, New York, 1993), p. 251.
- [60] R. G. Palmer, *Adv. Phys.* **31**, 669 (1982).
- [61] R. G. Palmer, in *Lectures in the Sciences of Complexity*, edited by D. L. Stein (Addison-Wesley, Reading, 1989), p. 275.
- [62] A. Möbius, A. Neklioudov, A. Díaz-Sánchez, K. H. Hoffmann, A. Fachat, and M. Schreiber, *Phys. Rev. Lett.* **79**, 4297 (1997).
- [63] J. Schneider, I. Morgenstern, and J. M. Singer, *Phys. Rev. E* **58**, 5085 (1998).
- [64] G. Dueck and T. Scheuer, *J. Comput. Phys.* **90**, 161 (1990).
- [65] M. Hasegawa, *Comput. Phys. Commun.* **182**, 229 (2011).