Dynamics and changing environments in highly optimized tolerance

Tong Zhou and J. M. Carlson
Department of Physics, University of California, Santa Barbara, California 93106
(Received 6 December 1999)

Highly optimized tolerance (HOT) is a mechanism for power laws in complex systems based on the robust design of systems in uncertain environments. Once the system, the environment, and the optimization scheme have been specified, the HOT state is fixed and corresponds to the set of measure zero (typically a single point) in the configuration space which minimizes a cost function U. Here we explore the U-dependent structures in configuration space which are associated with departures from the optimal state. We introduce dynamics, quantified by an effective temperature T, such that $T=0$ corresponds to the original HOT state, while $T\rightarrow\infty$ corresponds to completely random configurations. More generally, T defines the range in state space over which fluctuations are likely to be observed. In a fixed environment fluctuations always raise the average cost. However, in a time-dependent environment, mobile configurations can lower the average U because they adjust more efficiently to changes.

PACS number(s): 05.40.−a, 64.60.Ht, 05.65.+b, 87.17.−d

I. INTRODUCTION

Recently, highly optimized tolerance (HOT) was introduced as a mechanism which connects power-law distributions of events to the optimal design of a system subject to external perturbations [1–3]. HOT links robustness to the underlying origins of complexity in systems where design and/or evolution play an important role. HOT provides strikingly accurate fits to frequency vs size distributions for forest fires, power outages, and world wide web traffic [3]. One key feature of HOT systems is their special sensitivity to unexpected perturbations or systematic changes in the environment. In this paper we introduce a time-dependent environment, and derive conditions under which the original optimized state is replaced by an ensemble. The ensemble includes fluctuations about the original HOT state characterized by an effective temperature T.

We focus on a simple, one-dimensional forest fire model [4] for HOT considered previously in [1]. The optimal state minimizes a function U, associated with the expected loss due to a fire initiated by a single spark. The model also has an interpretation in terms of the design of world wide web sites, in which U is associated with the expected file transfer size. In Sec. II we define the model and calculate the density of states as a function of the cost function U. We replace the static description of the configuration, which consists of trees and firebreaks, to allow for dynamical fluctuations of the barriers, characterized by an effective temperature T. In Sec. III we study the response of the model to a changing environment, and derive an inequality describing when the increased robustness to environmental change outweighs design flaws introduced by fluctuations. We conclude in Sec. IV with a discussion of our results.

II. CONFIGURATION SPACE METRICS, DYNAMICS, AND FINITE TEMPERATURES

To date HOT systems have been described simply in terms of a static optimized configuration [1–3]. In this section we generalize this description to include fluctuations about the optimized state, which are characterized by an effective temperature T. We present our results in the context of the one-dimensional forest fire model [4], although the basic approach we outline is applicable to other systems as well.

In the configuration space, every point, i.e., each state of the system, has a corresponding value of U. The optimal configuration has measure zero and corresponds to the minimum $U\equiv U_0$. If the configuration space is N dimensional, it can be divided into $N-1$-dimensional subspaces, each of which consists of states with a fixed value of U. We define $V(U)$ to be the density of states, i.e., the relative number of configurations in each subspace, with $V(U_0)=0$.

A. One-dimensional forest fire model

The model consists of a continuous, one-dimensional line of length L, $0\leqslant x\leqslant L$. The line is subdivided into disjoint segments by a set of N cuts at positions $y_1<y_2<\ldots y_N$. The N-dimensional vector $\mathbf{y}$ uniquely defines the configuration. A probability distribution $p(x)$ specifies the likelihood the line is hit at position x. When a hit occurs at x, such that $y_i<x<y_{i+1}$, the associated cost or loss is the length $y_{i+1}-y_i$ of the segment containing x.

The analogy with the well-studied forest fire model [4] is made when the line is understood to be the continuum limit of a discrete one-dimensional percolation model [5] mapped into a finite interval. The segments represent contiguous occupied sites, or trees, and the cuts represent vacant sites, or firebreaks. Here we study the high density limit, in which all but a finite number of sites are occupied. The distribution $p(x)$ represents the spatially structured distribution of sparks which initiate fires and burn through connected clusters.

This model has also been studied in the context of world wide web site design [1,3], where cuts represent the division of the document into files, and $p(x)$ represents the probability a user will download information in the document up through the position x. Loss is associated with the cost of file transfer over the world wide web, which increases with file

size due to the increased delays.

We choose this model as our starting point because it is simple and tractable, yet rich enough to illustrate our basic results. In general, cuts can be thought of as resources which are allocated prior to a loss and serve to limit the damage. In the typical percolation model, sites are randomly occupied and configurations are characterized by their density. In designed configurations, the placements of the cuts or vacant sites are specifically chosen to minimize the expected loss, which we denote by U,

$$U(y_1,y_2,\ldots,y_N)=\sum_{i=1}^{N+1}(y_i-y_{i-1})\int_{y_i-1}^{y_i}p(x)dx, \quad (1)$$

where in the sum $y_i=0$ for $i=0$ and $y_i=L$ for $i=N+1$.

The HOT state minimizes U for a specific choice of $p(x)$ [1]. In the absence of uncertainty [i.e., $p(x)$ is a δ function at a known location], the HOT state is trivially associated with any configuration in which a cut is placed at the known location of the spark. Alternately, in the case of maximal uncertainty, $p(x)$ is spatially uniform, and the HOT state corresponds to equally spaced cuts. In this case, the size of the event is fixed at $L/(N+1)$. The most interesting cases correspond to nonuniform uncertainty, where $p(x)$ has regions of high and low probability. In that case, cuts are concentrated in regions of high $p(x)$, and sparse in regions of low $p(x)$. In [1] it was shown that for a broad class of $p(x)$, including distributions with Gaussian and exponential tails, configurations which minimize U have power-law distributions of events (see the Appendix). The probability distribution of events of size l is given by $P(l)\propto l^{-1}$.

In the absence of degeneracies associated with special choices of $p(x)$, the optimal state corresponds to a unique configuration, represented by a single point $\mathbf{y}_0$ in the N-dimensional configuration space $\mathbf{y}$. Overall, a very small portion of configurations will be fine tuned to give small U [Eq. (1)]. When $p(x)$ is sharply peaked [compared to $L/(N+1)$] with rapidly decaying tails, the value of U for most of the configurations is determined by the width of $p(x)$, since in that case U is determined nearly entirely by the length of the first segment.

In Fig. 1 we illustrate the density of states $V(U)$ as a function of $\delta U=U-U_0$ for a system consisting of $N=10$ cuts when $p(x)$ has exponential tails. To obtain this plot we first pick $C=5\times10^7$ configurations randomly and calculate the corresponding values of U. Our results are not sensitive to the specific choice of C as long as it is sufficiently large. This random sampling yields a good description of $V(U)$ near its maximum (the solid curve in Fig. 1). However, since these configurations are generated randomly, the much less likely configurations in the neighborhood of U_0 are not well represented. To resolve the structure near the HOT state, we locate the optimal configuration $\mathbf{y}_0$ by minimizing Eq. (1). Then we pick random relocations y_i of each cut within a prescribed neighborhood of its original placement y_{i0} defined by $y_{i0}-\epsilon_i\leqslant y_i\leqslant y_{i0}+\epsilon_i$, where $\epsilon_i=\epsilon/\sqrt{\partial^2U/\partial y_i^2}\big|_{\mathbf{y}=\mathbf{y}_0}$ is the change in y_i which produces a change of ϵ in U when all other cuts are held fixed. By increasing ϵ we generate a series of curves which extrapolate between the immediate vicinity of the HOT state and the initial random sampling.

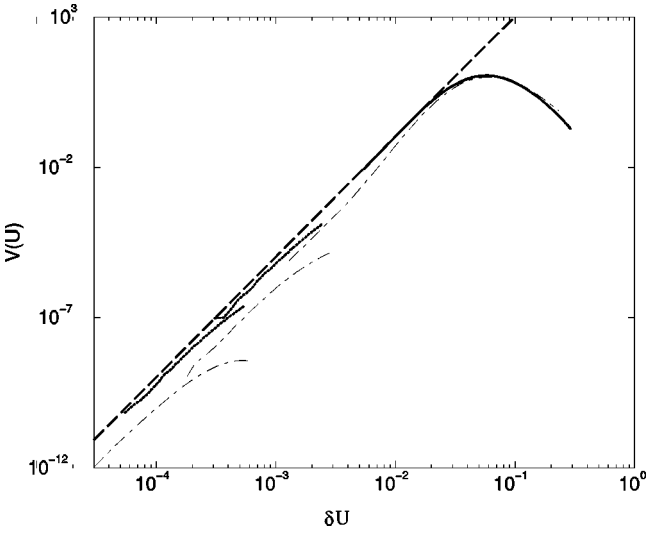

FIG. 1. The density of states $V(U)$. This plot is for a system of length $L=1$, with $N=10$ cuts and $p(x)=\alpha e^{-\alpha x}/(1-e^{-\alpha})$. Three curves are shown for $\alpha=5$: the heavy solid line is obtained from simple sampling while the two heavy dotted lines are from sampling from neighborhood of the optimal configuration. The heavy dashed line is for the power-law $(\delta U)^{(N/2)-1}$, where $N=10$ here. For comparison, three corresponding curves are shown for $\alpha=10$ in lighter dot-dashed lines.

For a particular neighborhood defined by ϵ, here assumed small, results for smaller values of U are well sampled, while for larger U inadequate sampling rounds the curve at the boundary of the neighborhood. The two heavy dotted lines in Fig. 1 are examples of our numerical results generated for two different values of ϵ. The lighter dot-dashed lines in that figure are corresponding numerical results for a different $p(x)$.

For the one-dimensional forest fire model the density of states $V(U)$ is relatively simple, exhibiting a minimum for the HOT state which is smoothly connected to a broad maximum which characterizes the most probable generic random configurations. In the neighborhood of the optimal state, the relationship between $V(U)$ and $\delta U\equiv U-U_0$ is well described by a power law, in which the exponent increases with N, but is independent of our choice of $p(x)$ (see below). In Fig. 1 this is illustrated by the fact that the slope of $\log V(U)$ vs $\log \delta U$ is independent of $p(x)$, for δU sufficiently small. The coefficient of the power law does depend on $p(x)$. For a given δU, steeper distributions $p(x)$ exhibit a smaller density of states.

The power-law form of $V(U)$ in the near neighborhood of the optimal state is quite general, and does not depend on our choice of $p(x)$ or the details of the forest fire model. Instead it is a property associated with the dimension of the configuration space, or, equivalently, the number of degrees of freedom available for the design. In the configuration space, if we move away from the optimal point $\mathbf{y}_0$ to $\mathbf{y}_0+\delta\mathbf{y}$ the first order change in U vanishes (by definition). Thus we have

$$\delta U=\frac{1}{2}\sum_{i,j}^{n}\frac{\partial^2U}{\partial y_i\partial y_j}(\delta y_i)(\delta y_j)=\frac{1}{2}\delta\mathbf{y}\cdot\nabla\nabla U\cdot\delta\mathbf{y}. \quad (2)$$

We define the eigenvectors of the matrix $\nabla\nabla U$ to be $\mathbf{z}_i$, and

the corresponding eigenvalues to be λ_i. Then rewriting $\delta\mathbf{y}$ in terms of $\mathbf{z}_i$ as $\delta\mathbf{y}=\Sigma c_i\mathbf{z}_i$, we have

$$\delta U=\sum \lambda_i c_i^2. \tag{3}$$

If we view $\mathbf{z}_i$ as the basis of a coordinate system, c_i's then form the coordinate of a point. So from Eq. (3) we see that fixed δU define a hyperellipse in the N-dimensional configuration space. The volume this ellipse encompasses is

$$V_c(U_0+\delta U)=\frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}+1\right)}\frac{(\delta U)^{N/2}}{\prod \sqrt{\lambda_i}}. \tag{4}$$

Because $\Pi\lambda_i$ is the determinant of the matrix $\nabla\nabla U$, we have

$$V(U_0+\delta U)=\frac{dV_c(U_0+\delta U)}{d(\delta U)}=\frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}\frac{(\delta U)^{(N/2)-1}}{\sqrt{\|\nabla\nabla U\|}}. \tag{5}$$

From this we see that near the optimal state $V(U)$ is described by a power law, where the exponent is determined by the dimensionality of the configuration space. This power law is shown as the dashed line in Fig. 1 and it agrees well with the numerical results.

Aside from numerical constants, there is only one nontrivial term in the prefactor of Eq. (5), $\sqrt{\|\nabla\nabla U\|}$, which is evaluated at $\mathbf{y}=\mathbf{y}_0$, and depends on the number of cuts and the probability distribution $p(x)$. Aside from special cases [e.g., uniform $p(x)$] this is a complicated function. However, for a fixed number of cuts, it satisfies a monotonicity condition. As shown in Fig. 1, it is smaller for $p(x)$ which exhibits more rapidly decaying tails—$V(U)$ satisfies the same power law for two different $p(x)$, but it is smaller in amplitude for larger α, i.e., steeper $p(x)$. This implies that for fixed δU, the density of states of the neighborhood within δU of U_0 is relatively smaller for steeper $p(x)$, reflecting relatively greater sensitivity of the optimal state.

B. Variability and effective temperatures

Consider the Laplace transform of $V(U)$ (the ''partition function''), which introduces the effective temperature $\beta=1/kT$:

$$Z_s(\beta)=\int \exp(-\beta U)V(U)dU=\int \exp(-\beta U)d^N y. \tag{6}$$

Here the subscript s refers to the fact that we only consider the spatial degrees of freedom.

This description is consistent with an ensemble of static configurations, each of which is weighted according to its value of U, or a system which evolves dynamically with time. The dynamics of the cuts could be described by a Brownian motion due to some noise, while the configuration still has a tendency towards low values of U,

$$\frac{\partial U}{\partial y_i}=-\eta_i\frac{dy_i}{dt}+F_i(t), \tag{7}$$

where $F_i(t)$ is a random force describing white noise and $-\eta_i(dy_i/dt)$ describes the dynamical friction. This is an equation for the diffusion process of the cuts in the field U. Its final state is the equilibrium state described by Eq. (6). In this case, the noise level is connected to temperature by

$$\langle F_i(t_1)F_i(t_2)\rangle=2\eta_i kT\delta(t_1-t_2). \tag{8}$$

This dynamical extension of the static forest fire model leads to Eq. (6). This expression is potentially very informative. It assigns to each configuration a statistical weight which depends on U, the minimum value of which is the target of design. When $T=0$, the configuration is limited to the optimal state, while when $T\to\infty$, U is no longer important and the configuration is totally random. If we decrease T from ∞ to 0, we realize the crossover from randomness to design.

We emphasize that Eq. (6) is independent of the derivation in terms of Brownian motion. Equation (6) leads to a concise mathematical representation of $V(U)$, just as the partition function in statistical mechanics leads to a corresponding $V(U)$. However, the form of U in Eq. (1) is different from traditional potential functions in statistical mechanics. Most importantly, it depends explicitly on spatial position x through the dependence on the hit distribution $p(x)$, which may have persistent structure which extends to large length scales. For a broad class of $p(x)$, including Cauchy, exponential, and Gaussian, minimization of the potential leads to HOT states which exhibit power-law distributions of events (see the Appendix). That is, power laws naturally arise in the ''ground state,'' and are not special features associated with criticality. Interestingly, while the positions of the cuts in the optimal are extremely sensitive to $p(x)$, the results of Sec. II A and the Appendix indicate that the statistical properties (the density of states and the exponents describing the power law) are independent of the details of $p(x)$.

C. Localization of the cuts

In confining our attention to a fixed, finite number of cuts (and then studying the limit of large N) we focus our attention on the high density limit of an underlying lattice model. However, in the typical $d=1$ percolation at the corresponding density, the cuts are randomly placed (corresponding to the broad maximum in Fig. 1), and there are no power laws in the distribution of events, even in the neighborhood of the $d=1$ critical point at $\rho=1$. After introducing fluctuations in the positions of the cuts, characterized by an effective temperature T, it is natural to ask whether there is a phase transition separating ''low-temperature'' designed states from ''high-temperature'' random ones. Although we started with a percolation model, we do not seek a percolation transition, since our fixed number of cuts prevents the system from percolating. Instead, we ask whether there is a localization transition as a function of temperature which bounds the amplitude of fluctuations of the positions of the cuts, as the number of cuts $N\to\infty$. Certainly for lower values of T the cuts are relatively less mobile on average than at higher T. However, not surprisingly, in the one-dimensional forest fire model the localization transition of the cuts is a simple, zero-temperature phase transition.

To see that the transition occurs at $T=0$, consider first the high temperature state. When $\beta=0$ $(T\to\infty)$, the above

model of a dynamical one-dimensional forest fire in a potential described by Eq. (1) reduces to a one-dimensional ideal gas, where the potential becomes irrelevant, and each microstate has the same probability. Thus the probability for the position of the ith cut being y_i is

$$P(y_i) \propto \frac{y_i^{i-1}}{(i-1)!} \frac{(L-y_i)^{N-i}}{(N-i)!}. \tag{9}$$

Thus $\langle y_i \rangle = i/N + 1L$ and

$$\langle (y_i - \langle y_i \rangle)^2 \rangle = \frac{i(N+1-i)}{N+2} \frac{L^2}{(N+1)^2}. \tag{10}$$

If we keep L/N a constant and let $N \to \infty$, then fluctuations are proportional to system size. This argument applies whenever the temperature is sufficiently high that U is not important. Clearly in this limit there is no structure.

On the other hand, for some sufficiently low temperature, fluctuations become localized. For this one-dimensional system, the fluctuations only become localized when T is strictly zero. To see this, consider a low but finite temperature. For simplicity, we consider a constant $p(x) = p_0$. For this case U [Eq. (1)] becomes

$$U(y_1, y_2, \ldots, y_N) = p_0 \sum_{i=1}^{N+1} (y_i - y_{i-1})^2. \tag{11}$$

The size of the typical fluctuation in each of the y_i is governed by when the corresponding change in U is of order kT. In analogy with arguments developed for melting transitions [6], we focus on the long-wavelength mode in which there is a linear increase in the displacement of the cuts (relative to the position which minimizes U) which starts at one end and peaks in the middle of the system, followed by a linear decrease approaching the opposite boundary. We define the displacement of the center cut to be d, so that the placement of the ith cut is $[i/(N/2)]d$ when $i < N/2$ and $[(N-i)/(N/2)]d$ when $i > N/2$. So the change in $y_{i+1} - y_i$ is $2d/N$ when $i < N/2$ and $-2d/N$ when $i > N/2$. The corresponding change in U from Eq. (11) is

$$\delta U = \frac{4p_0 d^2}{N}. \tag{12}$$

When δU is about kT, d is about $\sqrt{kTN/4p_0}$. So for any small but finite T, d goes to infinity in the thermodynamic limit.

In one dimension there are many configurations with U close to the minimum. There is no phase transition at finite temperature. Even when one y_i has a very big deviation from its mean position, the system is still able to accommodate the change. Thus no clear structure is preserved. This is true for any $p(x)$.

In higher dimensions [1–3] the HOT state corresponds to compact d-dimensional cellular regions, separated by $d-1$-dimensional boundaries. In that case, boundary fluctuations are associated with coherent modes of the $d-1$ dimensional barriers, involving of order N^{d-1} unoccupied sites in the discrete case. The melting transition will involve development of defects (holes) in the boundaries, allowing fires to spread between neighboring regions which were previously separate. This case is likely to be much more subtle than in $d=1$ where the boundaries are mobile but cannot disintegrate.

III. UNDERSPECIFIED AND CHANGING EXTERNAL ENVIRONMENTS

The coupling between the system (the line segment and cuts), and the external environment is represented by the distribution of sparks $p(x)$ which ignite fires. In the previous section we assumed $p(x)$ was known. Here we relax that assumption, and study the effects of a time dependent $p(x,t)$ on the HOT state. We derive conditions for which intrinsic variability in the state (modeled here as a finite effective temperature T) leads to a decrease in the expected loss. The key quantities which determine when noise is beneficial are the rate and amplitude of change of $p(x,t)$, the rate at which the configuration can adapt, and the sacrifice in the time average cost $\langle U \rangle$ associated with finite T fluctuations relative to the fully optimized state. Our results would apply generally to any statistical mechanics system described by a time varying potential. However, we want to point out that the trade-off between fluctuations and optimality is a key consideration for HOT.

To facilitate our analysis, we assume that changes in the environment are slow and continuous. This corresponds to the case where we have sufficient sampling to obtain a good estimate of $p(x,t)$ within a moving time window. This allows us to compute an instantaneous optimal configuration (the local equilibrium state), and model the evolution of the system as a finite-temperature relaxation towards this state. The time dependence of $p(x,t)$ results in a time dependent $U(\mathbf{y},t)$. Assuming local equilibrium, the partition function becomes

$$Z_s(t) = \int e^{-\beta U(\mathbf{y},t)} d\mathbf{y}, \tag{13}$$

and the statistical probability of a certain configuration $\mathbf{y}$ in the equilibrium state is

$$c_0(\mathbf{y},t) = \frac{e^{-\beta U(\mathbf{y},t)}}{\int e^{-\beta U(\mathbf{y},t)} d\mathbf{y}}. \tag{14}$$

The rate of change of the system is limited by the internal dynamics [e.g., Eqs. (7)], which typically prevents the system from equilibrating to the state characterized by Eq. (14). As a consequence, the system changes in response to environmental change, but never fully catches up.

We write the actual (nonequilibrium) probability of configuration $\mathbf{y}$ as $c(\mathbf{y},t)$. We assume that to a good approximation, the system evolves towards the instantaneous equilibrium distribution $c_0(\mathbf{y},t)$ with a fixed rate τ, i.e.,

$$\frac{\partial}{\partial t} c(\mathbf{y},t) = -\frac{1}{\tau} [c(\mathbf{y},t) - c_0(\mathbf{y},t)]. \tag{15}$$

Here τ is a time scale describing the speed with which the system adjusts to the environment.

We focus on the long-term behavior, when the effects of the initial state have decayed away. For large t, the solution to the above equation is

$$c(\mathbf{y},t)=\frac{e^{-t/\tau}}{\tau}\int_0^t c_0(\mathbf{y},t_1)e^{t_1/\tau}dt_1. \tag{16}$$

In order to obtain analytical results, we focus on the case in which the change in U is slow and small in amplitude. This allows us to focus on the dominant (Fourier-like) mode, which is characterized by an overall amplitude $N\epsilon$ of change and a frequency ω. We write

$$U(\mathbf{y},t)=U_b(\mathbf{y})\pm N\epsilon f(\mathbf{y}), \tag{17}$$

where $U_b(\mathbf{y})$ is stationary and captures most of $U(\mathbf{y},t)$ (b stands for ''background''), ϵ describes the magnitude of the average change per cut, which we assume to be small, and $f(\mathbf{y})$ describes the spatial dependence of the change. The equation takes the $+$ sign for the first half of the period, $2n\pi/\omega<t<(2n+1)\pi/\omega$, and the $-$ sign for the second half, $(2n+1)\pi/\omega<t<(2n+2)\pi/\omega$.

For simplicity, we consider the case that Z_s from Eq. (13) is time independent. This condition sets a constraint [obtained by plugging Eq. (17) into Eq. (13)] which is satisfied, e.g., when $U_b(\mathbf{y})$ is even and $f(\mathbf{y})$ is odd. This is typically satisfied because U_b is generally quadratic (even) to leading-order with respect to shifts in the positions of the cuts, while a linear (odd) f describes the leading-order effect of shifting weight from one side of each cut y_{i0} to the other side. However, this assumption is not essential for our basic conclusions.

Substituting into Eq. (16) for $c(\mathbf{y},t)$ we obtain

$$c(\mathbf{y},t)=e^{-\beta U_b(\mathbf{y})}\Bigg\{\cosh[\beta N\epsilon f(\mathbf{y})] \mp\left[1-\frac{2}{1+e^{-\pi/\omega\tau}}e^{-\Delta t/\tau}\right]\sinh[\beta N\epsilon f(\mathbf{y})]\Bigg\}/Z_s, \tag{18}$$

where the $-$ sign is for the first half of a period and the $+$ sign for the second half, and Δt is the time duration between t and the moment of the last change of $U(\mathbf{y},t)$.

This allows us to calculate the expectation value of U. Because it is time dependent, it is most useful to obtain the time average value over a period

$$\langle U\rangle=\frac{\omega}{2\pi}\int_0^{2\pi/\omega}dt\int U(\mathbf{y},t)c(\mathbf{y},t)d\mathbf{y} \tag{19}$$

$$=\frac{1}{Z_s}\int\cosh[\beta N\epsilon f(\mathbf{y})]U_b(\mathbf{y})e^{-\beta U_b(\mathbf{y})}d\mathbf{y}-\frac{N\epsilon}{Z_s}\left[1-\frac{2\omega\tau}{\pi}\times\tanh\left(\frac{\pi}{2\omega\tau}\right)\right]\int f(\mathbf{y})\sinh[\beta N\epsilon f(\mathbf{y})]e^{-\beta U_b(\mathbf{y})}d\mathbf{y}. \tag{20}$$

Finally, when Z_s is time independent, it can be calculated from Eq. (13),

$$Z_s=\int\cosh[\beta N\epsilon f(\mathbf{y})]e^{-\beta U_b(\mathbf{y})}d\mathbf{y}. \tag{21}$$

Equation (20) is the most general statement of our primary result—namely, $\langle U\rangle$ can be reduced by variability (finite T) in the presence of a changing environment. In order to see this more clearly, it is useful to simplify Eq. (20) further.

First, consider the case for $\epsilon=0$, i.e., no change in the environment. Then the second term in Eq. (20) vanishes and the first returns to the familiar form which for T sufficiently small is

$$\langle U_b(\mathbf{y})\rangle_{\epsilon=0}=U_0+\frac{N}{2}kT. \tag{22}$$

The factor $N/2$ is from the equipartition of energy—the N modes each contributes $kT/2$. It describes the sacrifice we make in $\langle U\rangle$ when we raise the noise level. Looking back at the expression for $V(U)$ in Eq. (5), we see the factor $N/2$ comes from the index for the power law, while the prefactor in that equation is not relevant in this calculation.

When $\epsilon>0$, the first term is modified by a term of order ϵ^2, which can be ignored for small ϵ. The second term in Eq. (20) which is linear in ϵ describes the beneficial effects of variability. To see it more clearly, we take $f(\mathbf{y})=\pm 1$. Then Eq. (20) simplifies to

$$\langle U\rangle=\langle U_b\rangle_{\epsilon=0}-N\epsilon\left[1-\frac{2\omega\tau}{\pi}\tanh\left(\frac{\pi}{2\omega\tau}\right)\right]\tanh\left(\frac{N\epsilon}{kT}\right). \tag{23}$$

The time scale τ depends on the intrinsic dynamics of the system. It is a key observation that higher temperature is associated with higher mobility of the dynamical degrees of freedom. This in turn leads to more rapid adjustment of the relative weights of the different configurations $c(\mathbf{y},t)$ to-

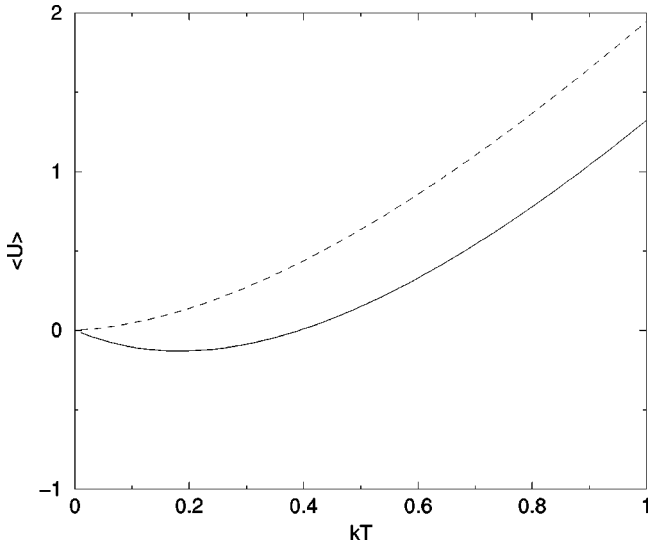

FIG. 2. The dependence of $\langle U\rangle$ on kT, calculated from Eq. (25). For the dashed line, Eq. (27) is not satisfied and $\langle U\rangle$ is a monotonically increasing function of T. In contrast, Eq. (27) is satisfied for the solid line, so that $\langle U\rangle$ exhibits a minimum for $T\neq 0$ so that fluctuations enhance performance.

wards the equilibria $c_0(\mathbf{y},t)$, which is reflected in a decreased value of τ. If there is not another time scale, we can only form τ in one way,

$$\tau = \frac{\pi}{2}\frac{B}{\sqrt{kT}}, \tag{24}$$

where the units of time arise from the square root of kT. Here B is a constant determined by the internal dynamics of the system and corresponds to the coefficient for the T-dependent time scale at which the system evolves towards equilibrium. τ is larger for larger values of B, and the system changes more slowly.

Now we have

$$\langle U\rangle = U_0 + \frac{N}{2}kT - N\epsilon\left[1 - \frac{B\omega}{\sqrt{kT}}\tanh\left(\frac{\sqrt{kT}}{B\omega}\right)\right]\tanh\left(\frac{N\epsilon}{kT}\right). \tag{25}$$

We can expand for small T and obtain

$$\langle U\rangle = U_0 + \left(\frac{N}{2} - \frac{N\epsilon}{3B^2\omega^2}\right)kT. \tag{26}$$

When we raise the noise level, there is a cost associated with fluctuations, but because the system is more flexible, we gain back some of the performance of the state in a changing environment due to increased adaptability associated with the dynamics. When the gain is greater than the cost, i.e., when

$$3B^2\omega^2 < 2\epsilon, \tag{27}$$

a certain level of noise can be beneficial. This is reflected by a minimum at finite T in the curve $\langle U(T)\rangle$ (Fig. 2). From this condition we see that when the change in the environments is too fast, the system will not be able to adjust effectively and the best solution is to keep still at the optimal state for the background U_b. Only when the change is sufficiently slow is it beneficial for the system to be mobile.

When $U(\mathbf{y},t)$ is time dependent, the finite T variability of the configurations about the HOT state leads to a sampling of some states with losses which are smaller than the average. However, it is not the intrinsic spread of the distribution of configurations at finite T which gives rise to lower values of the expected loss in Eq. (27). On average, such an effect does not lower $\langle U\rangle$, since aside from special choices of $p(x)$ finite T leads to as many configurations that raise U as lower U. Instead, it is the link between the amplitude of the fluctuations about the HOT state and the rate at which the distribution of configurations as a whole can adjust which leads to the nontrivial minimum. This link is rooted in the internal dynamics of the system which is ultimately responsible for both fluctuations about and relaxation towards the time-dependent optimal state.

IV. CONCLUSIONS

We have developed a methodology for systematically exploring neighborhoods of the HOT state, through incorporation of an effective temperature describing the mobility of the resources (cuts) in a potential described by the cost function U. Our method parallels traditional frameworks in statistical physics and dynamical systems. However, where statistical physics and dynamical systems typically focus on properties of the most probable configurations, HOT states are rare and specialized. They reflect high performance with respect to a distribution $p(x)$ of perturbations which the environment may impose on the system. Optimization of the resource allocations with respect to $p(x)$ leads to the robust yet fragile behavior of HOT systems. HOT systems outperform generic states in the situations they were designed for, but are simultaneously highly sensitive to design flaws and changes in the environment. This reflects an important tradeoff which is central to the majority of real complex systems where design and evolution play key roles.

Introducing finite-temperature fluctuations allows us to quantify an important *dynamical* tradeoff between losses (design flaws) associated with noisy fluctuations about an optimized state and gains associated with an increased ability of a system to adapt to changes in $p(x,t)$. When the mobility of resources is characterized by a single time scale, which can be related to the effective temperature T, we derive an inequality which separates cases in which noise improves the performance of the system from cases in which the performance is degraded. Our calculation assumes changes in the environment are relatively slow and small, so that a local (in time) optimal configuration can be defined. We compute the average cost $\langle U\rangle$ associated with the dominant frequency variations in the environment, assuming the relaxation of the system towards the local optimum can be described by a single time scale τ.

Three cases are of particular interest. First, if the dominant mode is a steady shift in demand (e.g., for transportation or communication systems), it corresponds to an infinite period characterizing the change in U. In this case by Eq. (27) it is always in the best interest of a system to incorporate some finite-temperature mobility and adapt.

Second, in some cases there is a semiregular pattern to the fluctuations associated with some nonzero dominant frequency of change. Examples include (noisy) weather and climate fluctuations which are coupled to ecological systems, fluctuations in spending and investment cycles which are coupled to a market economy, or daily variability in user demand which are coupled to transportation and communication systems as well as public utilities. In this case, because the patterns are somewhat predictable, time-dependent optimization can improve the performance by directly adjusting output and the allocation of resources to incorporate this variability. For example, many metropolitan areas adjust traffic patterns to account for the relative increases in incoming and/or outgoing traffic during morning and/or evening commutes. This scenario goes a step beyond the calculations performed here by anticipating the reallocation of resources in response to the changing environment. In this case there need not be a direct link between relaxation towards the optimal state and fluctuations, so Eq. (27) does not apply.

Finally, the most interesting case arises when the dominant mode is not *a priori* known. In some manmade technologies (e.g., algorithms that are used to optimize world wide web sites for maximum throughput), local optimization of the resource allocations is decoupled from random fluc-

tuations in the distribution of resourses, so Eq. (24) does not apply. However, in other cases it is the low level degrees of freedom, rather than an externally imposed optimization scheme, which are both constantly fluctuating and adapting to change in a manner which on average favors higher performance. This process is easiest to visualize in the context biology and ecology, where mutation leads to the next round of candidates subject to optimization through natural selection in a time-dependent environment. Our model represents a highly simplified description of this scenario. Equation (27) can be viewed as a condition on the frequency ω associated with change at which it becomes beneficial for the system to adapt to time-dependent conditions. For frequencies greater than $\omega=\sqrt{2\epsilon/3B^2}$ a system is better off optimizing for the average environment than it is making time-dependent adjustments. In any real complex, adaptive system, different frequency modes of a changing environment will be important at different levels in a hierarchical design.

ACKNOWLEDGMENTS

We wish to thank Carl Robert and John Doyle for many useful discussions. This work was supported by the David and Lucille Packard Foundation, NSF Grant No. DMR-9813752, and EPRI/DoD through the Program on Interactive Complex Networks.

APPENDIX

In this Appendix we solve the optimization problem for the one-dimensional forest fire model for arbitrary probability distribution function $p(x)$ of the initial spark. The forest is confined in the segment $0<x<1$. There are N cuts (fire breaks) which are located at $y_1, y_2, \ldots, y_N$. We assume N is big so that the intervals between neighboring cuts are small and $p(x)$ varies slowly in each interval.

The loss function to minimize is

$$U(y_1,y_2,\ldots,y_N)=\sum_{i=1}^{N+1}(y_i-y_{i-1})\int_{y_i-1}^{y_i}p(x)dx, \tag{A1}$$

where the sum $y_i=0$ for $i=0$ and $y_i=1$ for $i=N+1$. For the optimal positions of the y_i's, we have

$$\frac{\partial U}{\partial y_i}=0, \tag{A2}$$

which leads to

$$\int_{y_i}^{y_{i+1}}p(x)dx-\int_{y_{i-1}}^{y_i}p(x)dx=(2y_i-y_{i-1}-y_{i+1})p(y_i). \tag{A3}$$

In the above equation, because $p(x)$ is slowly varying in an interval, we can approximate it by keeping to the first-order terms in the Taylor expansion

$$p(x)=p(y_i)+\left(\frac{dp}{dx}\right)_{x=y_i}(x-y_i). \tag{A4}$$

Plug this into Eq. (A3) and we have

$$[(y_{i+1}-y_i)^2+(y_i-y_{i-1})^2]\frac{dp}{dy_i}=4(2y_i-y_{i-1}-y_{i+1})p(y_i). \tag{A5}$$

Let us view y_i as the value of some continuous function $y(s)$ at $s=i$. Then remembering the interval between y_i and y_{i+1} is small, i.e., $y(s)$ is a slow varying function of s, we can approximate

$$y_{i+1}-y_i=\left(\frac{dy(s)}{ds}\right)_{s=i} \tag{A6}$$

and

$$y_{i-1}+y_{i+1}-2y_i=\left(\frac{d^2y(s)}{ds^2}\right)_{s=i}. \tag{A7}$$

Then Eq. (A5) becomes

$$\frac{dp}{dy}\left(\frac{dy}{ds}\right)^2=-2p\frac{d^2y}{ds^2} \tag{A8}$$

so

$$\frac{d\ln p}{dy}=2\frac{d}{ds}\frac{1}{\frac{dy}{ds}} \tag{A9}$$

and

$$d\ln p=2\frac{dy}{ds}d\frac{1}{\frac{dy}{ds}}. \tag{A10}$$

So finally,

$$\frac{dy}{ds}=\frac{A}{\sqrt{p}}, \tag{A11}$$

where A is a constant.

We can proceed to solve function $y(s)$ and find the optimal positions for y_i's. But instead, let us look at the relation between the size of loss and its corresponding probability. The size of loss here is the length of an interval,

$$y_{i+1}-y_i=\left(\frac{dy(s)}{ds}\right)_{s=i}=\frac{A}{\sqrt{p}}. \tag{A12}$$

And the probability of this loss is

$$\int_{y_i}^{y_{i+1}}p(x)dx=(y_{i+1}-y_i)p=A\sqrt{p}. \tag{A13}$$

From the above two equations we see that $P(l)\propto l^{-1}$.

This power law is for arbitrary $p(x)$. Though the derivation depends on the assumption that N is big enough so that the intervals between cuts are small, this power law should still be a good approximation even for the situations when the assumption is not satisfied.

[1] J. M. Carlson and J. Doyle, Phys. Rev. E **60**, 1412 (1999).
[2] J. M. Carlson and J. Doyle, Phys. Rev. Lett. **84**, 2529 (2000).
[3] J. Doyle and J. M. Carlson, Phys. Rev. Lett. **84**, 5656 (2000).
[4] P. Bak, K. Chen, and C. Tang, Phys. Lett. A **147**, 290 (1990); B. Drossel and F. Schwabl, Phys. Rev. Lett. **69**, 1629 (1992).
[5] D. Stauffer, *Introduction to Percolation Theory* (Taylor, London, 1985).
[6] K. J. Strandburg, Rev. Mod. Phys. **60**, 161 (1988).