Surfing on (or drowning under?) the threshold

Memories from the ergodic phase: the awkward dynamics of spherical mixed p-spin models
Authors: G. Folena, S. Franz, F. Ricci-Tersenghi
arXiv:1903.01421

Recommended with a Commentary by Francesco Zamponi,
Laboratoire de Physique de l'Ecole Normale Supérieure, Paris

Consider the following mathematical problem. Given is a set of N real variables $\mathbf{w} \in \mathbb{R}^N$, and a Hamiltonian (energy) function $H[\mathbf{w}]$. The variables are initialized in equilibrium at temperature T_0, with probability $P_0[\mathbf{w}_0] \propto \exp\{-H[\mathbf{w}_0]/T_0\}$. The energy is then minimized by gradient descent,

$$\frac{\mathrm{d}w_i}{\mathrm{d}t} = -\frac{\partial H}{\partial w_i} \ , \qquad \mathbf{w}(t=0) = \mathbf{w}_0 \ , \qquad (1)$$

corresponding to an instantaneous "quench" to zero temperature. If the energy function is convex, e.g. quadratic, then the gradient descent always asymptotically converges to its absolute minimum, the ground state of the system. If, however, the energy function is "rough", i.e. it has a lot of local minima

Figure 1: An example of a rough energy function $H[\mathbf{w}]$, for $N = 2$.

and saddle points (Fig. 1), then (i) the gradient descent might remain trapped for very long time in the vicinity of saddles, where the gradient is very small, before eventually reaching a minimum, and (ii) a different minimum will be reached depending on the initial condition – hence, in general, the final minimum will not be the absolute one.

Rough energy functions are far from being exotic objects: they are routinely encountered in many important fundamental and practical problems. Some examples are: optimization problems such as the traveling salesman problem, i.e. the problem of finding the minimum path connecting a set of cities, where H is the distance traveled along a given path $\mathbf{w}$ [1]; glassy materials such as silicates, where H is the potential energy due to pair interactions of atoms in positions $\mathbf{w}$ [2]; a population of bacteria trying to maximize their fitness, where H is the negative of fitness and $\mathbf{w}$ is a coordinate in genotypic space [3]; and, of course, training a neural network to perform some task such as classification, where $\mathbf{w}$ are the neuronal weights and $H[\mathbf{w}]$ is the cost of misclassifying the training data [4].

Understanding the properties of the gradient descent dynamics is very important in all these applications. In particular, one would like to characterize the distribution $P_\infty[\mathbf{w}_\infty]$ of
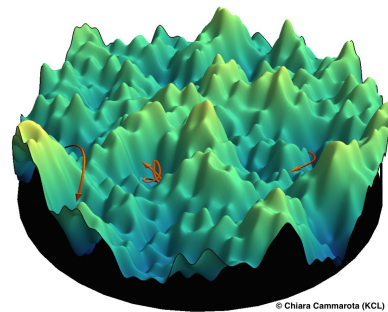
the final states reached at $t \to \infty$, as a function of the initial temperature T_0. Unfortunately, this requires solving the dynamics of the problem, which is often a very difficult technical task. One of the holy grails in the domain has been characterizing $P_\infty[\mathbf{w}_\infty]$, at least partially, *without* solving the dynamics explicitly, i.e. using information on $H[\mathbf{w}]$ only. For example, can one compute the average energy of the final states? And given that the gradient vanishes in the final state, does the next order in the expansion, i.e. the Hessian matrix $\frac{\partial^2 H}{\partial w_i \partial w_j}$, have some particular property?

A milestone result in this field has been the solution of the so-called *spherical pure p-spin model*, which is one of the simplest models of rough high-dimensional energy function [5]. For $p = 3$, the energy function can be written as a sum of $p = 3$-variable interactions

$$H[\mathbf{w}] = \sum_{ijk} J_{ijk} w_i w_j w_k + \mu \sum_i w_i^2 , \qquad (2)$$

where the couplings J_{ijk} are independent Gaussian variables with zero mean and variance $1/N^2$, and μ is a Lagrange multiplier used to keep $\mathbf{w}$ confined via the condition $\sum_i w_i^2 = N$.

Any choice of $p \geq 3$ leads to a very rough energy function, with a number of stationary points growing exponentially in N. The energy function of the p-spin model then displays a sharp topological transition when $N \to \infty$ [6, 7]. There exists a *threshold energy* e_{th}, such that, with probability one when $N \to \infty$, stationary points (i.e. zero-gradient points) of H are minima for $N e_{\text{GS}} \leq H[\mathbf{w}] < N e_{\text{th}}$, while they are saddles for $H[\mathbf{w}] > N e_{\text{th}}$. The argument to show this is pretty simple:

Figure 2: In the pure p-spin energy function, saddles (minima) dominate above (below) the threshold. States prepared at high T_0 loose memory of the initial condition and end up surfing on the threshold (blue). States prepared at low T_0 end up in a nearby minimum (red).

1. It can be shown that the Hessian matrix of Eq. (2), in a stationary point, is a real symmetric Gaussian matrix, with the diagonal shifted by the Lagrange multiplier μ. Its eigenvalue density $\rho(\lambda)$ is a Wigner semicircle [8] supported in $\lambda \in [\lambda_-, \lambda_+]$, with $\lambda_\pm = \mu \pm 2\sqrt{p(p-1)}$.

2. Furthermore, the Lagrange multiplier of Eq. (2) is related to the energy by $\mu = -pH/N = -pe$.

It follows that $\lambda_- = -pe - 2\sqrt{p(p-1)}$ is the lower edge of the spectrum. If $e < e_{\text{th}} = \sqrt{2(p-1)/p}$, then $\lambda_- > 0$ and all eigenvalues are positive: the stationary point is a minimum. If $e > e_{\text{th}}$, some eigenvalues are negative: it is a saddle. States with $e = e_{\text{th}}$ are *marginally stable*: they have flat directions associated to zero eigenvalues of the Hessian.

Furthermore, the dynamics of the spherical p-spin can be solved exactly. One can establish a set of *dynamical mean field equations* that describe the time evolution of the energy and of correlation functions during the gradient descent dynamics. While the equations cannot
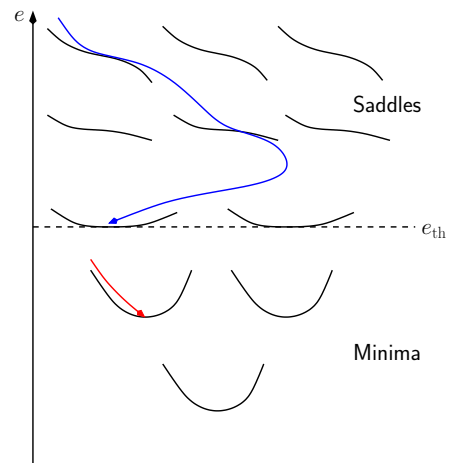
be solved analytically, a long-time asymptotic solution has been explicitly constructed [9]. This solution shows that, if the initial state is prepared at high enough temperature $T_0 > T_d$, then the energy asymptotically reaches the threshold with a non-trivial power-law behavior, $e(t) \sim e_{th} + At^{-2/3}$, irrespectively of the initial state; indeed, memory of the initial state is completely lost during the dynamics. The fact that the dynamics keeps "surfing" on the threshold, wandering around marginally stable states, is a very remarkable outcome of the asymptotic solution [9]. On the contrary, if the initial state is deep enough, $T_0 < T_d$, then it is already confined in the basin of attraction of a close-by minimum, to which the dynamics converges exponentially, $e(t) \sim e_\infty(T_0) + Be^{-t/\tau}$. The final energy $e_\infty(T_0)$ depends on the initial state and can be calculated by a simple thermodynamic construction [10]. Since its discovery, this scenario has become a paradigm for disordered systems with rough energy functions (Fig. 2).

The work of Folena et al. adds a new important ingredient to this paradigm. Inspired by previous work that detected some anomalies in its energy function [11], they investigated a small variation of Eq. (2), a *mixed p-spin model*, in which the Hamiltonian contains both 3-spin and 4-spin interactions. With this simple modification, the relation $\mu = -pe$ does not hold anymore, and the Lagrange multiplier is now independent of the energy. The Hessian still has a Wigner semicircle density, but the lower edge λ_- is now independent of the energy. Hence, marginally stable states can now, in principle, exist at any level of energy.

Figure 3: In the mixed p-spin energy function, saddles (minima) dominate above (below) the threshold, but rare saddles also exist below e_{th}, giving rise to a new, intermediate dynamical regime.

Next, Folena et al. solved numerically the dynamical mean field equation of the mixed p-spin model and found three distinct dynamical regimes. (i) For high $T_0 > T_{onset}$, the dynamics is memoryless and reaches a threshold level e_{th}, which is again a topological property of the energy function: at $e = e_{th}$, most of the stationary points are marginally stable, with $\lambda_- = 0$. This memoryless dynamics is described by the same asymptotic solution as for the pure model [9]. (ii) For $T_0 < T_{SF}$, the dynamics quickly reaches a minimum close enough to the initial state, as in the low-temperature behavior of the pure model [10]. (iii) A completely new regime is found for intermediate $T_0 \in [T_{SF}, T_{onset}]$: the dynamics asymptotically still wanders around marginally stable states, but at energy $e_\infty(T_0) < e_{th}$ where such marginal states are exponentially rare. Furthermore, memory of the initial condition is not lost, as manifested by the T_0-dependence of the final energy. While, as in the pure model, the first two regimes can be described solely in terms of properties of the energy function $H[\mathbf{w}]$, without any reference to the details of the dynamical evolution, the third regime is new, qualitatively different and poorly understood (*hic sunt leones*, in the authors' words). The final energy, and the properties of the final states, depend on the details of the dynamics, and for the moment no simple asymptotic solution has been found to describe the long time behavior.
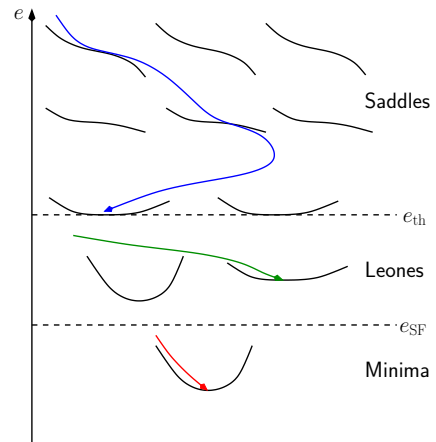
3

The new dynamical regime is certainly relevant for the study of glasses. In these materials, the existence of a regime in which the final energy depends on the initial temperature is well established numerically [2], but no microscopic theory was able to describe it. A more systematic comparison between mean field results and numerical simulations of realistic glass models will certainly shed additional light on the properties of glassy dynamics in this new regime. This new dynamical regime could also play an important role in many situations, from optimization to evolution and machine learning.

References

[1] M. Mézard, G. Parisi, M. Virasoro, *Spin glass theory and beyond*, World Scientific (1987).

[2] A. Heuer, *Exploring the potential energy landscape of glass-forming systems: from inherent structures via metabasins to macroscopic transport*, Journal of Physics: Condensed Matter **20**, 373101 (2008).

[3] H. Flyvbjerg, B. Lautrup, *Evolution in a rugged fitness landscape*, Physical Review A **46**, 6714 (1992).

[4] Y. LeCun, Y. Bengio, G. E. Hinton, *Deep learning*, Nature **521**, 436 (2015).

[5] T. Castellani, A. Cavagna, *Spin-glass theory for pedestrians*, Journal of Statistical Mechanics: Theory and Experiment **2005**, P05012.

[6] J. Kurchan, G. Parisi, M. Virasoro, *Barriers and metastable states as saddle points in the replica approach*, Journal de Physique I **3**, 1819 (1993).

[7] J. Kurchan, L. Laloux, *Phase space geometry and slow dynamics*, Journal of Physics A: Mathematical and General **29**, 1929 (1996).

[8] M. L. Mehta, *Random matrices*, Elsevier (2004).

[9] L. F. Cugliandolo, J. Kurchan, *Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model*, Physical Review Letters **71**, 173 (1993).

[10] S. Franz, G. Parisi, *Recipes for metastable states in spin glasses*, Journal de Physique I **5**, 1401 (1995).

[11] Y. Sun, A. Crisanti, F. Krzakala, L. Leuzzi, L. Zdeborová, *Following states in temperature in the spherical $s + p$-spin glass model*, Journal of Statistical Mechanics: Theory and Experiment, **2012**, P07002.