

Optimality and annealing path planning of dynamical analog solvers

Shu Zhou¹, K. Y. Michael Wong^{1,*}, Juntao Wang¹, David Shui Wing Hui², Daniel Ebler², and Jie Sun²

¹Department of Physics, The Hong Kong University of Science and Technology, Hong Kong SAR, China

²Theory Lab, Central Research Institute, 2012 Labs, Huawei Technologies Co. Ltd., Hong Kong SAR, China



(Received 17 March 2026; accepted 18 August 2026; published 3 September 2026)

Recently proposed analog solvers based on dynamical systems, such as Ising machines, are promising platforms for large-scale combinatorial optimization. Yet, given the heuristic nature of the field, there is very limited insight on optimality guarantees of the solvers, as well as how parameter schedules shape dynamics and outcomes. Here, we develop a dynamical mean-field framework to analyze Ising-machine dynamics for finding the ground-state energy of the Sherrington-Kirkpatrick (SK) model of spin glasses and identify mechanisms that enable rapid convergence to provably near-optimal energies. For a fixed target energy density E_c , we show that solutions are typically reached within $O(1)$ matrix vector multiplications, indicating constant time complexity. We further delineate theoretical limitations arising from different parameter-scheduling trajectories and demonstrate a pronounced benefit of temperature-only annealing for the coherent Ising machine. Building on these insights, we propose a general framework for designing optimized parameter schedules, thereby improving the practical effectiveness of Ising machines for complex optimization tasks. The superior performance of the dynamical solvers is illustrated by the attainment of the ground-state energy of the SK model.

DOI: [10.1103/m13m-ys11](https://doi.org/10.1103/m13m-ys11)

I. INTRODUCTION

Discrete optimization problems that can be formulated as Ising Hamiltonians have found broad applicability across various real-world domains, including information processing [1], route planning [2], protein folding [3], and machine learning [4]. The problem objective can be written as

$$\min_{\sigma \in \{-1, +1\}^N} \left(-\frac{1}{2} \sum_{ij} \sigma_i J_{ij} \sigma_j \right), \quad (1)$$

where the optimization variables are Ising spins $\sigma \in \{\pm 1\}^N$, which are coupled through the problem matrix $[J_{ij}]_{N \times N}$. The search for the spin configurations that minimize this Hamiltonian is widely regarded as computationally intractable for classical computers [5,6].

Recently, with the development of new computational paradigms such as quantum and quantum-inspired computing, there is an upsurge of interest in non-von-Neumann computing schemes to solve discrete optimization problems efficiently. Concretely, such schemes simulate the dynamics of classical or quantum variables, referred to as (analog) spins, and map them to their Ising counterparts via nonlinear activation functions. The Ising energy after mapping is referred to as the decoded energy. Various types of such devices

have emerged, leveraging different physical spin hardware [7]. Examples include quantum annealers based on superconducting circuits [8–11], digital systems utilizing special-purpose classical processors [12–17], and optical Ising machines employing pulse lasers [18–20]. Among these, the coherent Ising machine (CIM) stands out as a notable example, demonstrating superior scalability and efficiency [21,22]. Numerous CIM variations [7,23–26] and theoretical studies [27–29] have recently been proposed. Yet, despite advances in understanding steady states, the design of dynamical algorithms remains largely intuitive. The annealing schedule, crucial for guiding the system from high to low entropy via gain parameter adjustment [20], is often obstructed by metastable states [28] and gaps in spin distributions [29] that restrict necessary spin flips. To circumvent these convergence issues, a technique known as chaotic amplitude control (CAC) [27] was introduced, utilizing sophisticated parameter tuning and specially designed dynamics, and has shown promise in numerical scaling experiments [17]. Nevertheless, the systematic scaling of Ising-machine precision with network size remains unknown from both theoretical and experimental perspectives. This highlights the necessity for a systematic dynamical framework to evaluate performance—specifically, to determine how solution quality correlates with the machine’s dynamical evolution.

In this paper, we derive annealing schedules that are optimal within the dynamics of a given machine. Although previous studies have demonstrated that noise can be beneficial in various settings [30–33], the most widely adopted annealing protocol still varies only the linear gain a while maintaining weak noise throughout the evolution [7,20,27,28]. Our analysis shows that, for typical platforms such as CIMs, such gain-only schedules are generally

*Contact author: phkywong@ust.hk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

suboptimal and, in particular, do not lead to the fixed-accuracy polynomial-time scaling as revealed in recent algorithmic developments [34–36]. This highlights the importance of coordinated parameter scheduling, most notably involving temperature, for achieving robust and scalable optimization.

A key insight is the central role played by the density of near-zero soft spins during the dynamics. To quantify this effect, we introduce the concept of an effective gap, which connects the dynamical mean-field theory (DMFT) description to experimentally accessible dynamical activity and provides a practical criterion for schedule design. For CIMs, our analysis identifies a superior schedule that anneals only the noise, in contrast to common gain-only practice [7,20,27,28]. We further extend the framework to broader Ising-machine architectures and identify cases in which the theory instead recommends the gain-only practice during operation.

Our study of how the dynamics of Ising machines determines the accuracy of the output in optimization tasks is based on the search for ground-state energy of the Sherrington-Kirkpatrick (SK) model [37,38] of spin glasses. As a prototype of an optimization problem, the complex energy landscape of the SK model is typical of general combinatorial optimization problems [39–41], and the numerical value of its ground-state energy is well known for benchmarking [39–41]. Using the dynamical mean-field theory developed by Sompolinsky and Zippelius [42,43], we model the CIM using a soft-spin Ginzburg-Landau mean-field dynamics. This yields time-dependent expressions for the decoded energy across several Ising-machine protocols. Using these expressions, together with Monte Carlo and numerical experiments, we find that the decoded energy reaches the target value in an average of $O(1)$ steps of matrix vector multiplication, each with $O(N^2)$ operations. Furthermore, we found that the required precision of the CIM dynamics at finite time scales with the system size in the same manner as the finite-size scaling of the ground-state energy.

The significance of this result is highlighted by considering recent algorithmic developments, in which message-passing methods produce solutions that are arbitrarily close to the ground state with time complexity $C(\epsilon)N^2$, where ϵ is the fractional discrepancy with the ground state [34–36]. This demonstrates that despite the NP-hardness of Ising optimization in the worst case, near-optimal solutions are attainable in polynomial time for mean-field models. These schemes rely on oracle access to the Parisi functional minimizer. Although this minimizer has been extensively studied within the context of the SK model, such a dependency may limit the applicability of these methods to more general practical scenarios [44,45]. In this context, our results successfully capture the desired N -scaling behavior in physical Ising machines, thereby translating this theoretical advantage into experimentally relevant dynamics.

The remainder of this paper is organized as follows: In Sec. II, we introduce our methodology and derive the dynamic mean-field expression for the decoded Ising energy. Section III presents an analysis of the system’s time complexity, raising the fundamental question on the accessibility of a target energy for a given Ising machine. To address this, Secs. IV and V unveil the soft-hard spin dichotomy mechanism and

identify the dynamical barrier to evolution, which we define as the “effective gap.” Based on these findings, we propose an optimal annealing path through temperature annealing for CIMs that outperforms conventional approaches via gain annealing. In Sec. VI, we demonstrate that this framework can yield different optimal annealing paths for different systems and illustrate the phenomenon of barrier exposure. Finally, in Sec. VIII, utilizing the techniques developed in previous sections, we answer the question raised in Sec. III through asymptotic numerical analysis, providing evidence that Ising machines can access the ground state with an error margin $\epsilon \ll 1$ within polynomial time.

II. DYNAMICAL MEAN-FIELD THEORY

We first consider the example of CIM, which is governed by the following dynamical equation:

$$\frac{dx_i}{dt} = -x_i^3 + ax_i + \xi \sum_{j=1, j \neq i}^N J_{ij}x_j + h_i + \zeta_i, \quad (2)$$

where x_i represents the state of the i th spin, a is a gain coefficient, and ξ denotes the coupling strength [20]. Unless otherwise stated, we set $\xi = 1$ by absorbing it into the definition of the gain parameter. The noise term ζ_i is modeled as white noise characterized by a temperature T , satisfying $\langle \zeta_i(t) \rangle = 0$ and $\langle \zeta_i(t)\zeta_j(t') \rangle = 2T\delta_{ij}\delta(t - t')$. Additionally, we focus on Ising problems defined by the SK model, where the couplings J_{ij} are drawn from a Gaussian distribution with zero mean and variance J^2/N [37,38]. h_i is an external field set to be 0.

Such a system shares the mathematical structure of the time-dependent soft-spin Ginzburg-Landau model, analyzed by Sompolinsky and Zippelius to characterize the spin-glass phase of the SK model in equilibrium [42,43]. Referred to as the DMFT, this methodology traces the evolution of a generating functional to extract the dynamical variables of the system [46–48]. It was later applied to the dynamics of models such as the p -spin glass [49].

The Ising configuration is read out from the signs of the analog spins $s_i = \text{sgn}(x_i)$. The associated Ising energy is then $H_N := -\frac{1}{2N} \sum_{ij} \text{sgn}(x_i)J_{ij}\text{sgn}(x_j)$. We refer to this quantity as the decoded energy. Extending this framework, we incorporate a dummy source field into the generating functional; this field is conjugate to the decoded energy and serves to track its dynamical evolution.

Using DMFT [46–48] (see Appendix A for details), the mean-field effective expression for the CIM is obtained in the spirit of Refs. [42,43]:

$$\partial_t x(t) = -x^3(t) + ax(t) + J^2 \xi^2 \int_0^t dt' G(t, t') x(t') + \phi(t). \quad (3)$$

Here, ϕ represents a colored noise with $\langle \phi_i(t)\phi_i(t') \rangle = 2T\delta(t - t') + J^2 \xi^2 2Q(t, t')$. The term $Q(t, t') := \langle x(t)x(t') \rangle$ is the correlation function, which characterizes the mean similarity between the soft-spin states at times t and t' . Meanwhile, $G(t, t') := \langle \partial x(t)/\partial h(t') \rangle$ is the response function, describing how the system at time t responds to an infinitesimal perturbation applied at time t' . Due to causality, $G(t, t') = 0$ when

$t' > t$. In comparison with Eq. (2), this mapping replaces spatial couplings among soft spins with temporal correlations to their history. This construction extends naturally to other Ising machines such as those in Eqs. (8) and (F1). For instance, adding a preprocessing clipping function $\varphi(x) = \max(-1, \min(1, x))$ to Eq. (3) yields clipCIM [23], while removing the cubic term leads to simCIM [25,29] [i.e., Eq. (8)].

We also obtain a closed DMFT expression for the decoded energy

$$E_{\text{dec}}(t) = - \int_0^t \left\langle \frac{\partial \text{sgn}x(t)}{\partial h(t')} \right\rangle \langle \text{sgn}x(t)x(t') \rangle dt'. \quad (4)$$

This expression highlights the interplay between response and memory in the soft-spin dynamics. The decoded energy is evaluated from the final Ising variable $\text{sgn}x(t)$, whereas the response kernel describes its sensitivity to an infinitesimal field $h(t')$ applied to the earlier continuous trajectory. The factor $\langle \text{sgn}x(t)x(t') \rangle$ is therefore a mixed memory correlator between the final decoded spin and the earlier soft-spin amplitude, not a decoded-spin autocorrelation.

III. TIME DEPENDENCE OF THE DECODED ENERGY

We benchmark the theory by Monte Carlo simulations of both Eqs. (2) and (3). As shown in Figs. 1(a) and 1(b), the predictions for the diagonal of the correlation matrix Q and the early-time decoded energy are in excellent agreement. We note that the response function $\partial x(t)/\partial h(t')$ of individual nodes in CIM exhibit large variance in Fig. 1. Nevertheless, the Green's function $G(t, t')$ is well behaved after averaging $\partial x(t)/\partial h(t')$. The term $\langle \partial \text{sgn}x(t)/\partial h(t') \rangle$ in Eq. (4) has the same behavior. Hence, the decoded energy can be obtained directly from Eq. (4). As shown in Eq. (4) in the large- N limit, both the decay of the expected decoded energy toward its target and the expected time to reach it are independent of system size [see the blue line in Fig. 1(b)] for a fixed parameter set; the expected hitting time for a target decoded energy E_{dec} is also size independent, i.e., $t(E_{\text{dec}})_N \equiv t(E_{\text{dec}})$. In our numerical experiments, we confirmed that convergence of the dynamics can be achieved in $O(1)$ iterations. Each iteration involves matrix vector multiplication with $O(N^2)$ operations. Thus, the total number of operations required is $t(E_{\text{dec}})N^2$ [50].

Even with a size-independent hitting time, the accessibility of a target decoded energy E_{dec} remains nontrivial. In particular, a practical issue is to determine the accessible target decoded energy window that an Ising machine can attain within a limited number of iterations. From Fig. 1(b), for both the theoretical and numerical results obtained with fixed $a = 0$, $T = 10^{-5}$, $N = 2000$, a step size of 0.02, and 500 time steps, the lowest decoded energy reached remains well above the mean-field ground-state energy of the SK model. A pragmatic approach is to benchmark by sampling a sufficient number of long-time runs, as we will show later. Below, we first provide theoretical insights that reveal fundamental mechanisms and performance barriers of Ising machines.

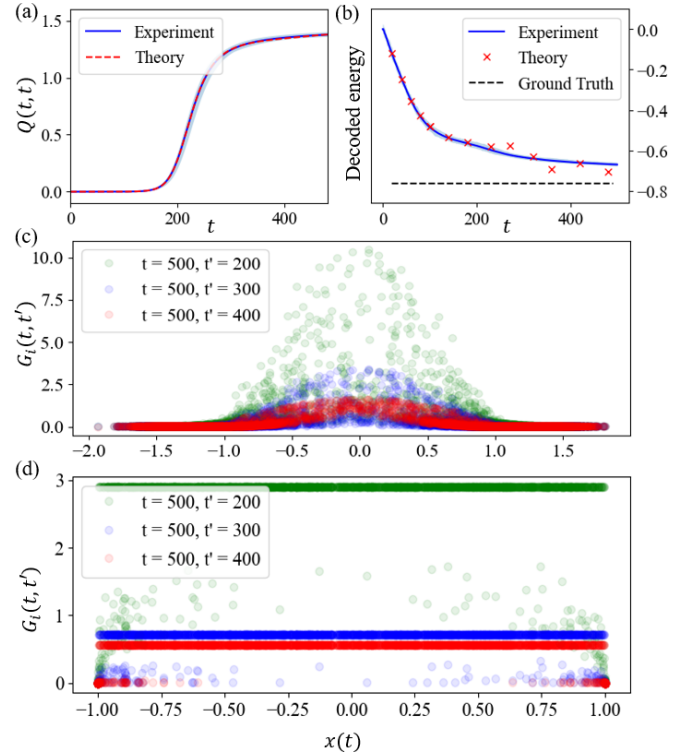


FIG. 1. (a) Comparison of the diagonal terms of the correlation function Q and the variance of soft spins measured in CIM experiments. (b) Comparison of the decoded energy obtained from Eq. (4) with experimental results from the CIM. The dashed line represents the ground-state energy of the SK model in the thermodynamic limit. (c) Dependence of the response function $\partial x(t)/\partial h(t')$ on the value of soft spins $x(t)$. (d) SimCIM's dependence of the response function $\partial x(t)/\partial h(t')$ on the value of the soft spins $x(t)$. Simulations from panels (a) to (c) were executed using the Euler-Maruyama method with a step size 0.02, $a = 0$, $T = 10^{-5}$, and $N = 2000$, while in panel (d), we use $a = -0.5$.

IV. SOFT AND HARD SPINS

An important element in the dynamics of the mean-field model is the history-dependent response function (or so called Onsager reaction term [51]) for each spin x_i , defined as the Green's function $G_i(t, t') := \partial x_i(t)/\partial h_i(t')$. In Fig. 1(c) [52], the heterogeneity of the responses increases with time. Figure 1(d) shows that the situation in simCIM (to be presented in a later section) is similar, except that bands of response functions appear due to the presence of bounds.

More significantly, Fig. 1(c) reveals a soft-hard dichotomy in the spin dynamics. We define a spin as soft when

$$|x_i(t)| < x^*(t),$$

and as hard when

$$|x_i(t)| \geq x^*(t).$$

For simCIM, $x_t^* \equiv 1$ naturally separates the near-zero soft-spin population from the large-amplitude hard-spin population. For CIM, where the amplitude distribution is broader and the separation is less sharply tied to a fixed amplitude scale, we define x_t^* through the contribution of large-amplitude spins to the response function. Specifically, x_t^* is related to a small

prescribed tolerance ϵ by

$$\int_0^t \int_{|x(t)| \geq x_t^*} \frac{\partial x(t)}{\partial h(t')} P(x(t), t) dx(t) dt' \leq \epsilon. \quad (5)$$

In this sense, hard spins are those whose contribution to the response is negligible, while soft spins carry most of the dynamical susceptibility.

With this definition, small-amplitude spins remain close to the origin, exhibit large response, and are therefore susceptible to perturbations and zero-crossing events. Large-amplitude spins, in contrast, have much weaker response and are unlikely to change sign, behaving as dynamically hardened variables.

This distinction is consistent with and complementary to the previous empirical “swing/trapped” observation of Ref. [53]. In that work, swing and trapped nodes were numerically observed from the bifurcation dynamics of finite-size continuous Ising-machine trajectories: Swing nodes remain near the origin and may oscillate between signs until later bifurcations determine them, whereas trapped nodes have already acquired large amplitudes and stable signs. Here, the soft-hard classification is formulated within the DMFT framework in the thermodynamic limit and is characterized by the amplitude distribution, response function, and zero-crossing activity. It therefore provides a mean-field dynamical explanation of the same phenomenology, while showing that the separation emerges self-consistently from the collective dynamics even at fixed gain.

These observations lead us to interpret the dynamical process as follows: Spins with large amplitudes will first stabilize, reducing the dynamics to a subspace spanned by the smaller-amplitude (soft) spins. These soft spins can then evolve rapidly and in parallel, consistent with the elevated response values observed in Fig. 1. Notably, this dichotomy can be even more pronounced in Ising machines that employ clipping mechanism $\varphi(x)$: Once hard spins hit the boundary, they become effectively responseless and make zero contributions to the Onsager reaction term [see Fig. 1(d) and Eq. (9)].

V. GAP OPENING

In the above discussion, we argued that the final approach to the target decoded energy is dominated by the dynamics of soft spins within the reduced subspace. In the extreme limit where soft spins vanish altogether, the system freezes and the decoded energy stops evolving. At equilibrium, we refer to this phenomenon as “gap opening” as discussed in our previous work [29] or “hard phase” according to Ref. [28]. Dynamically, this behavior can be further understood via Eq. (4). As derived in Appendix B, one step zero-crossing rate $\mu(t)$ can be approximately expressed as

$$\sqrt{\Delta t} \mu(t) = \frac{\sqrt{T}}{\sqrt{\pi}} P(0). \quad (6)$$

The scaling of $\sqrt{\Delta t}$ appears because, in the presence of a white-noise component, the measured zero-crossing rate at finite time resolution scales as $\mu(t) \propto 1/\sqrt{\Delta t}$. Multiplying by $\sqrt{\Delta t}$ removes this trivial time-resolution dependence. Note that the right-hand side solely depends on the intrinsic characterization of the model [such as T and $P(0)$], and dependence

on simulation parameters (such as Δt) is decoupled. Thus, $\sqrt{\Delta t} \mu(t)$ becomes a more intrinsic quantity for comparing contour levels. Hereafter, it will be referred to as the effective gap.

When a gap opens in the distribution at a specific instant t' , the spin density at $x(t') = 0$ vanishes, rendering further spin flips impossible. Consequently, the response term $\langle \partial \text{sgn} x(t) / \partial h(t') \rangle = 2 \langle \delta(x(t)) \partial x(t) / \partial h(t') \rangle$ also vanishes for all subsequent times $t > t'$, as the trajectory $x(t)$ no longer contributes to the Dirac delta distribution at 0. Under these conditions, the decoded energy ceases to decrease. Therefore, maintaining a finite density of near-zero spins throughout the evolution is crucial to avoid dynamical stagnation and ensure continuous energy reduction.

Equilibrium analysis shows that, in the thermodynamic limit, a true gap (vanishing zero-crossing rate) opens only in the noiseless limit, $\beta \rightarrow \infty$. In practice, however, CIM dynamics always operate with a small but nonzero noise level, either via deliberate injection or imposed by hardware constraints [20]. Under these conditions, any gapped spin-amplitude distribution is smoothed: Noise continually replenishes the zero-crossing rate, which therefore never strictly vanishes. As a result, a rigorously defined gap does not arise in dynamics, and the updating of Ising states never fully halts.

Even so, for finite systems the evolution can become too slow to matter, i.e., a practical infinity for $t(E_{\text{dec}})$ for desired values of E_{dec} . While updates persist in principle, the density of near-zero spins can drop to negligible levels, rendering the dynamics effectively frozen. We use the ensemble zero-crossing rate $\sqrt{\Delta t} \mu(t)$ as the primary indicator of the effective gap. Thus, a sufficiently small $\mu(t)$ signals the onset of the effective gap, where meaningful evolution has essentially ceased and provides a sufficient condition for the practical inaccessibility of a target energy E_{dec} .

To connect the effective gap with quantities accessible in simulations, we refer to the timescale $1/(N\mu(t))$ as the time budget τ_{budget} , which is the timescale taken by the simulation model to flip one spin [54]. When the time budget becomes of the order 1 or longer, the dynamics is effectively frozen. The time budget is related to the effective gap via

$$\tau_{\text{budget}} = \frac{\sqrt{\Delta t}}{N[\sqrt{\Delta t} \mu(t)]}. \quad (7)$$

Figure 2 shows the contours of the effective gaps for selected small values in the (a, T) space. These contours identify regions where spin flips are statistically unlikely within the available computational budget. The plot is obtained for CIM at quasiequilibrium, based on our theoretical calculations. [55] We numerically verify this trend of the effective gap in Fig. 2. Specifically, using the values of $P(0)$ derived in the one-step replica-symmetry-breaking (1RSB) approximation, we expect that in our simulation setting, the effective gap at which the dynamics is frozen lies in the range between 10^{-6} and 10^{-5} . This agrees well with simulations searching for the freezing points of annealing in Fig. 10.

The contour plots elucidate the discrepancy between zero-temperature equilibrium theory and experimental observations, while providing practical guidance for parameter scheduling in Ising machines. Our analysis particularly

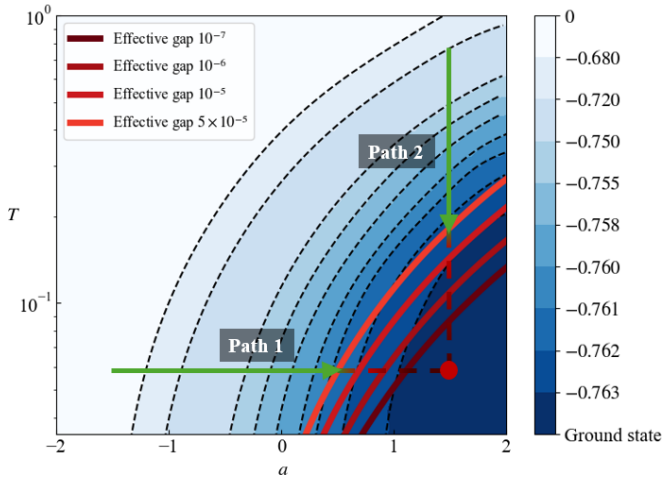


FIG. 2. The decoded energy contours and effective gap lines of the CIM. Path 1 represents the conventional strategy of adjusting the parameter a , while Path 2 represents annealing solely by temperature. Both paths aim at the same termination point (the red dot) by schedule but are prematurely halted by the effective gap. Notably, Path 2 reaches a lower decoded energy than Path 1 when they are halted at the same effective gap contour.

demonstrates that, for CIMs with a fixed deeply buried target point, slow temperature annealing constitutes a better strategy.

A common CIM implementation initializes the parameter a with a small value where the energy landscape is sufficiently simple to facilitate finding the global minimum at the beginning. The parameter a is subsequently increased gradually [7], while maintaining constant low noise. It is commonly assumed that sufficiently slow adjustment of a enables the Ising machine to preserve its ground state despite increasing landscape complexity, consistent with adiabatic evolution principles [20,27].

However, equilibrium analysis reveals that the ground state of the CIM can only be reached at significantly large a and $T = 0$ [53]. As demonstrated by Ref. [28], in the noiseless case, the parameter a schedule should terminate at -0.45 for the SK model in CIMs, corresponding to the true gap, where all near-zero soft spins vanish. While noise enables Ising machines to progress beyond this threshold, the effective gap nevertheless impedes further reduction in decoded energy. Specifically, once the system traverses the effective gap boundary, the timescale required to maintain the ground state or target energy E_{dec} becomes computationally prohibitive, scaling at least as $O(1/N\mu)$. Beyond this critical point, further increasing parameter a yields no performance enhancement in practice and instead leads to a longer convergence time.

The gap-formation process shown in Fig. 3, together with the single-run trajectories in Fig. 9, illustrates this difference. Suppose the target decoded energy is 10^{-2} above the ground-state energy. For gain annealing, as shown in Figs. 3(a) and 9(b), the gap opens at around $a = 0$. Figure 9(a) shows that the CIM has not reached the target decoded energy. In contrast, for temperature annealing, as shown in Figs. 3(b) and 9(b), the spin density at $x = 0$ remains nonzero until the CIM reaches the target decoded energy, preserving a finite probability of further sign changes and equilibration.

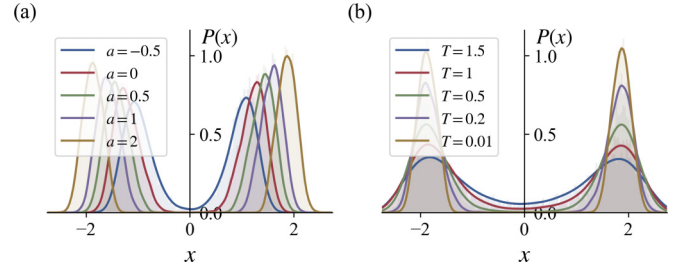


FIG. 3. An illustration of the distributional evolution of CIM. Panel (a) fixes $T = 0.01$ and anneals a ; panel (b) fixes $a = 2$ and anneals T .

With the help of the effective gap contours in the phase space, we are able to design a more effective annealing schedule that involves not only adjusting the parameter a but also the temperature. From the previous analysis, one can overlay the effective gap contours with theoretically achievable decoded energy contours. As shown in Fig. 2, we observe that the energy contours are more slanted compared to the effective gap contours. This indicates that traditional strategies involving horizontal adjustments of parameters (modifying a alone) at a low T are more likely to be halted early by the effective gap. In contrast, vertical strategies (temperature annealing) are affected at a later stage and offer a more reliable path to optimization [56].

This result is consistent with numerical experiments. As shown in Fig. 4 for the SK model, we compare three parameter-tuning strategies: (1) increasing a to a_e at a fixed,

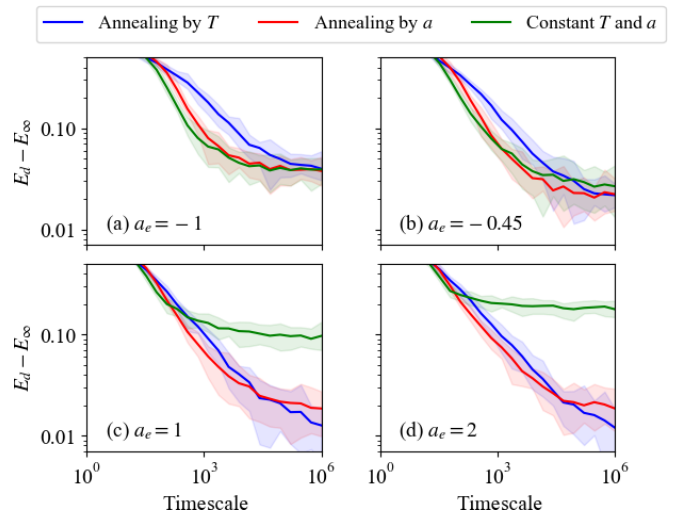


FIG. 4. Comparison of decoded energy profiles obtained via different parameter-tuning strategies across varying timescales. Data points represent the mean value of the decoded energy from the final step across 20 independent runs with the same annealing protocol for SK models of $N = 2000$. Detailed parameter settings can be found in Appendix C. The dashed line represents the ground truth. a_e marks the terminal point of a . Panel (a) corresponds to the gapless regime, panel (b) aligns with the gap opening at zero temperature, and panels (c) and (d) represent gap-opened regime. Shaded areas indicate the range of maximum and minimum values observed. The ground state E_{∞} is set to be -0.7632 .

low temperature (Path 1 in Fig. 2), (2) annealing the temperature while keeping a_e high and constant (Path 2 in Fig. 2), and (3) maintaining both T and a_e as constants. Each strategy was independently tested with varying total run times, as represented on the x axis. Longer timescales yield smoother parameter adjustments and improved performance.

In Figs. 4(a) and 4(b), the terminal points lie outside the region enclosed by the effective gap contour of $O(\frac{1}{N})$, avoiding early termination by the contour. Consequently, with sufficiently long run times, adjusting either T or a achieves comparable performance. In contrast, in Figs. 4(c) and 4(d), early termination by the effective gap causes T annealing to outperform a tuning on longer timescales. More details about experiment settings and results are provided in Appendix C.

VI. EXPOSED VERSUS HIDDEN DECODED ENERGY

As described in the previous section, if the state of the target decoded energy is hidden in the region enclosed by the effective gap contour in the space of a and T , then the annealing process, either by annealing the gain or temperature or both, cannot access the target state, and we observe that for CIM, temperature annealing ends up at a lower decoded energy on the effective gap contour. On the other hand, if the state with the target decoded energy is exposed by the effective gap contour, i.e., located at the boundary or outside the hidden region, then scheduling a can achieve comparable or even better performance than scheduling T .

We demonstrate such a behavior in simCIM [25], which has been implemented on digitized hardware. Its dynamics is governed by

$$\frac{dx_i}{dt} = a\varphi(x_i) + \sum_j J_{ij}\varphi(x_j), \quad (8)$$

where the clipping function is defined as $\varphi(x) = x$ for $|x| \leq 1$ and $\text{sgn}x$ otherwise, and the cubic saturation term is avoided in the dynamics. The DMFT expression of simCIM can be derived in the same manner; it reads

$$\partial_t x(t) = ax(t) + J^2 \xi^2 \int_{t_0}^t dt' G(t, t') x(t') + \phi(t), \quad (9)$$

$$x(t) = \varphi(x(t)).$$

The decoded energy expression is also consistent with Eq. (4), while the soft-hard spin classification and the response function distribution are shown in Fig. 1(d).

In simCIM, the optimal decoded energy is located at the gapless-binary transition point $(a, T) = (0, 0)$ [29], implying that it lies exactly on the effective gap contour with $P(0) = 0$. Envisaging the adoption of an effective gap contour with $P(0)$ approaching 0 in numerical experiments, it is possible that the annealed states can get arbitrarily close to the optimal point. In such cases, appropriately annealing by either temperature or gain can lead to decoded energies approaching optimum, as shown in Fig. 5(a).

Apparently, this picture of an exposed optimal state is inconsistent with the gap-opening scenario presented in the previous section, which is common in many Ising machines. This is because normally the spins of the optimal state have a binary distribution separated by a wide gap, implying that

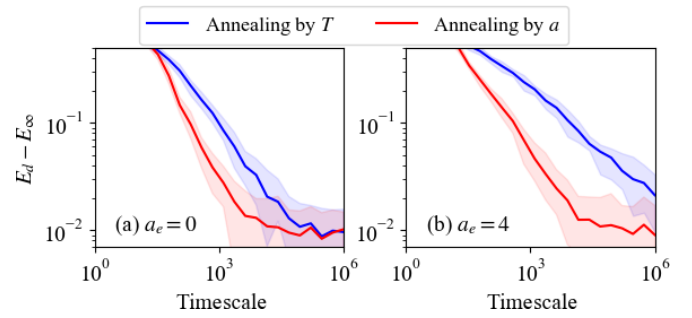


FIG. 5. Comparison of decoded energy profiles obtained via different parameter-tuning strategies across varying timescales for simCIM. Each data point consists of 20 independent trials conducted under different total running timescales, as represented on the x axis.

$P(0)$ should have vanished in the vicinity of the optimal state in the space of a and T .

How can the optimal state in simCIM be exposed by the effective gap? Remarkably, gap formation in simCIM goes through an alternative scenario, in which the spin distribution in the neighborhood of $x = 0$ is continuously reduced when a increases, instead of disappearing discontinuously. From equilibrium analysis [29], we establish that at zero temperature, near-zero soft spins in the simCIM system vanish only when the ground state becomes accessible, corresponding to $(a, T) = (0, 0)$ [57]. Prior to this point, a finite zero-crossing rate persists even in the absence of noise. We conceptually illustrate the relationship between the decoded energy contours and the effective gap contour for simCIM in Fig. 6. Consequently, near $(a, T) = (0, 0)$, the effective gap boundary closely coincides with the lowest decoded energy contour, as illustrated at the terminal of Path 1 and Path 2 in Fig. 6. This alignment implies that slow annealing of a yields comparable performance to temperature annealing when terminating at $(a, T) = (0, 0)$, as evidenced by Fig. 5(a). However, for termination at $a > 0$ and $T = 0$, continuous annealing of a provides no performance enhancement. Conversely, temperature annealing exhibits reduced efficacy due to the effective gap emergence [see Path 3 and the blue curve in Fig. 5(b)].

Although the DMFT derivation above is exact only for the mean-field SK structure, we also test whether the same qualitative effective-gap mechanism remain useful

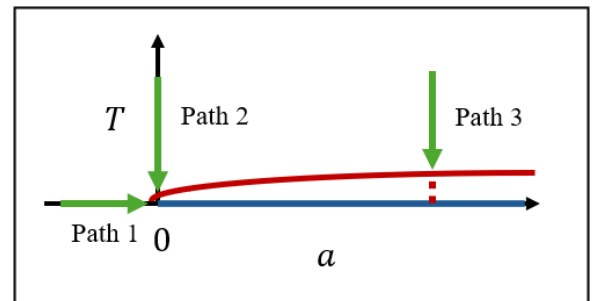


FIG. 6. Schematic plot of the decoded energy contour at the ground state (blue line) and an effective gap contour (red line) of simCIM. The effective gap contour is closely aligned with the ground-state contour near the origin.

empirically beyond this setting. Numerical evidence on the Gset benchmark is presented in Appendix E.

VII. SCALING FUNCTION OF DECODED ENERGY

We have discussed the planning of optimal pathways for annealing in Ising machines. The remaining issue is how the resultant decoded energy depends on time. Remarkably, for a family of models, we note that the finite-time decoded energy of a model approaches its asymptotic value at the same rate as the family of models in the large- N limit. This implies that knowing the convergence behavior of the large- N models is useful for estimating the convergence behavior of the models of arbitrary size.

To see this, we note that Eq. (4), derived from DMFT, implies that the decoded energy decay is independent of system size. While strictly valid only in the thermodynamic limit ($N \rightarrow \infty$), we observe that the energy-density evolution is largely insensitive to finite-size effects for large N . This justifies the approximation

$$E_{\text{dec}}(t, N) - E_{\text{dec}}(\infty, N) \approx \lim_{N \rightarrow \infty} (E_{\text{dec}}(t, N) - E_{\text{dec}}(\infty, N)), \quad (10)$$

where $E_{\text{dec}}(t, N)$ denotes the decoded energy obtained by the Ising machine at time t for a system of size N . Essentially, the relaxation of the decoded energy toward equilibrium in a finite-size system approximates the behavior of the system in the thermodynamic limit. Consequently, in conjunction with the well-characterized scaling behavior of the SK model, one can investigate the solution quality at finite time for arbitrary N . The finite-size scaling of the equilibrium ground-state energy for the SK model follows an $N^{-2/3}$ power law, as derived theoretically [58] and numerically [39–41]:

$$E_{\text{dec}}(\infty, N) \approx E_{\infty} + AN^{-2/3}, \quad (11)$$

where $E_{\infty} = -0.7632$ represents the ground-state energy of the infinite-size SK model and $A = 0.701$ [41,59].

Here, we apply simCIM, generally believed to outperform CIM [24,29], by running 30 trials for each of 100 relaxations of SK instances at various system sizes, using Path 2 as described in Fig. 6.

As shown in Fig. 7, the decoded energies for given values of t have linear relationships with $N^{-2/3}$, sharing the same slope A with the finite-size scaling of the ground-state energy. Extrapolating the linear relations to the intercept, we obtain the thermodynamic limit of the decoded energies of finite times. We observed that the average and minimum values of $E_{\text{dec}}(t, \infty)$ versus t can be approximated as power laws. In particular,

$$E_{\text{dec}}(t, \infty) \approx E_{\infty} + Bt^{-\xi}. \quad (12)$$

As shown in Fig. 8, $E_{\text{dec}}(t, \infty)$ has a very good power-law fit with t after subtracting the bias term E_{∞} . For average values of $E_{\text{dec}}(t, \infty)$, we obtain $E_{\infty} = -0.7626 \pm 0.0003$, $B = 4.657$, and $\xi = 0.713$, whereas for minimum values of $E_{\text{dec}}(t, \infty)$, we get $E_{\infty} = -0.7632 \pm 0.0003$, $B = 4.656$, and $\xi = 0.763$. Thus, Eqs. (11) and (12) allow us to estimate the decoded energy at finite time for arbitrary N . Equivalently, if $\epsilon = E_{\text{dec}}(t, \infty) - E_{\infty}$ denotes the target energy-density

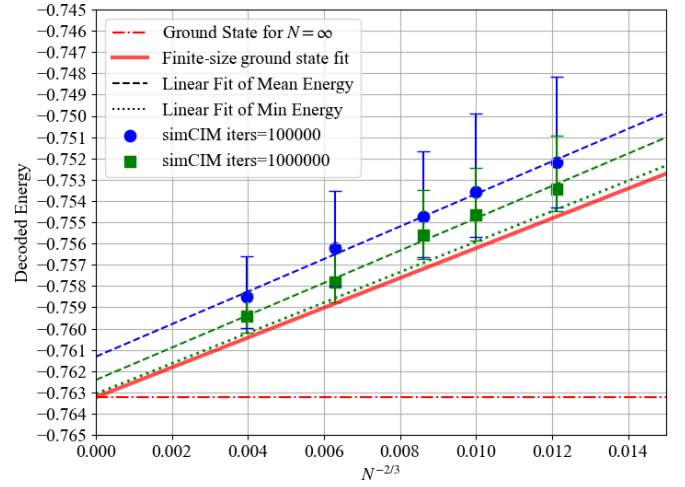


FIG. 7. Energies obtained by simCIM using the Euler-Maruyama method (step size 0.1) with 100 000 and 1 000 000 iterations. For each system size, $N = 750, 1000, 1250, 2000$, and 4000 , we evaluate 200 independent SK instances and perform 30 runs per instance. Data points report the decoded energy averaged over the 30 runs and then averaged across instances. Error bars show, for each instance, the mean of the minimum and maximum decoded energies observed over the 30 runs. The red line denotes the mean ground-state energy with finite-size corrections reported in Ref. [59]. The green dotted line indicates the fit for the expected minimum value after 10^6 iterations, yielding an extrapolated infinite-size value of -0.76307 ± 0.00018 .

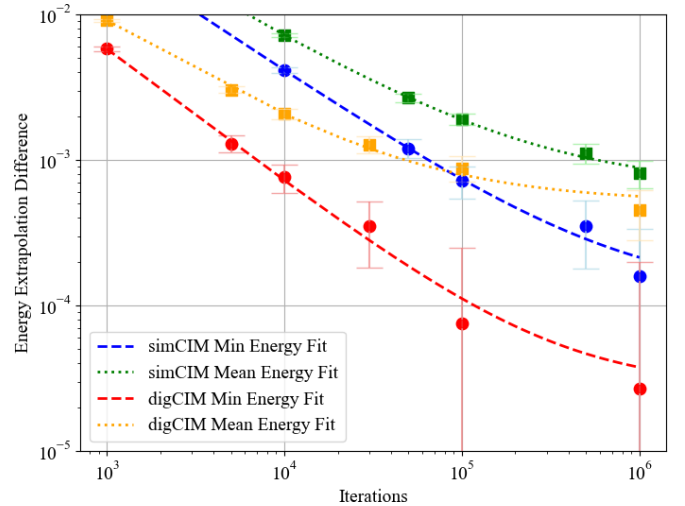


FIG. 8. The decoded energy $\lim_{N \rightarrow \infty} E_{\text{dec}}(t, N) \equiv E_{\text{dec}}(t, \infty)$, extrapolated to the infinite-size limit ($N \rightarrow \infty$) at various iteration steps for simCIM and digCIM. Experimental parameters are consistent with Figs. 7 and 12. Squares indicate the mean energy, extrapolated from the average of 30 trials across 200 independent SK instances. Dots represent the minimum energy, extrapolated from the best result obtained in 30 trials. Error bars indicate the standard error of the extrapolated values. Solid lines indicate fits to Eq. (12) for each dataset. Data are offset by -0.76323 [39,59], a slightly lower estimate of the SK ground-state energy, rather than by the individual fitted bias parameters. The y axis represents the deviation from this reference value on a logarithmic scale.

error, Eq. (12) gives $t(\epsilon) \sim \epsilon^{-1/\xi}$. The corresponding SK computational cost is therefore $\text{Cost}(N, \epsilon) = C_{\text{sim/dig}}(\epsilon)N^2 \sim N^2\epsilon^{-1/\xi}$, rather than a uniform $O(N^2)$ cost as $\epsilon \rightarrow 0$; the $O(N^2)$ statement applies at fixed target accuracy.

Finally, applying the scaling relations of the dynamical cases enables us to compute the ground-state energy of the Ising model at the thermodynamic limit with accuracy matching state-of-the-art measurements [39–41,59]. This is achieved by directly reading the value of E_∞ from Eq. (12) for fitting the minimum values of $E_{\text{dec}}(t, \infty)$, as finding the minimal decoded energy in practice is a more common decoding method than finding the mean value.

The value of $E_\infty = -0.763\,17 \pm 0.000\,29$ is close to the numerical estimate of the ground-state energy of -0.7632 [39–41,59] and the theoretical estimate of $-0.763\,17$ [60]. It is also numerically consistent with the results of Ref. [34], where a message-passing algorithm is claimed to achieve near-optimal solutions arbitrarily close to the ground state in $O(N^2)$ time.

An even more accurate estimate of the ground-state energy is provided by digCIM [29] with shorter convergence time and higher success probability. The estimate by digCIM is $-0.763\,17 \pm 0.000\,15$ (see Appendix F). Besides our work, several experimental variants of Ising machines report superior performance, including digCIM [29], dSB [26], and CAC [27]. Appendix G provides a phenomenological comparison with dSB and CAC. In these examples, $P(0)$ remains finite while the decoded energy decreases, suggesting that these mechanisms may help sustain soft-spin activity and postpone dynamical hardening.

VIII. CONCLUSION

In summary, we analyzed the behaviors of dynamical optimization solvers by deriving their dynamical mean-field descriptions. To obtain a theoretically meaningful description for finite-size systems, we introduced the concept of an effective gap, which can impede further energy reduction. We proposed using the density of near-zero soft spins as an indicator for the onset of this effective gap. By comparing decoded-energy contours with the emergence of the gap, we designed an optimal parameter-scheduling strategy for Ising machines. Notably, in contrast to the conventional approach of annealing the gain, the optimal schedule for CIM annealing is through annealing the temperature. We also showed that this framework naturally generalizes to other platforms. On

the other hand, our analysis of simCIM indicates that if the target decoded energy is exposed by the effective gap contour, annealing the parameter a is also effective. Across all cases, our theoretical predictions are in strong agreement with numerical experiments. We mention in passing that temperature annealing is also more advantageous than gain annealing in digCIM. We have shown that since system states in the regime of negative gain are isomorphic with each other in digCIM [29], annealing the gain offers little help for convergence to the optimal solution, in contrast with temperature annealing.

Building on this, we argued that Ising machines can reach near-optimal energies within time complexity $t(E_{\text{dec}})N^2$, so their fixed-accuracy N scaling has the same SK N^2 form as rigorously analyzed message-passing methods [35], while the accuracy dependence is set empirically by $t(\epsilon) \sim \epsilon^{-1/\xi}$ rather than by a proven matching bound. We further identified two key ingredients underlying this performance: the fixing of hard spins and simultaneous updating soft spins. In addition, we clarified the pivotal role of the zero-crossing rate of soft spins in driving improvements of the decoded energy.

This study focused on gradient-descent dynamics. In machines using Hamiltonian dynamics with momentum, we expect similar asymptotic time-complexity advantages but with smaller constant factors, and extra parameter scheduling; exploring this regime is an interesting direction for future work.

ACKNOWLEDGMENT

This work was supported in part by Huawei through a CSTT project.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: DMFT CALCULATION

We first reformulate Eq. (2) as a Martin-Siggia-Rose-Janssen-De Dominicis (MSRJD) functional integral [46,61–63]. The resulting action closely follows the dynamical formulation of Sompolinsky and Zippelius [42,43], with the additional source term $l_e(t)$ coupled to the decoded Ising energy. As shown below, this source field allows us to generate the decoded-energy observable within the DMFT formalism and introduces the corresponding additional order parameters:

$$Z[t|t_0] = \int Dx D\hat{x} \prod_i \exp \left\{ \int dt \left[l_i(t)x_i(t) + l_e(t) \sum_{i<j} J_{ij} \text{sgn}x_i(t) \text{sgn}x_j(t) + i\hat{x}_i(t) \left(-\partial_t x_i(t) - x_i^3(t) + ax_i(t) + \xi \sum_j J_{ij} x_j(t) \right) - T\hat{x}_i^2(t) - \frac{1}{2}(3x_i^2(t) - a) \right] \right\}. \quad (\text{A1})$$

Here, $\hat{x}_i(t)$ is the MSRJD auxiliary response field. It enforces the stochastic equation of motion in the functional-integral representation and, physically, correlations involving $\hat{x}_i(t)$ generate linear response functions of the trajectory responding to infinitesimal perturbations. The fields $l_i(t)$ and $l_e(t)$ are source fields conjugate to $x_i(t)$ and to the decoded Ising energy $\sum_{i<j} J_{ij} \text{sgn}x_i(t) \text{sgn}x_j(t)$, respectively. In particular, $l_e(t)$ is introduced as a generating field that enables the extraction of $E_{\text{dec}}(t)$

from the dynamical action. Finally, the term

$$\exp \left[-\frac{1}{2} \int dt (3x_i^2(t) - a) \right]$$

is the Jacobian associated with transforming the stochastic path integral from noise variables to dynamical variables in the Stratonovich convention [63,64].

By integrating over the coupling matrix and leveraging the properties of the SK model, we reorganize the terms to obtain the disorder average:

$$\begin{aligned} \mathbb{E}_{J_{ij}}(Z[t|t_0]) &= \int Dx D\hat{x} \prod_i \exp(S_i^0) \times \prod_{i<j} \exp\left(\frac{J^2}{2N} \Sigma_{ij}^2\right), \\ S_i^0 &= \int dt \left\{ l_i(t)x_i(t) + i\hat{x}_i(t)[- \partial_t x_i(t) - x_i^3(t) + ax_i(t)] - \frac{1}{2}[3x_i^2(t) - a] - T\hat{x}_i^2(t) \right\}, \\ \Sigma_{ij} &= \int dt [\xi i\hat{x}_i(t)x_j(t) + \xi i\hat{x}_j(t)x_i(t) + l_e(t)\text{sgn}x_i(t)\text{sgn}x_j(t)]. \end{aligned} \quad (\text{A2})$$

The interaction terms can be decoupled using the Hubbard-Stratonovich transformation. After reorganization, we obtain

$$\begin{aligned} \prod_{i<j} \exp\left(\frac{J^2}{2N} \Sigma_{ij}^2\right) &= C \left(\prod_{k=1}^7 \int dq_k(t_1, t_2) \right) \exp \left\{ -NJ^2 \int dt_1 \int dt_2 \left[\frac{\xi^2}{2} q_1 q_2 + \frac{\xi}{2} q_3 a_4 + q_5 q_6 + \frac{1}{4} q_7^2 \right. \right. \\ &\quad + J^2 \sum_i \left(\frac{\xi^2}{2} q_1 i\hat{x}_i(t_1)i\hat{x}_i(t_2) + \frac{\xi^2}{2} q_2 x_i(t_1)x_i(t_2) + \frac{\xi^2}{2} q_3 i\hat{x}_i(t_1)x_i(t_2) + \frac{\xi^2}{2} q_4 x_i(t_1)i\hat{x}_i(t_2) \right. \\ &\quad \left. \left. + l_e(t_1)q_5 \text{sgn}x_i(t_1)i\hat{x}_i(t_2) + q_6 \text{sgn}x_i(t_1)x_i(t_2) + \frac{1}{2}\sqrt{l_e(t_1)l_e(t_2)}q_7 \text{sgn}x_i(t_1)\text{sgn}x_i(t_2) \right) \right] \right\}, \end{aligned} \quad (\text{A3})$$

where C is a constant generated by the transformations and does not affect subsequent calculations. The auxiliary variables q_k are introduced to decouple quadratic terms such as $i\hat{x}(t), x_j(t)$ and can be validated via the saddle-point conditions, e.g., $\partial \mathbb{E}(Z)/\partial q_k = 0$. Here, D denotes the Gaussian measure. After enforcing causality and integrating out the auxiliary variables, the only order parameters with direct physical meaning are the correlation functions (q_1 and q_5) and the response functions (q_3, q_4 , and q_6). Since we can always set the auxiliary variables l_e and $\hat{x}$ to zero, the remaining auxiliary order parameters can likewise be taken to vanish. From the saddle-point equations, their explicit forms are given by

$$\begin{aligned} q_1(t_1, t_2) &= \langle x(t_1)x(t_2) \rangle, \\ q_3(t_1, t_2) &= \langle x(t_1)i\hat{x}(t_2) \rangle, \quad \text{for } t_1 > t_2, \\ q_4(t_1, t_2) &= \langle i\hat{x}(t_1)x(t_2) \rangle, \quad \text{for } t_1 < t_2, \\ q_5(t_1, t_2) &= \langle \text{sgn}x(t_1)x(t_2) \rangle, \\ q_6(t_1, t_2) &= l_e(t_1)\langle \text{sgn}x(t_1)i\hat{x}(t_2) \rangle, \quad \text{for } t_1 > t_2. \end{aligned} \quad (\text{A4})$$

The first three terms can be used to recover the mean-field dynamics [Eq. (3)], as derived by Sompolinsky and Zippelius [43]. Meanwhile, the latter order parameter provides the dynamical mean-field expression for the decoded energy [Eq. (4)], as follows:

$$\begin{aligned} E_{\text{dec}}(t) &= \frac{\partial \mathbb{E}_{J_{ij}}(Z[t|t_0])}{\partial l_e(t)} \\ &= -J^2 \int_0^t dt' \langle \text{sgn}x(t)x(t') \rangle \langle \text{sgn}x(t)i\hat{x}(t') \rangle. \end{aligned} \quad (\text{A5})$$

APPENDIX B: THE ZERO-CROSSING RATE

Consider the dynamical equation [Eq. (3)] when the Ising machine approaches equilibrium, which becomes

$$\frac{dx}{dt} \approx (a + \chi)x + \phi(t), \quad (\text{B1})$$

where the susceptibility χ is the steady-state value of the integral $\int dt' G(t, t')$. In discrete-step dynamics, the dynamics becomes

$$x(t + \Delta t) = x(t) + \Delta t[a_{\text{eff}}x(t) + \sqrt{\langle \phi(t)^2 \rangle}z(t)], \quad (\text{B2})$$

where we further denote the effective gain $a_{\text{eff}} = a + \chi$, and $z(t)$ is a Gaussian variable of mean 0 and unit variance. Suppose the spin at time t is $x(t) = -\epsilon$. Then, the condition that $x(t + \Delta t)$ changes sign is

$$-\epsilon + \Delta t(-a_{\text{eff}}\epsilon + \sqrt{\langle \phi(t)^2 \rangle}z(t)) > 0. \quad (\text{B3})$$

This implies that

$$z(t) > \frac{(1 + a_{\text{eff}}\Delta t)\epsilon}{\Delta t\sqrt{\langle \phi(t)^2 \rangle}} \approx \frac{\epsilon}{\Delta t\sqrt{\langle \phi(t)^2 \rangle}}. \quad (\text{B4})$$

Averaging over the Gaussian distribution of $z(t)$, the average transition rate becomes

$$R(\epsilon) = \frac{P(-\epsilon)}{\Delta t} \int_{\epsilon/(\Delta t\sqrt{\langle \phi(t)^2 \rangle}}^{\infty} Dz. \quad (\text{B5})$$

Noting that $P(-\epsilon) \approx P(0)$ and integrating over non-negative values of ϵ , the zero-crossing rate (in positive or negative

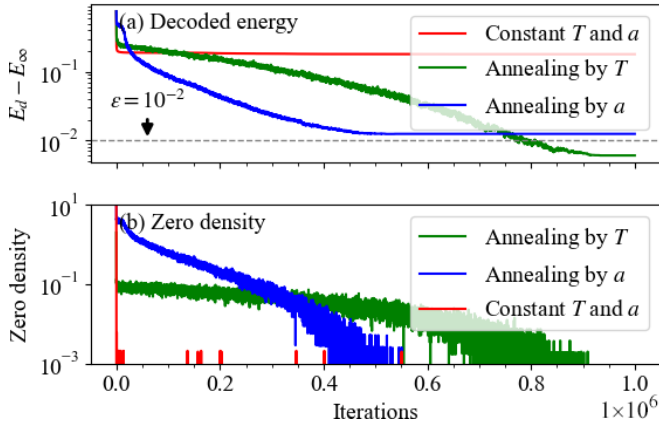


FIG. 9. Trajectory of the decoded energy and $P(0)$ from a single CIM experiment with $a_e = 2$. Due to the results being numerically sampled from a single trial, the curve shown exhibits fluctuations.

direction) becomes

$$\mu(t) = \frac{P(0)}{\Delta t} \int_0^\infty D_{zz} \Delta t \sqrt{\langle \phi(t)^2 \rangle} = \frac{\sqrt{\langle \phi(t)^2 \rangle}}{\sqrt{2\pi}} P(0). \quad (\text{B6})$$

The above equation is applicable to spin transitions from the negative to the positive side. The expression for transitions from positive to negative is identical. Hence, we have shown that the zero-crossing rate is proportional to $P(0)$:

$$\langle \phi(t)^2 \rangle \simeq \frac{2T}{\Delta t} + 2J^2 \xi^2 Q(t, t). \quad (\text{B7})$$

The dominant term of the zero-crossing rate then becomes

$$\mu(t) \simeq \sqrt{\frac{T}{\pi \Delta t}} P(0) \quad (\text{B8})$$

for small Δt and nonzero T .

APPENDIX C: NUMERICAL EXPERIMENT SETTING

The annealing schedule for the temperature T is defined as $T(\tau) = T_0 - (T_0 - T_e)\tau/S$, where $\tau = 1, 2, \dots$, and $S = 10^6$ represents discrete time steps. The initial and final temperatures are set to $T_0 = 2$ and $T_e = 0.01$, respectively, while the drive parameter is fixed at $a = a_e$. For the linear annealing of a (red curve), the parameter starts at $a_0 = -2$ and evolves to various final values, a_e , as shown in the different subfigures. During this process, the temperature is held constant at $T = 0.01$. The annealing schedule for a is given by $a(\tau) = a_0 - (a_0 - a_e)\tau/S$. For the constant parameter scenario, both a and T are fixed at $a = 2$ and $T = 0.01$, respectively, over the same timescale.

All experiments were performed independently 20 times using the Euler-Maruyama method with a time step of $\Delta t = 0.05$, following the same annealing protocol for SK models with $N = 10\,000$ spins.

The evolution trajectory of the decoded energy and its numerical value of $P(0)$ for a single experiment is shown in Fig. 9. As predicted, the zero-crossing rate for annealing by a quickly drops to zero, causing the decoded energy to plateau. In contrast, annealing by T delays the emergence

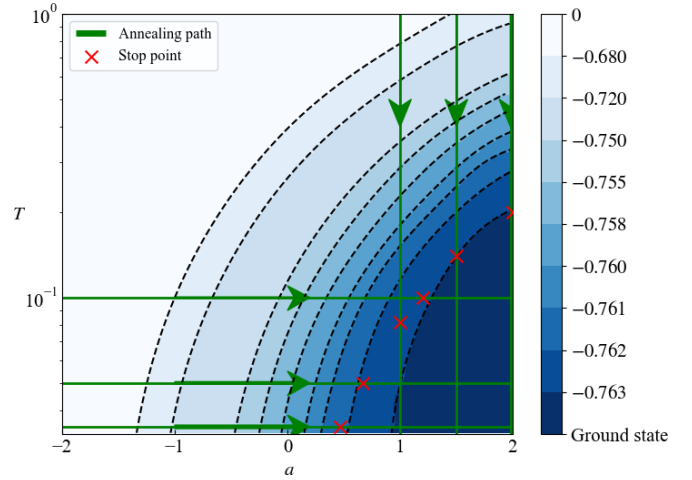


FIG. 10. Numerical verification of the effective gap. The green lines show the annealing trajectories for both a annealing and T annealing, using the protocols described in Appendix C. Red crosses mark representative points where the near-zero density $P(0)$ vanishes numerically for systems of size $N = 10\,000$.

of the effective gap, allowing it to gradually suppress other strategies and achieve the best performance.

APPENDIX D: NUMERICAL VERIFICATION OF EFFECTIVE GAPS

According to the schedule protocols discussed above, and as illustrated in Fig. 9, one can estimate the parameter region in which the zero-crossing rate vanishes. We performed finite-size simulations with $N = 10\,000$ along several annealing trajectories and marked the points where the zero-crossing rate becomes statistically indistinguishable from zero. These points define a numerical dynamical barrier, shown in Fig. 10. Specifically, we considered a annealing from $a = -2$ to $a = 2$ at fixed temperatures $T = 0.1, 0.05$, and 0.035 , and T annealing from $T = 1$ to $T = 0$ at fixed gains $a = 1, 1.5$, and 2 . The numerically obtained dynamical barrier is consistent with the theoretical zero-density contours shown in Fig. 2.

Using this dynamical barrier as a guide for trajectory design, we note that temperature annealing alone is not sufficient when the gain a is too small, because the system then lacks a strong enough deterministic drive toward low-energy configurations. Conversely, gain annealing can benefit from enhanced exploration at sufficiently large T . However, this strategy requires a delicate choice of temperature: If T is too large, thermal fluctuations can wash out the energy landscape of the original Ising problem.

APPENDIX E: Gset EXPERIMENTS

The concepts of energy barrier and related strategies, such as temperature annealing, are applicable beyond the SK model; we observe similar phenomena in general benchmarks and datasets. Here, we use the G1 and G6 instances from the popular Max-Cut benchmark dataset Gset [65] as examples.

As shown in Fig. 11, similar to the SK model, temperature annealing in the CIM benefits from the delayed appearance

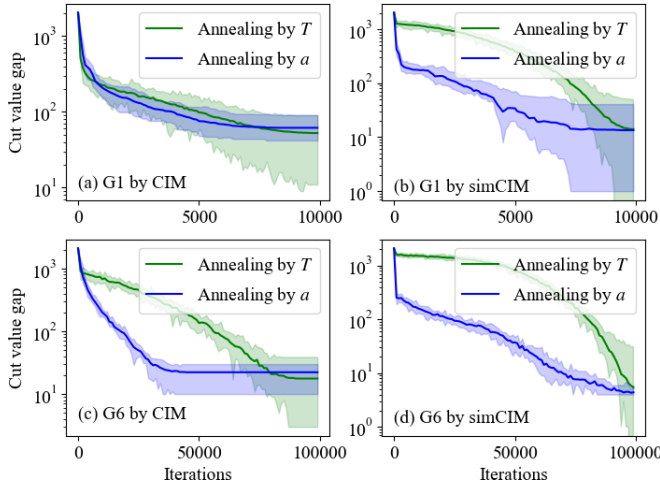


FIG. 11. Comparison of cut value profiles obtained via different parameter-tuning strategies for CIM and simCIM on G1 and G6. The solid lines represent the average over 20 independent trials, while the shaded areas indicate the range between the maximum and minimum observed values. The y axis shows the difference between the obtained energy and the known best solution.

of the dynamical barrier, resulting in better performance. In contrast, for simCIM, which benefits from the exposure of the optimal point, both temperature annealing and linear gain (a) annealing yield the same mean energy, outperforming the standard CIM. Notably, although the mean values for simCIM are identical, temperature annealing exhibits larger fluctuations. This leads to worse maximum values but better minimum values (increasing the chance of obtaining the exact solution on G1 and G6). This potentially suggests that in practice, given computational resources for multiple trials and postselection of the best result, temperature annealing remains the superior choice.

The physical picture accounting for the above observations is the same as that in the SK model. For CIM, this can be seen from the dynamical equation of the individual nodes. While different network structures and optimization objectives provide different external driving terms, the combination of the linear gain and the cubic saturation in the internal driving term algebraically creates bifurcated spin values, resulting in the soft-hard dichotomy and the effective-gap picture. Indeed, simulation results for G1 in the Supplemental Material of Ref. [29] showed that the gap-opened phase exists in G1.

For simCIM, due to the absence of the cubic saturation term, the linear gain term directly drives the spins toward the threshold values for positive gains, again resulting in the soft-hard dichotomy and the effective-gap picture. Furthermore, this direct drive leads to the simultaneous appearance of the gapless and binary phases, and hence the exposure of the optimal state as described in Sec. VI. These are also confirmed by simulation results for G1 in the Supplemental Material of Ref. [29].

We note, however, that this should be interpreted as an empirical observation rather than as a rigorous extension of the SK analysis to general sparse or structured graphs. In particular, G1 and G6 are not sparse in the asymptotic sense relevant here: The probability of a nonzero edge does not scale

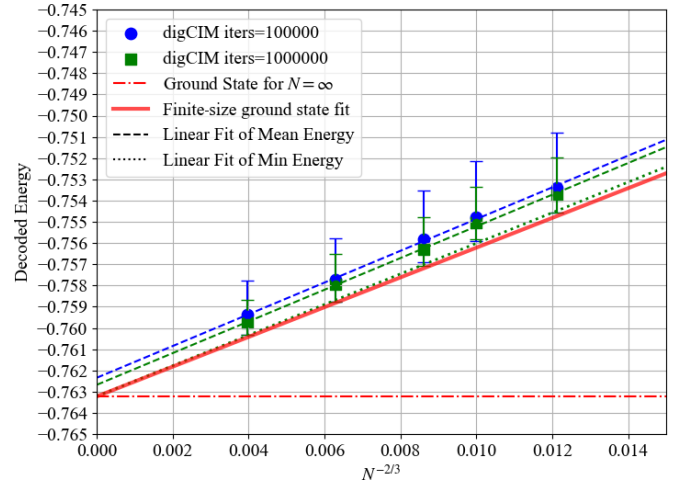


FIG. 12. Energies obtained by digCIM using the Euler-Maruyama method (step size 0.1) with 100 000 and 1 000 000 iterations. It sets $a = -5$, with temperature annealing from 0.5 to 0. For each system size, $N = 750, 1000, 1250, 2000$, and 4000, we evaluate 200 independent SK instances and perform 30 runs per instance. Data points report the minimum decoded energy averaged over the 30 runs and then averaged across instances. Error bars show, for each instance, the mean of the minimum and maximum decoded energies observed over the 30 runs. The red line denotes the mean ground-state energy with finite-size corrections reported in Ref. [59]. The green dotted line indicates the fit for the expected minimum value after 10^6 iterations, yielding an extrapolated infinite-size value of $-0.763\,203 \pm 0.000\,17$.

as $O(1/N)$, but remains of order $O(1)$. The case for sparse or structured Max-Cut instances remains to be explored.

We also observe that the spectra of G1 and G6 approximately follow the semicircle law after normalization, making their large-scale spectral properties closer to those of the SK coupling matrix. This provides a possible universality-based explanation for the observed agreement, but we do not regard it as evidence for exact universality across general sparse or structured Max-Cut instances.

APPENDIX F: DigCIM RESULTS

DigCIM takes a form similar to simCIM:

$$\frac{dx_i}{dt} = a\varphi(x_i) + \xi \sum_{j=1, j \neq i}^N J_{ij} \text{sgn} x_j + \zeta_i. \quad (\text{F1})$$

In our previous work [29], we showed that this dynamics attains the SK ground state uniformly for any negative a . It also exhibits superior dynamical properties. As shown in Fig. 12, it also obeys the finite-size scaling relation as in Eq. (11). Similarly, in Fig. 8, we plot the average and minimum values of $E_{\text{dec}}(t, \infty)$ versus t , confirming the scaling relations in Eq. (12), where $E_{\infty, \text{ave}} = -0.762\,45 \pm 0.000\,18$, $B_{\text{ave}} = 3.644$, $\xi_{\text{ave}} = 0.879$, $E_{\infty, \text{min}} = -0.763\,17 \pm 0.000\,15$, $B_{\text{min}} = 3.641$, and $\xi_{\text{min}} = 0.934$. Hence, applying the scaling relations enables us to compute the ground-state energy of the Ising model at the thermodynamic limit even more accurate than simCIM.

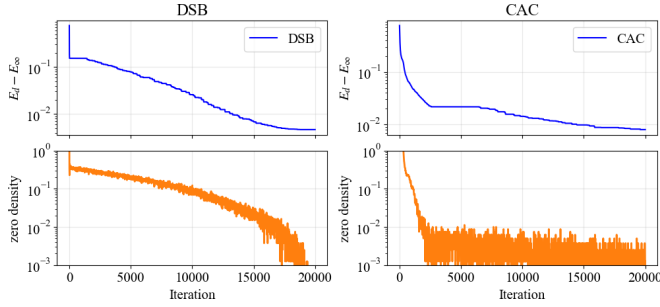


FIG. 13. Trajectory of the decoded energy and $P(0)$ from single-run experiments using dSB (discretized SBM) and chaotic amplitude control (CAC). The parameter settings for dSB follow the recommendations of Ref. [26], while those for CAC follow Ref. [17].

Comparing the solution quality of simCIM and digCIM under limited computational resources in Fig. 8, we see that, although they share the same asymptotic time complexity, digCIM benefits from a superior proportionality constant of the power law, resulting in a faster rate of energy decay.

APPENDIX G: SOFT-SPIN ACTIVITY IN ADVANCED DYNAMICAL SOLVERS

There are several popular dynamical solvers with more sophisticated mechanisms than the gradient-based systems discussed above, and many of them have demonstrated promising performance. Although a rigorous theoretical analysis of these solvers is beyond the scope of this work, our framework can still be used to numerically reveal common dynamical features and provide a brief qualitative discussion.

One example is the simulated bifurcation machine (SBM) introduced in Ref. [26]. SBM evolves the dynamics by introducing momentum, which effectively plays a role analogous to temperature: It helps the system to continue exploring the configuration space by maintaining the effective gap and preserving a large population of soft spins until late in the evolution. Since SBM is based on a Hamiltonian structure similar to the systems discussed in the main text, its dynamics is expected to exhibit related phenomena. This expectation is supported by the similarity between the temperature-annealing trajectory in Fig. 9 and the SBM results shown in Fig. 13.

Another example is the CAC solver introduced in Ref. [27]. CAC dynamically suppresses amplitude heterogeneity and

destabilizes shallow local minima by introducing an additional error dynamics that modulates the effective coupling strength ξ . From the perspective of the present framework, this destabilization mechanism can also maintain persistent spin flipping and preserve a finite soft-spin population as shown in Fig. 13, albeit through a mechanism different from temperature annealing or momentum-induced exploration. Thus, while CAC is not explicitly described by the gradient-based DMFT equations analyzed in the main text, its empirical success is consistent with the broader picture that robust optimization requires avoiding premature hardening of the spin amplitudes and maintaining dynamical activity near the origin.

APPENDIX H: A RELATED WORK

Recently, we became aware of a parallel work analyzing the CIM's energy landscape for the SK model with ferromagnetic coupling via the laser gain parameter a [66]. That study provided the geometry and dynamics of annealed CIM optimization in SK-related models with hidden or planted solutions. It emphasized the geometric structure of the annealed landscape, particularly in planted models.

They showed that low-energy global minima develop soft modes that can guide optimization dynamics, and that these minima eventually become rigid, terminating further optimization gains from gain annealing. In contrast, the present work focuses on the unplanted SK optimization problem and develops a DMFT framework for the decoded Ising energy itself. In particular, our main practical conclusion is that temperature-only annealing can outperform conventional gain annealing in CIM/digCIM for a fixed deep target point in phase space.

Nevertheless, the two analyses are consistent. Their observation that useful global minima possess soft modes, but eventually become rigid under further annealing, is closely related to our finding that optimization requires maintaining a finite density of near-zero, responsive soft spins. In our terminology, the loss of such soft degrees of freedom corresponds to entering the effective-gap region, where zero-crossing activity is suppressed and decoded-energy improvement stalls. Thus, their landscape-level picture of soft modes and rigidity provides a complementary geometric interpretation of the dynamical bottleneck identified here.

- [1] M. V. LeVine and H. Weinstein, Aim for allostery: Using the Ising model to understand information processing and transmission in allosteric biomolecular systems, *Entropy* **17**, 2895 (2015).
- [2] S. Bao, M. Tawada, S. Tanaka, and N. Togawa, Multi-day travel planning using Ising machines for real-world applications, in *IEEE International Intelligent Transportation Systems Conference (ITSC)* (IEEE, Piscataway, NJ, 2021), pp. 3704–3709.

- [3] K. Ooka and M. Arai, Accurate prediction of protein folding mechanisms by simple structure-based statistical mechanical models, *Nat. Commun.* **14**, 6338 (2023).
- [4] J. D. Bryngelson and P. G. Wolynes, Spin glasses and the statistical mechanics of protein folding, *Proc. Natl. Acad. Sci. USA* **84**, 7524 (1987).
- [5] F. Barahona, On the computational complexity of Ising spin glass models, *J. Phys. A: Math. Gen.* **15**, 3241 (1982).

- [6] S. Arora, E. Berger, E. Hazan, G. Kindler, and M. Safra, On non-approximability for quadratic programs, in *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)* (IEEE Computer Society Press, Los Alamitos, CA, 2005), pp. 206–215.
- [7] N. Mohseni, P. L. McMahon, and T. Byrnes, Ising machines as hardware solvers of combinatorial optimization problems, *Nat. Rev. Phys.* **4**, 363 (2022).
- [8] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
- [9] M. Johnson, M. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. Berkley, J. Johansson, P. Bunyk, E. Chapple, C. Enderud, J. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, and G. Rose, Quantum annealing with manufactured spins, *Nature (London)* **473**, 194 (2011).
- [10] S. Boixo, T. F. Rønnow, S. Isakov, Z. Wang, D. Wecker, D. Lidar, J. Martinis, and M. Troyer, Evidence for quantum annealing with more than one hundred qubits, *Nat. Phys.* **10**, 218 (2014).
- [11] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: Methods and implementations, *Rep. Prog. Phys.* **83**, 054401 (2020).
- [12] Y. Salathé, M. Mondal, M. Oppliger, J. Heinsoo, P. Kurpiers, A. Potočnik, A. Mezzacapo, U. Las Heras, L. Lamata, E. Solano, S. Filipp, and A. Wallraff, Digital quantum simulation of spin models with circuit quantum electrodynamics, *Phys. Rev. X* **5**, 021027 (2015).
- [13] J. Chou, S. Bramhavar, S. Ghosh, and W. Herzog, Analog coupled oscillator based weighted Ising machine, *Sci. Rep.* **9**, 14786 (2019).
- [14] M. Yamaoka, C. Yoshimura, M. Hayashi, T. Okuyama, H. Aoki, and H. Mizuno, A 20k-spin Ising chip to solve combinatorial optimization problems with CMOS annealing, *IEEE J. Solid-State Circuits* **51**, 303 (2015).
- [15] M. Aramon, G. Rosenberg, E. Valiante, T. Miyazawa, H. Tamura, and H. Katzgraber, Physics-inspired optimization for quadratic unconstrained problems using a digital annealer, *Front. Phys.* **7**, 48 (2019).
- [16] K. Yamamoto, K. Kawamura, K. Ando, N. Mertig, T. Takemoto, M. Yamaoka, H. Teramoto, A. Sakai, S. Takamaeda-Yamazaki, and M. Motomura, Statica: A 512-spin 0.25M-weight annealing processor with an all-spin-updates-at-once architecture for combinatorial optimization with complete spin-spin interactions, *IEEE J. Solid-State Circuits* **56**, 165 (2020).
- [17] T. Leleu, F. Khoiratee, T. Levi, R. Hamerly, T. Kohno, and K. Aihara, Scaling advantage of chaotic amplitude control for high-performance combinatorial optimization, *Commun. Phys.* **4**, 266 (2021).
- [18] A. Marandi, Z. Wang, K. Takata, R. Byer, and Y. Yamamoto, Network of time-multiplexed optical parametric oscillators as a coherent Ising machine, *Nat. Photon.* **8**, 937 (2014).
- [19] Y. Yamamoto, K. Aihara, T. Leleu, K.-I. Kwarabayashi, S. Kako, M. Fejer, K. Inoue, and H. Takesue, Coherent Ising machines—Optical neural networks operating at the quantum limit, *npj Quantum Inf.* **3**, 49 (2017).
- [20] Z. Wang, A. Marandi, K. Wen, R. Byer, and Y. Yamamoto, A coherent Ising machine based on degenerate optical parametric oscillators, *Phys. Rev. A* **88**, 063853 (2013).
- [21] T. Inagaki, Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi, P. McMahon, T. Umeki, K. Enbutsu, O. Tadanaga, H. Takenouchi, K. Aihara, K.-I. Kwarabayashi, K. Inoue, S. Utsunomiya, and H. Takesue, A coherent Ising machine for 2000-node optimization problems, *Science* **354**, 603 (2016).
- [22] T. Honjo, T. Sonobe, K. Inaba, T. Inagaki, T. Ikuta, Y. Yamada, T. Kazama, K. Enbutsu, T. Umeki, R. Kasahara, K. ichi Kwarabayashi, and H. Takesue, 100,000-spin coherent Ising machine, *Sci. Adv.* **7**, eabh0952 (2021).
- [23] F. Böhm, T. Vaerenbergh, G. Verschaffelt, and G. Van der Sande, Order-of-magnitude differences in computational performance of analog Ising machines induced by the choice of nonlinearity, *Commun. Phys.* **4**, 149 (2021).
- [24] F. Böhm, G. Verschaffelt, and G. Van der Sande, A poor man's coherent Ising machine based on opto-electronic feedback systems for solving optimization problems, *Nat. Commun.* **10**, 3538 (2019).
- [25] E. S. Tiunov, A. E. Ulanov, and A. I. Lvovsky, Annealing by simulating the coherent Ising machine, *Opt. Express* **27**, 10288 (2019).
- [26] H. Goto, K. Endo, M. Suzuki, Y. Sakai, T. Kanao, Y. Hamakawa, R. Hidaka, M. Yamasaki, and K. Tatsumura, High-performance combinatorial optimization based on classical mechanics, *Sci. Adv.* **7**, eabe7953 (2021).
- [27] T. Leleu, Y. Yamamoto, P. McMahon, and K. Aihara, Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity, *Phys. Rev. Lett.* **122**, 040607 (2019).
- [28] A. Yamamura, H. Mabuchi, and S. Ganguli, Geometric landscape annealing as an optimization principle underlying the coherent Ising machine, *Phys. Rev. X* **14**, 031054 (2024).
- [29] S. Zhou, K. Y. M. Wong, J. Wang, D. S. W. Hui, D. Ebler, and J. Sun, Phase analysis of Ising machines and their implications on optimization, *Phys. Rev. Res.* **08**, 033147 (2026).
- [30] F. Böhm, D. Alonso-Urquijo, G. Verschaffelt, and G. Van der Sande, Noise-injected analog Ising machines enable ultrafast statistical sampling and machine learning, *Nat. Commun.* **13**, 5847 (2022).
- [31] T. Kanao and H. Goto, Simulated bifurcation assisted by thermal fluctuation, *Commun. Phys.* **5**, 153 (2022).
- [32] K. Lee, S. Chowdhury, and K. Y. Camsari, Noise-augmented chaotic Ising machines for combinatorial optimization and sampling, *Commun. Phys.* **8**, 35 (2025).
- [33] H. Goto, R. Hidaka, and K. Tatsumura, Edge-of-chaos-enhanced quantum-inspired algorithm for combinatorial optimization, *Phys. Rev. Appl.* **25**, 044011 (2026).
- [34] A. E. Alaoui and A. Montanari, Algorithmic thresholds in mean field spin glasses, *arXiv:2009.11481*.
- [35] A. Alaoui, A. Montanari, and M. Sellke, Optimization of mean-field spin glasses, *Ann. Probab.* **49** (2021).
- [36] A. Montanari, Optimization of the Sherrington-Kirkpatrick Hamiltonian, *SIAM J. Comput.* **54**, FOCS19-1 (2025).
- [37] D. Sherrington and S. Kirkpatrick, Solvable model of a spin-glass, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [38] S. Kirkpatrick, Frustration and ground-state degeneracy in spin glasses, *Phys. Rev. B* **16**, 4630 (1977).
- [39] S. Boettcher, Extremal optimization for Sherrington-Kirkpatrick spin glasses, *Eur. Phys. J. B* **46**, 501 (2005).

- [40] S.-Y. Kim, S. J. Lee, and J. Lee, Ground-state energy and energy landscape of the Sherrington-Kirkpatrick spin glass, *Phys. Rev. B* **76**, 184412 (2007).
- [41] M. Palassini, Ground-state energy fluctuations in the Sherrington-Kirkpatrick model, *J. Stat. Mech.* (2008) P10005.
- [42] H. Sompolinsky and A. Zippelius, Dynamic theory of the spin-glass phase, *Phys. Rev. Lett.* **47**, 359 (1981).
- [43] H. Sompolinsky and A. Zippelius, Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses, *Phys. Rev. B* **25**, 6860 (1982).
- [44] A. Auffinger and W.-K. Chen, The Parisi formula has a unique minimizer, *Commun. Math. Phys.* **335**, 1429 (2015).
- [45] M. Talagrand, The Parisi formula, *Ann. Math.* **163**, 221 (2006).
- [46] P. C. Martin, E. D. Siggia, and H. A. Rose, Statistical dynamics of classical systems, *Phys. Rev. A* **8**, 423 (1973).
- [47] C. Chow and M. Buice, Path integral methods for stochastic differential equations, *J. Math. Neurosci.* **5**, 8 (2015).
- [48] D. Grzetic, R. Wickham, and A.-C. Shi, Statistical dynamics of classical systems: A self-consistent field approach, *J. Chem. Phys.* **140**, 244907 (2014).
- [49] L. F. Cugliandolo and J. Kurchan, Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model, *Phys. Rev. Lett.* **71**, 173 (1993).
- [50] The nonmonotonicity observed in the theoretical results at later evolutionary stages in Fig. 1 is attributed to the finite sampling inherent in the Monte Carlo method. With increasing time, the response matrix grows quadratically, while the near-zero density diminishes. This leads to increased variance in the estimator $\langle \partial \text{sgn} x(t) / \partial h(t') \rangle$, resulting in statistical fluctuations.
- [51] M. Oppen and D. Saad, *Advanced Mean Field Methods: Theory and Practice* (MIT Press, Cambridge, MA, 2001).
- [52] Because $G_i(t, t')$ depends on the full trajectory, the figure appears as a scatter plot.
- [53] J. Wang, D. Ebler, K. Y. M. Wong, D. S. W. Hui, and J. Sun, Bifurcation behaviors shape how continuous physical dynamics solves discrete Ising optimization, *Nat. Commun.* **14**, 2510 (2023).
- [54] S. O. Rice, Mathematical analysis of random noise, *Bell Syst. Tech. J.* **23**, 282 (1944).
- [55] We utilized the 1RSB formula derived in our previous work to generate the curves [29]. The theoretical values may exhibit a small bias relative to the true full-RSB results, but they are sufficient to capture the overall trend. More precise but qualitatively similar results can be obtained by running Monte Carlo simulations according to Eq. (3) for sufficiently long durations and large samples.
- [56] This also suggests choosing the gain at the point where the decoded energy contour is tangent to the effective gap contour.
- [57] The result $(a, T) = (0, 0)$ includes the full-RSB correction. See Ref. [29].
- [58] T. Aspelmeier, A. Billoire, E. Marinari, and M. Moore, Finite size corrections in the Sherrington-Kirkpatrick model, *J. Phys. A: Math. Theor.* **41**, 205005 (2008).
- [59] S. Boettcher, Simulations of ground state fluctuations in mean-field Ising spin glasses, *J. Stat. Mech.* (2010) P07002.
- [60] R. Oppermann, M. J. Schmidt, and D. Sherrington, Double criticality of the Sherrington-Kirkpatrick model at $t = 0$, *Phys. Rev. Lett.* **98**, 127201 (2007).
- [61] H.-K. Janssen, On a Lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties, *Z. Phys. B* **23**, 377 (1976).
- [62] C. De Dominicis, Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques, *J. Phys. Colloq.* **37**, C1-247 (1976).
- [63] C. De Dominicis and L. Peliti, Field-theory renormalization and critical dynamics above T_c : Helium, antiferromagnets, and liquid-gas systems, *Phys. Rev. B* **18**, 353 (1978).
- [64] R. Bausch, H.-K. Janssen, and H. Wagner, Renormalized field theory of critical dynamics, *Z. Phys. B* **24**, 113 (1976).
- [65] Y. Ye, The Gset dataset, 2003, <http://web.stanford.edu/%7Eyye/yye/Gset/>.
- [66] F. Ghimenti, A. Sriram, A. Yamamura, H. Mabuchi, and S. Ganguli, Geometry and dynamics of annealed optimization in the coherent Ising machine with hidden and planted solutions, *Phys. Rev. E* **113**, 054123 (2026).