

The geometry and dynamics of annealed optimization in the coherent Ising machine with hidden and planted solutions

Federico Ghimentì,¹ Adithya Sriram,² Atsushi Yamamura (山村 篤志),¹ Hideo Mabuchi,¹ and Surya Ganguli¹

¹*Department of Applied Physics, Stanford University, Stanford, CA 94305, USA*

²*Department of Physics, Stanford University, Stanford, CA 94305, USA*

The coherent Ising machine (CIM) is a nonconventional hardware architecture for finding approximate solutions to large-scale combinatorial optimization problems. It operates by annealing a laser gain parameter to adiabatically deform a high-dimensional energy landscape over a set of soft spins, going from a simple convex landscape to the more complex optimization landscape of interest. We address how the evolving energy landscapes guides the optimization dynamics against problems with hidden planted solutions. We study the Sherrington-Kirkpatrick spin-glass with ferromagnetic couplings that favor a hidden configuration by combining the replica method, random matrix theory, the Kac-Rice method and dynamical mean field theory. We characterize energy, number, location, and Hessian eigenspectra of global minima, local minima, and critical points as the landscape evolves. We find that low energy global minima develop soft-modes which the optimization dynamics can exploit to descend the energy landscape. Even when these global minima are aligned to the hidden configuration, there can be exponentially many higher energy local minima that are all unaligned with the hidden solution. Nevertheless, the annealed optimization dynamics can evade this cloud of unaligned high energy local minima and descend near to aligned lower energy global minima. Eventually, as the landscape is further annealed, these global minima become rigid, terminating any further optimization gains from annealing. We further consider a second optimization problem, the Wishart planted ensemble, which contains a hidden planted solution in a landscape with tunable ruggedness. We describe CIM phase transitions between recoverability and non-recoverability of the hidden solution. Overall, we find intriguing relations between high-dimensional geometry and dynamics in analog machines for combinatorial optimization.

CONTENTS

		B. Dynamical mean field theory of the annealed optimization dynamics	20
I. Statistical physics for analog computing	2	IX. Wishart Planted Ensemble	22
II. Model	4	A. Low laser gain	24
III. Heuristics and phase diagram for global minima	5	B. Resolvent and single-spin distribution	24
IV. Stability of global minimum at low laser gain	6	C. Convex phase	25
V. Computation of the resolvent of the Hessian	8	D. Dynamics	27
A. Lower edge of the spectrum	10	X. Outlook	28
VI. The structure of global minima and their Hessian eigenspectra	10	Acknowledgments	29
A. Convex phase	10	A. Derivation of Eq. (25): Computing the resolvent	29
B. Spin-glass phase and rigidity transition	12	B. Saddle point equations	31
VII. Proliferation of paramagnetic local minima and critical points	14	C. Replica symmetric ansatz	32
A. Phase diagram of most abundant minima and critical points	16	D. Stability of the replica symmetric solution	33
B. A portrait of critical points at large laser gain	17	E. Derivation of Eq. (44)	33
C. Order parameters as reactivities of the grand-potential	18	F. Outline of the derivation of the full replica symmetry breaking equations	33
VIII. Annealed optimization dynamics of the CIM	19	G. Replica calculation of the grand-potential	37
A. Numerical simulations	19	H. Dynamical mean field theory	47
		I. Wishart planted ensemble	50

I. STATISTICAL PHYSICS FOR ANALOG COMPUTING

The solution of combinatorial optimization problems [1] is central to many domains of mathematics and science, ranging from information theory and computer science to robotics and circuit design. The development of algorithms and technologies capable of reducing the computational cost of finding solutions is thus of fundamental importance. In recent years, unconventional hardware architectures [2–5] that find approximate solutions through the nonlinear dynamical evolution of analog systems, have been developed and implemented to tackle NP-hard problems. The potential advantage of these approaches over conventional digital systems lies in exploiting the exotic analog dynamics they can implement, as well as in the possibility of scaling up these devices to unprecedented sizes. Among these architectures, the coherent Ising machine (CIM) [6] stands out, as in the last decade the number of degrees of freedom implemented in a single device has grown from a hundred [7] to thousands [8] to tens of thousands [9].

The CIM [10] is a heuristic solver for quadratic, unconstrained binary optimization (QUBO) problems, involving the minimization of an energy function over N pairwise interacting binary spins. Many problems of practical interest can be formulated in this way, including the traveling salesman problem, the partitioning problem and the graph coloring problem [11]. Rather than minimizing an energy function over a set of discrete variables, the CIM minimizes an energy function over a set of continuous degrees of freedom, or soft spins, within an optical network. Moreover, as a certain laser gain parameter is slowly annealed over time, the single spin energy landscape evolves from a single well quartic potential, to a soft, wide double well potential, to a sharp or rigid double well potential. This in turn induces the high dimensional CIM energy landscape geometry over all spins to also adiabatically evolve, as the soft spins are forced to become binary spins. The final endpoint of this evolution is mapped back to a binary spin configuration yielding an approximate solution to the original optimization problem of interest.

Despite the proliferation of a diverse set of technological advances [10, 12–16], we currently lack a theoretical framework to understand the success and failure modes of the CIM, and of analog computing systems in general. On the other hand, from the perspective of statistical physics, the problem of solving combinatorial optimization problems is equivalent to the problem of finding ground-state energy configurations in high-dimensional disordered systems. Indeed an arsenal of tools from spin-glass theory [17–20] have been developed and applied, not only to optimization problems, but also to information theory [21–23], machine learning [24–26], and ecol-

ogy and evolution [27–31]. The techniques developed yield insights into the structure of ground states and metastable states of the energy landscape of disordered systems, as well as their dynamical properties. These techniques, including the replica method, cavity method, Kac-Rice formula, and dynamic mean field theory, have only recently started to find applications and extensions to the realm of analog computing devices [32]. In particular, we still do not have an understanding of how the CIM energy landscape geometry and dynamics changes under annealing, when the connectivity of the system is non-random enough so as to bias low energy configurations towards a hidden, or planted configuration originating from structure in the connectivity.

In this work, we leverage techniques from the statistical physics of disordered systems to understand the energy landscape geometry and the efficacy of the CIM in solving two prototypical NP-hard problems with such hidden, planted configurations. We first consider the case of the Sherrington-Kirkpatrick (SK) model with additional rank 1 ferromagnetic coupling [33], and we present a phase diagram for the properties of the global minima and most abundant local minima as the laser gain annealing parameter and the strength of the ferromagnetic coupling are varied.

Our main finding for this model is schematically depicted in Fig. 1. We unveil a rich phase diagram, for both global and local minima. At low value of the laser gain, the ground state of the CIM is either a paramagnet or a ferromagnet, depending on the strength of the ferromagnetic coupling. Further increase of the laser gain increases the ruggedness of the landscape, leading to a transition into a spin-glass phase. The latter can be classified into two distinct categories, depending on the features of the nonlinear excitations around global minima. Indeed, for intermediate values of laser gain and ferromagnetic coupling, the spectrum of the Hessian at global minima is gapless, indicating many soft modes, and the support of the distribution of the soft spin population spans the whole real axis, with a nonzero value of its density at the origin. We call this phase a ‘soft’ spin-glass. Upon further increase of the laser gain, the spectrum of the Hessian at global minima develops a gap away from 0, and the support of the distribution of the population of soft spins splits into two disconnected domains. This corresponds to a rigid spin-glass phase, where in global minima, there is an absence of soft modes and all spins are strongly committed to finite nonzero values. We show that the soft-to-rigid transitions occurs at a value of the laser gain for which the effective energy of a mean-field model that we derive becomes non-convex.

Furthermore, by a combination of dynamical mean field theory and numerical simulations, we discover that the two spin-glass phases have differential impacts on the efficacy of the annealed optimization dynamics. Indeed, the properties of the critical points visited during the annealing process are very similar to those of global minima. Thus in the soft spin-glass phase, linear excitations

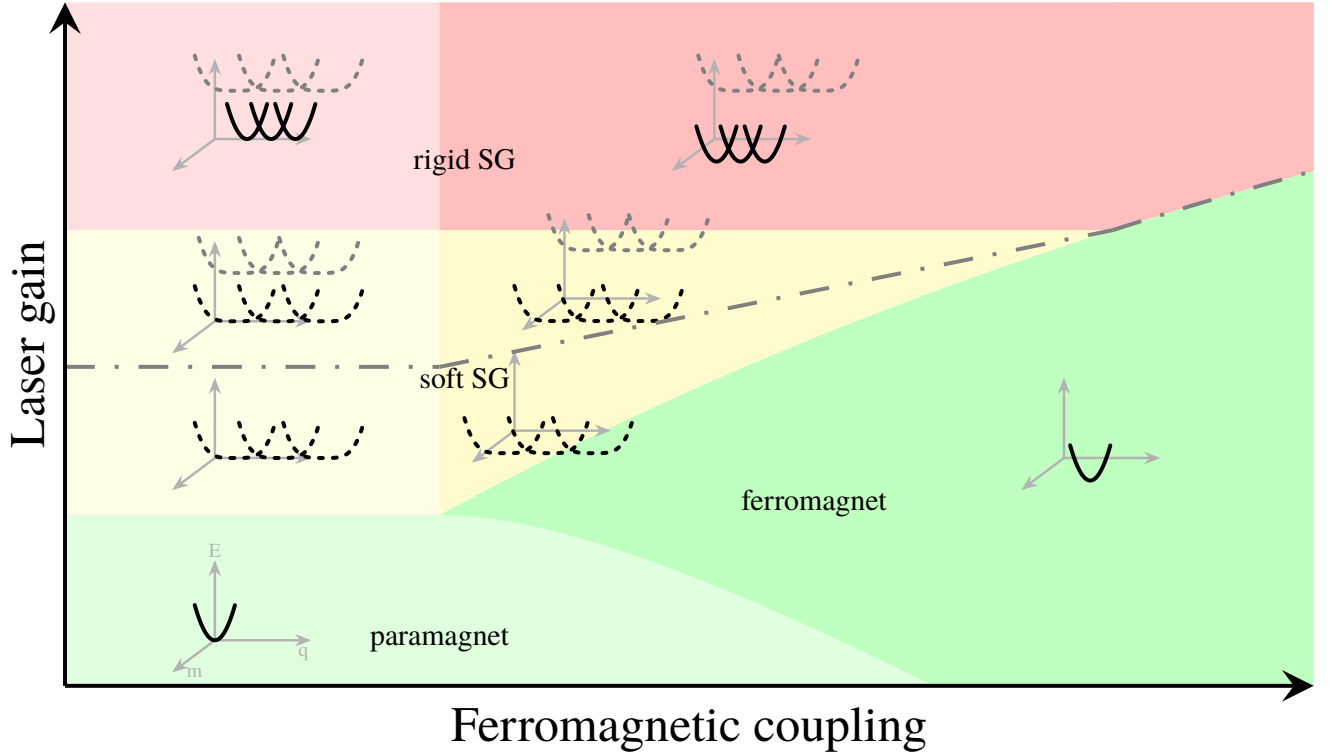


FIG. 1. **Schematic phase diagram for global minima and most abundant local minima of the coherent Ising machine.** In each phase, we summarize the properties of global (black curves) and local (grey curves) minima by their positions in a three-dimensional coordinate system indicating their typical energy E , square distance from the origin q , and the net magnetization m . Additionally, we indicate whether either minima are soft (wide dashed curves) or rigid (sharp solid curves). At low laser gain and ferromagnetic coupling (Sec. IV), the system is in a paramagnetic phase, with a single rigid global minimum located at the origin (light green region). If the ferromagnetic coupling is large enough, further increase of the laser gain leads to a transition of the Baik-Ben Arous Peché type toward a ferromagnetic phase (dark green region, Sec. VI A). As the laser gain increases further, both the paramagnetic and the ferromagnetic phases become unstable toward a spin-glass (SG) phase (yellow and red regions). This overall SG phase can be subdivided into two phases, according to the behavior of the spectrum of the Hessian at global minima (analyzed in Sec. VI B). For intermediate values of laser gain and ferromagnetic couplings, the lower edge of the Hessian at global minima touches the origin yielding a 'soft' spin-glass (light paramagnetic and dark ferromagnetic yellow regions). At larger values of the laser gain or ferromagnetic coupling, the spectrum of the Hessian at global minima is gapped away from the origin, yielding a rigid spin-glass (light paramagnetic and dark ferromagnetic red regions). Using a replicated Kac-Rice calculation (Sec. VII), we further add to this picture by studying the properties of the most abundant local minima across the phase diagram. This yields a further subdivision of the soft spin-glass phase by a critical line (gray dash-dotted line) into two sub-phases. Above the critical line (but not below), local minima proliferate at an exponential rate with the size of the system. Interestingly, these minima are paramagnetic for *any* value of the ferromagnetic coupling, unlike the global minima, exist at higher energy than the global minima, and they are soft, with a gapless Hessian eigenspectrum. Furthermore, these exponentially many higher energy paramagnetic local minima persist into rigid spin-glass phase, and they retain their soft character there, despite the fact that global minima become rigid (light and dark red regions). Numerical simulations and dynamical mean field theory (Sec. VIII) demonstrate that the annealed optimization dynamics of the CIM visits spin configurations that are much more similar to global minima than to the most abundant local minima. Importantly, annealing across the soft SG phase allows the CIM to evade higher energy local minima and reach better approximate solutions to the original combinatorial optimization problem, by exploiting soft modes of near global minima to descend. But when the annealed optimization dynamics enters the rigid SG phase, the dynamics finally terminates and the energy of the solution to the combinatorial optimization problem does not decrease further.

corresponding to soft modes with small Hessian eigenvalues are exploited by the CIM to flip the sign of uncommitted spins with values near 0, allowing the CIM to reach approximate solutions of the QUBO problem with lower and lower energies as the annealing proceeds.

On the other hand, in the rigid spin-glass phase, further annealing does not yield improved solutions. Here, no soft modes exist to be exploited, all the spins are committed to nonzero values, and the CIM gets stuck in a single configuration. The analytical determination of the

boundaries between the soft and rigid spin-glass phases and the importance of this soft to rigid phase transition in governing the efficacy of the CIM annealed optimization dynamics, is one of the main technical results of this work.

We also push our analysis of the CIM energy landscape geometry beyond the structure of global minima. Indeed, we characterize, by means of a replicated Kac-Rice calculation, the properties of the most abundant local minima and critical points of the CIM as a function of laser gain and the ferromagnetic coupling. We identify a linear phase transition boundary within the soft spin-glass phase (dashed-dotted line in the yellow regions of 1, separating a phase where local minima and critical points grow subexponentially with system size below the line, and a phase where they proliferate exponentially with system size above the line. Interestingly, the properties of typical critical points and local minima are significantly different from those visited by the dynamics, both in terms of average energy and magnetization. In particular, in the phase where their number grows at an exponential rate with the system size, the most abundant local minima are always paramagnetic with soft, gapless Hessian eigenspectra. Overall, the picture that emerges from our analysis is that as the laser gain is annealed from low to high, the CIM descends through the energy landscape by exploiting the soft modes in the vicinity of global minima, thereby passing beneath an exponentially large cloud of proliferating higher energy local minima when they first appear.

Last, but not least, we study the global minima of the CIM with the Wishart planted ensemble [34], a model with a predefined ground state and tunable ruggedness. The main results of our analysis of this model are presented in the phase diagram in Fig. 9. At low laser gain, the CIM has a single global minimum at the origin, independently of the ruggedness of the original combinatorial optimization problem. If the parameter controlling the ruggedness of the combinatorial optimization problem is sufficiently large, increasing the laser gain leads initially to a ferromagnetic phase, where full recovery of the planted solution is possible. However, in this phase, spectral methods are equally effective in retrieving the planted solution. If the laser gain is increased further, the global minima of the CIM undergoes a spin-glass transition, no matter whether the planted solution in the combinatorial optimization problem was originally in a hard-to-retrieve or easy-to-retrieve phase. In the easy-to-retrieve phase, spectral initialization of the CIM allows recovery of the planted solution, despite the spin-glass structure of the global minima. In the hard-to-retrieve phase, at large value of the laser gain, the annealing dynamics of the CIM obtains approximate solutions with a lower energy compared to a spectral method.

Our paper is organized as follows. In Sec. II we introduce the model of the CIM with a Sherrington-Kirkpatrick connectivity matrix. In Sec. III we give a heuristic argument for the importance of soft modes in

the annealing process and present the phase diagram for the global minima of the CIM. After a first look at the properties of global minima for small values of the annealing parameter in Sec. IV, we present a replica calculation of the resolvent of the Hessian at global minima, and illustrate how the lower edge of the spectrum can be determined in Sec. V. In Sec. VI we study the properties of the convex and spin-glass phases of the system. In Sec. VII we address the properties of the most abundant local minima of the system through a replicated Kac-Rice calculation, and identify a phase transition toward a phase where typical paramagnetic minima proliferate. In Sec. VIII we address the dynamics of the model, showing that the properties of the ground state are a good proxy for the performance of the system. In Sec. IX, we move to the case of the Wishart planted ensemble, and through a replica calculation and numerical simulations of the annealing process we identify convex and spin-glass phases for the CIM applied to this problem. Finally we present conclusions and future directions in Sec. X.

II. MODEL

The combinatorial optimization problem of interest in this work is finding the ground state of the Ising energy function E_{Ising}

$$E_{\text{Ising}}(\mathbf{s}) = -\frac{1}{2} \mathbf{s} \cdot \mathbf{J}_{\text{Ising}} \mathbf{s}, \quad (1)$$

defined over N binary spin variables $\mathbf{s} \in \{-1, 1\}^N$. The connectivity matrix $\mathbf{J}_{\text{Ising}}$ depends on the specific problem at hand, but it is usually a large matrix with entries of different signs, for which the energy landscape of E_{Ising} can be highly rugged. Instead of directly searching for the ground state of Eq. (1), the CIM minimizes the energy $E(\mathbf{x}, a)$ of a system of N soft spins x_i . These soft spins are obtained by measuring the phase differences between optical parametric oscillators in an optical network with respect to a given laser pumping frequency. The softness of each spin is controlled by the strength of the laser gain a , which is adiabatically changed during the optimization process. The total energy $E(\mathbf{x}, a)$, is given by

$$E(\mathbf{x}, a) \equiv \sum_{i=1}^N E_I(x_i, a) - \frac{1}{2} \mathbf{x} \cdot \mathbf{J}_{\text{Ising}} \mathbf{x}. \quad (2)$$

The laser gain strength controls the shape of the single-site internal energy $E_I(x, a)$ defined as

$$E_I(x, a) \equiv \frac{1}{4} x^4 - \frac{a}{2} x^2. \quad (3)$$

When $a < 0$, $E_I(x, a)$ is a convex function of x , with a single minimum located at the origin. When $a > 0$, the function E_I is a double-well potential, with two minima located at $\pm\sqrt{a}$, separated by an energy barrier of height $\frac{1}{2}a^2$. For a very low and negative value of the laser gain,

$a \ll 0$, at low energy, the soft spins are localized around the origin, while for a large and positive value of the laser gain, $a \gg 0$, the soft spins are akin to binary variables, as the single site energy E_I favors them to reside within one of the two wells. For a fixed value of a , the CIM dynamics can be approximated by minimization of the energy through the gradient descent dynamics

$$\tau \frac{d}{dt} \mathbf{x}(t) = -\nabla E(\mathbf{x}(t), a), \quad (4)$$

where τ is a microscopic timescale of the system and where $\nabla = [\partial_{x_1}, \partial_{x_2}, \dots, \partial_{x_N}]^T$ is the gradient operator in the space of the soft spin variables. The system is initialized from a random configuration at a low value of the laser gain $a \ll 0$. The laser gain is then increased adiabatically, with a timescale much slower than the one required by the system to converge to a local minimum through the gradient descent dynamics in Eq. (4). The annealing of the laser gain proceeds further up to a final value of $a \gg 0$, where the single-site energy E_I given by Eq. (3) constrains the spins x_i to localize around the positions $\pm\sqrt{a}$. At the end of the annealing process, the final configuration of the soft-spin system is transformed into an Ising spin configuration through the mapping $s_i = \text{sgn } x_i$. The CIM exploits the soft nature of the continuous spins x_i to reach low energy configurations of the optimization problem given by Eq. (1).

We work with a connectivity matrix $\mathbf{J}_{\text{Ising}}$ for which the ground state of the corresponding combinatorial optimization problem has a partial or complete alignment with a preferential direction $\mathbf{t}$. The latter is chosen, without loss of generality, to be $\mathbf{t} \equiv [1, \dots, 1]^T$. The first model we consider is given by $\mathbf{J}^{\text{Ising}} = \mathbf{J}^{\text{SK}}$, where $\mathbf{J}^{\text{SK}}$ is the connectivity matrix of the Sherrington-Kirkpatrick spin-glass with a finite ferromagnetic alignment J_0 [33]. It is obtained summing a rank-1 matrix and a random matrix $\mathbf{J}$,

$$\mathbf{J}^{\text{SK}} = \frac{J_0}{N} \mathbf{t} \otimes \mathbf{t} + \frac{1}{\sqrt{N}} \mathbf{J}. \quad (5)$$

The matrix $\mathbf{J}$ is drawn from the Gaussian orthogonal ensemble [35]: it is a symmetric matrix whose entries are i.i.d Gaussian variables with zero mean and variance $\langle J_{ij}^2 \rangle_{\mathbf{J}} = 1 + \delta_{ij}$. We denote by $\langle \dots \rangle_{\mathbf{J}}$ the average over different realizations of the disordered couplings $\mathbf{J}$. The strength J_0 of the rank-1 perturbation favors ferromagnetic ordering, and it competes against the disordered interactions in the system, encoded by $\mathbf{J}$. For $J_0 < 1$, the ground state of the Ising energy in Eq. (1) with $\mathbf{J}^{\text{Ising}} = \mathbf{J}^{\text{SK}}$ is paramagnetic, while for $J_0 > 1$, the ground state is ferromagnetic, or partially aligned along the direction $\mathbf{t}$ [18, 36]. The resulting Ising model is a paradigmatic example in the theory of spin-glasses [17]. It presents a rugged energy landscape, with the number of minima close to the ground state growing exponentially with the size of the system. These low energy minima are organized in a complex hierarchical structure,

whose description, achieved by means of the full replica symmetry breaking theory, is a landmark achievement in the theory of disordered systems [37, 38]. An efficient message-passing algorithm to find solutions arbitrarily close to the ground-state has been recently developed [39], but the Sherrington-Kirkpatrick spin-glass constitutes nevertheless an important target on which to derive and test our understanding of theoretical principles governing the performance and evolving energy landscape geometry of the CIM.

In what follows, we denote by $E^{\text{SK}}(\mathbf{x}, a)$ the energy of the CIM when the connectivity matrix is given by $\mathbf{J}^{\text{SK}}$, and we refer to this model as CIM-SK. Thus we focus on the energy $E^{\text{SK}}(\mathbf{x}, a)$ given by

$$E^{\text{SK}}(\mathbf{x}, a) \equiv \sum_{i=1}^N E_I(x_i, a) - \frac{1}{2} \mathbf{x} \cdot \mathbf{J}^{\text{SK}} \mathbf{x}, \quad (6)$$

with $\mathbf{J}^{\text{SK}}$ given by Eq. (5). The main idea we put forward in this work is that the distribution of eigenvalues of the Hessian of E^{SK} around low energy minima is a good proxy for determining when the annealing process is effective. In the following Section, we provide a heuristic argument motivating this claim.

III. HEURISTICS AND PHASE DIAGRAM FOR GLOBAL MINIMA

Let us consider the annealing dynamics from a perspective complementary to the gradient descent given by Eq. (4), by looking at how its fixed points move as the laser gain increases. Consider a critical point $\mathbf{x}_c$ of the CIM-SK, such that $\nabla E^{\text{SK}}(\mathbf{x}, a)|_{\mathbf{x}=\mathbf{x}_c} = \mathbf{0}$. We can track the motion of the critical point as the laser gain is infinitesimally changed by taking the derivative of the stationary condition with respect to a , thus obtaining

$$\mathbf{H}^{\text{SK}}(\mathbf{x}_c, a) \frac{d\mathbf{x}_c}{da} = \mathbf{x}_c(a), \quad (7)$$

where we introduced the Hessian matrix of the CIM-SK energy $\mathbf{H}^{\text{SK}}(\mathbf{x}, a)$, defined as

$$H_{ij}^{\text{SK}}(\mathbf{x}, a) \equiv \frac{\partial^2 E^{\text{SK}}(\mathbf{x}, a)}{\partial x_i \partial x_j}. \quad (8)$$

Equation (7) describes the motion of a critical point during the annealing process, assuming that critical points respond linearly to an infinitesimal change of the laser gain a . This is true if the Hessian $\mathbf{H}^{\text{SK}}(\mathbf{x}_c, a)$ is invertible. However, if the Hessian is not invertible, as is the case if its spectrum contains a zero eigenvalue, then Eq. (7) no longer holds, and we expect the critical point to jump discontinuously to a new configuration as the laser gain increases. Under gradient descent dynamics, the system then flows to a novel critical point. This sudden change might involve a change in the sign of some of the soft spins x_i , possibly allowing to reach a lower Ising energy.

The argument above suggests that an interesting proxy for the effectiveness of the annealing process is given by the distribution of eigenvalues of the Hessian at the critical points visited by the CIM during the annealing dynamics. In general, the relationship between the geometry of the energy landscape of a disordered system and its dynamics during a rapid quench or a slow annealing process is a broad, long-standing and challenging question, for which unambiguous answers have been provided mainly in pure spin-glass models on spheres [40]. To further characterize the performance of the CIM and its connection with the geometry of its changing landscape, we investigate the properties of the deepest energy minima of the model. These minima are typically unreachable on a reasonable timescale, but after their characterization we can ask whether they leave some trace at higher energy levels, which can affect the performance of the CIM.

In the following, we analyze the properties of global minima and the spectrum of the associated Hessian as a function of the laser gain a and the ferromagnetic alignment strength J_0 . The main result of our analysis is presented in Fig. 2. In panel (a), we reveal a phase diagram of the global minima of the system. The phase diagram exhibits four different regions, characterized by different properties of the distribution $P^{\text{SK}}(x)$ of single spin values, and the spectrum of the Hessian $\rho^{\text{SK}}(\lambda)$. In region $\textcircled{\text{I}}$, there is a single, paramagnetic global minimum: the energy landscape is convex, and the spin distribution is concentrated around the origin. In region $\textcircled{\text{II}}$, the system is a ferromagnet, with two global minima related by an inversion symmetry, and the global minimum of the system is ferromagnetic. The single spin distribution is skewed (panel (c)). In region $\textcircled{\text{III}}$ and $\textcircled{\text{IV}}$ the global minima have a spin-glass structure, with a rugged energy landscape, but the properties of the Hessian and of the single spin distribution at the global minima are different in the two phases. In phase $\textcircled{\text{III}}$, the single spin distribution has a nonzero probability density at the origin, and the spectrum of the Hessian is gapless (panel (d)). We refer to this phase as a non-rigid spin-glass. We designate the boundary between the convex phase and the non-rigid spin-glass phase as a increases by the curve $a_{\text{sg}}(J_0)$, the critical value of a at which the transition occurs for a given J_0 . In phase $\textcircled{\text{IV}}$, the spectrum of the Hessian is gapped away from the origin, and the single spin distribution has zero probability density at the origin (panel (e)). We refer to this phase as a rigid spin-glass. We denote by $a_r(J_0)$ the value of the laser gain at which this rigidity transition first occurs as a increases. Together these results, discussed further below, begin to justify our schematic description introduced in Fig. 1 with regions $\textcircled{\text{I}}$, $\textcircled{\text{II}}$, $\textcircled{\text{III}}$, and $\textcircled{\text{IV}}$, corresponding to the light green, dark green, light and dark yellow, and light and dark red regions in Fig. 1 respectively.

Interestingly, upon increasing the laser gain at fixed J_0 , we see that two scenarios are possible, depending on the strength of the ferromagnetic coupling J_0 . At lower

values of J_0 , we have $a_{\text{sg}}(J_0) < a_r(J_0)$: as a increases, the system can transition from a convex phase, $\textcircled{\text{I}}$ or $\textcircled{\text{II}}$, to a non-rigid spin-glass phase $\textcircled{\text{III}}$, and then into the rigid spin-glass phase $\textcircled{\text{IV}}$. Alternatively, for larger values of J_0 , we can have $a_{\text{sg}}(J_0) = a_r(J_0)$: the system transitions from a convex phase $\textcircled{\text{II}}$ to a rigid-spin-glass phase $\textcircled{\text{IV}}$. We next describe how we obtain these results.

IV. STABILITY OF GLOBAL MINIMUM AT LOW LASER GAIN

Let us consider the case where the laser gain is very low, $a \ll 0$. In this case, the single-site energy E_I in Eq. (3) is convex and dominates over the interaction term. We thus expect the system to have a single critical point at the origin, $\nabla E^{\text{SK}}(\mathbf{x} = \mathbf{0}, a) = \mathbf{0}$, which is also the global minimum. The Hessian $H_{ij}^{\text{SK}}(\mathbf{x}, a)$ of the CIM-SK is

$$\begin{aligned} H_{ij}^{\text{SK}}(\mathbf{x}, a) &= \delta_{ij} \partial_x^2 E_I(x_i, a) - J_{ij}^{\text{SK}} \\ &= \delta_{ij} (3x_i^2 - a) - J_{ij}^{\text{SK}} \end{aligned} \quad (9)$$

The typical shape of the landscape in the vicinity of a configuration $\mathbf{x}$ can be understood from the distribution of the eigenvalues $\rho_{\mathbf{x}}^{\text{SK}}(\lambda)$ of $\mathbf{H}^{\text{SK}}(\mathbf{x}, a)$, obtained by averaging over realizations of the connectivity matrix $\mathbf{J}$,

$$\rho_{\mathbf{x}}^{\text{SK}}(\lambda) \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \langle \delta(\lambda_i(\mathbf{x}) - \lambda) \rangle_{\mathbf{J}}, \quad (10)$$

where $\{\lambda_i(\mathbf{x})\}_{i=1}^N$ is the set of eigenvalues of $\mathbf{H}^{\text{SK}}(\mathbf{x}, a)$. When evaluated at $\mathbf{x} = \mathbf{0}$, The Hessian becomes

$$\begin{aligned} \mathbf{H}^{\text{SK}}(\mathbf{0}, a) &= -a\mathbf{I} - \mathbf{J}^{\text{SK}} \\ &= -a\mathbf{I} - \frac{1}{\sqrt{N}}\mathbf{J} - \frac{J_0}{N}\mathbf{t} \otimes \mathbf{t}. \end{aligned} \quad (11)$$

In the second equality we have decomposed the Hessian into a sum of a diagonal matrix, a random matrix from the Gaussian orthogonal ensemble and a rank-1 perturbation. The spectral properties of this class of random matrix have been widely studied. For $J_0 = 0$, the spectrum $\rho_{\mathbf{0}}^{\text{SK}}(\lambda)$ is given by Wigner semicircle law [41] centered around $-a$,

$$\rho_{\mathbf{0}}^{\text{SK}}(\lambda) = \frac{\sqrt{(2-a-\lambda)(2+a+\lambda)}}{2\pi}, \quad (12)$$

with support $\lambda \in [-2-a, 2-a]$. For the origin to be a minimum, the lower edge of the support of the spectrum must be positive. This condition is violated for $a > -2$. Above this laser gain value, the origin is no longer a minimum of the system, and we expect a bifurcation in the energy landscape leading to new minima. When $J_0 > 0$, the ferromagnetic coupling can be thought of as a signal, while the random Gaussian couplings can be thought of as noise. The competition between these two elements,

studied using the low temperature limit of the Hessian eigenspectrum $\rho^{\text{SK}}(\lambda)$, defined as

$$\begin{aligned}\rho^{\text{SK}}(\lambda) &\equiv \lim_{N \rightarrow \infty} \frac{1}{N} \int d\mathbf{x} \left\langle P_{\text{B}}^{\text{SK}}(\mathbf{x}) \sum_{i=1}^N \delta(\lambda_i(\mathbf{x}) - \lambda) \right\rangle_{\mathbf{J}} \\ &\equiv \lim_{N \rightarrow \infty} \left\langle \frac{1}{N} \sum_{i=1}^N \delta(\lambda_i(\mathbf{x}) - \lambda) \right\rangle_{\mathbf{x}, \mathbf{J}}^{\text{SK}}\end{aligned}\quad (16)$$

and taking the limit $\beta \rightarrow \infty$. The second equality defines the average over the Boltzmann distribution $\langle \dots \rangle_{\mathbf{x}}^{\text{SK}} \equiv \int d\mathbf{x} \dots P_{\text{B}}^{\text{SK}}(\mathbf{x})$. The spectrum $\rho^{\text{SK}}(\lambda)$ can be obtained from the resolvent $G^{\text{SK}}(z)$, defined as

$$\begin{aligned}G^{\text{SK}}(z) &\equiv \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \text{Tr} [z - \mathbf{H}^{\text{SK}}(\mathbf{x})]^{-1} \right\rangle_{\mathbf{x}, \mathbf{J}}^{\text{SK}} \\ &= \int d\lambda \frac{\rho^{\text{SK}}(\lambda)}{z - \lambda}.\end{aligned}\quad (17)$$

This formula can be inverted to express the spectrum $\rho(\lambda)$ as a function of the resolvent [35]

$$\rho^{\text{SK}}(\lambda) = \lim_{\epsilon \rightarrow 0} \frac{G^{\text{SK}}(\lambda - i\epsilon) - G^{\text{SK}}(\lambda + i\epsilon)}{2\pi i}, \quad (18)$$

so that by computing the resolvent in the low temperature limit we can obtain the spectrum of the Hessian of global minima of the CIM machine. We next compute this resolvent.

V. COMPUTATION OF THE RESOLVENT OF THE HESSIAN

We compute the resolvent following the approach of [43, 44]. We first note that from its definition in Eq. (17), the resolvent can be rewritten as¹

$$\begin{aligned}G^{\text{SK}}(z) &= \lim_{N \rightarrow \infty} \frac{1}{N} \frac{d}{dz} \langle \text{Tr} \log [z - \mathbf{H}^{\text{SK}}(\mathbf{x}, a)] \rangle_{\mathbf{x}, \mathbf{J}}^{\text{SK}} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \frac{d}{dz} \lim_{l \rightarrow 0} \partial_l \langle \log Z_l^{\text{SK}} \rangle_{\mathbf{J}},\end{aligned}\quad (19)$$

where we have introduced a modified partition function Z_l^{SK} ,

$$\begin{aligned}Z_l^{\text{SK}} &\equiv \int d\mathbf{x} \exp \left[-\beta E^{\text{SK}}(\mathbf{x}, a) \right. \\ &\quad \left. + l \text{Tr} \log [z - \mathbf{H}^{\text{SK}}(\mathbf{x}, a)] \right].\end{aligned}\quad (20)$$

To compute the average over the realizations of the disordered couplings, we resort to the replica trick:

$$\langle \log Z_l^{\text{SK}} \rangle_{\mathbf{J}} = \lim_{n \rightarrow 0} \frac{1}{n} \log \langle (Z_l^{\text{SK}})^n \rangle_{\mathbf{J}}. \quad (21)$$

The limit $n \rightarrow 0$ is approached by performing an analytic continuation from an integer n . For integer n , we can interpret Z_l^n as the partition function of n replicas of the original system, all with the same realization of the disordered couplings $\mathbf{J}$:

$$\begin{aligned}\langle (Z_l^{\text{SK}})^n \rangle_{\mathbf{J}} &= \int_{\mathbf{x}^\alpha} \left\langle \exp \left[\sum_{\alpha=1}^n \left(-\beta E^{\text{SK}}(\mathbf{x}^\alpha, a) \right. \right. \right. \\ &\quad \left. \left. \left. + l \text{Tr} \log [z - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] \right) \right] \right\rangle_{\mathbf{J}},\end{aligned}\quad (22)$$

where $\int_{\mathbf{x}^\alpha} \equiv \int \prod_{\alpha=1}^n d\mathbf{x}^\alpha$ is an integral over the configurations of the n different replicas. The term on the second line of Eq. (22) can be rewritten as

$$\begin{aligned}&l \sum_{\alpha} \text{Tr} \log [z - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] \\ &= 2l \log \prod_{\alpha} \sqrt{\det [z - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)]} \\ &= \log \left\{ \int_{\phi^\alpha} \exp \left[\frac{i}{2} \sum_{\alpha} \phi^\alpha \cdot [z - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] \phi^\alpha \right] \right\}^{-2l},\end{aligned}\quad (23)$$

where the symbol $\int_{\phi^\alpha} \equiv \prod_{\alpha} \frac{d\phi^\alpha}{\sqrt{2\pi i}}$ is a shorthand notation for the integral over a set of n bosonic fields $\{\phi^\alpha\}_{\alpha=1}^N$ of dimension N . Substituting Eq. (23) and Eq. (22) into the expression of the resolvent in Eq. (19) we obtain

$$\begin{aligned}
G^{\text{SK}}(z) &= \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{Nn} \partial_l \frac{d}{dz} \log \int_{\mathbf{x}^\alpha} \left\langle e^{-\beta \sum_\alpha E(\mathbf{x}^\alpha, a)} \left(\int_{\phi^\alpha} \exp \left[\frac{i}{2} \sum_\alpha \phi^\alpha \cdot [z - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] \phi^\alpha \right] \right)^{-2l} \right\rangle_{\mathbf{J}} \\
&= - \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \lim_{N \rightarrow \infty} \frac{2}{Nn} \partial_l \frac{d}{dz} \log \int_{\mathbf{x}^\alpha} \left\langle e^{-\beta \sum_\alpha E(\mathbf{x}^\alpha, a)} \left(\int_{\phi^\alpha} \exp \left[\frac{i}{2} \sum_\alpha \phi^\alpha \cdot [z - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] \phi^\alpha \right] \right)^l \right\rangle_{\mathbf{J}} \quad (24) \\
&= - \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \lim_{N \rightarrow \infty} \frac{2}{Nn} \partial_l \frac{d}{dz} \log \int_{\mathbf{x}^\alpha, \phi^{\alpha\rho}} \left\langle \exp \left[-\beta \sum_{\alpha=1}^n E(\mathbf{x}^\alpha, a) + \frac{i}{2} \sum_{\alpha=1}^n \sum_{\rho=1}^l \phi^{\alpha\rho} \cdot [z \mathbf{1} - \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] \phi^{\alpha\rho} \right] \right\rangle_{\mathbf{J}}.
\end{aligned}$$

In the first equality, we assumed as customary when resorting to the replica trick, that the limits $n, l \rightarrow 0$ and $N \rightarrow \infty$ commute. In the second equality, we have treated l as a dummy variable, rescaling it by a factor 2 and changing its sign. In the third equality, we have expanded the integral over the bosonic variables as an integral over a new set of replicated fields $\phi^{\alpha\rho}$ for $\alpha = 1, \dots, n$ and $\rho = 1, \dots, l$, with n and l integers whose limit to 0 is eventually taken by means of an analytic continuation. We used a shorthanded notation for the integral over these replicated variables, $\int_{\mathbf{x}^\alpha, \phi^{\alpha\rho}} \equiv \int \prod_{\alpha=1}^n \prod_{\rho=1}^l \frac{d\mathbf{x}^\alpha d\phi^{\alpha\rho}}{\sqrt{2\pi i}}$. The different replicated soft spins $\mathbf{x}^\alpha$ and bosonic fields $\phi^{\alpha\rho}$ share the same realizations of the quenched disordered couplings $\mathbf{J}$. This induces correlations among the different replicas. Such correlations are described by a set of order parameters $\mathbf{\Gamma}$, that can be inserted in the expression of the resolvent after averaging over the disorder and using some manipulations detailed in Appendix A. As a result, we obtain

$$\begin{aligned}
G^{\text{SK}}(z) &= - \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \lim_{N \rightarrow \infty} \frac{2}{Nn} \\
&\quad \times \partial_l \frac{d}{dz} \log \int d\mathbf{\Gamma} e^{N S^{\text{SK}}[\mathbf{\Gamma}]}. \quad (25)
\end{aligned}$$

Where the set of parameters $\mathbf{\Gamma} = \{\mathbf{Q}, \mathbf{m}, \mathbf{P}, \mathbf{T}, \mathbf{M}\}$ contains an $n \times n$ matrix $\mathbf{Q}$, an n dimensional vector $\mathbf{m}$, a $n \times n \times l$ tensor $\mathbf{T}$ and a $n \times n \times l \times l$ tensor $\mathbf{P}$. In the limit of large N , the integral in Eq. (25) is dominated by a saddle-point:

$$G^{\text{SK}}(z) = - \lim_{n \rightarrow 0} \frac{d}{dz} \lim_{l \rightarrow 0} \frac{2}{n} \partial_l \sup_{\mathbf{\Gamma}} S^{\text{SK}}[\mathbf{\Gamma}]. \quad (26)$$

The extremization of the action S yields a set of self-consistent equations that determine the order parameters $\mathbf{\Gamma}$. These saddle point equations are evaluated in Appendix B within the assumption that there are no correlations among the replicated bosonic fields and the

replicated soft spins. As we will see shortly, this ansatz amounts to neglecting correlations between the distribution of soft-spins in global minima and the Hessian matrix of global minima, across different realizations of the disorder. We show in Appendix B 1 that this approximation is stable against a small perturbation of the order parameters at the saddle point. We obtain the following self consistent equation for the resolvent,

$$G^{\text{SK}}(z) = \int dx \frac{P^{\text{SK}}(x)}{z - \partial_x^2 E_I(x, a) - G^{\text{SK}}(z)}. \quad (27)$$

This equation is equivalent to the Pastur self-consistent condition [45] for the resolvent of the sum of a Wigner matrix $\mathbf{J}$ and a diagonal matrix with independent, identically distributed elements equal to $\partial_x^2 E_I(x, a)$, where the value of x for each diagonal entry is drawn i.i.d. from $P^{\text{SK}}(x)$, which is the mean field single-spin probability distribution, defined as

$$P^{\text{SK}}(x) \equiv \lim_{n \rightarrow 0} \frac{1}{n} \sum_{\alpha=1}^n \langle \delta(x - x^\alpha) \rangle_{\text{MF}}^{\text{SK}}. \quad (28)$$

The average $\langle \dots \rangle_{\text{MF}}^{\text{SK}}$ is an average over a set of n replicated single-site soft spins x^α , and it reads

$$\langle \dots \rangle_{\text{MF}}^{\text{SK}} = \frac{1}{Z_{\text{MF}}^{\text{SK}}} \int \prod_{\alpha=1}^n d\mathbf{x}^\alpha e^{-\beta E_{\text{MF}}^{\text{SK}}(\{\mathbf{x}^\alpha\})}, \quad (29)$$

where the mean field energy $E_{\text{MF}}^{\text{SK}}$ and the mean field partition function $Z_{\text{MF}}^{\text{SK}}$ are, respectively,²

$$\begin{aligned}
E_{\text{MF}}^{\text{SK}} &= \sum_{\alpha} E_I(x^\alpha, a) - \frac{\beta}{2} \sum_{\alpha, \beta} Q_{\alpha\beta} x^\alpha x^\beta \\
&\quad - J_0 \sum_{\alpha} m_{\alpha} x^\alpha \\
Z_{\text{MF}}^{\text{SK}} &= \int \prod_{\alpha=1}^n d\mathbf{x}^\alpha e^{-\beta E_{\text{MF}}^{\text{SK}}(\{\mathbf{x}^\alpha\})}. \quad (30)
\end{aligned}$$

¹We use here the symbol z as a shorthanded notation for $z + i0^+$, the limiting value of a complex number with real part z and a vanishingly small positive complex part.

²Note that, with a slight abuse of notation, we are using the symbol β to denote both the inverse temperature and a replica index.

The overlap matrix $Q_{\alpha\beta}$ and the magnetization vector m_α are determined by the saddle point equations

$$Q_{\alpha\beta} = \langle x^\alpha x^\beta \rangle_{\text{MF}}^{\text{SK}} \quad m_\alpha = \langle x^\alpha \rangle_{\text{MF}}^{\text{SK}}. \quad (31)$$

The matrix $Q_{\alpha\beta}$ contains the correlations among different replicas, while the vector m_α expresses the average magnetization of each replica. Equations (27) to Eq. (31) contain, in principle, everything needed to determine the spectral distribution $\rho^{\text{SK}}(\lambda)$, once the single-spin distribution P^{SK} is known. In order to determine the latter, an explicit structure of the overlap matrices has to be chosen. Before doing this, however, let us discuss how the position of the edge of the spectrum can be determined using the self-consistent equation for the resolvent.

A. Lower edge of the spectrum

In this Section we discuss, following [46, 47], how to determine the lower edge of the spectrum of the Hessian from the knowledge of the resolvent $G^{\text{SK}}(z)$ given in Eq. (27). We introduce a function $g(z) \equiv G^{\text{SK}}(z) - z$. Equation (27) can be inverted to express z as a function of g , namely,

$$z = - \left[g + \int dx \frac{P^{\text{SK}}(x)}{g + \partial_x^2 E_I(x, a)} \right] \equiv z(g). \quad (32)$$

The support of the spectrum $\rho^{\text{SK}}(\lambda)$ is defined as the set of values $\lambda + i\epsilon$ for which $G^{\text{SK}}(\lambda + i\epsilon)$ has a nonzero imaginary part in the limit $\epsilon \rightarrow 0$. Outside this region, $G^{\text{SK}}(\lambda + i\epsilon)$ is real. Therefore, the support of $\rho^{\text{SK}}(\lambda)$ can be obtained as the complement of the image of $z(g)$ for real g . In particular, the lower edge of the spectrum is determined by the maximum attained by $z(g)$. From its definition, we see that $z(g)$ is defined in the interval $g \in (-\infty, -E''_{I,\text{max}}] \cup [-E''_{I,\text{min}}, \infty)$, where $E''_{I,\text{max}}, E''_{I,\text{min}}$ are respectively the maximum and the minimum value attained by the function $\partial_x^2 E_I(x, a)$, with a fixed, in the region where $P^{\text{SK}}(x) \neq 0$. We expect the maximum of $z(g)$ to lie in the $[-E''_{I,\text{min}}, \infty)$ branch. The lower edge $\lambda_{\text{min}}^{\text{SK}}$ of the spectrum of $\rho^{\text{SK}}(\lambda)$ is thus given by

$$\lambda_{\text{min}}^{\text{SK}} = z(g^*) \quad g^* \equiv \arg \max_{g \in [-E''_{I,\text{min}}, \infty)} z(g). \quad (33)$$

In particular, if the function $z(g)$ reaches a maximum in the interval $(-E''_{I,\text{min}}, \infty)$, then g^* satisfies the condition

$$\left. \frac{dz(g)}{dg} \right|_{g=g^*} = -1 + \int dx \frac{P^{\text{SK}}(x)}{(g^* + \partial_x^2 E_I(x, a))^2} = 0. \quad (34)$$

The important interplay of this condition with the nature of the spin-glass phase of the CIM-SK will be discussed below. Once $P^{\text{SK}}(x)$ is determined by means of replica theory, we can determine where the lower edge of the spectrum of the Hessian lies.

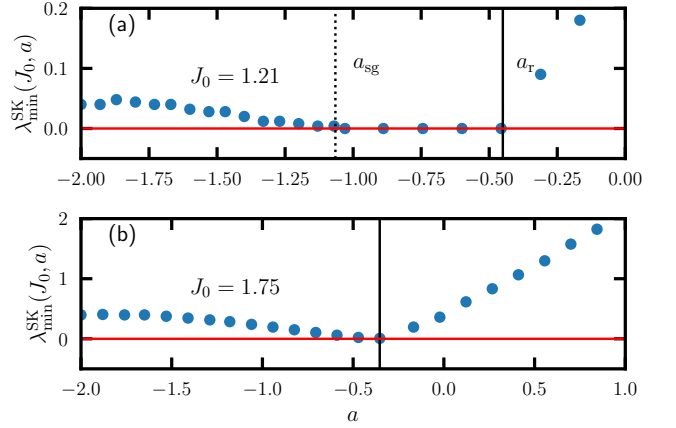


FIG. 3. Lower edge of the spectrum of the Hessian of global minima in the annealed landscape. We plot $\lambda_{\text{min}}^{\text{SK}}(J_0, a)$ as a function of the laser gain a for two different values of the mean connectivity J_0 . Panel (a): in the replica-symmetric phase, the lower edge of the spectrum decreases as the laser gain increases, until it hits zero (horizontal solid red line) at the spin-glass transition value $a_{\text{sg}}(J_0)$ (vertical dotted black line). Upon further increase of the laser gain in the spin-glass phase, the lower edge of the Hessian remains zero, as long as the laser gain is below the rigidity transition value a_r (vertical solid black line). For $a > a_r$, the spectrum is gapped. In panel (b), the spin-glass and the rigidity transition occur at the same value of the laser gain, $a_r = a_{\text{sg}}$. The spectrum has a gap both above and below this value.

VI. THE STRUCTURE OF GLOBAL MINIMA AND THEIR HESSIAN EIGENSPECTRA

A. Convex phase

To compute the probability density of the soft-spin population at global minima $P^{\text{SK}}(x)$ from the mean field calculation presented in Sec. V, we need to make an ansatz about the structure of the overlap matrix $Q_{\alpha\beta}$ of the system. The simplest structure is a replica symmetric ansatz, invariant under permutations of the different replicas,

$$Q_{\alpha\beta} = q_d \delta_{\alpha\beta} + q_o (1 - \delta_{\alpha\beta}) \quad m_\alpha = m. \quad (35)$$

The diagonal part of $Q_{\alpha\beta}$ contains the overlap of a replica with itself, while its off-diagonal elements contain overlaps among different replicas. The difference $q_d - q_o$ is a measure of the width, in phase space, of the thermodynamic state considered. As the temperature decreases toward 0, we expect the states occupied by the different replicas of the system to shrink toward the global minimum, which implies that $q_d - q_o \rightarrow 0$. It is thus convenient to consider an $O(1)$ rescaled overlap difference,

$$\Delta\tilde{q} \equiv \beta(q_d - q_o).^3$$

A calculation detailed in Appendix C shows that the saddle point equation for the order parameters q_o , $\Delta\tilde{q}$ and m are given by

$$\begin{aligned}\Delta\tilde{q} &= \overline{\langle x^2 \rangle_{\text{RS}}^{\text{SK}} - (\langle x \rangle_{\text{RS}}^{\text{SK}})^2} \\ q_o &= \overline{(\langle x \rangle_{\text{RS}}^{\text{SK}})^2} \\ m &= \overline{\langle x \rangle_{\text{RS}}^{\text{SK}}},\end{aligned}\quad (36)$$

where the replica-symmetric mean field average $\langle \dots \rangle_{\text{RS}}^{\text{SK}} = \frac{1}{Z_{\text{eff}}^{\text{SK}}(h)} \int dx \dots e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)}$ involves an effective partition function $Z_{\text{eff}}^{\text{SK}}$ and an effective energy $E_{\text{eff}}^{\text{SK}}$, which read, respectively,

$$\begin{aligned}Z_{\text{eff}}^{\text{SK}}(h) &\equiv \int dx e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)} \\ E_{\text{eff}}^{\text{SK}}(x, h) &\equiv E_I(x, a + \Delta\tilde{q}) - (\sqrt{q_o}h + J_0m)x \\ &= \frac{1}{4}x^4 - \frac{a + \Delta\tilde{q}}{2}x^2 - (\sqrt{q_o}h + J_0m)x,\end{aligned}\quad (37)$$

and the average $\overline{\dots} \equiv \int \frac{dh}{\sqrt{2\pi}} \dots e^{-h^2/2}$ is an average over the realizations of a zero mean unit variance Gaussian field h . The single spin probability distribution in the convex phase $P_{\text{RS}}^{\text{SK}}(x)$ is instead given by

$$P_{\text{RS}}^{\text{SK}}(x) = \overline{e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)} [Z_{\text{eff}}^{\text{SK}}(h)]^{-1}}, \quad (38)$$

As a result of the mean field analysis, the single spin distribution reduces to a Boltzmann distribution for a single-site soft spin system, with effective energy $E_{\text{eff}}^{\text{SK}}$. The rescaled overlap difference $\Delta\tilde{q}$ acts as a shift in the laser gain of the single-site energy E_I , while the overlap among different replicas q_o determines the coupling strength to the quenched random Gaussian field h . Finally, the mean connectivity J_0 determines the coupling strength of x to the magnetization m of the system. Note that here the magnetization m is the mean of the soft-spin distribution, and can therefore take values larger than 1.

In the low temperature limit and for a given realization of the random field h , the replica-symmetric partition function is dominated by the minimum $x^*(h)$ of the energy $E_{\text{RS}}^{\text{SK}}$. The saddle-point equations then become

$$\begin{aligned}\Delta\tilde{q} &= \overline{[\partial_x^2 E_{\text{eff}}^{\text{SK}}|_{x=x^*(h)}]^{-1}}, \\ q_o &= \overline{x^*(h)^2}, \\ m &= \overline{x^*(h)}.\end{aligned}\quad (39)$$

This set of self consistent equations can be solved numerically through iterative methods for different values

of the laser gain a and the ferromagnetic strength J_0 . The single spin probability distribution becomes

$$\begin{aligned}P_{\text{RS}}^{\text{SK}}(x) &= \overline{\delta(x - x^*(h))} \\ &= \frac{1}{\sqrt{2\pi q_o}} |\partial_x^2 E_I(x, a + \Delta\tilde{q})| e^{-\frac{(x^*(h))^2}{2}},\end{aligned}\quad (40)$$

where the field $h^*(x)$ satisfies the stationary condition $\partial_x E_{\text{eff}}^{\text{SK}}(x, h)|_{h=h^*(x)} = 0$, i.e.,

$$\sqrt{q_o}h^*(x) + J_0m = x[x^2 - (a + \Delta\tilde{q})]. \quad (41)$$

When $a + \Delta\tilde{q} \leq 0$, the effective energy $E_{\text{eff}}^{\text{SK}}(x, h)$ is a convex function of x for *all* realizations of the random field h . The support of $P_{\text{RS}}^{\text{SK}}$ then extends over the whole real axis. If instead $a + \Delta\tilde{q} > 0$, the interval $(-\sqrt{a + \Delta\tilde{q}}, \sqrt{a + \Delta\tilde{q}})$ is excluded from the support of $P_{\text{RS}}^{\text{SK}}(x)$.

The saddle point equations are complemented with a study of the stability of the replica symmetric solution. This can be done expanding the action S in Eq. (25) around the replica symmetric overlap matrix and magnetization. The replica symmetric solution remains an extremum of the action S as long as the largest eigenvalue of the Hessian remains negative. This eigenvalue, known as the replicon $\tilde{\Lambda}_{\text{R}}^{\text{SK}}$, has a known structure, and it evaluates to (details in App. D)

$$\tilde{\Lambda}_{\text{R}}^{\text{SK}} = -1 + \overline{[\partial_x^2 E_{\text{eff}}^{\text{SK}}(x, a)|_{x=x^*(h)}]^{-2}}. \quad (42)$$

The replica symmetric solution is stable as long as $\tilde{\Lambda}_{\text{R}}^{\text{SK}} < 0$. A positive replicon eigenvalue signals a non-linear instability of the system under the perturbation generated by a small random external field applied to the soft spins. Using the saddle point equations and the stability condition, we can determine the phase boundaries shown in Fig. 2. This is what we do next.

1. Phase boundary of convex paramagnetic phase

We first recover the results obtained in Sec. IV concerning the stability of the paramagnetic, convex phase. In the paramagnetic phase, we have $m = q_o = 0$. The single spin distribution becomes a Dirac delta centered around the origin, $P_{\text{RS}}^{\text{SK}}(x) = \delta(x)$. The remaining order parameter $\Delta\tilde{q}$ is determined from the saddle point equation

$$\Delta\tilde{q} = -\frac{1}{a + \Delta\tilde{q}}. \quad (43)$$

In App. E, we show that in order for the paramagnetic solution to be stable, we must have

$$\Delta\tilde{q} \leq \min[1, J_0^{-1}]. \quad (44)$$

Moreover, to ensure that $m = q_o = 0$, the paramagnetic, replica symmetric energy $E_{\text{eff}}^{\text{SK}} = \frac{1}{4}x^4 - \frac{1}{2}(a + \Delta\tilde{q})x^2$ must admit a minimum at $x = 0$. This implies

$$a + \Delta\tilde{q} \leq 0. \quad (45)$$

³The scaling $q_d - q_o \propto \beta^{-1}$ is then self-consistently verified through the saddle point equations.

Combining together Eq. (44) and Eq. (45) we obtain that the convex paramagnetic state is stable as long as

$$a \leq \begin{cases} -2 & \text{if } J_0 \leq 1 \\ -J_0 - \frac{1}{J_0} & \text{if } J_0 > 1, \end{cases} \quad (46)$$

which is equivalent to the condition $\lambda_{0,m} > 0$, with $\lambda_{0,m}$ given by Eq. (14), which was obtained through random matrix theory. This analysis allows us to conclude that for $J_0 < 1$, the system has a transition into a spin-glass phase at $a_{\text{sg}}(J_0) = -2$. This is the horizontal dotted line in Fig. 2 (a). For $J_0 > 1$, the second case in Eq. (46) yields a transition line toward a ferromagnetic state. This is the dashed line $a_{\text{BBP}}(J_0)$ in Fig. 2 (a). We address the stability of this ferromagnetic phase next.

2. Stability of the ferromagnetic phase

When the replica symmetric solution is ferromagnetic, its boundary with the spin-glass phase has to be computed numerically, solving the saddle point Equations (39). The replica symmetric solution can become unstable through two different pathways. In the first pathway, the replicon eigenvalue $\tilde{\Lambda}_{\text{R}}^{\text{SK}}$ changes sign continuously, similarly to what happens at the spin-glass transition from the paramagnetic phase. At the transition point, the effective energy $E_{\text{eff}}^{\text{SK}}(x, h)$ is a convex function of x for all the realizations of the quenched random field h . In this case, as discussed more in detail in Sec. VIB, in the spin-glass phase the support of the single-spin distribution covers the whole real axis, and the lower edge of the Hessian is located at the origin. This corresponds to the soft spin-glass phase. In the second pathway, the effective energy $E_{\text{eff}}^{\text{SK}}(x, h)$ changes its convexity while the replicon eigenvalue is still negative. When this happens, the Hessian of the replica-symmetric energy $E_{\text{RS}}^{\text{SK}}(x, h)$ vanishes when evaluated at its minimum $x_m(h)$ for some values of h , while the replicon eigenvalue in Eq. (42) jumps discontinuously from a negative value to ∞ , and so the convex phase becomes unstable. We argue in Sec. VIB that the nonconvexity of the effective energy in the spin-glass phase implies that the support of the single-spin distribution splits into two disconnected domains, and that the spectrum of the Hessian at global minima is gapped away from the origin. These are the signature of the rigid spin-glass phase. The expression of $E_{\text{eff}}^{\text{SK}}$ in Eq. (37), shows that the convex to non-convex transition occurs when

$$a + \Delta\tilde{q} = 0. \quad (47)$$

This condition is used to determine the stability boundary between the ferromagnetic phase (II) and the rigid spin-glass phase (IV). The above discussion concludes the identification of the boundaries of the convex phase of the CIM-SK model.

3. Hessian at the spin-glass transition

At both spin-glass transitions, the spectral distribution of the eigenvalues of the Hessian for global minima is gapless. To see this, let us consider the two pathways through which the replica symmetric solution becomes unstable. In the first pathway, the replicon eigenvalue given by Eq. (42) approaches zero continuously at the transition. The condition $\tilde{\Lambda}_{\text{R}}^{\text{SK}} = 0$ implies that

$$-1 + \int dx \frac{P_{\text{RS}}^{\text{SK}}(x)}{(\partial_x^2 E_I(x, a) - \Delta\tilde{q})^2} = 0, \quad (48)$$

where we made use of the definition of the replica symmetric mean field energy given by Eq. (37). Equation (48) can be compared with the extremization condition of $z(g)$ in Eq. (34), allowing us to conclude that $g^* = -\Delta\tilde{q}$. Using Eq. (33), the lower edge of the spectrum $\lambda_{\text{min}}^{\text{SK}}$ is thus given by

$$\begin{aligned} \lambda_{\text{min}}^{\text{SK}} &= z(-\Delta\tilde{q}) \\ &= \Delta\tilde{q} - \int dx \frac{P_{\text{RS}}^{\text{SK}}(x)}{\partial_x^2 E_I(x, a) - \Delta\tilde{q}} = 0. \end{aligned} \quad (49)$$

In the second equality, we made use of the saddle point equation for $\Delta\tilde{q}$, given by Eq. (36). Equation (49) is consistent with the random matrix analysis of the Hessian at the origin performed in Sec. IV.

In the second pathway to instability of the convex phase, the replicon eigenvalue jumps discontinuously from a negative value to ∞ when $a + \Delta\tilde{q} = 0$. In this case, the minimum value taken by the Hessian of the single site energy E_I is $-E''_{I,\text{min}} = a$. Equation (34) does not admit a solution, and we thus locate the lower bound of the spectrum as $z(g = -E''_{I,\text{min}})$. When $a + \Delta\tilde{q} = 0$, we obtain from Eq. (41) that $-E''_{I,\text{min}} = a$. The position of the lower edge of the spectrum is thus given by

$$\begin{aligned} \lambda_{\text{min}}^{\text{SK}} &= z(a) \\ &= -a - \int dx \frac{P(x)}{\partial_x^2 E_I(x, a) + a} \\ &= \Delta\tilde{q} - \int dx \frac{P(x)}{\partial_x^2 E_I(x, a) - \Delta\tilde{q}} = 0. \end{aligned} \quad (50)$$

In the third equality, we used the fact that $a + \Delta\tilde{q} = 0$, while in the final passage we used again the saddle point equation for $\Delta\tilde{q}$ given by Eq. (36).

We thus see that when approaching the spin-glass transition from the replica symmetric phase, both pathways lead to the same fate for the lower edge of the spectrum of the Hessian. At the edge of the convex phase, soft modes arise around the global minima, leading to an instability that eventually brings the system into a spin-glass phase, which we now address.

B. Spin-glass phase and rigidity transition

The glassy phase of the CIM-SK is described through the full replica symmetry breaking solution [17–20, 48].

The phase space of the system shatters into many pure states, organized into an ultrametric structure. The inter-replica overlap q_o is replaced by a continuous, monotonically increasing function $q(y)$, defined on the interval $[0, 1]$, which is the inverse of the cumulative distribution of overlaps among the different states of the system. The overlap difference $\Delta\tilde{q}$ is now replaced by $\Delta\tilde{q} = \beta(q_d - q(1))$. In Appendix F we outline how the saddle point equations for $q(y)$, $\Delta\tilde{q}$ and m are obtained. These saddle point equations read

$$\begin{aligned} q(y) &= \int dh P(y, h) [\partial_h f(y, h)]^2 \\ \Delta\tilde{q} &= \int dh P(1, h) \partial_h^2 f(1, h) \\ m &= \int dh P(1, h) \partial_h f(1, h). \end{aligned} \quad (51)$$

The function $f(y, h)$ describes a hierarchy of free energies of the system, computed across different levels of the ultrametric tree across which different states are arranged. The field h encodes a form of quenched disorder stemming from the overlap among different pure states. While in the replica-symmetric solution this field is Gaussian, in the spin-glass phase a hierarchy of probability distributions $P(y, h)$ arises, and the distribution of internal fields $P(y = 1, h)$ is no longer Gaussian. The function $f(y, h)$ and $P(y, h)$ satisfy the partial differential equations

$$\begin{aligned} \partial_y f(y, h) &= -\frac{1}{2} \frac{dq(y)}{dy} [\beta y (\partial_y f(y, h))^2 + \partial_h^2 f(y, h)] \\ \partial_y P(y, h) &= \frac{1}{2} \frac{dq(y)}{dy} [\partial_h^2 P(y, h) \\ &\quad - 2\beta \partial_h (P(y, h) \partial_h f(y, h))], \end{aligned} \quad (52)$$

together with the boundary conditions

$$\begin{aligned} f(1, h) &= \beta^{-1} \log \int dx e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)} \\ P(0, h) &= \frac{1}{\sqrt{2\pi q(0)}} e^{-\frac{h^2}{2q(0)}}, \end{aligned} \quad (53)$$

where the single-spin, mean field energy $E_{\text{eff}}^{\text{SK}}(x, h)$ is defined in Eq. (37). In the low-temperature limit, we show in App. F that the single-spin distribution $P^{\text{SK}}(x)$ is obtained by evaluating the probability to realize a quenched field h that minimizes the effective energy $E_{\text{eff}}^{\text{SK}}(x, h)$, namely,

$$P_{\text{fRSB}}^{\text{SK}}(x) = |\partial_x^2 E_I(x, a + \Delta\tilde{q})| P(1, h^*(x)), \quad (54)$$

where $h^*(x)$ is determined by the stationary condition in Eq. (41). Using Eq. (54) we can determine the resolvent and the spectrum of the Hessian in the spin-glass phase. The numerical integration of the full replica symmetry breaking equations at low temperature follows the methods developed in [49], and it is discussed in Appendix F.

The full-replica symmetry breaking spin-glass phase is marginally stable in the space of all possible structures that the overlap matrix $\mathbf{Q}$ can take. We report for completeness a derivation of this widely appreciated fact in Appendix F, where we compute the replicon eigenvalue $\tilde{\lambda}_{\text{R}}^{\text{SK}}$ and we show that

$$\tilde{\lambda}_{\text{R}}^{\text{SK}} = -1 + \int dh P(1, h) [\partial_h^2 f(1, h)]^2 = 0. \quad (55)$$

Using the marginal stability of the spin-glass phase, and the possible change in convexity of the effective energy $E_{\text{eff}}^{\text{SK}}$, we can study the properties of the linear excitations around the spin-glass global minima.

1. Hessian in spin-glass phase

The behavior of the lower edge of the spectrum of the Hessian in the spin-glass phase is determined by an interplay between the shape of the effective single-site potential $E_{\text{eff}}^{\text{SK}}(x, h)$ and the marginal stability of the spin-glass phase. If the effective single site energy $E_{\text{eff}}^{\text{SK}}(x, h)$ is a convex function of x for all the values of h , (i.e. when $a + \Delta\tilde{q} < 0$), soft modes appear in the spin-glass phase, and the lower edge of the spectrum of the Hessian for global minima touches the origin. If instead $E_{\text{eff}}^{\text{SK}}(x, h)$ is a nonconvex function of x , as it is the case when $a + \Delta\tilde{q} > 0$, the spectrum of the Hessian is gapped away from the origin.

When $a + \Delta\tilde{q} < 0$ we show in Appendix F that the marginal stability condition in Eq. (55) becomes

$$\tilde{\lambda}_{\text{R}}^{\text{SK}} = -1 + \int dx \frac{P_{\text{fRSB}}^{\text{SK}}(x)}{[\partial_x^2 E_I(x, a) - \Delta\tilde{q}]^2} = 0. \quad (56)$$

This condition implies that the maximum of $z(g)$, defined in Eq. (32), occurs for $g^* = -\Delta\tilde{q}$. The marginal stability condition is, in fact, equivalent to the stationary condition $z'(g)|_{g=-\Delta\tilde{q}} = 0$, as can be seen by comparing Eq. (56) with Eq. (34). We thus obtain for the lower edge of the spectrum

$$\begin{aligned} \lambda_{\text{min}}^{\text{SK}} &= z(-\Delta\tilde{q}) \\ &= -\Delta\tilde{q} + \int dx \frac{P_{\text{fRSB}}^{\text{SK}}(x)}{\partial_x^2 E_I(x, a) - \Delta\tilde{q}} = 0, \end{aligned} \quad (57)$$

where the second equality follows from the zero-temperature limit of the saddle point equation for $\Delta\tilde{q}$ in Eq. (51). The spectrum of the Hessian is therefore gapless.

If $a + \Delta\tilde{q} > 0$, we show in Appendix F that the marginal stability condition is modified, and that the following inequality holds,

$$-1 + \int dx \frac{P_{\text{fRSB}}^{\text{SK}}(x)}{[\partial_x^2 E_I(x, a) - \Delta\tilde{q}]^2} < 0. \quad (58)$$

Combining this fact with Eq. (34) implies that $z'(g)|_{g=-\Delta\tilde{q}} > 0$. On the other hand, the saddle point equations for $\Delta\tilde{q}$ yield $z(g = -\Delta\tilde{q}) = 0$. Therefore, the maximum of $z(g)$ must be larger than 0. We thus obtain that

$$\lambda_{\min}^{\text{SK}} = z(g^*) > z(-\Delta\tilde{q}) = 0, \quad (59)$$

and therefore the spectrum of the Hessian is gapped. Establishing the existence of gapped and gapless spin-glass phases in the CIM is one of the main results of this work. Our derivation highlights how this phenomenon stems from an interplay between a change of the shape of the effective single site energy and the nature of the spin-glass phase in the CIM-SK.

We now turn to studying the properties of the most abundant, local minima of the evolving energy landscape of the CIM.

VII. PROLIFERATION OF PARAMAGNETIC LOCAL MINIMA AND CRITICAL POINTS

In this Section we study the number and the properties of the most abundant minima of the CIM as the laser gain increases. This is done by means of a combination of the Kac-Rice formula and the replica approach [50–53].⁴ We denote by $\mathcal{N}^{\text{SK}}(r, \mathbf{J})$ the number of critical points with intensive index r , defined as the index of a critical point (number of negative eigenvalues of its Hessian) divided by N , in the coherent Ising machine for a given realization of the disorder $\mathbf{J}$. The complexity $\Sigma^{\text{SK}}(r, \mathbf{J})$ is defined as the exponential rate at which $\mathcal{N}^{\text{SK}}(r, \mathbf{J})$ grows as the size of the system increases:

$$\mathcal{N}^{\text{SK}}(r, \mathbf{J}) \propto e^{N\Sigma^{\text{SK}}(r, \mathbf{J})}, \quad (60)$$

where the proportionality factor accounts for subexponential contributions. While the number of critical points fluctuates with different realizations of the disorder, we expect the complexity to concentrate around its average over different realizations of the connectivity $\mathbf{J}$. We are thus interested in computing the disorder-averaged complexity $\Sigma^{\text{SK}}(r)$, defined as

$$\Sigma^{\text{SK}}(r) = \lim_{N \rightarrow \infty} \frac{1}{N} \langle \log \mathcal{N}^{\text{SK}}(r, \mathbf{J}) \rangle_{\mathbf{J}}. \quad (61)$$

The number of critical points $\mathcal{N}^{\text{SK}}(r, \mathbf{J})$ is counted using the Kac-Rice formula [54],

$$\mathcal{N}^{\text{SK}}(r, \mathbf{J}) = \int d\mathbf{x} \left(\prod_{i=1}^N \delta(\partial_i E^{\text{SK}}(\mathbf{x}, a)) \right) \times |\det \mathbf{H}^{\text{SK}}(\mathbf{x}, a)| \delta(\mathcal{I}(\mathbf{x}) - Nr). \quad (62)$$

⁴In the literature on disordered systems, the most abundant minima are also referred to as ‘typical’. However, in this work, we will denote by ‘typical’ the points visited the gradient descent dynamics of the CIM, which is discussed in Sec. VIII and Sec. VIIB.

The Dirac delta enforces $\mathbf{x}$ to be a critical point of the CIM energy, while the quantity $|\det \mathbf{H}^{\text{SK}}(\mathbf{x}, a)|$ is a phase space volume factor. The extensive index $\mathcal{I}(\mathbf{x})$ is the number of unstable directions around the critical point $\mathbf{x}$. The number of critical points can thus be regarded as the partition function of a ‘microcanonical’ distribution, whose measure is given by the integrand in Eq. (62). Within this framework, the complexity is thus analogous to the entropy of critical points in the system. An equivalent description, which we adopt here, employs a ‘grand-canonical’ ensemble. We introduce the grand-potential $\Omega^{\text{SK}}(\mu)$, defined as

$$\begin{aligned} \Omega^{\text{SK}}(\mu) &\equiv - \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \log \int d\mathbf{x} \left(\prod_{i=1}^N \delta(\partial_i E(\mathbf{x})) \right) \right. \\ &\quad \times |\det \mathbf{H}^{\text{SK}}(\mathbf{x}, a)| e^{\mu \mathcal{I}(\mathbf{x})} \left. \right\rangle_{\mathbf{J}} \\ &\equiv - \lim_{n \rightarrow \infty} \frac{1}{N} \langle \log Z_{\Omega} \rangle_{\mathbf{J}}, \end{aligned} \quad (63)$$

where the last line defines the grand canonical partition function Z_{Ω}^{SK} . The quantity μ is a chemical potential associated to the index of the critical points of the system. As $\mu \rightarrow -\infty$, the integrand in Eq. (63) concentrates around local minima of the energy landscape. The complexity is recovered from the grand-potential by means of a Legendre transform as

$$\Sigma^{\text{SK}}(r) = \inf_{\mu} [-\mu r - \Omega^{\text{SK}}(\mu)]. \quad (64)$$

This extremization condition yields a relationship between the intensive index r and the derivative of the grand-potential,

$$r = - \frac{\partial \Omega^{\text{SK}}(\mu)}{\partial \mu}. \quad (65)$$

In particular, when focusing on minima, we have $r = 0$. Since the grand-potential is a convex function of the chemical potential we obtain, for the complexity of minima, that

$$\Sigma(r = 0) = - \lim_{\mu \rightarrow -\infty} \Omega^{\text{SK}}(\mu). \quad (66)$$

The evaluation of the grand-potential in the limit of large negative chemical potential allows us to compute the complexity of minima of the system.

On the other hand, evaluating $\Omega^{\text{SK}}(\mu = 0)$ gives access to the number of most abundant critical points, irrespective of the value of their intensive index r . This can be seen using the inverse of the Legendre transform in Eq. (64), evaluated at $\mu = 0$, which yields

$$-\Omega^{\text{SK}}(\mu = 0) = \sup_r \Sigma^{\text{SK}}(r). \quad (67)$$

The index of the most abundant critical points can be computed using the extremization condition in Eq. (65),

obtaining

$$r^* \equiv \arg \max_r \Sigma^{\text{SK}}(r) = - \left. \frac{\partial \Omega^{\text{SK}}(\mu)}{\partial \mu} \right|_{\mu=0} \quad (68)$$

The discussion above highlights the two cases we are interested in: the complexity of the most abundant minima, and the complexity of the most abundant critical points, obtained respectively from the evaluation of the grand-potential in Eq. (63) for $\mu \rightarrow -\infty$ and $\mu = 0$, respectively.

The grand-potential is computed using the replica trick and averaging over the disorder. The details of this calculation are given in Appendix G. The main difficulty presented by the computation arises from the

presence of correlations between the gradient of the energy $\nabla E^{\text{SK}}(\mathbf{x}, a)$ and the determinant of the Hessian in Eq. (63) across different realizations of the disorder. Because of the presence of the single site energy E_I , these correlations are particularly difficult to analyze. In order to make progress, we neglect these correlations. and we verify this approximation *a posteriori* by means of numerical experiments on small size systems, as discussed in Sec. VII B. To enforce the stationary condition of the configurations contributing to the partition function in Eq. (63), new order parameters, coupling different replicas, need to be introduced. These order parameters are then determined self-consistently by means of saddle point equations. The replica-symmetric solution reads

$$\begin{aligned} \Omega^{\text{SK}}(\mu) &= \frac{1}{2} [(\Delta A + A_o)^2 - A_o^2 + 2A_o t_R + (\Delta C + C_o)(\Delta q + q_o) - C_o q_o] + v m - \overline{\overline{\log Z_{\text{RS}, \Omega}(h_1, h_2)}} \\ Z_{\text{RS}, \Omega}(h_1, h_2) &= \int dx \frac{|\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)|}{\sqrt{2\pi \Delta q}} \exp \left[\frac{1}{2} \Delta C x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) - \frac{(\Delta A x - \partial_x E_{\Omega, \text{eff}}(x, h_1))^2}{2\Delta q} \right] \\ E_{\Omega, \text{eff}}(x, h_1) &\equiv E_I(x, a + t_R) - (J_0 m + \sqrt{q_o} h_1) x \\ F(h_1, h_2) &\equiv v + \sqrt{C_o - \frac{A_o^2}{q_o}} h_2 + \frac{A_o}{\sqrt{q_o}} h_1. \end{aligned} \quad (69)$$

The function $\mathcal{I}_{\text{MF}}(x) \equiv \Theta(-\partial_x^2 E_{\Omega, \text{eff}}(x, h_1))$, with Θ the Heaviside step function, is the index of a mean field effective energy $E_{\Omega, \text{eff}}(x, h_1)$. The double overline $\overline{\overline{\cdot}} \equiv \int \frac{dh_1 dh_2}{2\pi} \dots e^{-\frac{1}{2}(h_1^2 + h_2^2)}$ denotes an average over the random Gaussian fields h_1, h_2 . Note that the field h_1 stems from the overlap among different replicas of the system. Its physical origin is akin to the random Gaussian field appearing in the global minima calculation in Sec. VI. The field h_2 stems instead from overlaps among a given replica and the Lagrange multipliers that enforce the replicated system to occupy a critical point. The 8 order parameters $\Delta A, A_o, \Delta C, C_o, \Delta q, q_o, m, v$ are determined by the saddle point equations

$$\begin{aligned} \Delta A &= \frac{1}{2\Delta q} \left(\overline{\langle x \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}} - \overline{\langle x \rangle_{\Omega} \langle \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}} \right) - \frac{1}{2} t_R \\ A_o &= -\frac{1}{\Delta q} \left(\Delta A q_o - \overline{\langle x \rangle_{\Omega} \langle \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}} \right) \\ \Delta q &= \overline{\langle x^2 \rangle_{\Omega}} - \overline{\langle x \rangle_{\Omega}^2} \\ q_o &= \overline{\langle x \rangle_{\Omega}^2} \\ \Delta C &= -\frac{1}{\Delta q} + \frac{\Delta A^2}{\Delta q} + \frac{1}{\Delta q^2} \left[-2\Delta A \left(\overline{\langle x \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}} - \overline{\langle x \rangle_{\Omega} \langle \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}} \right) + \overline{\langle (\partial_x E_{\Omega, \text{eff}})^2 \rangle_{\Omega}} - \overline{\langle \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}^2} \right] \\ C_o &= \frac{1}{\Delta q^2} \left(\Delta A^2 q_o - 2\overline{\langle x \rangle_{\Omega} \langle \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}} + \overline{\langle \partial_x E_{\Omega, \text{eff}} \rangle_{\Omega}^2} \right) \\ v &= \frac{J_0}{\Delta q} \left(\overline{\langle \partial_x E_{\Omega, \text{eff}}(x, h_1) \rangle_{\Omega}} - m \Delta A \right) \\ m &= \overline{\langle x \rangle_{\Omega}} \\ t_R &= \overline{\left\langle \frac{1}{|\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)|} \right\rangle_{\Omega}}, \end{aligned} \quad (70)$$

where $\langle \dots \rangle_{\Omega}$ denotes an average over the single spin distribution $P_{\Omega}^{\text{SK}}(x, h_1, h_2)$, defined as

$$P_{\Omega}^{\text{SK}}(x, h_1, h_2) \propto |\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)| \exp \left[\frac{1}{2} \Delta C x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) - \frac{(\Delta A x - \partial_x E_{\Omega, \text{eff}}(x, h_1))^2}{2\Delta q} \right]. \quad (71)$$

Upon marginalizing over the random Gaussian fields h_1 and h_2 , this expression yields the mean-field probability distribution of the population of soft spins at the critical points of the N dimensional system, where each critical point is being weighted by an exponential factor $e^{\mu \mathcal{I}}$, which thus skews the distribution toward appropriate value of the index $\mathcal{I}$ through the chemical potential μ . The order parameters q_d, q_o are the overlap of a replica with itself and with another replica, respectively, in an assumed replica symmetric ansatz. The parameter m is the magnetization of the system, and v is a Lagrange multiplier that, when different from zero, enforces the magnetized configurations to be critical points. The order parameters $\Delta C, C_o$ contain the overlaps among different replicas of auxiliary fields that enforce the stationary condition in the Kac-Rice formula. Similarly, ΔA and A_o involve the cross-overlaps among auxiliary fields and replicated soft spin variables.

The single spin distribution at the most abundant critical points selected by the chemical potential μ can be rewritten as a weighted sum of Dirac delta distributions centered around the critical points of the effective mean field energy $E_{\Omega, \text{eff}}$, defined in Eq. (69). We show in App. G that, for any continuous test function $f(x)$, we have

$$\langle f(x) \rangle_{\Omega} = \int dx f(x) P_{\Omega}(x, h_1, h_2) \propto \left\langle \sum_{y \in \text{Crt}[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} e^{\frac{1}{2} \left(\Delta C - \frac{\Delta A^2}{\Delta q} \right) y^2 + y F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(y) + \frac{\Delta A y}{\Delta q} h_0} f(y) \right\rangle_{h_0}, \quad (72)$$

where the brackets $\langle \dots \rangle_{h_0}$ denote an average over the realizations of a new random field h_0 , which follows a Gaussian distribution of zero mean and variance Δq . The order parameters $\Delta A, A_o, \Delta C, C_o$, and v have the effect of reweighting each critical point. When these order parameters are zero, the average over the single-spin distribution reduces to

$$\langle f(x) \rangle_{\Omega} \propto \left\langle \sum_{y \in \text{Crt}[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} e^{\mu \mathcal{I}_{\text{MF}}(y)} f(y) \right\rangle_{h_0}, \quad (73)$$

where we see that only the chemical potential controls the relative weight of minima, saddle points and maxima in the computation of the averages.

The order parameters $\Delta A, \Delta C, A_o, C_o$, and v play a special role with respect to the complexity of the energy landscape. In App. G, we show that these order parameters are self-consistently equal to 0 when the effective energy $E_{\Omega, \text{eff}}(x, h)$ is a convex function of x for *all* the realizations of the field h . From Eq. (69), we see that this condition is satisfied when $a + t_R < 0$. In this phase, we show in App. G that the grand-potential is $\Omega^{\text{SK}}(\mu) = 0$ for all values of the chemical potential μ . Thus the number of most abundant critical points of any index grows at a sub-exponential rate with the size of the system. Moreover, from Eq. (65) we see that the intensive index of the most abundant critical points is 0, which implies that the most abundant critical points are minima. The grand-potential in Eq. (63) could thus have been computed dropping the absolute value from the contribution coming from the determinant, which can then be computed by introducing fermionic and bosonic fields that share a supersymmetry [52, 55–58]. We thus refer to the phase where $\Delta A = A_o = \Delta C = C_o = v = 0$ as the supersymmetric phase.

The supersymmetric phase breaks down when the effective energy $E_{\Omega, \text{eff}}(x, h)$ becomes a nonconvex function of x for some values of h , i.e. when $a + t_R > 0$. We

then have $\Omega(0) \neq 0$. Thus the number of critical points increases at an exponential rate with the size of the system. Moreover, the most abundant critical points are no longer minima, as can be seen from Eq. (68) which yields, within the replica-symmetric phase

$$r^* = \overline{\langle \mathcal{I}_{\text{MF}}(\mathbf{x}) \rangle}_{\Omega} \Big|_{\mu=0} > 0, \quad (74)$$

where the average is meant to be evaluated when the chemical potential μ is zero.

The discussion above implies that the line where $a + t_R = 0$ in the (a, J_0) plane determines the boundary separating a supersymmetric phase with a subexponential number of critical points, and a supersymmetry-broken phase where critical points proliferate at an exponential rate in the size of the system. We parametrize this phase boundary as $a_{\Sigma}(J_0)$. In the supersymmetry-broken phase the number of saddles is exponentially larger than the number of minima.

A. Phase diagram of most abundant minima and critical points

The saddle point equations in Eq. (70) are solved numerically by iteration, in the cases $\mu \rightarrow -\infty$ and $\mu = 0$, for different values of a and J_0 . The results are represented in the phase diagram in Fig. 5. By checking the change in sign of $a + t_R$, we can draw the boundary separating the supersymmetric and the supersymmetry-broken phases. In the supersymmetric phase, the grand-potential is independent from μ , so we show the results obtained for $\mu \rightarrow -\infty$. In this phase, we have $\Sigma(r) = 0$ for all intensive indexes r . The mean value of the individual soft spin distribution at the critical points is 0 for $J_0 < 1$, and grows monotonically as J_0 or a are increased when $J_0 > 1$. For large values of J_0 , the boundary of the supersymmetric phase overlaps with the boundary, displayed in Fig. 2, between the ferromagnetic phase (II) and the rigid spin-glass phase (IV).

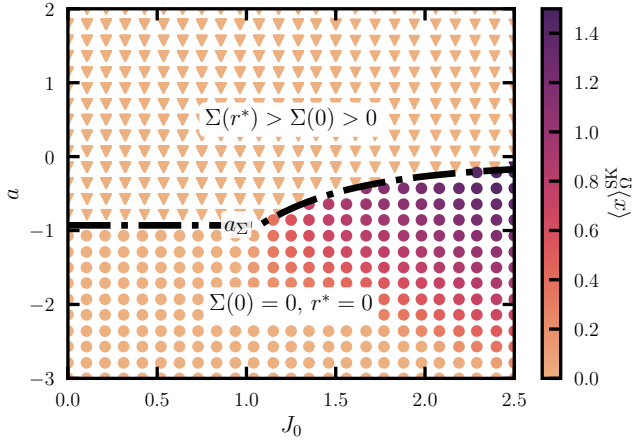


FIG. 4. **Proliferation of paramagnetic critical points** Phase diagram of the most abundant minima and critical points of the CIM-SK, as obtained from the replicated Kac-Rice calculation. The heat map is the mean of the soft spin variable $\langle x \rangle_{\Omega}^{\text{SK}}$, obtained for $\mu \rightarrow -\infty$. The dash-dotted line identifies a critical value $a_{\Sigma^*}(J_0)$ which separates a region where the complexity is zero from a region with an exponentially large (in the size of the system) number of paramagnetic local minima and critical points. In this region, the index of the most abundant critical points, r^* , is larger than 0, i.e., the most abundant critical points are saddles. The mean value of the soft spin distribution is zero both for the most abundant minima and for the most abundant critical points.

When supersymmetry is broken, the complexity of the most abundant critical points and of the minima is larger than 0. The numerical solution of the saddle point equation shows that the mean value of the soft spin distribution in this phase is 0 both for the most abundant critical point and for the case of minima. Thus at the supersymmetry-breaking transition, an exponentially large ‘cloud’ of *paramagnetic* critical points proliferate across the energy landscape. Interestingly, this picture holds also for moderately large values of the ferromagnetic coupling J_0 . The properties of the most abundant local minima for the coherent Ising machine in the absence of ferromagnetic coupling have been studied in [32]. In particular, it has been shown that in the supersymmetry-broken phase the spectrum of the Hessian around these local minima is gapless. Therefore, we conclude that also in the presence of a ferromagnetic coupling the Hessian at the most abundant local minima has a gapless spectrum, independent of the strength of J_0 .

We further study the stability of the paramagnetic solution in the supersymmetry-breaking phase against fluctuations of the mean spin value m and of the associated Lagrange multiplier v . Our analysis, carried out in App. G, shows that the paramagnetic solution is stable against these perturbations in the supersymmetry-broken phase.

The determination of the phase boundary across which the complexity becomes different from 0, and paramag-

netic critical points proliferate across the landscape, is also one of the main findings of this work. It justifies the dashed-dotted phase boundary in the yellow soft spin-glass region in Fig. 1. We next complement this finding with an illustrative visualization of the properties of the critical points at large laser gain, and with an interpretation of a subset of the supersymmetry-breaking order parameters as reactivities of the grand-potential against certain specific perturbations of the single site energy of the coherent Ising machine.

B. A portrait of critical points at large laser gain

We can compare our theory to numerical experiments as follows. For a fixed value of J_0 and a , we solve the saddle point equations (70) numerically while scanning through different values of the chemical potential μ . For each value of μ , we compute theoretical predictions for the self-overlap $q_d(\mu)$, the mean value of the soft-spin distribution $m(\mu)$, and the intensive energy $e(\mu)$ of the most abundant critical point with fixed intensive index $r(\mu)$. The intensive energy can be derived from the stationary condition of the gradient descent dynamics, Eq. (4), and it reads

$$e(\mu) = \overline{\overline{E_I(x, a)}_{\Omega}} - \frac{1}{2} \overline{\overline{x \partial_x E_I(x, a)}_{\Omega}}. \quad (75)$$

We then construct parametric plots of how each of these predictions vary with the intensive index and we compare these theoretical predictions against experimental measurements obtained from direct sampling and enumeration of critical points in small size systems, following a procedure described in App. G. We can compute pairs of properties of experimentally sampled critical points and display them in scatter plots, while comparing to theory.

A theory-experiment comparison of this type is shown in Fig. 5 for two different values of the ferromagnetic coupling J_0 , at a large value of the laser gain a . The numerical results show good agreement with the theoretical predictions, and a visual understanding of the properties of the critical points can be obtained. For both values of the ferromagnetic coupling, the intensive energy of the critical points decreases as their self-overlap increases. This observation demonstrates that the lower the energy of a critical point, the further away it is from the origin in the soft-spin configuration space. This trend is similar both without (Fig. 5a) and with (Fig. 5c) a strong ferromagnetic coupling J_0 . The main difference in the two cases arises when considering parametric scatter plots of the intensive energy as a function of the magnetization. In the absence of ferromagnetic coupling (Fig. 5b), the cloud of critical points in finite size systems has the shape of a cone, with its spread increasing at lower energies and indices. In the presence of strong ferromagnetic coupling (Fig. 5d) the bottom of the cone bifurcates into an arrowhead, whose low energy tails consist of magnetized minima. However, the most abundant critical points for

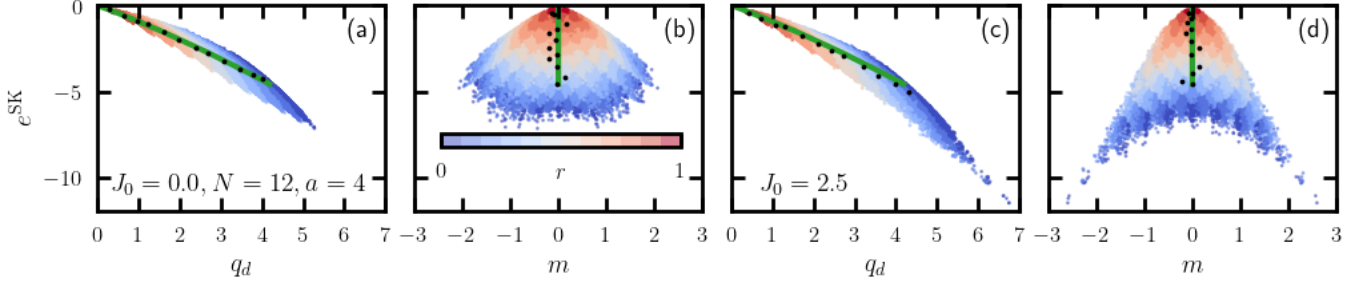


FIG. 5. **The most abundant critical points are paramagnetic.** Scatter plots of critical points sampled from systems of small size $N = 12$, at two different values of the ferromagnetic bias J_0 , for values of the laser gain a such that the ground state of the system is in a spin-glass. The green line is the prediction from the Kac-Rice theory, the color of the points is the intensive index r of the critical points, and the black dots are the properties of the most abundant critical points at fixed values of the intensive index r . Panels (a-b): scatter plots of the intensive energy as a function of the self-overlap $q_d \equiv \frac{1}{N} \sum_i \langle x_i^2 \rangle_{\mathbf{J}}$ (mean squared distance of a soft-spin configuration from the origin), and of the intensive energy as a function of the mean of the soft spin distribution $m \equiv \frac{1}{N} \sum_i \langle x_i \rangle_{\mathbf{J}}$ for $J_0 = 0$. Panels (c-d): plots of the same quantities, but obtained for a system with large ferromagnetic coupling J_0 .

any given index are paramagnetic, both with and without a strong ferromagnetic coupling J_0 .

Overall, this theory-experiment comparison reveals a qualitative illustration of the salient properties of critical points in the supersymmetry broken phase, and shows that our replicated Kac-Rice calculation faithfully predicts the geometry of the landscape of the coherent Ising machine, at the level of most abundant critical points, even for modest system sizes. A main summary is that at large laser gain a , for both weak and strong ferromagnetic coupling J_0 , ferromagnetic minima occur at lower energy, but they are exponentially rarer than the most abundant minima, which are instead paramagnetic and occur at higher energies.

C. Order parameters as reactivities of the grand-potential

We conclude this section by addressing how some supersymmetry-breaking order parameters appearing in the Kac-Rice saddle point equations (70) can be interpreted as derivatives of the grand-potential $\Omega^{\text{SK}}(\mu)$ with respect to perturbations of the single-site energy of the CIM. These perturbations are implemented by substituting the contribution from each single-site energy $E_I(x_i, a)$ in Eq. (2) with a modified version $E_I(x, a, \epsilon)$, given by

$$E_I(x_i, a, \epsilon) = \frac{1}{4}x_i^4 - \frac{1}{2}(a + \epsilon_A)x_i^2 - \epsilon_v J_0 x_i + \sqrt{2\epsilon_C} g_i x_i. \quad (76)$$

The vector $\epsilon \equiv [\epsilon_A, \epsilon_v, \epsilon_C]^T$ contains the strength of the different types of perturbations performed on the CIM energy. They correspond respectively to changing the laser gain by an amount ϵ_A , applying an external magnetic field of intensity $\epsilon_v J_0$, and applying a set of external random i.i.d Gaussian fields g_i , with mean 0 and variance

1. In App. G we detail how these perturbations modify the grand-potential, and obtain the relations

$$\begin{aligned} \Delta A + A_o &= -\partial_{\epsilon_A} \Omega^{\text{SK}}(\mu) \Big|_{\epsilon=0} \\ \Delta C &= -\partial_{\epsilon_C} \Omega^{\text{SK}}(\mu) \Big|_{\epsilon=0} \\ v &= -\partial_{\epsilon_v} \Omega^{\text{SK}}(\mu) \Big|_{\epsilon=0}, \end{aligned} \quad (77)$$

which connect the order parameters to susceptibilities of the grand-potential Ω^{SK} under external perturbations of the single-site energy of the coherent Ising machine. In light of this analysis, in the supersymmetry-broken phase the number of critical points grows substantially upon external perturbation of the landscape. In Sec. VII A, we found that $v = 0$ for the most abundant critical points even in the supersymmetry-broken phase. This implies that the landscape has a low reactivity to linear perturbations by means of an external magnetic field.

The analysis carried out so far implies a fundamental discordance in the structure of the energy landscape at different energy levels: for large enough values of the ferromagnetic coupling J_0 and the laser gain a , global minima (and also rare low energy local minima) of the system are magnetized, but they are screened by exponentially many more higher energy most abundant local minima that are paramagnetic in nature. Given this landscape complexity, it is not clear how the dynamics of CIM energy minimization will then proceed, while the energy landscape geometry is annealed, by slowly increasing the laser gain. To address this important question, we now turn to directly analyzing the dynamics of the annealing process.

VIII. ANNEALED OPTIMIZATION DYNAMICS OF THE CIM

A. Numerical simulations

We first simulate the landscape annealing dynamics of the CIM for a system of $N = 5 \times 10^3$ spins. We focus on the case $J_0 > 1$, since the behavior of the annealing dynamics for $J_0 = 0$ has been explored in [32], and for $J_0 < 1$ we saw that the global minima of the energy landscape are paramagnetic. For a given value of J_0 , we first sample an instance of the connectivity matrix $\mathbf{J}$, and we quench the system at a fixed value of the laser gain a in the convex ferromagnetic region of the phase diagram, i.e. phase (II) in Fig. 2. This sets the initial condition at time $t = 0$. We then simulate the annealing process by integrating numerically Eq. (4) using Euler's method [59] with a time-step $\Delta t = 0.15$. The laser gain $a = a(t)$ is slowly increased during the dynamics following the protocol

$$a(t) = a(0) + \dot{a}t, \quad (78)$$

with a fixed rate $\dot{a}$. As the dynamics proceeds, we track the minimum eigenvalue of the Hessian of the system $\lambda_{\min}^{\text{SK}}(a(t))$, with the entries of the Hessian given by Eq. (9). We also track the intensive Ising energy of the system $e_{\text{Ising}}^{\text{SK}}(a(t)) = -\frac{1}{2N} \mathbf{s}(t) \cdot \mathbf{J} \mathbf{s}(t)$, with $s_i(t) = \text{sgn}[x_i(t)]$. These measurements are then averaged over 50 independent realizations of the quenched couplings $\mathbf{J}$. The effectiveness of the annealing process is compared against a spectral method, where the ground state configuration is approximated by $\mathbf{s}_{\text{spectral}}$. The latter is obtained by looking for the real eigenvector with the largest eigenvalue of $\mathbf{J}$, and taking the sign of its entries. The approximated ground state is then used to obtain a spectral approximation to the energy $e_{\text{spectral}}^{\text{SK}} \equiv -\frac{1}{2N} \mathbf{s}_{\text{spectral}} \cdot \mathbf{J} \mathbf{s}_{\text{spectral}}$. Results for an annealing rate $\tau \dot{a} = 10^{-4}$ and two representative values of J_0 are shown in Fig. 6.

At small values of the laser gain, the annealing process takes place in the convex phase. The spectrum of the Hessian is gapped, and the gap monotonically decreases as the laser gain increases. Interestingly, the Ising energy of the system decreases too, as the system aligns more and more toward the ferromagnetic solution of the combinatorial optimization problem. This trend continues until the laser gain reaches the spin-glass transition value a_{sg} . The dynamics of the system becomes qualitatively different depending on whether the laser gain value for the rigidity transition, a_r , is identical to or different from the spin-glass transition point a_{sg} .

If $a_{\text{sg}} \neq a_r$, as in Fig. 6(a-b), the minimum value of the Hessian remains close to zero in the interval $[a_{\text{sg}}, a_r]$. By looking at individual trajectories, represented as semi-transparent lines in the figure, we observe that the lower edge of the spectrum can sometimes change sign. During these sign changes, the Ising energy of the system decreases in sudden steps. This phenomenon corresponds

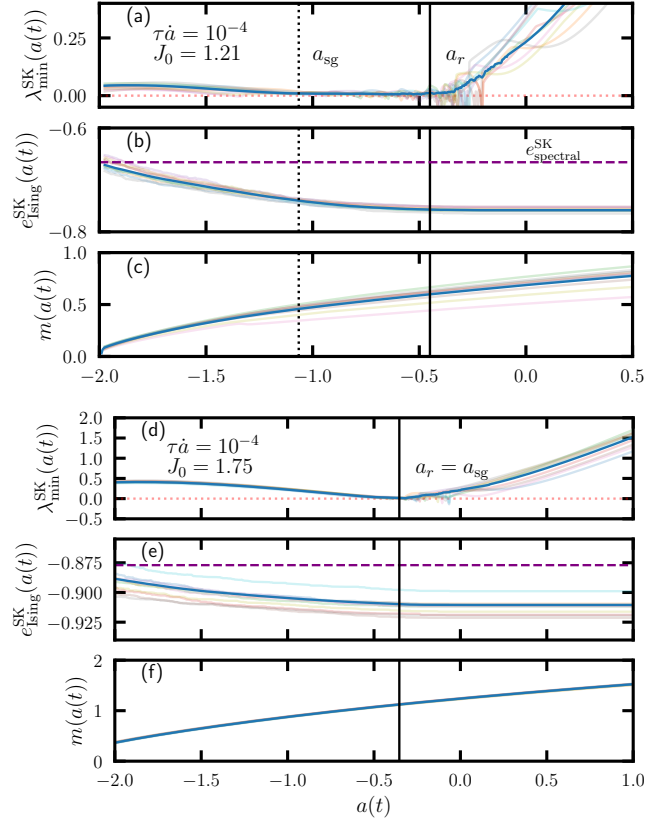


FIG. 6. Annealing dynamics and rigidity transition

Plots of the minimum value of the spectrum of the Hessian $\lambda_{\min}^{\text{SK}}(a(t))$ and the Ising energy $e_{\text{Ising}}^{\text{SK}}(a(t))$ during the annealed dynamics of the coherent Ising machine for two different values of J_0 . The solid blue curves represent an average over 50 independent runs, while the 10 curves in the background are individual trajectories, included for illustrations. The horizontal dashed purple line denotes the average of the spectral approximation to the Ising energy across the different realizations of $\mathbf{J}$. The vertical dotted and solid ones denote the values of a_{sg} and a_r at which a spin-glass and a rigidity transition occur, respectively. Panel (a): parametric plot of $\lambda_{\min}^{\text{SK}}(a(t))$ for $J_0 = 1.21$. The spin-glass transition and the rigidity transition take place at two different values of the laser gain, a_{sg} and a_r , respectively. Panel (b): Parametric plot of the Ising energy $e_{\text{Ising}}^{\text{SK}}(a(t))$ for $J_0 = 1.21$. Panel (c): Parametric plot of the mean value of the soft spins $m(a(t))$ during the annealing process. Panel (d): Parametric plot of $\lambda_{\min}^{\text{SK}}(a(t))$ for $J_0 = 1.75$, where the rigidity transition and the spin-glass transition occur at the same value of the laser gain, $a_r = a_{\text{sg}}$. Panel (e): Parametric plot of the Ising energy $e_{\text{Ising}}^{\text{SK}}(a(t))$ for $J_0 = 1.75$. Panel (f): parametric plot of the mean value of the soft spin population $m(a(t))$ during the annealing process.

to the bifurcation of the current local stable minimum, where the system resides, into an unstable saddle. When such a bifurcation from stable minimum to unstable saddle occurs, the soft spins escape the saddle through its unstable direction(s), toward new configurations in which some of the spins may change their sign, yielding a lower

Ising energy. In the average trajectory, these jumps are smoothed out in time, and the Ising energy gradually decreases as the laser gain increases. When $a \gtrsim a_r$, the spectral gap starts to increase with the laser gain, and the Ising energy stops decreasing, reaching a plateau value. The system gets trapped in a minimum with a gapped spectrum of the Hessian, from which it cannot escape by an adiabatic increase of the laser gain.

In the second case, when $a_{sg} = a_r$, as in Fig. 6 (c-d), the typical value of the lower edge of the spectrum of the Hessian touches the origin and then starts rising again for $a \approx a_{sg}$, and the Ising energy immediately enters its plateau region. Some instability transitions are observed at the level of individual trajectories, but the average value of the edge of the spectrum increases monotonically with a . In this situation, annealing beyond the convex phase yields marginal reductions in the Ising energy reached by the system. Interestingly, however, in both cases considered the annealing process yields Ising configurations with lower energies compared to the spectral solution, even in the convex phase.

Finally, we observe that for both values of J_0 , the mean value of the population of soft spins during the annealing dynamics, $m(a(t))$, is different from 0. The critical point in the vicinity of which the dynamics takes place are thus *not* one of the most abundant ones identified by means of the Kac-Rice calculation in Sec. VII. We observe that $m(a(t))$ grows monotonically with a . This can be understood by observing that as the laser gain increases, the separation between the two wells in the single site energy E_I , defined in Eq. (3), increases too. The dynamics thus takes place within a region of the phase space whose center is located further and further away from the origin as the laser gain increases, leading to larger m .

The rigidity transition values we identified by means of our static calculation in Sec. VIB holds for the global minima of the system. These configurations may be different from the typical minima encountered by the annealed optimization dynamics. However, the fact that the position of the lower edge of the spectrum of the Hessian increases with a above the rigidity transition demonstrates that the properties of global minima are a good proxy of the properties of the typical minima explored by the CIM during landscape annealing. At the end of the annealing schedule displayed in Fig. 6, we measure the energy of the coherent Ising machine per spin $e^{SK}(a(t))$. The final values obtained in the cases investigated are $e^{SK}(a(t) = 1) \approx -1.72 \pm 0.01$ for $J_0 = 1.21$, and $e^{SK}(a(t) = 1) \approx -2.16 \pm 0.02$ for $J_0 = 1.75$. These quantities are very close to the value of the ground state energy for the corresponding values of the laser gain and ferromagnetic coupling that can be computed from the analysis in Sec. VI, which are respectively $e_{gs}^{SK} \approx -1.83$ and $e_{gs}^{SK} \approx -2.16$.

Motivated by the above observations, we compare the properties of the typical soft-spin configurations visited during the annealing processes with the properties of global minima, most abundant local minima, and typi-

cal configurations reached through fast quenches at fixed laser gain a . Our results are shown in Fig. 7.

We compare the CIM energy, the mean value of the soft spin population, and the self-overlap obtained during the annealed process. As soon as the complexity of local minima becomes nonzero, the properties of the most abundant minima become considerably different from the properties of the points visited by both quenched and annealed dynamics. The average CIM energy, the soft-spin population mean value, as well as the self-overlap q_d at any given value of the laser gain in the annealing or the quench process are considerably closer to the values obtained for global minima rather than local minima. This difference becomes particularly striking when the mean-value of the soft-spin population $m(a)$ is considered, as this quantity is identically zero at all the values of the laser gain for the most abundant local minima, and increases with the laser gain both for global minima and as the annealed dynamics unfolds. Interestingly both the annealed dynamics and quenches at fixed laser gain reach very similar configurations from the point of view of CIM energy, magnetization and mean squared value of the soft spin population. This result suggests that for intermediate values of the laser gain, the annealed optimization dynamics can successfully evade the exponential proliferation of most abundant higher energy local minima, and pierce down to much lower energy levels, to visit spin configurations more similar to lower energy global minima.

To deepen our understanding of the dynamics of the soft spins across the rigidity transition, we develop a dynamical mean field theory (DMFT) which describes the annealing dynamics of the CIM in the thermodynamic limit, and we use it to describe how the probability distribution of the soft spins evolves during the annealing process.

B. Dynamical mean field theory of the annealed optimization dynamics

In the thermodynamic limit, the annealed landscape optimization dynamics of the CIM can be described using dynamical mean field theory (DMFT), involving an effective Langevin equation for the dynamics of a soft spin of the system during the annealing process. The derivation of the effective process using the dynamical cavity method [60] is presented in App. H. Here we give an intuitive understanding of the terms involved. Let us consider the dynamics of a single tagged spin in the system. Since the number of interactions per spin grows linearly with system size, the interaction of the tagged spin with the rest of the system acts as a small perturbation to the dynamics of a system of $N - 1$ spins, from which the tagged spin has been excluded (i.e. the cavity system of $N - 1$ spins). Therefore, the interactions among the tagged spin and the rest of the system are expanded using linear response theory, and decomposed into a ran-

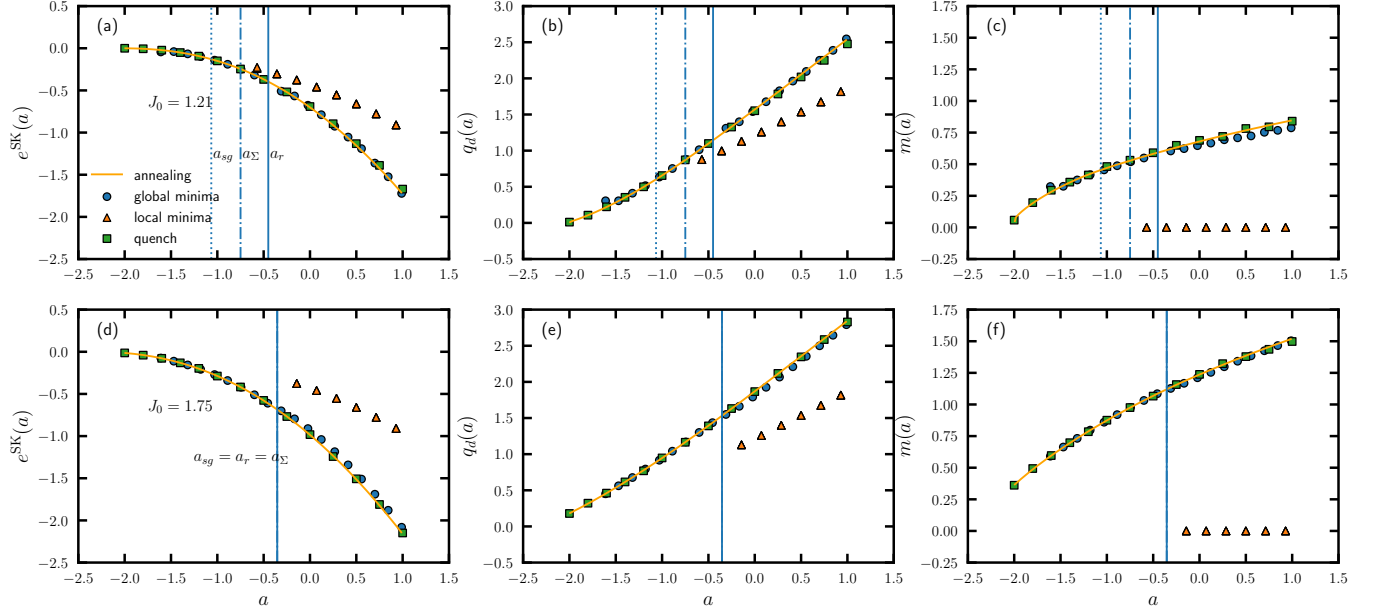


FIG. 7. Configurations visited by quench and annealed dynamics are more similar to global minima than most abundant local minima. The top and bottom row correspond to two different values of J_0 , respectively. In each panel, we show the behavior of relevant average quantities as a function of the laser gain for: annealed dynamics (orange solid line); global minima (blue circles); local minima (red triangles); terminal state of a quenched dynamics at fixed a (green squares). The average quantities of interest are: the CIM-energy per spin $e^{\text{SK}}(a)$ (panels (a) and (d)); the mean squared value of the soft spin population $q_a(a)$ (panels (b) and (e)); the mean value of the soft spin population $m(a)$ (panels (c) and (f)). The vertical blue lines signal the values of a at which the spin-glass transition (a_{sg} , dotted line), the supersymmetry breaking transition (a_{Σ} , dot-dashed line), and the rigidity transition (a_r , solid line) take place, respectively. The orange solid line is obtained from numerical simulations of the annealing dynamics for a system of $N = 5000$ spins and an annealing rate $\dot{a}\tau = 10^{-4}$ (averaged over 50 independent trajectories). The blue circles are obtained from the numerical integration of the full replica symmetry breaking solution for the properties of global minima. The red triangles are obtained from the numerical solution of the Kac-Rice calculation carried out in Sec. VII for the properties of the most abundant local minima. The green squares are obtained from the terminal configuration of a quench of the dynamics at fixed a , with an initialization from the neighborhood of the eigenvector with the largest eigenvalue of the matrix $\mathbf{J}$ (averaged over 10 independent trajectories). Overall the panels show from left to right that the annealed dynamics explores spin configurations more similar to global minima, with lower energy (left), higher distance from the origin (middle), and higher magnetization (right) than that of the most abundant local minima.

dom force, $\xi(t)$, and a response term that encodes the composite effect of the the perturbation that the tagged spin exerts on the cavity and the resultant back reaction of the cavity onto the tagged spin. Because of the ferromagnetic coupling J_0 , the tagged spin is also subject to a magnetic field of strength $J_0 m(t)$, where $m(t)$ is the mean value of the soft-spin population as the dynamics takes place. In the thermodynamic limit, the random force $\xi(t)$ has Gaussian statistics due to the central limit theorem. The single-spin effective process $x(t)$ thus reads

$$\tau \dot{x}(t) = -\partial_x E_I(x(t), a(t)) + J_0 m(t) + \int_0^t d\tau R(t, \tau) x(\tau) + \xi(t). \quad (79)$$

The noise $\xi(t)$ is Gaussian with zero mean and temporal correlations defined as

$$\langle \xi(t) \xi(t') \rangle_{\mathbf{J}} = C(t, t'). \quad (80)$$

Because of the fully connected, mean field nature of the model, the knowledge of the statistics of the sin-

gle spin dynamics allows us to self-consistently determine the noise autocorrelation function $C(t, t')$, the response function $R(t, t')$ and the magnetization $m(t)$, through the conditions

$$\begin{aligned} R(t, t') &= \left\langle \frac{\delta x(t)}{\delta \xi(t')} \right\rangle_{\mathbf{J}} \\ C(t, t') &= \langle x(t) x(t') \rangle_{\mathbf{J}} \\ m(t) &= \langle x(t) \rangle_{\mathbf{J}}, \end{aligned} \quad (81)$$

Equations (79), (80) and Eq. (81) are a closed set of equations which constitute the DMFT of the CIM. They can be solved numerically for arbitrary initial conditions and annealing schedule through an iterative scheme described in Appendix H. Once these equations are solved, the time dependent single-spin distribution $P^{\text{SK}}(x, t)$ at time t is obtained by running the dynamics given by Eq. (79) under different realizations of the noise $\xi(t)$. In Fig. 8, we show how the probability distribution of soft spins evolves at fixed $J_0 = 1.75$, as we anneal the system across

the rigidity transition. The initial condition $P(x, t = 0)$ is taken to be the single spin distribution at the global minimum for $a(t = 0) = -1$, which lies in the ferromagnetic, replica-symmetric region of the phase diagram in Fig. 2. In this initial condition, the distribution of spins is skewed toward the positive region of the x axis, because of the influence of the ferromagnetic alignment, and the value of the probability density at the origin is finite. As the annealing dynamics proceeds, the value of $P(0, t)$ decreases, until it becomes zero approximately at the rigidity transition. For $a(t) \gtrsim a_r$, the single-spin probability distribution becomes gapped, and the width of the gap increases with a . This result supports, and rationalizes, the observation that above the rigidity transition the annealing process yields no improvement to the Ising energy of the system: as the laser gain increases, flipping a spin by a small increase of a becomes harder and harder. Moreover, the dynamical mean field theory successfully predicts the evolution of the single-spin probability distribution for a moderate system size of $N = 2000$ spins.

In summary, we have developed a framework to identify the rigidity transition in the ground states of the CIM (Sec. VIB), and we have provided evidence that the line across which this transition occurs is a good indicator of the boundary below which the annealing process allows to reach configurations with lower Ising energy (Sec. VIII and Sec. VIIB). Moreover, the properties of the configurations visited during both quenched and annealed dynamics are much more similar to those of global minima rather than those of most abundant local minima. The strength of the ferromagnetic coupling J_0 determines whether the rigidity transition occurs within the spin-glass phase or at the boundary between the convex and the spin-glass phase.

In the next section, we turn to a model where a specific ground state is engineered into a tunable, rugged energy landscape, and ask the question of how effective the CIM is in recovering such a planted solution.

IX. WISHART PLANTED ENSEMBLE

The Wishart planted ensemble (WPE) [34] is a fully connected, binary spin model with a known ground state, called the planted solution, and a tunable parameter that controls how hard it is to find the planted ground state through sampling methods such as parallel tempering [61]. The connectivity matrix $\mathbf{J}^{\text{WPE}}$ is tuned in such a way that the ground state of the Ising problem is given by the direction $\mathbf{t} = [1, \dots, 1]^T$, namely

$$\mathbf{t} \in \arg \min_{\mathbf{s} \in \{-1, 1\}^N} (-\mathbf{s} \cdot \mathbf{J}^{\text{WPE}} \mathbf{s}). \quad (82)$$

This is achieved by defining $\mathbf{J}^{\text{WPE}}$ as

$$\begin{aligned} \mathbf{J}^{\text{WPE}} &= \tilde{\mathbf{J}} - \text{diag } \tilde{\mathbf{J}} \\ \tilde{\mathbf{J}} &= -\frac{1}{N} \sum_{\mu=1}^M \mathbf{w}^\mu \otimes \mathbf{w}^\mu. \end{aligned} \quad (83)$$

The diagonal part of $\mathbf{J}^{\text{WPE}}$ is set to zero to hinder the recovery of the ground state using a spectral method like the one described in Sec. VIIB. The matrix $\tilde{\mathbf{J}}$ is an $N \times N$ real-valued matrix, composed by a sum of outer products of M random N -dimensional patterns $\{\mathbf{w}^\mu\}_{\mu=1}^M$. The patterns $\mathbf{w}^\mu$ are correlated random variables, defined as follows so all of them are orthogonal to $\mathbf{t}$:

$$\begin{aligned} \mathbf{w}^\mu &\equiv \Sigma^{1/2} \mathbf{z}^\mu \\ \Sigma^{1/2} &\equiv \sqrt{\frac{N}{N-1}} \left[\mathbf{1} - \frac{1}{N} \mathbf{t} \otimes \mathbf{t}^T \right], \end{aligned} \quad (84)$$

with $\{\mathbf{z}^\mu\}_{\mu=1}^M$ a set of M i.i.d. Gaussian vectors with zero mean and identity covariance. The matrix $\Sigma^{1/2}$ projects (and scales) any vector onto the orthogonal complement of $\mathbf{t}$. The definition of $\mathbf{w}^\mu$ implies

$$\mathbf{w}^\mu \cdot \mathbf{t} = 0, \quad (85)$$

and this orthogonality condition implies $\tilde{\mathbf{J}} \mathbf{t} = \mathbf{0}$. In this way, the connectivity matrix $\mathbf{J}^{\text{WPE}}$, defined in Eq. (83), satisfies Eq. (82). In fact, for any binary spin configuration $\mathbf{s} \in \{-1, 1\}^N$, we have

$$-\mathbf{s} \cdot \mathbf{J}^{\text{WPE}} \mathbf{s} = \frac{1}{N} \sum_{\mu=1}^M |\mathbf{w}^\mu \cdot \mathbf{s}|^2 + \text{Tr } \tilde{\mathbf{J}} \geq \text{Tr } \tilde{\mathbf{J}}. \quad (86)$$

The last inequality becomes tight when $\mathbf{s} = \mathbf{t}$, which implies that $\mathbf{t}$ is a ground state of the system. The Wishart planted ensemble can be seen as an anti-Hopfield model [62] with correlated patterns: the connectivity matrix $\mathbf{J}^{\text{WPE}}$ stores M patterns, which are correlated in such a way as to be all orthogonal to the planted direction $\mathbf{t}$. The minus sign in the definition of $\mathbf{J}^{\text{WPE}}$ ensures that the spin configurations that minimize the Ising energy $-\mathbf{s} \cdot \mathbf{J}^{\text{WPE}} \mathbf{s}$ are repelled by the stored patterns, instead of being attracted to them, as in the case of the traditional Hopfield model [63, 64]. The relevant parameter determining the hardness of finding a ground state is the ratio between the dimensionality of the spin configuration, N , and the number of patterns M . We denote this ratio by $\alpha \equiv M/N$, with larger values of α indicating easier optimization problems, as we describe next.

The phase diagram of this model with Ising spins and a finite temperature has been studied in [34], assuming a replica-symmetric structure. For $\alpha > 1$, at high temperatures the global minimum of the free energy is paramagnetic. As the temperature is reduced below a critical value $T_c(\alpha)$ the system crosses a first order phase transition toward a state where a local paramagnetic minimum coexists with the global, ferromagnetic minimum of

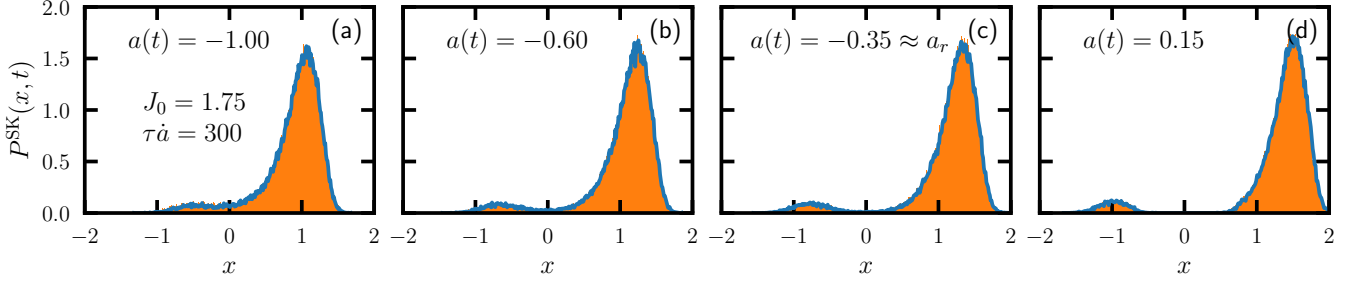


FIG. 8. **Dynamical mean field theory of the annealing process of the coherent Ising machine.** Plots of the DMFT prediction for the single spin probability distribution $P(x, t)$ (blue lines) for the annealed dynamics of the CIM with $J_0 = 1.75$, for which $a_r = a_{\text{sg}}$. The orange histograms are obtained from the numerical integration of Eq. (4) for 50 independent realizations of the disorder for a system of $N = 2000$ spins. The different panels show snapshots of $P(x, t)$ at different stages of the annealing, depending on the value of the laser gain with respect to a_r . The system is initialized at the global minimum for $a(0) = -1$ and is then gradually annealed with a constant rate $\tau\dot{a}$.

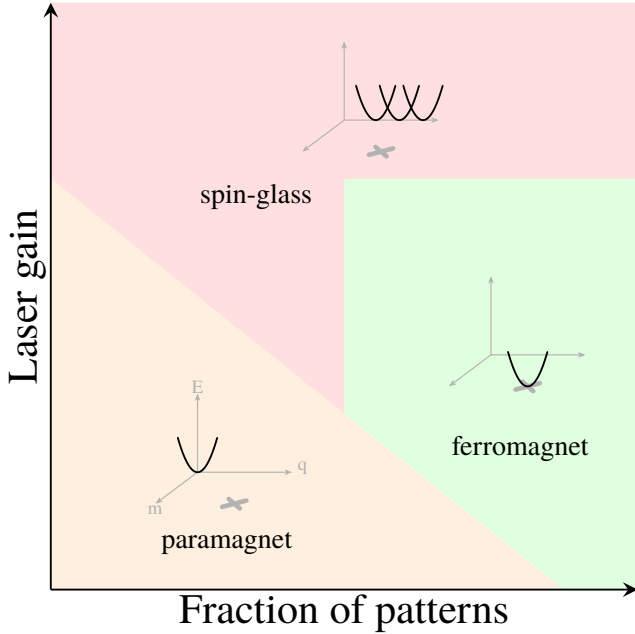


FIG. 9. **Schematic phase diagram of global minima of the coherent Ising machine with Wishart planted ensemble.** Our mean field theory predicts the existence of three different regions, depending on the strength of the laser gain and fraction of patterns encoded in the connectivity matrix. At low laser gain, and below a critical threshold of the fraction of pattern, the system is a paramagnet, with the soft spin distribution concentrating at the origin, and the planted solution is not recovered. The boundary region of the paramagnetic phase is determined in Sec. IX A. At intermediate value of the laser gain, if the fraction of patterns is above a critical threshold, the system is a ferromagnet (Sec. IX C), and full recovery of the planted solution is possible. At large value of the laser gain, the system is in a spin-glass phase, and the gradient descent dynamics from uninformed initial condition does not allow to recover the planted solution (Sec. IX D).

the free energy. Upon further lowering the temperature below a value $T_m(\alpha)$, another transition occurs toward a state where the free energy is convex again, and the global minimum largely overlaps with the planted direction $\mathbf{t}$. In this phase the ground state can be attained easily through Monte Carlo sampling methods. On the other hand, for $\alpha < 1$, only the first order transition at $T_c(\alpha)$ is present. Below the critical temperature, the ferromagnetic state is the global minimum of the free energy, but its basin of attraction is very small, while the paramagnetic state is a stable local minimum with a large basin of attraction. The metastability of the paramagnetic state at all temperatures below $T_c(\alpha)$ for $\alpha < 1$ hinders the retrieval of the planted solution.

In the following, we study the properties of global minima of the CIM with the Wishart planted random connectivity matrix $\mathbf{J}^{\text{WPE}}$. The soft spins $\mathbf{x}$ perform a gradient descent dynamics of the energy function E^{WPE} , defined as

$$E^{\text{WPE}}(\mathbf{x}) \equiv \sum_{i=1}^N E_I(x_i, a) - \frac{1}{2} \mathbf{x} \cdot \mathbf{J}^{\text{WPE}} \mathbf{x}, \quad (87)$$

where the single site energy $E_I(x, a)$ is defined in Eq. (3). We refer to this model as a CIM-WPE. We study the properties of the global minima of the CIM-WPE energy as the laser gain a and the fraction of patterns α stored in $\mathbf{J}^{\text{WPE}}$ is varied. As is done for the CIM-SK, we first address the properties of the Hessian $H_{ij}^{\text{WPE}}(\mathbf{x}) \equiv \frac{\partial^2 E^{\text{WPE}}(\mathbf{x})}{\partial x_i \partial x_j}$. The probability density of its eigenvalues averaged over different realizations of the disorder, $\rho_{\mathbf{x}}^{\text{WPE}}(\lambda)$, is defined as

$$\rho_{\mathbf{x}}^{\text{WPE}}(\lambda) = \lim_{N \rightarrow \infty} \left\langle \frac{1}{N} \sum_{i=1}^N \delta(\lambda_i^{\text{WPE}}(\mathbf{x}) - \lambda) \right\rangle_{\mathbf{z}^\mu}, \quad (88)$$

where $\{\lambda_i^{\text{WPE}}(\mathbf{x})\}_{i=1}^N$ is the set of eigenvalues of the Hessian matrix $\mathbf{H}^{\text{WPE}}(\mathbf{x})$, and $\langle \dots \rangle_{\mathbf{z}^\mu}$ is an average over the realization of the Gaussian random variables $\mathbf{z}^\mu$, appearing in Eq. (84). To evaluate the spectral distribution of

global minima of the energy E^{WPE} , we can average $\rho_{\mathbf{x}}^{\text{WPE}}$ over the Boltzmann distribution $P_{\text{B}}^{\text{WPE}}(\mathbf{x})$, defined as

$$\rho_{\text{B}}^{\text{WPE}}(\mathbf{x}) = \frac{e^{-\beta E^{\text{WPE}}(\mathbf{x})}}{\int d\mathbf{x}' e^{-\beta E^{\text{WPE}}(\mathbf{x}')}}, \quad (89)$$

with β the inverse temperature. In the limit $\beta \rightarrow \infty$, this distribution concentrates around global minima of E^{WPE} . We can then consider the spectral distribution $\rho^{\text{WPE}}(\lambda)$, defined as

$$\rho^{\text{WPE}}(\lambda) \equiv \lim_{N \rightarrow \infty} \left\langle \frac{1}{N} \sum_{i=1}^N \delta(\lambda_i^{\text{WPE}}(\mathbf{x}) - \lambda) \right\rangle_{\mathbf{x}, \mathbf{z}^\mu}, \quad (90)$$

where the average $\langle \dots \rangle_{\mathbf{x}}$ denotes an average of the soft spin configurations $\mathbf{x}$ over the Boltzmann distribution for the CIM-WPE, $P_{\text{B}}^{\text{WPE}}$. The spectral distribution $\rho^{\text{WPE}}(\lambda)$ can be equivalently obtained from the knowledge of the resolvent $G^{\text{WPE}}(z)$, defined as

$$G^{\text{WPE}}(z) \equiv \int d\lambda \frac{\rho^{\text{WPE}}(\lambda)}{z - \lambda}. \quad (91)$$

Once $G^{\text{WPE}}(z)$ is known, it can be inverted to determine the spectral distribution ρ^{WPE} , through the same procedure described in Sec. V for the CIM-SK. To gather some intuition, we start by looking at low values of the laser gain a .

A. Low laser gain

For low values of the laser gain $a \ll 0$, the single-site energy of the CIM, E_I in Eq. (87), dominates over the interaction term. As in the CIM-SK, we expect the origin to be the global minimum of the energy of the system. The Hessian $\mathbf{H}^{\text{WPE}}(\mathbf{x})$ evaluated at the origin is

$$\mathbf{H}^{\text{WPE}}(\mathbf{0}) = -\mathbf{J}^{\text{WPE}} - a\mathbf{1}. \quad (92)$$

From the definition of $\mathbf{J}^{\text{WPE}}$ given by Eq. (83), we see that an eigenvector of $\mathbf{H}^{\text{WPE}}(\mathbf{0})$ is given by $\mathbf{t}/\sqrt{N}$. The associated eigenvalue is

$$\frac{1}{N} \mathbf{t} \cdot \mathbf{H}^{\text{WPE}}(\mathbf{0}) \mathbf{t} = \frac{1}{N} \text{Tr} \tilde{\mathbf{J}} - a \approx -\alpha - a, \quad (93)$$

where in the first equality we used the definition of $\mathbf{J}^{\text{WPE}}$, given by Eq. (83), and in the second we observed that in the large N limit the quantity $\text{Tr} \tilde{\mathbf{J}}$ concentrates around its mean value $-M$ by the central limit theorem. This eigenvalue becomes negative when

$$a > -\alpha, \quad (94)$$

which yields a lower bound to the region in the (α, a) plane where the origin is no longer a stable minimum of the CIM-WPE energy. From Eq. (92), we see that the spectrum of the Hessian around the origin follows

the same distribution of the spectrum of the connectivity matrix $-\mathbf{J}^{\text{WPE}}$ with the support shifted by an amount a . The spectral distribution of $-\mathbf{J}^{\text{WPE}}$ has been computed in [34]. It follows a Marchenko-Pastur law with a persistent Dirac delta distribution due to the presence of the planted solution $\mathbf{t}$, and a shift of the support by an amount α since $\mathbf{J}^{\text{WPE}}$ is constructed by subtracting from the matrix $\tilde{\mathbf{J}}$ its diagonal elements. Overall, the spectral distribution around the origin is

$$\rho_{\mathbf{0}}^{\text{WPE}}(\lambda) = \begin{cases} \alpha \delta(\lambda + a + \alpha) + \rho_{\text{MP}}(\lambda + a + \alpha) & \alpha \leq 1 \\ \frac{1}{N} \delta(\lambda + a + \alpha) + \rho_{\text{MP}}(\lambda + a + \alpha) & \alpha > 1 \end{cases}, \quad (95)$$

where $\rho_{\text{MP}}(\lambda)$ is the bulk of the Marchenko-Pastur law, which reads

$$\rho_{\text{MP}}(\lambda) \equiv \frac{\sqrt{(\lambda - \lambda_-)(\lambda_+ - \lambda)}}{2\pi\lambda} \quad (96)$$

$$\lambda_{\pm} = (1 \pm \sqrt{\alpha})^2, \quad \lambda \in [\lambda_-, \lambda_+]$$

and, with a slight abuse of notation, we have included the persistent Dirac delta at the origin for $\alpha > 1$, even though its weight vanishes in the thermodynamic limit. This mass is responsible for the instability of the origin as the laser gain increases above $-\alpha$.

B. Resolvent and single-spin distribution

To determine the resolvent of the CIM-WPE in regions where the global minimum of the system is different from the origin, we assume that we can neglect the correlations between the matrix elements of $\mathbf{J}^{\text{WPE}}$ and the location of the global minimum of the CIM energy. This assumption is justified by the calculation performed above for the SK model. Within this approximation, the Hessian $\mathbf{H}^{\text{WPE}}(\mathbf{x})$ is given by a random diagonal perturbation to $\mathbf{J}^{\text{WPE}}$. The resolvent of $\mathbf{J}^{\text{WPE}}$ is the resolvent of a Wishart matrix [35], shifted by an amount $\alpha\mathbf{1}$ due to the subtraction of the diagonal element in Eq. (83). The self-consistent equation satisfied by the resolvent of $G^{\text{WPE}}(z)$ is then given by [46]

$$G^{\text{WPE}}(z) = \int d\mathbf{x} \frac{P^{\text{WPE}}(\mathbf{x})}{z - \partial_x^2 E_I(\mathbf{x}, a + \alpha) - \frac{\alpha}{1 - G^{\text{WPE}}(z)}}, \quad (97)$$

where $P^{\text{WPE}}(\mathbf{x})$ is the probability distribution of a soft spin at the global minima of the CIM-WPE. It is obtained by computing the probability to find any spin in the system in a given configuration $\mathbf{x}$ in the thermodynamic limit,

$$\lim_{N \rightarrow \infty} \left\langle \frac{1}{N} \sum_{i=1}^N \delta(y_i - x) \right\rangle_{\mathbf{y}}^{\text{WPE}} = P^{\text{WPE}}(x). \quad (98)$$

This computation is performed in Appendix I using the replica trick. The final result is

$$P^{\text{WPE}}(x) = \lim_{n \rightarrow 0} \frac{1}{n} \sum_{\alpha=1}^n \langle \delta(x - x^\alpha) \rangle_{\text{MF}}^{\text{WPE}}. \quad (99)$$

The mean field average $\langle \dots \rangle_{\text{MF}}^{\text{WPE}}$ is over a set of n replicated soft spins,

$$\langle \dots \rangle_{\text{MF}}^{\text{WPE}} = \frac{1}{Z_{\text{MF}}^{\text{WPE}}} \int \prod_{\alpha=1}^n dx^\alpha e^{-\beta E_{\text{MF}}^{\text{WPE}}(\{x^\alpha\})}, \quad (100)$$

with the limit $n \rightarrow 0$ to be eventually taken by means of an analytic continuation. The mean field energy $E_{\text{eff}}^{\text{WPE}}$ is⁵

$$E_{\text{MF}}^{\text{WPE}} = \sum_{\beta=1}^n E_I(x^\alpha, a + \alpha) - \frac{\alpha\beta}{2} \sum_{\alpha,\beta} r_{\alpha\beta} x^\alpha x^\beta - \alpha \sum_{\alpha=1}^n \hat{m}_\alpha x^\alpha. \quad (101)$$

and the mean field partition function $Z_{\text{MF}}^{\text{WPE}}$ is

$$Z_{\text{MF}}^{\text{WPE}} = \int \prod_{\alpha=1}^n dx^\alpha e^{-\beta E_{\text{MF}}^{\text{WPE}}(\{x^\alpha\})}. \quad (102)$$

We see from Eq. (101) that the laser gain is shifted by an amount equal to the fraction of stored patterns α . The order parameters $r_{\alpha\beta}$, $\hat{m}_\alpha$ arise as a consequence of the Wishart statistics of the connectivity matrix, and they are related to the matrix of overlaps $\mathbf{Q}$ among the different replicas and a magnetization vector $\mathbf{m}$ encoding the overlap of each replica with the direction $\mathbf{t}$. The saddle point equations determining the different order parameters are

$$\begin{aligned} \beta r_{\alpha\beta} &= -\frac{\partial}{\partial \beta Q_{\alpha\beta}} \text{Tr} \log[\mathbf{1}_n + \beta \mathbf{Q} - \beta \mathbf{m} \otimes \mathbf{m}] \\ \hat{m}_\alpha &= -\frac{1}{2} \frac{\partial}{\partial \beta m_\alpha} \text{Tr} \log[\mathbf{1}_n + \beta \mathbf{Q} - \beta \mathbf{m} \otimes \mathbf{m}] \\ Q_{\alpha\beta} &= \langle x^\alpha x^\beta \rangle_{\text{MF}}^{\text{WPE}} \\ m_\alpha &= \langle x^\alpha \rangle_{\text{MF}}^{\text{WPE}}. \end{aligned} \quad (103)$$

These can be solved once an ansatz for the structure of the order parameters in replica space is chosen. In what follows, we study the replica-symmetric ansatz.

C. Convex phase

The replica symmetric ansatz for the CIM-WPE is

$$\begin{aligned} r_{\alpha\beta} &= \delta_{\alpha\beta} r_d + (1 - \delta_{\alpha\beta}) r_o \\ Q_{\alpha\beta} &= \delta_{\alpha\beta} q_d + (1 - \delta_{\alpha\beta}) q_o \\ m_\alpha &= m \\ \hat{m}_\alpha &= \hat{m}. \end{aligned} \quad (104)$$

Since we work in the low temperature limit $\beta \rightarrow \infty$, it is convenient to define the rescaled gaps

$$\begin{aligned} \Delta \tilde{r} &\equiv \beta(r_d - r_o) \\ \Delta \tilde{q} &\equiv \beta(q_d - q_o). \end{aligned} \quad (105)$$

In Appendix I we compute the single-spin probability $P_{\text{RS}}^{\text{WPE}}(x)$ in the replica symmetric phase, starting from Eq. (99). The final result is

$$P_{\text{RS}}^{\text{WPE}}(x) = e^{-\beta E_{\text{eff}}^{\text{WPE}}(x, h)} \left[\int dx' e^{-\beta E_{\text{eff}}^{\text{WPE}}(x', h)} \right]^{-1}, \quad (106)$$

where $\overline{\dots} \equiv \int dh \dots \frac{1}{\sqrt{2\pi}} e^{-h^2/2}$ is an average over the realizations of the Gaussian quenched disorder h . The replica-symmetric mean field energy $E_{\text{eff}}^{\text{WPE}}(x, h)$ reads

$$E_{\text{eff}}^{\text{WPE}}(x, h) \equiv E_I(x, a + \alpha + \alpha \Delta \tilde{r}) - (\sqrt{\alpha r_o} h + \alpha \hat{m}) x. \quad (107)$$

The laser gain in the mean field energy $E_{\text{eff}}^{\text{WPE}}$ is shifted by an amount $\alpha + \alpha \Delta \tilde{r}$, while the order parameter $\alpha \hat{m}$ acts as a magnetic field induced by the presence of the correlated patterns orthogonal to $\mathbf{t}$. In the limit $\beta \rightarrow \infty$, the single spin probability distribution $P_{\text{RS}}^{\text{WPE}}(x)$ concentrates, for each realization of the disorder h , around the minimum $x^*(h)$ of the mean field energy $E_{\text{eff}}^{\text{WPE}}(x, h)$. The saddle point equations Eq. (103) then become

$$\begin{aligned} \Delta \tilde{q} &= \overline{[\partial_x^2 E_{\text{eff}}^{\text{WPE}}(x^*(h), h)]^{-1}} \\ q_o &= \overline{(x^*(h))^2} \\ m &= \overline{x^*(h)} \\ \Delta \tilde{r} &= -\frac{1}{1 + \Delta \tilde{q}} \\ r_o &= \frac{q_o - m^2}{(1 + \Delta \tilde{q})^2} \\ \hat{m} &= \frac{m}{1 + \Delta \tilde{q}}. \end{aligned} \quad (108)$$

In App. H, we obtain a condition for the stability of the replica-symmetric solution based on the sign of the replicon $\tilde{\Lambda}_{\text{R}}^{\text{WPE}}$, which reads

$$\tilde{\Lambda}_{\text{R}}^{\text{WPE}} \equiv -1 + \alpha \Delta \tilde{r}^2 \overline{[\partial_x^2 E_{\text{eff}}^{\text{WPE}}(x^*(h), h)]^{-2}} < 0. \quad (109)$$

The saddle point equations given by Eq. (108) can be solved numerically, while checking the stability condition

⁵Note that, with a slight abuse of notation, we are using the symbol α to denote both the fraction of patterns stored in $\mathbf{J}^{\text{WPE}}$ and a replica index, and we are using the symbol β to denote both the inverse temperature and a replica index.

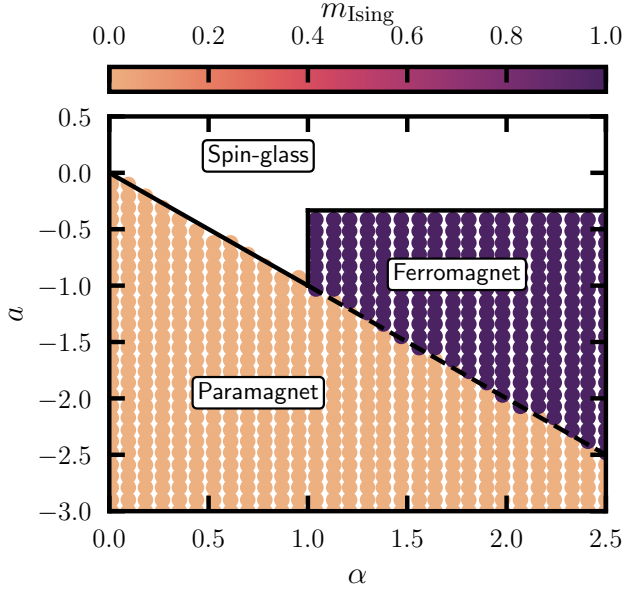


FIG. 10. **Phase diagram of global minima for the coherent Ising machine with Wishart planted ensemble.** The color-code is the Ising magnetization of the global minimum of the CIM-WPE as the laser gain a and the fraction of patterns α are changed. We identify a convex paramagnetic phase, where the soft spins concentrate at the origin; a ferromagnetic phase, where $m_{\text{Ising}} = 1$ the planted solution $\mathbf{t}$ is fully recovered, and a spin-glass phase (white region in the phase diagram). The phase diagram has been obtained by numerical integration of the saddle point equations in Eq. (103), and the phase boundaries are obtained analytically, as discussed in Sec. IX C and Sec. IX C.

in Eq. (109), through an iterative method. Once the order parameters are known, we can obtain the Ising magnetization of the CIM-WPE at global minima from the single spin distribution $P_{\text{RS}}^{\text{WPE}}(x)$, according to the formula

$$m_{\text{Ising}} = \int_0^\infty dx [2P_{\text{RS}}^{\text{WPE}}(x)] - 1, \quad (110)$$

where single spin distribution is evaluated in the low temperature limit $\beta \rightarrow \infty$.

The resulting phase diagram is displayed in Fig. 10. For $a < -\alpha$ the global minimum is paramagnetic. As anticipated in Sec. IX A, the single site energy dominates over the interaction term of the CIM-WPE. As a increases, two different scenarios occur depending on the fraction of patterns α stored into the connectivity matrix $\mathbf{J}^{\text{WPE}}$. If $\alpha < 1$, the replicon eigenvalue $\tilde{\Lambda}_{\text{R}}^{\text{WPE}}$ in Eq. (109) becomes positive for $a > -\alpha$. The replica symmetric solution is no longer stable, and the system enters a spin-glass phase. If instead $\alpha > 1$, the paramagnetic minimum becomes unstable for $a > -\alpha$, and the global minimum of the system becomes ferromagnetic. The Ising magnetization m_{Ising} jumps discontinuously from 0 to 1 upon crossing this boundary, meaning that the CIM fully retrieves the planted ground state $\mathbf{t}$. Interestingly,

even for $\alpha > 1$, upon further increase of the laser gain, the system undergoes a spin-glass transition at $a = 1/3$. The presence of a spin-glass phase suggests that for $\alpha > 1$ the annealing process needs to be run from small values of the laser gain in order to successfully retrieve the ground state. We next discuss how these phases can be derived analytically from the saddle point equations in Eq. (108), and comment on the form taken by the spectrum of the Hessian and the single-spin distribution.

1. Convex paramagnetic phase

For $a < -\alpha$, we assume a paramagnetic ansatz $r_o = q_o = m = \hat{m} = 0$. The single-spin distribution $P_{\text{RS}}^{\text{WPE}}(x)$ is then a Dirac delta distribution centered around the origin,

$$P_{\text{RS}}^{\text{WPE}}(x) = \delta(x). \quad (111)$$

The saddle point equations for $\Delta\tilde{q}$ and $\Delta\tilde{r}$ then become

$$\begin{aligned} \Delta\tilde{q} &= -\frac{1}{a + \alpha + \alpha\Delta\tilde{r}} \\ \Delta\tilde{r} &= -\frac{1}{1 + \Delta\tilde{q}}, \end{aligned} \quad (112)$$

while the stability condition in Eq. (109) becomes

$$-1 + \frac{\alpha\Delta\tilde{q}^2}{(1 + \Delta\tilde{q})^2} < 0, \quad (113)$$

The order parameter $\Delta\tilde{q}$ can be obtained by solving Eq. (112), and it reads

$$\Delta\tilde{q} = \frac{-1 - a + \sqrt{(1 - a)^2 - 4\alpha}}{a + \alpha}. \quad (114)$$

As $a \rightarrow -\alpha$, the overlap gap $\Delta\tilde{q}$ diverges, and the replicon eigenvalue tends to $\tilde{\Lambda}_{\text{R}}^{\text{WPE}} \rightarrow -a + \alpha$. If $\alpha > 1$, we then obtain that the replicon eigenvalue changes sign continuously as a passes through the values $-\alpha$, and the paramagnetic phase becomes unstable. If $\alpha < 1$, the mean field energy $E_{\text{eff}}^{\text{WPE}}$ becomes nonconvex at the transition. The replicon eigenvalue jumps then discontinuously to ∞ . The condition for the convexity of the effective single-site energy $E_{\text{eff}}^{\text{WPE}}(x, h)$ is $a + \alpha + \alpha\Delta\tilde{r} < 0$, or, using the saddle point equations Eq. (108),

$$a + \frac{\alpha\Delta\tilde{q}}{1 + \Delta\tilde{q}} < 0. \quad (115)$$

Substituting Eq. (114) into the equation above, we obtain that the convex paramagnetic phase is no longer stable when $a > -\alpha$, and that a transition toward a spin-glass phase occurs for $\alpha < 1$.

2. Ferromagnetic minimum

For $\alpha > 1$, the numerical solution of the saddle point equations shows that $r_o = 0$, while q_o , m and $\hat{m}$ are different from zero. This type of solution implies that the single spin distribution is a Dirac delta centered around the point x^* that minimizes the replica-symmetric mean field energy $E_{\text{eff}}^{\text{WPE}}(x)$ in Eq. (107), which no longer depends on the quenched random field h :

$$\begin{aligned} P_{\text{RS}}^{\text{WPE}}(x) &= \delta(x - x^*) \\ x^{*3} - (a + \alpha(\Delta\tilde{r} + 1))x^* - \alpha\hat{m} &= 0. \end{aligned} \quad (116)$$

Since $x^* \neq 0$, the Ising magnetization of the system becomes $m_{\text{Ising}} = 1$. The coherent Ising machine fully recovers the planted ground state in this phase.

The instability boundary of the ferromagnetic phase can be found analytically using our replica calculation. We first observe that, when $\alpha > 1$ the replicon eigenvalue in Eq. (109) can jump discontinuously from a negative value to ∞ when the energy function $E_{\text{eff}}^{\text{WPE}}$ in Eq. (107) becomes non convex, i.e. when the stability condition in Eq. (115) is violated. To determine the laser gain value for which the instability takes place, we need to find x^* . This can be done by observing that, if $r_o = 0$, the saddle point equations Eq. (108) yields $\hat{m} = \frac{x^*}{1 + \Delta\tilde{q}}$. Substituting this expression in the stationary condition in Eq. (116) we obtain $x^* = \sqrt{a + \alpha}$, which is well-defined in the region $a > -\alpha$. The saddle point equation Eq. (108) for $\Delta\tilde{q}$ becomes

$$\Delta\tilde{q} = \frac{1}{3x^{*2} - a - \alpha\Delta\tilde{q}/(1 + \Delta\tilde{q})}. \quad (117)$$

Substituting the known value of x^* allows to solve for $\Delta\tilde{q}$, obtaining

$$\Delta\tilde{q} = \frac{1 - 2a - 3\alpha - \sqrt{1 + 4a + 4a^2 + 2\alpha + 12a\alpha + 9\alpha^2}}{4(a + \alpha)}. \quad (118)$$

Substituting this expression into the convexity condition in Eq. (115) and solving for a allows us to conclude that the ferromagnetic solution is stable when $\alpha > 1$ and the laser gain a obeys the bounds

$$-\alpha < a < -\frac{1}{3}. \quad (119)$$

This is the purple region shown in Fig. 10, and we have concluded our discussion of the stability boundary of the ferromagnetic solution.

The spectrum of the Hessian of the ferromagnetic minima in the replica-symmetric phase is given by a Marchenko-Pastur distribution shifted by an amount $2(a + \alpha)$. This can be seen by substituting the single spin probability distribution of Eq. (116) into the expression

for the resolvent given by Eq. (91). The result is

$$\begin{aligned} G^{\text{WPE}}(z) &= \frac{1}{z - 3x^{*2} - a - \alpha - \frac{\alpha}{1 - G^{\text{WPE}}(z)}} \\ &= \frac{1}{z - 2(a + \alpha) - \frac{\alpha}{1 - G^{\text{WPE}}(z)}}, \end{aligned} \quad (120)$$

which is the self-consistent equation for the resolvent of a Wishart matrix with parameter α , perturbed by a constant diagonal matrix $2(a + \alpha)\mathbf{1}$.

After discussing the properties of global minima of the CIM-WPE, we turn to a study annealing dynamics of the system and the degree to which the planted ground state can be recovered.

D. Dynamics

We perform numerical simulation of the annealing process of the CIM-WPE by integrating numerically the gradient descent dynamics

$$\tau \frac{d}{dt} \mathbf{x}(t) = -\nabla E^{\text{WPE}}(\mathbf{x}), \quad (121)$$

for a system of $N = 2000$ spins. For a fixed value of α , we first initialize the system with a quench at $a = a(0)$ within the paramagnetic, replica symmetric phase. We then run the annealing dynamics, using the same protocol of Sec. VIII. During the annealing process, we track the position of the lower edge of the spectrum of the Hessian of the system $\lambda_{\text{min}}^{\text{WPE}}(t)$ in its current configuration, its Ising magnetization $m_{\text{Ising}}(t)$ and its Ising energy $e_{\text{Ising}}^{\text{WPE}}(t)$. The process is repeated for 50 independent realizations of the disordered couplings $\mathbf{J}^{\text{WPE}}$. The results are displayed in Fig. 11, where we show results for the representative value of $\alpha = 0.8$. The lower edge of the spectrum of the Hessian remains close to the origin in an interval of values above the spin-glass transition value $a_{\text{sg}} = -\alpha$. In the region where the lower edge of the Hessian remains close to the origin, the Ising energy of the system decreases through configurational changes similar to the ones witnessed for the Sherrington-Kirkpatrick model in Sec. VIII. The Ising magnetization of the system, shown in the inset of Fig. 11(a) remains close to zero, indicating that the ground state of the system is far from being recovered. The Ising energy found through the annealing dynamics is lower than the one obtained with a spectral method, represented by a purple dashed line in Fig. 11(b). This analysis suggests that for $\alpha < 1$, there is a paramagnetic spin-glass phase in the CIM-WPE, where the system finds lower energy solutions thanks to the annealing dynamics but struggles to align with the planted solution.

To investigate the spin-glass phase for $\alpha > 1$, we simulate fast quenches from a random initial conditions for a density of stored patterns $\alpha = 1.5$. The soft spins at initial times are independently sampled from a Gaussian distribution of standard deviation 0.1. We perform quenches for two different values of the laser gain, at

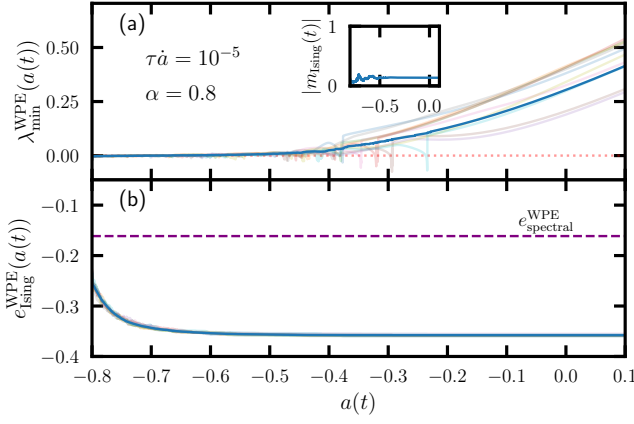


FIG. 11. **Annealing dynamics of the coherent Ising machine in the Wishart planted ensemble** Panel (a): lower edge of the spectrum of the Hessian as a function of the laser gain $a(t)$. The solid blue line represent the average over different realizations of the dynamics and the connectivity matrix $\mathbf{J}^{\text{WPE}}$, while the semitransparent lines are random individual trajectories. The lower edge remains close to the value of zero (signaled by the red dotted line) at the beginning of the annealing process and starts increasing around $a(t) \approx 0.4$. In the inset, we show the Ising magnetization $m_{\text{Ising}}(t)$ as a function of the laser gain. Panel (b): Ising energy of the system during the annealing process. The dashed purple line is the estimate of the ground state energy obtained from a spectral method.

$a = -1.0$ and $a = 0.0$, which correspond to the ferromagnetic and spin-glass states in the phase diagram in Fig. 10, respectively. The time evolution of the Ising energy and Ising magnetization during the quenches are shown in Fig. 12. When $a < a_{\text{sg}}$, the Ising magnetization quickly converges to 1, and the ground state is completely retrieved. The energy of the ground state matches that found using a spectral method, as expected when $\alpha > 1$. When $a > a_{\text{sg}}$, the Ising magnetization at long times becomes considerably smaller, and the Ising energy recovered by the coherent Ising machine is greater than the ground-state energy. These results demonstrate the importance of initializing the annealing process from low values of the laser gain in the coherent Ising machine.

X. OUTLOOK

In this work, we employ a variety of techniques from the statistical physics of disordered systems, including the replica method, random matrix theory, Kac-Rice formulas, and dynamical mean field theory, to gain insights into how the evolving geometry of a high dimensional energy landscape impacts the efficacy of the annealed optimization dynamics of the coherent Ising machine. We have characterized the low-energy minima of the coherent Ising machine when tackling the Sherrington Kirkpatrick model with ferromagnetic coupling, and found a region

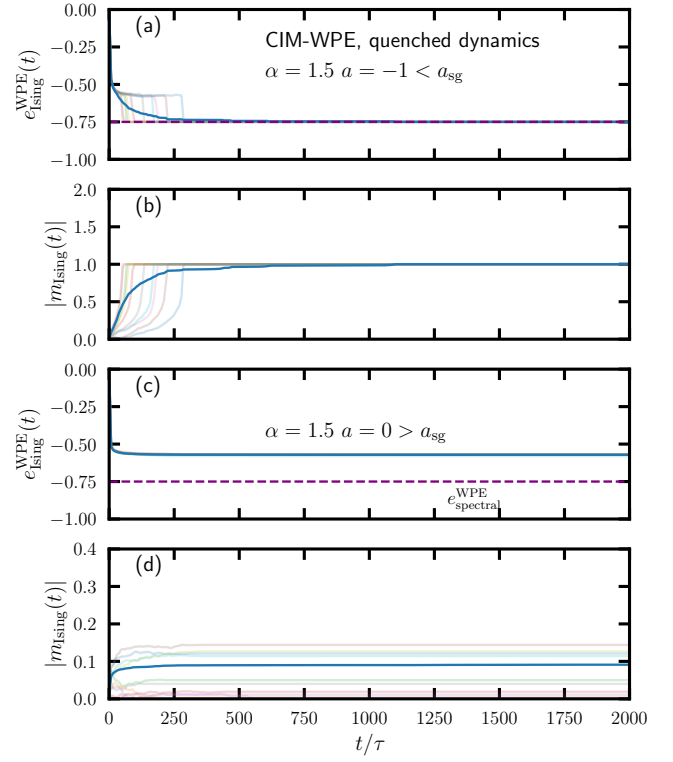


FIG. 12. **Quench dynamics of the CIM-WPE in the convex and in the spin-glass phase.** Plots of the Ising energy $e_{\text{Ising}}(t)$ and of the Ising magnetization $m_{\text{Ising}}(t)$ as a function of time during a quench from random initial condition at fixed density of patterns $\alpha = 1.5$, for two values of the laser gain. Panels (a-b): results for a quench in the convex ferromagnetic phase, $a < a_{\text{sg}}$. After an initial transient, the ground state is successfully recovered. Panels (c-d): results for a quench in the spin-glass phase, $a > a_{\text{sg}}$. at long times, the Ising configuration of the system displays only a partial overlap with the ground state.

in the space of possible laser gain and ferromagnetic coupling where the global minima present soft modes. The appearance of these soft modes is tied with the nature of the spin-glass phase of the system and with the shape of the single site energy of the CIM. Our study of the annealed optimization dynamics shows that these soft modes are exploited to reach lower energies of the associated combinatorial optimization problem during the annealing process, yielding a mechanistic explanation for the efficacy of the CIM. Also, through a replicated Kac-Rice calculation, we showed that the annealing dynamics reaches states with a net nonzero magnetization in regions of the landscape where the most abundant critical points and minima are instead paramagnetic. We highlight here some venues of research to be explored building on our work.

The first venue concerns the design and evaluation of novel Ising machine architectures. Soft modes appear as a crucial element in the CIM that determine the effectiveness of the annealing process. As these soft modes

are tied to the shape of the single-site energies, it would be interesting to study the impact of higher order self-interactions on the appearance and lifetime of the soft modes, as well as the effect of heterogeneous modulations of the laser gain across the different units, mimicking mean field models of supercooled liquids [44, 47, 65].

The second venue concerns the study of alternative dynamics for the CIM. Here, low-energy Ising configurations are found using a gradient descent dynamics. What are limits and possibilities when nongradient dynamics are employed [66–69] is an interesting avenue of research that could be tackled using the dynamical mean field theory employed here. The latter could be combined with high-dimensional optimal control [70], to design annealing schedules subject to a fixed energy and time budget.

A third avenue points at sharpening even further our understanding of the geometry of the annealed landscape and its relation with the dynamics. The Kac-Rice calculation presented here addresses the properties of the most abundant critical points, independent of their energy. The construction of a full map of the critical points and a characterization of the linear excitations around them, conditioned on any possible energy level, is a challenging task, both from the numerical and the analytical standpoints, that future efforts could address.

ACKNOWLEDGMENTS

We thank Aditya Mahadevan and Pierfrancesco Urbani for very insightful discussions. We are grateful to

J  r  mie Klinger and Aditya Mahadevan for useful feedbacks on an early version of the manuscript. We thank the Stanford Sherlock cluster for computing resources. F.G. acknowledges support from a postdoctoral fellowship of the Stanford’s Leinweber Institute of Theoretical Physics. A.S. acknowledges funding from the National Science Foundation Graduate Research Fellowship, A.Y. acknowledges support from the Masason foundation. S.G. thanks the Simons foundation and a Schmidt foundation polymath award for support. This work was supported by the National Science Foundation, under the awards CCF-2423832 and CCF-1918549. During the finalization of this manuscript, we became aware of a parallel work [71] addressing the efficacy of the coherent Ising machine in the absence of ferromagnetic coupling, and how its performance can be improved by suitable modifications of the CIM energy landscape.

Appendix A: Derivation of Eq. (25): Computing the resolvent

We start by substituting the expression of the CIM-SK energy $E^{\text{SK}}(\mathbf{x}, a)$, given by Eq. (6) into the integral of Eq. (24), which thus reads

$$\begin{aligned} \int_{\mathbf{x}^\alpha, \phi^{\alpha\rho}} \left\langle \exp \left[-\beta \sum_{\alpha=1}^n \sum_{i \leq j}^N \left(\delta_{ij} E_I(x_i^\alpha, a) - \frac{J_{ij}}{\sqrt{N}} x_i^\alpha x_j^\alpha \right) + \frac{\beta J_0}{2N} \sum_{\alpha=1}^n \left(\sum_{i=1}^N x_i^\alpha \right)^2 \right. \right. \\ \left. \left. + \frac{i}{2} \sum_{\alpha=1}^n \sum_{\rho=1}^l \sum_{i,j}^N \phi_i^{\alpha\rho} \left[(z - \partial_x^2 E_I(x_i^\alpha, a)) \delta_{ij} + \frac{J_{ij}}{\sqrt{N}} + \frac{J_0}{2N} \right] \phi_j^{\alpha\rho} \right] \right\rangle_{\mathbf{J}}, \end{aligned} \quad (\text{A1})$$

where we recall that the quantity z is a shorthand notation for $z + i0^+$, the limiting value to the real axis of a complex variable with positive imaginary part. This choice ensures the convergence of the integrals over the bosonic variables $\phi_i^{\alpha\rho}$. The Gaussian average $\langle \dots \rangle_{\mathbf{J}}$ over the realizations of the disorder is evaluated using the integral

$$\langle e^{\gamma J_{kl}} \rangle_{\mathbf{J}} = \int \prod_{i \leq j} \frac{dJ_{ij} e^{-\frac{1}{2(1+\delta_{ij})} J_{ij}^2}}{\sqrt{2\pi}} e^{\gamma J_{kl}} = e^{\frac{\gamma^2}{2}(1+\delta_{kl})}, \quad (\text{A2})$$

valid for any $\gamma \in \mathbb{C}$, and pair of indices k, l . In the large N limit, the difference between the variance of diagonal and off-diagonal elements can be neglected, as its contribution in the exponential of Eq. (A2) is of order $O(1)$, while we anticipate the leading contributions to grow linearly with N . Upon averaging over the disorder, the integral in Eq. (A2) becomes

$$\begin{aligned} \int_{\mathbf{x}^\alpha, \phi^{\alpha\rho}} \exp \left[-\beta \sum_{\alpha} \sum_{i=1}^N E_I(x_i^\alpha, a) + \frac{N\beta^2}{4} \sum_{\alpha, \beta} \left(\frac{\mathbf{x}^\alpha \cdot \mathbf{x}^\beta}{N} \right)^2 + \frac{N\beta J_0}{2} \sum_{\alpha} \left(\frac{\mathbf{x}^\alpha \cdot \mathbf{t}}{N} \right)^2 + \frac{iN\beta}{2} \sum_{\alpha, \beta, \rho} \left(\frac{\mathbf{x}^\alpha \cdot \phi^{\beta\rho}}{N} \right)^2 \right. \\ \left. - \frac{N}{4} \sum_{\alpha, \beta, \rho, \sigma} \left(\frac{\phi^{\alpha\rho} \cdot \phi^{\beta\sigma}}{N} \right)^2 + \frac{iN J_0}{2} \sum_{\alpha, \rho} \left(\frac{\phi^{\alpha\rho} \cdot \mathbf{t}}{N} \right)^2 + \frac{i}{2} \sum_{\alpha, \rho} \sum_i (\phi_i^{\alpha\rho})^2 (z - \partial_x^2 E_I(x_i^\alpha, a)) \right], \end{aligned} \quad (\text{A3})$$

where the indices α, β run from 1 to n , while the indices ρ, σ run from 1 to l , and we recall that $\mathbf{t}$ is a N -dimensional vector with all entries equal to one, $\mathbf{t} \equiv [1, 1, \dots, 1]^T$. We note that, with a slight ambiguity, the symbol β denotes both the inverse temperature and the index of the replicated soft spin. The quadratic terms can be linearized by means of a Hubbard-Stratonovich transformation

$$e^{\frac{\gamma}{2N}x^2} = \int_{-\infty}^{\infty} \frac{du}{\sqrt{2\pi/N\gamma}} e^{-\frac{N\gamma}{2}u^2 + \gamma ux}, \quad (\text{A4})$$

with γ a real positive quantity. This transformation introduces integrals over new variables which, as shown below, concentrate around the order parameters of the system in the thermodynamic limit. For instance, the term involving coupling among replicas of the soft spin system $\mathbf{x}^\alpha$ is linearized by introducing an $n \times n$ matrix

$\mathbf{Q}$, thus becoming

$$\exp \left[\frac{N\beta^2}{4} \sum_{\alpha,\beta} \left(\frac{\mathbf{x}^\alpha \cdot \mathbf{x}^\beta}{N} \right)^2 \right] = \int \prod_{\alpha \leq \beta} \frac{dQ_{\alpha\beta}}{\sqrt{2\pi/(N\beta^2)}} \\ \times \exp \left[-\frac{N\beta^2}{4} \sum_{\alpha,\beta} Q_{\alpha\beta}^2 + \frac{\beta^2}{2} \sum_{\alpha,\beta} Q_{\alpha\beta} (\mathbf{x}^\alpha \cdot \mathbf{x}^\beta) \right]. \quad (\text{A5})$$

Similar manipulations can be done for the other terms. This leads to the introduction of an $n \times n \times l$ tensor $\mathbf{T}$, a $n \times l$ matrix $\mathbf{M}$, an $n \times n \times l \times l$ tensor $\mathbf{P}$, and an n -dimensional vector $\mathbf{m}$. We denote this set of new variables by $\mathbf{\Gamma} = \{\mathbf{Q}, \mathbf{m}, \mathbf{T}, \mathbf{P}, \mathbf{M}\}$. From Eq. (A5), we see that the Hubbard-Stratonovich transformation decouples spins and fields belonging to different sites within a single replica, at the expense of introducing couplings between spins at the same site of distinct replicas. The integral $\int_{\mathbf{x}^\alpha, \phi^{\alpha\rho}}$ thus factorizes into a product of N integrals over the $n \times l$ dimensional space of single-site, replicated soft spins x^α and fields $\phi^{\alpha\rho}$. Equation (A3) thus becomes

$$\int d\mathbf{\Gamma} \exp \left[-\frac{N\beta^2}{4} \sum_{\alpha,\beta} Q_{\alpha\beta}^2 - \frac{N\beta J_0}{2} \sum_{\alpha} m_{\alpha}^2 - \frac{iN\beta}{2} \sum_{\alpha,\beta,\rho} T_{\alpha\beta\rho}^2 - \frac{N}{4} \sum_{\alpha,\beta,\rho,\sigma} P_{\alpha\beta\rho\sigma}^2 - \frac{iNJ_0}{2} \sum_{\alpha,\rho} M_{\alpha\rho}^2 \right] \\ \times \left\{ \int \prod_{\alpha=1}^n \prod_{\rho=1}^l \frac{dx^\alpha d\phi^{\alpha\rho}}{\sqrt{2\pi i}} \exp \left[-\beta \sum_{i=1}^N E_I(x_i^\alpha) + \frac{\beta^2}{2} \sum_{\alpha,\beta} Q_{\alpha\beta} x^\alpha x^\beta + \beta J_0 \sum_{\alpha} m_{\alpha} x^\alpha + i\beta \sum_{\alpha,\beta,\rho} T_{\alpha\beta\rho} x^\alpha \phi^{\beta\rho} \right. \right. \\ \left. \left. - \frac{i}{2} \sum_{\alpha,\beta,\rho,\sigma} \phi^{\alpha\rho} \phi^{\beta\sigma} P_{\alpha\beta\rho\sigma} + iJ_0 \sum_{\alpha,\rho} M_{\alpha\rho} \phi^{\alpha\rho} + \frac{i}{2} \sum_{\alpha,\rho} (\phi^{\alpha\rho})^2 (z - \partial_x^2 E_I(x^\alpha, a)) \right] \right\}^N \\ \equiv \int d\mathbf{\Gamma} e^{N S^{\text{SK}}[\mathbf{\Gamma}]}. \quad (\text{A6})$$

We neglect the normalization factors in the differential $d\mathbf{\Gamma} \equiv \prod_{\alpha \leq \beta}^n \prod_{\rho \leq \sigma}^l dQ_{\alpha\beta} dm_{\alpha} dT_{\alpha\beta\rho} dP_{\alpha\beta\rho\sigma} dM_{\alpha\rho}$, which contribute only with subleading order in N to the exponential. The last line defines the action $S[\mathbf{\Gamma}, z]$ as

$$S^{\text{SK}}[\mathbf{\Gamma}, z] \equiv -\frac{\beta^2}{4} \sum_{\alpha,\beta} Q_{\alpha\beta}^2 - \frac{\beta J_0}{2} \sum_{\alpha} m_{\alpha}^2 - \frac{i\beta}{2} \sum_{\alpha,\beta,\rho} T_{\alpha\beta\rho}^2 - \frac{1}{4} \sum_{\alpha,\beta,\rho,\sigma} P_{\alpha\beta\rho\sigma}^2 - \frac{iJ_0}{2} \sum_{\alpha,\rho} M_{\alpha\rho}^2 + \log Z_{x,\phi}^{\text{SK}} \\ Z_{x,\phi}^{\text{SK}} = \int \prod_{\alpha} \prod_{\rho} \frac{dx^\alpha d\phi^{\alpha\rho}}{\sqrt{2\pi i}} \exp \left[-\beta \sum_{\alpha=1}^n E_I(x^\alpha, a) + \frac{\beta^2}{2} \sum_{\alpha,\beta} Q_{\alpha\beta} x^\alpha x^\beta + \beta J_0 \sum_{\alpha} m_{\alpha} x^\alpha + i\beta \sum_{\alpha,\beta,\rho} T_{\alpha\beta\rho} x^\alpha \phi^{\beta\rho} \right. \\ \left. - \frac{i}{2} \sum_{\alpha,\beta,\rho,\sigma} \phi^{\alpha\rho} \phi^{\beta\sigma} P_{\alpha\beta\rho\sigma} + iJ_0 \sum_{\alpha,\rho} M_{\alpha\rho} \phi^{\alpha\rho} + \frac{i}{2} \sum_{\alpha,\rho} (\phi^{\alpha\rho})^2 (z - \partial_x^2 E_I(x^\alpha, a)) \right] \\ \equiv \int \prod_{\alpha} \prod_{\rho} \frac{dx^\alpha d\phi^{\alpha\rho}}{\sqrt{2\pi i}} e^{-\beta E_{x,\phi}^{\text{SK}}[x^\alpha, \phi^{\alpha\rho}]} \quad (\text{A7})$$

$Z_{x,\phi}^{\text{SK}}$ is a mean field partition function of the system, computed in the space of the single-site replicated soft

spins x^α and fields $\phi^{\alpha\rho}$.

Appendix B: Saddle point equations

In this section we evaluate the saddle point equations $\partial_{\mathbf{\Gamma}} S^{\text{SK}}[\mathbf{\Gamma}, z] = \mathbf{0}$ for the action $S^{\text{SK}}[\mathbf{\Gamma}, z]$ given by Eq. (A7). Taking the derivative with respect to the different order parameters we obtain

$$\begin{aligned} Q_{\alpha\beta} &= \langle x^\alpha x^\beta \rangle_{x,\phi}^{\text{SK}} & m_\alpha &= \langle x^\alpha \rangle_{x,\phi}^{\text{SK}} \\ T_{\alpha\beta\rho} &= \langle x^\alpha \phi^{\beta\rho} \rangle_{x,\phi}^{\text{SK}} & M_{\alpha\rho} &= \langle \phi^{\alpha\rho} \rangle_{x,\phi}^{\text{SK}} \\ P_{\alpha\beta\rho\sigma} &= -i \langle \phi^{\alpha\rho} \phi^{\beta\sigma} \rangle_{x,\phi}^{\text{SK}}, \end{aligned} \quad (\text{B1})$$

where the average $\langle \dots \rangle_{x,\phi}$ is performed over the replicated soft spin variables and bosonic fields, and is defined as

$$\langle \dots \rangle_{x,\phi}^{\text{SK}} \equiv \frac{1}{Z_{x,\phi}^{\text{SK}}} \int \mathbf{x}^\alpha, \phi^{\alpha\rho} \dots e^{-\beta E_{x,\phi}^{\text{SK}}[\{\mathbf{x}^\alpha, \phi^{\alpha\rho}\}]} \quad (\text{B2})$$

The saddle point equations are evaluated using the following ansatz for the order parameter encoding the correlations of the bosonic fields,

$$P_{\alpha\beta\rho\sigma} = \Pi(z) \delta_{\alpha\beta} \delta_{\rho\sigma} \quad M_{\alpha\rho} = 0 \quad T_{\alpha\beta\rho} = 0. \quad (\text{B3})$$

We have thus assumed that the replicated bosonic fields decouple from the replicated soft spins, and that there are no correlations among bosonic fields in different replicas. Using this structure, the saddle point equations for the overlap matrix $Q_{\alpha\beta}$ and the magnetization m_α become

$$Q_{\alpha\beta} = \langle x^\alpha x^\beta \rangle_{\text{MF}}^{\text{SK}} \quad m_\alpha = \langle x^\alpha \rangle_{\text{MF}}^{\text{SK}} \quad (\text{B4})$$

while the saddle point equation for $\Pi(z)$ reads

$$\Pi(z) = \left\langle \frac{1}{z - E_I''(x^\alpha, a) - \Pi(z)} \right\rangle_{\text{MF}}^{\text{SK}}. \quad (\text{B5})$$

The mean field average $\langle \dots \rangle_{\text{MF}}^{\text{SK}}$ is an average over the replicated soft spin variables:

$$\langle \dots \rangle_{\text{MF}}^{\text{SK}} \equiv \frac{1}{Z_{\text{MF}}^{\text{SK}}} \int \prod_{\alpha} d\mathbf{x}^\alpha e^{-\beta E_{\text{MF}}^{\text{SK}}[\{\mathbf{x}^\alpha\}]}, \quad (\text{B6})$$

and the mean field energy is defined as

$$\begin{aligned} E_{\text{MF}}^{\text{SK}} &= \sum_{\alpha} E_I(x^\alpha) - \frac{\beta}{2} \sum_{\alpha,\beta} Q_{\alpha\beta} x^\alpha x^\beta \\ &\quad - J_0 \sum_{\alpha} m_\alpha x^\alpha \\ Z_{\text{MF}}^{\text{SK}} &= \int \prod_{\alpha=1}^n d\mathbf{x}^\alpha e^{-\beta E_{\text{MF}}^{\text{SK}}[\{\mathbf{x}^\alpha\}]}. \end{aligned} \quad (\text{B7})$$

We now have all the elements to compute the resolvent in Eq. (25). Taking the derivative with respect to z of the action we obtain

$$\begin{aligned} G^{\text{SK}}(z) &= - \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \frac{i}{n} \partial_l \sum_{\alpha,\beta,\rho,\sigma} \langle \phi^{\alpha\rho} \phi^{\beta\sigma} \rangle_{x,\phi} \\ &= \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \frac{1}{n} \partial_l \sum_{\alpha,\beta,\rho,\sigma} P_{\alpha\beta\rho\sigma} \\ &= \lim_{n \rightarrow 0} \lim_{l \rightarrow 0} \frac{1}{n} \partial_l n l \Pi(z) \\ &= \lim_{n \rightarrow 0} \Pi(z). \end{aligned} \quad (\text{B8})$$

In the second line, we used the saddle point equation in Eq. (B1), and in the third line we used our ansatz for the order parameter $P_{\alpha\beta\rho\sigma}$ given by Eq. (B3). The resolvent can be thus identified with the quantity $\Pi(z)$. This implies that the resolvent satisfies a self-consistent equation

$$G^{\text{SK}}(z) = \lim_{n \rightarrow 0} \left\langle \frac{1}{z - \partial_x^2 E_I(x^\alpha, a) - G^{\text{SK}}(z)} \right\rangle_{\text{MF}}^{\text{SK}}, \quad (\text{B9})$$

which can be rewritten, using the Dirac delta distribution, as

$$\begin{aligned} G^{\text{SK}}(z) &= \lim_{n \rightarrow 0} \frac{1}{n} \sum_{\alpha=1}^n \int dx \frac{\langle \delta(x - x^\alpha) \rangle_{\text{MF}}^{\text{SK}}}{z - \partial_x^2 E_I(x, a) - G^{\text{SK}}(z)} \\ &= \int dx \frac{P^{\text{SK}}(x)}{z - \partial_x^2 E_I(x, a) - G^{\text{SK}}(z)}. \end{aligned} \quad (\text{B10})$$

The second line defines the single spin probability distribution $P^{\text{SK}}(x)$ as

$$P^{\text{SK}}(x) = \lim_{n \rightarrow 0} \frac{1}{n} \sum_{\alpha=1}^n \langle \delta(x - x^\alpha) \rangle_{\text{MF}}^{\text{SK}}. \quad (\text{B11})$$

Equation (B10) corresponds to Eq. (26) in the main text.

1. Stability of the ansatz of Eq. (B3)

The ansatz chosen in Eq. (B3) is valid as long as the action $S^{\text{SK}}[\mathbf{\Gamma}, z]$ has a maximum at $\mathbf{\Gamma} = \mathbf{\Gamma}^*$, with $\mathbf{\Gamma}^*$ the set of order parameters satisfying the saddle-point equations, Eq. (B3) and Eq. (B5). To check this condition, we expand the action around its saddle-point value:

$$\begin{aligned} S^{\text{SK}}[\mathbf{\Gamma}^* + \delta\mathbf{\Gamma}, z] &\approx S^{\text{SK}}[\mathbf{\Gamma}^*, z] \\ &\quad + \frac{1}{2} \sum_{\nu,\nu'} \frac{\partial^2 S^{\text{SK}}[\mathbf{\Gamma}, z]}{\partial \Gamma_\nu \partial \Gamma_{\nu'}} \bigg|_{\mathbf{\Gamma}=\mathbf{\Gamma}^*} \delta\Gamma_\nu \delta\Gamma_{\nu'}. \end{aligned} \quad (\text{B12})$$

The set of parameters $\mathbf{\Gamma}^*$ maximizes the action as long as the eigenvalues of the Hessian matrix $\frac{\partial^2 S^{\text{SK}}[\mathbf{\Gamma}, z]}{\partial \Gamma_\nu \partial \Gamma_{\nu'}} \bigg|_{\mathbf{\Gamma}=\mathbf{\Gamma}^*}$ are all negative. Since within our ansatz in Eq. (B3) the bosonic fields and the soft spin variables are decoupled, we restrict our study on the subspace spanned the order parameters $P_{\alpha\beta\rho\sigma}, T_{\alpha\beta\rho}, M_{\alpha\rho}$. Within our ansatz, the only non zero entries of the Hessian matrix in Eq. (B12) are

$$\begin{aligned}
\frac{\partial^2 S^{\text{SK}}}{\partial P_{\alpha\beta\rho\sigma} \partial P_{\alpha'\beta'\rho'\sigma'}} &= -\frac{1}{2} [\delta_{\alpha\alpha'} \delta_{\beta\beta'} \delta_{\rho\rho'} \delta_{\sigma\sigma'} + \delta_{\alpha\beta'} \delta_{\beta\alpha'} \delta_{\rho\sigma'} \delta_{\sigma\rho'}] \\
&\quad - \frac{1}{2} [\langle \phi^{\alpha\rho} \phi^{\beta\sigma} \phi^{\alpha'\rho'} \phi^{\beta'\sigma'} \rangle_{\text{MF}}^{\text{SK}} - \langle \phi^{\alpha\rho} \phi^{\beta\sigma} \rangle_{\text{MF}}^{\text{SK}} \langle \phi^{\alpha'\rho'} \phi^{\beta'\sigma'} \rangle_{\text{MF}}^{\text{SK}}] \\
\frac{\partial^2 S^{\text{SK}}}{\partial T_{\alpha\beta\rho} \partial T_{\alpha'\beta'\rho'}} &= i\delta_{\alpha\alpha'} \delta_{\beta\beta'} \delta_{\rho\rho'} - \beta^2 \langle x^\alpha x^{\alpha'} \phi^{\beta\rho} \phi^{\beta'\rho'} \rangle_{\text{MF}}^{\text{SK}} = i\delta_{\alpha\alpha'} \delta_{\beta\beta'} \delta_{\rho\rho'} - \beta^2 Q_{\alpha\alpha'} \langle \phi^{\beta\rho} \phi^{\beta'\rho'} \rangle_{\text{MF}}^{\text{SK}} \\
\frac{\partial^2 S^{\text{SK}}}{\partial M_{\alpha\rho} \partial M_{\alpha'\rho'}} &= i\delta_{\alpha\alpha'} \delta_{\beta\beta'} \delta_{\rho\rho'} - J_0 \langle \phi^{\alpha\rho} \phi^{\alpha'\rho'} \rangle_{\text{MF}}^{\text{SK}}.
\end{aligned} \tag{B13}$$

The real parts of the terms involving partial derivatives with respect to $T_{\alpha\beta\rho}$ or $M_{\alpha\rho}$ are negative, and fluctuations within these subspace are thus irrelevant. We focus therefore on the block of the Hessian involving fluctuations of the tensor $P_{\alpha\beta\rho\sigma}$. The only nonzero entries in this block are

$$\begin{aligned}
\frac{\partial^2 S^{\text{SK}}}{\partial P_{\alpha\beta\rho\sigma} \partial P_{\alpha\beta\rho\sigma}} &= -\frac{1}{2} [1 + \langle (\phi^{\alpha\rho}) \rangle_{\text{MF}}^{\text{SK}}]^2 \\
\frac{\partial^2 S^{\text{SK}}}{\partial P_{\alpha\alpha\rho\rho} \partial P_{\alpha\alpha\rho\rho}} &= -1 \\
+ \int dx \frac{P^{\text{SK}}}{[z - \partial_x^2 E_I(x, a) - G^{\text{SK}}(z)]^2} &= 0,
\end{aligned} \tag{B14}$$

where the $\alpha \neq \beta$ and $\rho \neq \sigma$. The term on the first line of Eq. (B14) has a negative definite real part, while the term on the second line is zero, as it can be verified by taking the derivative of the saddle point equation for the resolvent in Eq. (B9) with respect to G^{SK} . Our ansatz is therefore marginally stable against longitudinal fluctuations in the space of the tensor $P_{\alpha\alpha\rho\rho}$. Motivated by past work in the analysis of the ground states of disordered systems [72], we assume that no instability develops when considering perturbations of higher order.

Appendix C: Replica symmetric ansatz

In this Section we show how to obtain the saddle point equations in the replica symmetric case, given by

$$\begin{aligned}
Q_{\alpha\beta} &= q_d \delta_{\alpha\beta} + q_o (1 - \delta_{\alpha\beta}) \\
m_\alpha &= m.
\end{aligned} \tag{C1}$$

Plugging Eq. (C1) into the mean-field partition function given by Eq. (B7), we obtain

$$\begin{aligned}
Z_{\text{MF}}^{\text{SK}} &= \int \prod_{\alpha=1}^n dx^\alpha \exp \left[-\beta \sum_{\alpha} E_I(x^\alpha, a) \right. \\
&\quad - \frac{\beta^2}{2} \Delta q \sum_{\alpha} (x^\alpha)^2 + \frac{\beta^2}{2} q_o \left(\sum_{\alpha} x^\alpha \right)^2 \\
&\quad \left. + J_0 m \sum_{\alpha} x^\alpha \right]
\end{aligned} \tag{C2}$$

We now use the rescaled overlap gap $\Delta\tilde{q} \equiv \beta(q_d - q_o)$ and perform a Hubbard-Stratonovich transformation, obtaining

$$\begin{aligned}
Z_{\text{MF}}^{\text{SK}} &= \int \prod_{\alpha=1}^n dx^\alpha \exp \left[-\beta \sum_{\alpha} E_I(x^\alpha, a) \right. \\
&\quad \left. - \frac{\beta}{2} \Delta\tilde{q} \sum_{\alpha} (x^\alpha)^2 - \beta(\sqrt{q_o}h - J_0 m)x^\alpha \right] \\
&\equiv \left(\int dx e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)} \right)^n \\
&\equiv (\overline{Z_{\text{RS}}^{\text{SK}}(h)})^n,
\end{aligned} \tag{C3}$$

where $\overline{\dots} \equiv \int \frac{dh}{\sqrt{2\pi}} \dots e^{-h^2/2}$ is an average over the realization of a random Gaussian field h , which acts as a quenched disorder in the system. In the second line, we have observed that the partition function factorizes among different replicas and we defined the replica symmetric energy function

$$E_{\text{eff}}^{\text{SK}}(x, h) \equiv E_I(x, a + \Delta\tilde{q}) - (\sqrt{q_o}h + J_0 m)x. \tag{C4}$$

The last line of Eq. (C3) defines the replica symmetric partition function $Z_{\text{RS}}^{\text{SK}}(h)$ for a given realization of the Gaussian field h . We compute in a similar fashion the single spin distribution in Eq. (B11) within the replica symmetric phase,

$$\begin{aligned}
P^{\text{SK}}(x) &= \lim_{n \rightarrow 0} \frac{1}{n} \sum_{\alpha=1}^n \int \prod_{\beta=1}^n dx^\beta \delta(x - x^\alpha) \\
&\quad \times e^{-\beta E_{\text{RS}}^{\text{SK}}[\{x^\beta\}]} [Z_{\text{MF}}^{\text{SK}}]^{-1} \\
&= \lim_{n \rightarrow 0} e^{-\beta E_{\text{RS}}^{\text{SK}}(x, z)} (\overline{Z_{\text{RS}}^{\text{SK}}(h)})^{n-1} \\
&= e^{-\beta E_{\text{RS}}^{\text{SK}}(x, z)} [Z_{\text{RS}}^{\text{SK}}(h)]^{-1} \\
&\equiv P_{\text{RS}}^{\text{SK}}(x),
\end{aligned} \tag{C5}$$

which is Eq. (38) of the main text.

Appendix D: Stability of the replica symmetric solution

The replica symmetric solution is stable as long as the order parameters that satisfy the saddle point equations in Eq. (36) are an extremum of the action S^{SK} . This requirement breaks down when the Hessian of the action, introduced in Eq. (B12), acquires a positive eigenvalue along some direction related to a perturbation of the overlap matrix $Q_{\alpha\beta}$ and to the magnetization m_α . The Hessian of the action involves the following components,

$$\begin{aligned} \frac{\partial^2 S^{\text{SK}}}{\partial Q_{\alpha\beta} \partial Q_{\gamma\epsilon}} &= -\frac{\beta^2}{2} (\delta_{\alpha\beta} \delta_{\gamma\epsilon} + \delta_{\alpha\gamma} \delta_{\beta\epsilon}) \\ &+ \frac{\beta^4}{2} \left(\overline{\langle x^\alpha x^\beta x^\gamma x^\epsilon \rangle_{\text{RS}}^{\text{SK}}} - \overline{\langle x^\alpha x^\beta \rangle_{\text{RS}}^{\text{SK}}} \overline{\langle x^\gamma x^\epsilon \rangle_{\text{RS}}^{\text{SK}}} \right) \\ \frac{\partial^2 S^{\text{SK}}}{\partial Q_{\alpha\beta} \partial m_\gamma} &= \beta^3 J_0 \left[\overline{\langle x^\alpha x^\beta x^\gamma \rangle_{\text{RS}}^{\text{SK}}} - \overline{\langle x^\alpha x^\beta \rangle_{\text{RS}}^{\text{SK}}} \overline{\langle x^\gamma \rangle_{\text{RS}}^{\text{SK}}} \right] \\ \frac{\partial^2 S^{\text{SK}}}{\partial m_\alpha \partial m_\beta} &= -\beta J_0 \delta_{\alpha\beta} \\ &+ (\beta J_0)^2 \left[\overline{\langle x^\alpha x^\beta \rangle_{\text{RS}}^{\text{SK}}} - \overline{\langle x^\alpha \rangle_{\text{RS}}^{\text{SK}}} \overline{\langle x^\beta \rangle_{\text{RS}}^{\text{SK}}} \right]. \end{aligned} \quad (\text{D1})$$

The first eigenvalue that changes sign at the instability transition is known as the replicon, $\Lambda_{\text{R}}^{\text{SK}}$. Its expression depends on the generic structure of the overlap matrix $Q_{\alpha\beta}$, rather than the specific model at hand, and it is given by [73, 74]

$$\begin{aligned} \Lambda_{\text{R}}^{\text{SK}} &= 2 \frac{\partial^2 S^{\text{SK}}}{\partial Q_{\alpha\beta} \partial Q_{\alpha\beta}} - 4 \frac{\partial^2 S^{\text{SK}}}{\partial Q_{\alpha\beta} \partial Q_{\beta\gamma}} + 2 \frac{\partial^2 S^{\text{SK}}}{\partial Q_{\alpha\beta} \partial Q_{\gamma\epsilon}} \\ &= \beta^2 \left[-1 + \beta^2 \left(\overline{\langle x^2 \rangle_{\text{MF}}^{\text{SK}}} - \overline{\langle x \rangle_{\text{MF}}^{\text{SK}}}^2 \right) \right]. \end{aligned} \quad (\text{D2})$$

The replica-symmetric saddle point is stable as long as $\Lambda_{\text{R}}^{\text{SK}} < 0$. In the low temperature limit $\beta \rightarrow \infty$, it is convenient to consider a rescaled version of the replicon eigenvalue $\tilde{\Lambda}_{\text{R}}^{\text{SK}} \equiv \beta^{-2} \Lambda_{\text{R}}^{\text{SK}}$. A saddle point evaluation of the replica-symmetric average in Eq. (D2) yields the following stability condition for the replica symmetric phase,

$$\tilde{\Lambda}_{\text{R}}^{\text{SK}} = -1 + [\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(h))]^{-2} < 0, \quad (\text{D3})$$

where $x^*(h)$ is the minimum of the replica symmetric mean field energy $E_{\text{eff}}^{\text{SK}}$ in Eq. (C4) for a fixed realization of the quenched field h . We thus recover Eq. (42) of the main text.

Appendix E: Derivation of Eq. (44)

In this Appendix we compute the stability boundary of the convex paramagnetic phase for the CIM-SK. Starting from the replica-symmetric saddle point equations

Eq. (36), we consider the case where $m = q_o = 0$. The only relevant order parameter is $\Delta\tilde{q}$, which can be computed from the saddle point equations as

$$\Delta\tilde{q} = -\frac{1}{a + \Delta\tilde{q}}. \quad (\text{E1})$$

The stability of this paramagnetic phase is determined by the condition $\tilde{\Lambda}_{\text{R}}^{\text{SK}} < 0$ in Eq. (42), which for $m = 0, q_o = 0$ yields

$$\Delta\tilde{q}^2 < 1 \implies \Delta\tilde{q} < 1. \quad (\text{E2})$$

We also need to impose the paramagnetic state to be stable against perturbation of the magnetization, i.e. that $\frac{\partial^2 S^{\text{SK}}}{\partial m_\alpha \partial m_\beta} < 0$. From Eq. (D1) we obtain

$$\Delta\tilde{q} < J_0^{-1}. \quad (\text{E3})$$

Combining Eq. (E2) and Eq. (E3) we obtain

$$\Delta\tilde{q} < \min[1, J_0^{-1}]. \quad (\text{E4})$$

This is Eq. (44) of the main text.

Appendix F: Outline of the derivation of the full replica symmetry breaking equations

In this Appendix we outline how to derive the full replica symmetry breaking equation for the CIM-SK, presented in Sec. VIB. When the replica-symmetric solution becomes unstable, a new ansatz for the overlap matrix $\mathbf{Q}$ is required. Different replicas are no longer equivalent, and their overlaps organize in a hierarchical structure [17, 18, 37, 48]. All the rows of $\mathbf{Q}$ are the same up to a permutation of their elements. The description of the first row of $\mathbf{Q}$ is thus enough to characterize the overlap matrix. The first row is divided in $k+2$ bands of different lengths. We label each band by an index $i = 0, \dots, k+1$. The band with index $k+1$ corresponds to the first diagonal entry of the matrix $\mathbf{Q}$. For $i \leq k+1$, the i th band contains $m_i - m_{i+1}$ elements, all equal to a given overlap value q_i . The set $\{m_i\}_{i=0}^{k+1}$ is ordered in the following way

$$1 = m_{k+1} < m_k < \dots < m_1 < m_0 = n. \quad (\text{F1})$$

The overlap matrix becomes

$$\begin{aligned} Q_{\alpha\beta} &= q_d \delta_{\alpha\beta} \\ &+ (1 - \delta_{\alpha\beta}) \left[q_0 + \sum_{i=1}^k (q_i - q_{i-1}) I_{\alpha\beta}^{m_i} \right], \end{aligned} \quad (\text{F2})$$

where $I_{\alpha\beta}^{m_i}$ is an $n \times n$ covered along its diagonal by n/m_i block matrices of size $m_i \times m_i$, with all the entries equal to 1. When the limit $n \rightarrow 0$ is taken by means of an analytic continuation, the order of the size of the bands is inverted, so that

$$0 = m_0 < m_1 < \dots < m_k < m_{k+1} = 1. \quad (\text{F3})$$

Using this structure, the mean field partition function $Z_{\text{MF}}^{\text{SK}}$ becomes [18]

$$\lim_{n \rightarrow 0} \frac{1}{n} \log Z_{\text{MF}}^{\text{SK}} = \beta \int \mathcal{D}h_0 f_1(h_0 \sqrt{q_0}), \quad (\text{F4})$$

where $\mathcal{D}h_0 \equiv \frac{dh_0}{\sqrt{2\pi}} e^{-\frac{h_0^2}{2}}$ is the Gaussian measure. The function $f_1(h)$ is the first member of a hierarchy of $k+1$ 'free energies', which satisfy a recursion relation,

$$f_i(h) = \frac{1}{\beta m_i} \log \int \mathcal{D}h_i e^{\beta m_i f_{i+1}(h + \sqrt{q_i - q_{i-1}} z_i)}, \quad (\text{F5})$$

together with the boundary condition

$$f_{k+1}(h) = \frac{1}{\beta} \log \int dx e^{-\beta E_{\text{RS}}^{\text{SK}}(x, h)}, \quad (\text{F6})$$

from which we recognize f_{k+1} as the free energy of a single soft spin system with energy $E_{\text{eff}}^{\text{SK}}$. The field h acts as a form of quenched disorder, but its distribution is no longer Gaussian, because of the hierarchical structure of the states of the system.

The saddle point equations for the overlaps q_i , q_d and the magnetization m can be computed as derivatives of the logarithm of the partition function in Eq. (F4). We obtain

$$\begin{aligned} q_i &= \int dh P_i(h) [\partial_h f_i(h)]^2 \\ \Delta \tilde{q} &= \int dh P_k(h) [\partial_h^2 f_{k+1}(h)] \\ m &= \int dh P_k(h) \partial_h f_{k+1}(h). \end{aligned} \quad (\text{F7})$$

The function $P_i(h)$ is the distribution of internal fields h acting at the level i of the hierarchy. It is defined as the integrated response of the function f_1 to a change of the free energy at the level $i+1$ in the hierarchy,

$$P_i(h) \equiv \int \mathcal{D}h_0 \frac{\delta f_1(h_0 \sqrt{q_0})}{\delta f_{i+1}(h)}, \quad (\text{F8})$$

where $\frac{\delta f_1(h_0)}{\delta f_{i+1}(h)}$ is a functional derivative. The distribution of internal fields satisfies a recursive relation too, namely,

$$\begin{aligned} P_i(h) &= \int \mathcal{D}h_i P_{i-1}(h - \sqrt{q_i - q_{i-1}} h_i) \\ &\times e^{\beta m_i (f_{i+1}(h) - f_i(h - \sqrt{q_i - q_{i-1}} h_i))}, \end{aligned} \quad (\text{F9})$$

with the boundary condition

$$P_0(h) = \frac{1}{\sqrt{2\pi q_0}} e^{-\frac{h^2}{2q_0}}. \quad (\text{F10})$$

The single-spin distribution P^{SK} becomes instead

$$P^{\text{SK}}(x) = \int dh P_k(h) \frac{e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)}}{Z_{\text{eff}}^{\text{SK}}(h)}. \quad (\text{F11})$$

The full replica symmetry breaking is obtained when the number of levels k in the hierarchy is sent to infinity. The bands in the hierarchical matrix $\mathbf{Q}$ form then a continuum, and the overlaps are described by a continuous function $q(y)$, with $y \in [0, 1]$, which for finite k is defined as

$$q(y) = q_i \quad m_i \leq y < m_{i+1}. \quad (\text{F12})$$

The same continuum limit is taken for the hierarchical free energies $f_i(h)$ and for the internal field distribution $P_i(h)$, which become continuous functions $f(y, h)$ and $P(y, h)$ respectively. The recursive relations given by Eq. (F5) and Eq. (F9) become a backward and a forward Fokker-Planck equations, respectively,

$$\begin{aligned} \dot{f}(y, h) &= -\frac{\dot{q}(y)}{2} [f'' + \beta y (f')^2] \\ \dot{P}(y, h) &= \frac{\dot{q}(y)}{2} [P'' - 2\beta y (P f')'], \end{aligned} \quad (\text{F13})$$

where we used a shorthand notation for the derivatives, $\dot{f}(y, h) = \partial_y f(y, h)$ and $f'(y, h) = \partial_h f(y, h)$. The boundary conditions of the partial differential equations in Eq. (F13) are

$$\begin{aligned} f(y=1, h) &= \frac{1}{\beta} \log \int dx e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)} \\ P(y=0, h) &= \frac{1}{\sqrt{2\pi q(0)}} e^{-\frac{h^2}{2q(0)}}. \end{aligned} \quad (\text{F14})$$

The saddle point equations in the continuum limit are

$$\begin{aligned} q(y) &= \int dh P(y, h) [f'(y, h)]^2 \\ \Delta \tilde{q} &= \int dh P(1, h) f''(1, h) \\ m &= \int dh P(1, h) f'(1, h), \end{aligned} \quad (\text{F15})$$

and the single spin distribution is

$$P_{\text{RSB}}^{\text{SK}}(x) = \int dh P(1, h) \frac{e^{-\beta E_{\text{eff}}^{\text{SK}}(x, h)}}{Z_{\text{eff}}^{\text{SK}}(h)}. \quad (\text{F16})$$

1. Marginal stability

The replicon in the full replica symmetry breaking phase reads, from its expression in Eq. (D2),

$$\tilde{\Lambda}_{\text{R}}^{\text{SK}} = -1 + \int_{-\infty}^{\infty} dh P(1, h) [f''(1, h)]^2. \quad (\text{F17})$$

We now show that $\tilde{\Lambda}_{\text{R}}^{\text{SK}} = 0$ in the spin-glass phase. Taking the derivative with respect to y in the saddle point

equation for $q(y)$ Eq. (F15), yields

$$\begin{aligned} \dot{q}(y) &= \int dh \dot{P}(y, h) [f'(y, h)]^2 \\ &+ 2 \int dh P(y, h) \dot{f}'(y, h) f'(y, h). \end{aligned} \quad (\text{F18})$$

Substituting Eq. (F13) into the expression above and integrating by parts several times we obtain

$$\dot{q}(y) = \dot{q}(y) \int dh P(y, h) [f''(y, h)]^2, \quad (\text{F19})$$

which implies, for $y = 1$,

$$-1 + \int dh P(1, h) [f''(1, h)]^2 = 0. \quad (\text{F20})$$

From the expression of the replicon in Eq. (F17) we thus conclude that

$$\tilde{\Lambda}_R^{\text{SK}} = 0 \quad (\text{F21})$$

in the spin-glass phase.

2. Low temperature solution

In the low temperature limit $\beta \rightarrow \infty$ the boundary condition for the function $f(y, h)$ is evaluated through a saddle point method, yielding

$$f(y = 1, h) = -E_{\text{eff}}^{\text{SK}}(x^*(h), h), \quad (\text{F22})$$

where $x^*(h)$ is the minimizer of the energy $E_{\text{eff}}^{\text{SK}}(x, h)$ for a given realization of the internal field h . This minimizer satisfies the equation

$$x^*(h)^3 - (a + \Delta\tilde{q})x^*(h) - J_0m = h. \quad (\text{F23})$$

If $a + \Delta\tilde{q} \leq 0$, then Eq. (F23) has a unique solution. The higher order derivatives of $f(y = 1, h)$ can be obtained

from Eq. (F22) as

$$\begin{aligned} f'(y = 1, h) &= x^*(h) \\ f''(y = 1, h) &= \partial_h x^*(h) = \frac{1}{3x^*(h)^2 - a - \Delta\tilde{q}}. \end{aligned} \quad (\text{F24})$$

The saddle point equations Eq. (F15) then become

$$\begin{aligned} q(y) &= \int dh P(1, h) [x^*(h)]^2 \\ \Delta\tilde{q} &= \int dh \frac{P(1, h)}{3x^*(h)^2 - a - \Delta\tilde{q}} \\ m &= \int dh P(1, h) x^*(h), \end{aligned} \quad (\text{F25})$$

while the marginal stability condition is

$$\tilde{\Lambda}_R^{\text{SK}} = -1 + \int dh \frac{P(1, h)}{(3x^*(h)^2 - a - \Delta\tilde{q})^2} = 0. \quad (\text{F26})$$

Finally, a saddle-point evaluation of the single spin probability distribution in Eq. (F16) yields

$$\begin{aligned} P_{\text{RSB}}^{\text{SK}}(x) &\propto |3x^2 - a - \Delta\tilde{q}| \\ &\times P(1, h = x^3 - (a + \Delta\tilde{q})x + J_0m). \end{aligned} \quad (\text{F27})$$

For $a + \Delta\tilde{q} > 0$, the effective energy $E_{\text{eff}}^{\text{SK}}$ is nonconvex. Equation (F22) is still valid, provided that $x^*(h)$ is the global minimum of $E_{\text{eff}}^{\text{SK}}(x, h)$. However, the global minimum $x^*(h)$ exhibits a discontinuous jump at $h = -J_0m$. The integral in marginality condition given by Eq. (F21) must be carefully handled around the discontinuity as the zero-temperature limit is taken. As discussed in [65] for a related model of mean-field disordered systems, let us consider two distinct regimes, depending on the strength of the quenched disorder h with respect to the temperature T . A first regime is obtained when $h + J_0m \gg T$, far away from the discontinuity, while a second one is given by $h + J_0m \sim T$. The first derivative of the free energy in Eq. (F22) is split into the two regimes as

$$f'(y = 1, h) = \begin{cases} x^*(h) & \text{if } |h + J_0m| \gg T \\ x^*(h) \tanh(\beta(h + J_0m)x^*(h)) & \text{if } |h + J_0m| \sim T, \end{cases} \quad (\text{F28})$$

The integral in the marginal stability condition in Eq. (F21) is then evaluated in the limit $T \rightarrow 0$ by separately

considering the two regimes. The computation reads

$$\begin{aligned}
& \int_{-\infty}^{\infty} dh P(1, h) [f''(1, h)]^2 = \int_{-\infty}^{\infty} dh P(1, h - J_0 m) [f''(1, h - J_0 m)]^2 \\
& = \int_{h \in (-\infty, -T) \cup (T, \infty)} dh P(1, h - J_0 m) [f''(1, h - J_0 m)]^2 + \int_{-T}^T dh P(1, h - J_0 m) [f''(1, h - J_0 m)]^2 \\
& = \int_{-\infty}^{\infty} dh \frac{P(1, h)}{\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(h))^2} + \int_{-T}^T dh \beta^2 P(1, h - J_0 m) [x^*(h - J_0 m)]^4 (1 - \tanh(\beta h))^2 \\
& + \int_{-T}^T dh P(1, h - J_0 m) \frac{[x^*(h - J_0 m)]^2 (1 - \tanh(\beta h))^2 (\beta h)^2}{\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(h - J_0 m))} \\
& = \int_{-\infty}^{\infty} dh \frac{P(1, h)}{\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(h))^2} + \int_{-1}^1 d\tilde{h} \beta P(1, T\tilde{h} - J_0 m) [x^*(T\tilde{h} - J_0 m)]^4 (1 - \tanh(\tilde{h}))^2 \\
& + \int_{-1}^1 d\tilde{h} T P(1, T\tilde{h}) \frac{[x^*(T\tilde{h} - J_0 m)]^2 (1 - \tanh(\tilde{h}))^2 \tilde{h}^2}{\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(T\tilde{h} - J_0 m))} \\
& = \int_{-\infty}^{\infty} dh \frac{P(1, h)}{\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(h))^2} + \int_{-1}^1 d\tilde{h} \tilde{P}(1, \tilde{h}) [x^*(J_0 m)]^4 (1 - \tanh(\tilde{h}))^2 \\
& + T^2 \int_{-1}^1 d\tilde{h} \tilde{P}(1, \tilde{h}) \frac{[x^*(J_0 m)]^2 (1 - \tanh(\tilde{h}))^2 \tilde{h}^2}{\partial^2 E_{\text{eff}}^{\text{SK}}(x^*(J_0 m))} \\
& = \int_{-\infty}^{\infty} dh \frac{P(1, h)}{(3x^*(h)^2 - a - \Delta\tilde{q})^2} + \int_{-1}^1 d\tilde{h} \tilde{P}(1, \tilde{h}) (a + \Delta\tilde{q})^4 (1 - \tanh(\tilde{h}))^2
\end{aligned} \tag{F29}$$

In the first equality, we shifted the integration variable by $J_0 m$. In the second equality, we split the integration domain into two different regions, isolating a region around the origin of width $\sim T$. In the third equality, we explicitly computed the quantity $f''(y = 1, h)$. In the fourth equality, a rescaled field was introduced $T\tilde{h} \equiv h$. In the fifth equality, we introduced a rescaled and shifted distribution of fields $T\tilde{P}(1, \tilde{h}) = P(1, T\tilde{h} + J_0 m)$. Finally, in the last equality, we substituted for the explicit expression of $x^*(h)$ and its derivatives. The marginal stability condition in Eq. (F21) thus reads

$$\tilde{\Lambda}_{\text{R}}^{\text{SK}} = -1 + \int_{-\infty}^{\infty} dh \frac{P(1, h)}{(3x^*(h)^2 - a - \Delta\tilde{q})^2} + \int_{-1}^1 d\tilde{h} \tilde{P}(1, \tilde{h}) (a + \Delta\tilde{q})^4 (1 - \tanh(\tilde{h}))^2 = 0. \tag{F30}$$

Notice that the third contribution to the replicon is positive. In order for the marginal stability condition of the full-replica symmetry breaking solution with respect to perturbations in the space of overlap matrices to be satisfied, the internal field distribution must decay to zero around $J_0 m$ at least as fast as linearly, i.e. $P(1, h - J_0 m) \sim |h|$ for $h \ll 1$. Moreover, when $a + \Delta\tilde{q} > 0$ the marginality condition implies

$$-1 + \int dh \frac{P(1, h)}{(3x^*(h)^2 - a - \Delta\tilde{q})^2} < 0. \tag{F31}$$

This inequality can equivalently be expressed through a change of variable from the internal field distribution to the soft spin distribution as

$$-1 + \int dx \frac{P_{\text{RSB}}^{\text{SK}}(x)}{(3x^2 - a - \Delta\tilde{q})^2} < 0, \tag{F32}$$

which is Eq. (58) in the main text.

3. Numerical solution of full replica symmetry breaking equations

To solve the full replica-symmetry breaking equations at zero temperature, we adopt a method similar to the one developed in [49]. We discretize the continuous solution into k steps. Since we are interested in the low temperature limit of the full replica-symmetry breaking equations, we adopt a rescaled size of the blocks $\tilde{m}_i \equiv \beta m_i$, for $i = 0, \dots, k+1$. The rescaled variables $\tilde{m}_i$ are ordered in the following way, which can be read from Eq. (F3),

$$0 < \tilde{m}_0 < \tilde{m}_1 < \dots < \tilde{m}_k < \tilde{m}_{k+1} = \beta \rightarrow \infty. \tag{F33}$$

Starting from an initial guess of the discretized Parisi function q_i , we iteratively solve the recursion relation for the functions $\tilde{f}_i(h) \equiv \tilde{m}_i f(h)$,

$$\tilde{f}_i(h) = \log \int \mathcal{D}w_i e^{\frac{\tilde{m}_i}{\tilde{m}_{i+1}} \tilde{f}_{i+1}(h+w_i)}, \tag{F34}$$

where $\mathcal{D}w_i \equiv \frac{dw_i}{\sqrt{2\pi(q_i - q_{i-1})}} e^{-\frac{w_i^2}{2(q_i - q_{i-1})}}$ is a Gaussian measure of variance $q_i - q_{i-1}$. The boundary condition of the recursion relation is

$$\tilde{f}_{k+1}(h) = -\tilde{m}_{k+1} \min_x E_{\text{eff}}^{\text{SK}}(x, h). \quad (\text{F35})$$

Note that the use of the rescaled variable $\tilde{f}_i$ requires the computation of the ratio $\tilde{m}_i/\tilde{m}_{i+1}$, which is typically more well-behaved than the quantities $\tilde{m}_i$. We also compute the derivatives $\tilde{f}'_i(h)$, using a recursion relation obtained from Eq. (F34),

$$\begin{aligned} \tilde{f}'_i(h) &= \frac{\tilde{m}_i}{\tilde{m}_{i+1}} \int \mathcal{D}w_i \tilde{f}'_i(h + w_i) \\ &\times e^{\frac{\tilde{m}_i}{\tilde{m}_{i+1}} \tilde{f}_{i+1}(h + w_i) - \tilde{f}_i(h)}, \end{aligned} \quad (\text{F36})$$

together with the boundary condition

$$\tilde{f}'_{k+1}(h) = \tilde{m}_{k+1} x^*(h), \quad (\text{F37})$$

with $x^*(h) \equiv \arg \min_x E_{\text{eff}}^{\text{SK}}(x, h)$. We then compute the probability distributions of the internal fields $P_i(h)$ using the recursion relations

$$\begin{aligned} P_i(h) &= \int \mathcal{D}w_i P_{i-1}(h - w_i) \\ &\times e^{\frac{\tilde{m}_i}{\tilde{m}_{i+1}} \tilde{f}_{i+1}(h - w_i) - \tilde{f}_i(h - w_i)}, \end{aligned} \quad (\text{F38})$$

with the boundary condition given by Eq. (F10). If $q_0 < 0.01$, we replace the function P_0 by a Dirac delta centered at the origin. Once the distribution of internal fields are known, we compute the order parameters from the low temperature limit of the saddle point equations Eq. (F7), which read

$$\begin{aligned} q_i &= \frac{1}{\tilde{m}_i^2} \int dh P_i(h) [f'_i(h)]^2 \\ \Delta\tilde{q} &= \int dh \frac{P_k(h)}{3x^*(h)^2 - a - \Delta\tilde{q}} \\ m &= \int dh P_k(h) x^*(h). \end{aligned} \quad (\text{F39})$$

The position of the bands $\tilde{m}_i$ is updated too, using the formula [49]

$$\begin{aligned} \tilde{m}_i &= \frac{-B_i - \sqrt{B_i^2 - 4A_i C_i}}{2A_i} \\ A_i &= \frac{1}{4}(q_i^2 - q_{i-1}^2) \\ B_i &= -\frac{1}{\tilde{m}_{i+1}(1 - \delta_{ik}) + \delta_{ik}} \int dh P_{i-1}(h) \\ &\times \int \mathcal{D}w_i \tilde{f}_i(h - w_i) e^{\frac{\tilde{m}_i}{\tilde{m}_{i+1}} \tilde{f}_{i+1}(h - w_i) - \tilde{f}_i(h - w_i)} \\ C_i &= \int dh P_{i-1}(h) \tilde{f}_i(h), \end{aligned} \quad (\text{F40})$$

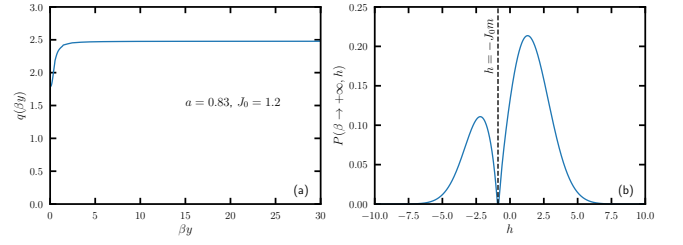


FIG. 13. **Overlap function and distribution of internal fields in the spin-glass phase.** Panel (a): overlap function $q(\beta y)$ as a function of the rescaled quantity βy . Due to the presence of a spontaneous magnetization, the value of q at the origin $q(0)$ is different from 0. As $\beta y \rightarrow \infty$, the overlap function tends to a constant. Panel (b): distribution of internal fields $P(\beta y \rightarrow \infty, h)$. The distribution approaches 0 linearly in $|h - J_0 m|$ for $h \approx -J_0 m$.

with the boundaries $\tilde{m}_0 = 0$ and $\tilde{m}_{k+1} \gg 1$ held fixed. The steps described can be iterated until the desired level of convergence is achieved. To build Fig. 2 (a), we iterated the numerical scheme 20 times with $k = 25$, which is enough to obtain convergence of the quantities m and $a + \Delta\tilde{q}$ with a degree of accuracy of $\sim 1\%$. At large values of the laser gain, the initial condition for the order parameters $q_i, \Delta\tilde{q}, m$ was taken from results obtained at slightly larger laser gains. To obtain the curves for $P^{\text{SK}}(x)$, $\rho^{\text{SK}}(\lambda)$ and $\lambda_m(a, J_0)$ in Fig. 2(b-e) and Fig. 3 the number of iterations was increased to 50, and the number of steps to $k = 35$. Moreover, instead of updating the quantity $\Delta\tilde{q}$ during each iteration, we instead fixed a value of the effective laser gain $a_{\text{eff}} \equiv a + \Delta\tilde{q}$. After convergence, we computed $\Delta\tilde{q}$ and, consequently, a . We show in Fig. 13 a representative result of the numerical iteration.

Appendix G: Replica calculation of the grand-potential

This Appendix is devoted to the computation of the grand-potential in Eq. (63), namely,

$$\Omega^{\text{SK}}(\mu) = -\lim_{N \rightarrow \infty} \frac{1}{N} \langle \log Z_{\Omega}^{\text{SK}} \rangle_{\mathbf{J}}, \quad (\text{G1})$$

where the grandcanonical partition function Z_{Ω}^{SK} is given by

$$\begin{aligned} Z_{\Omega}^{\text{SK}} &\equiv \int d\mathbf{x} \left(\prod_{i=1}^N \delta(\partial_i E^{\text{SK}}(\mathbf{x}, a)) \right) \\ &\times |\det \mathbf{H}^{\text{SK}}(\mathbf{x}, a)| e^{\mu \mathcal{I}(\mathbf{x})} \\ &= \int_{\mathbf{u}, \mathbf{x}} e^{-\mathbf{u} \cdot \nabla E^{\text{SK}}(\mathbf{x}, a) + \mu \mathcal{I}(\mathbf{x})} |\det \mathbf{H}^{\text{SK}}(\mathbf{x}, a)|. \end{aligned} \quad (\text{G2})$$

In the second line, we used an integral representation of the Dirac delta distribution. The integral $\int_{\mathbf{u}, \mathbf{x}} \equiv \int_{-i\infty}^{+i\infty} \frac{d\mathbf{u}}{(2\pi i)^N} \int d\mathbf{x}$ is over a $2N$ dimensional space of soft

spin variables $\mathbf{x}$ and auxiliary fields $\mathbf{u}$. The average of the disorder is computed, as in Sec. V through the replica

$$\beta\Omega^{\text{SK}}(\mu) = -\lim_{n \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{Nn} (\langle (Z_\Omega^{\text{SK}})^n \rangle_{\mathbf{J}} - 1). \quad (\text{G3})$$

The disorder-averaged, replicated grandcanonical partition function $\langle Z_\Omega^{\text{SK}} \rangle_{\mathbf{J}}$ being

$$\begin{aligned} \langle (Z_\Omega^{\text{SK}})^n \rangle_{\mathbf{J}} &= \int_{\mathbf{x}^\alpha, \mathbf{u}^\alpha} \left\langle \exp \left[\sum_{\alpha=1}^n \left(\sum_{i=1}^N u_i^\alpha \left(-\partial_x E_I(x_i^\alpha, a) + \frac{J_0}{N} \sum_{j=1}^N x_j^\alpha + \frac{1}{\sqrt{N}} \sum_{j=1}^N J_{ij} x_j^\alpha \right) \right) + \mu \sum_{\alpha} \mathcal{I}(\mathbf{x}^\alpha) \right] \right. \\ &\quad \left. \times \prod_{\alpha} |\det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)| \right\rangle_{\mathbf{J}}. \end{aligned} \quad (\text{G4})$$

The integral $\int_{\mathbf{x}^\alpha, \mathbf{u}^\alpha} \equiv \int \prod_{\alpha=1}^n \frac{d\mathbf{x}^\alpha d\mathbf{u}^\alpha}{(2\pi i)^N}$ is over the space of n replicated $2N$ dimensional systems of soft spins $\mathbf{x}^\alpha$ and auxiliary fields $\mathbf{u}^\alpha$. The absolute value of the determinant of $\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)$ is correlated with the auxiliary fields $\mathbf{u}^\alpha$ and with the index $\mathcal{I}(\mathbf{x}^\alpha)$ of different replicas. A full treatment of the grand-potential requires taking these correlations into account, for instance by means of conditional averages, as done in [50] for spiked tensor models. In our case, however, the analysis is hindered by the single-site term $\partial_x E_I(x, a)$. To proceed further, we make an approximation by decoupling the gradient of the CIM-SK energy and the determinant of the Hessian among different realizations of the disorder, that is

$$\begin{aligned} \langle Z_\Omega^{\text{SK}} \rangle_{\mathbf{J}} &\approx \int_{\mathbf{x}^\alpha, \mathbf{u}^\alpha} \left\langle \exp \left[\sum_{\alpha=1}^n \left(\sum_{i=1}^N u_i^\alpha \left(-\partial_x E_I(x_i^\alpha, a) + \frac{J_0}{N} \sum_{j=1}^N x_j^\alpha + \frac{1}{\sqrt{N}} \sum_{j=1}^N J_{ij} x_j^\alpha \right) \right) \right] \right\rangle_{\mathbf{J}} \\ &\quad \times \left\langle \prod_{\alpha} e^{\mu \mathcal{I}(\mathbf{x}^\alpha)} |\det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)| \right\rangle_{\mathbf{J}}. \end{aligned} \quad (\text{G5})$$

Moreover, we neglect the correlations between different replicas when computing the index and the absolute value of the determinant of the Hessian, i.e.,

$$\begin{aligned} &\left\langle \prod_{\alpha} e^{\mu \mathcal{I}(\mathbf{x}^\alpha)} |\det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)| \right\rangle_{\mathbf{J}} \\ &\approx \prod_{\alpha} \left\langle e^{\mu \mathcal{I}(\mathbf{x}^\alpha)} \right\rangle_{\mathbf{J}} \langle |\det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)| \rangle_{\mathbf{J}}. \end{aligned} \quad (\text{G6})$$

We expect these approximation to allow us to capture the properties of most typical critical points, when no conditioning on their magnetization or energy is taken into account.

1. First contribution

We consider the first average over the disorder in Eq. (G5), which can be computed in a similar way as done in App. A. We obtain

$$\begin{aligned} &\left\langle \exp \left[\sum_{\alpha=1}^n \left(\sum_{i=1}^N u_i^\alpha \left(-\partial_x E_I(x_i^\alpha, a) + \frac{J_0}{N} \sum_{j=1}^N x_j^\alpha + \frac{1}{\sqrt{N}} \sum_{j=1}^N J_{ij} x_j^\alpha \right) \right) \right] \right\rangle_{\mathbf{J}} \\ &= \exp \left[-\sum_{\alpha} \sum_i u_i^\alpha \partial_x E_I(x_i^\alpha, a) + J_0 \sum_{\alpha} \sum_i u_i^\alpha \left(\frac{1}{N} \sum_{j=1}^N x_j^\alpha \right) + \frac{1}{4N} \sum_{i,j} \left(\sum_{\alpha} u_i^\alpha x_j^\alpha + u_j^\beta x_i^\beta \right)^2 \right] \\ &= \exp \left[-\sum_{\alpha} \sum_i u_i^\alpha \partial_x E_I(x_i^\alpha, a) + N J_0 \sum_{\alpha} \left(\frac{\mathbf{u}^\alpha \cdot \mathbf{t}}{N} \right) \left(\frac{\mathbf{x}^\alpha \cdot \mathbf{t}}{N} \right) + \frac{N}{2} \sum_{\alpha, \beta} \left[\left(\frac{\mathbf{u}^\alpha \cdot \mathbf{u}^\beta}{N} \right) \left(\frac{\mathbf{x}^\alpha \cdot \mathbf{x}^\beta}{N} \right) + \left(\frac{\mathbf{u}^\alpha \cdot \mathbf{x}^\beta}{N} \right)^2 \right] \right], \end{aligned} \quad (\text{G7})$$

where we recall that $\{\mathbf{x}^\alpha, \mathbf{u}^\alpha\}_{\alpha=1}^n$ is a set of N -dimensional vectors representing the soft spin and auxiliary field configurations of each replica. We then use the Hubbard-Stratonovich transform introduced in Eq. (A4) to linearize the quadratic term $\left(\frac{\mathbf{u}^\alpha \cdot \mathbf{x}^\beta}{N}\right)^2$, and we insert the following set of Dirac deltas, together with their integral representations, for each distinct pairs of replica indexes α, β ,

$$\begin{aligned}\delta(NQ_{\alpha\beta} - \mathbf{x}^\alpha \cdot \mathbf{x}^\beta) &= \int_{-\infty}^{\infty} \frac{d\lambda_{\alpha\beta}}{2\pi i} e^{-N\lambda_{\alpha\beta}Q_{\alpha\beta} + \lambda_{\alpha\beta}(\mathbf{x}^\alpha \cdot \mathbf{x}^\beta)} \\ \delta(Nm_\alpha - \mathbf{x}^\alpha \cdot \mathbf{t}) &= \int_{-\infty}^{\infty} \frac{dv_\alpha}{2\pi i} e^{-Nv_\alpha m_\alpha + v_\alpha(\mathbf{x}^\alpha \cdot \mathbf{t})}.\end{aligned}\quad (\text{G8})$$

Equation (G7) thus becomes

$$\begin{aligned}& \int \prod_{\alpha=1}^n \frac{dv_\alpha dm_\alpha}{2\pi i} \prod_{\alpha \leq \beta} \frac{dQ_{\alpha\beta} d\lambda_{\alpha\beta}}{2\pi i} \prod_{\alpha, \beta} dw_{\alpha\beta} \exp \left[-N \left(\mathbf{m} \cdot \mathbf{v} + \text{Tr} \left[\lambda \mathbf{Q} + \frac{1}{2} \mathbf{w} \mathbf{w}^T \right] \right) \right] \\ & \times \exp \left[- \sum_{\alpha} \sum_i u_i^\alpha \partial_x E_I(x_i^\alpha, a) + \sum_{\alpha, \beta} \lambda_{\alpha\beta} (\mathbf{x}^\alpha \cdot \mathbf{x}^\beta) + J_0 \sum_{\alpha} (\mathbf{u}^\alpha \cdot \mathbf{t}) m_\alpha + \sum_{\alpha} v_\alpha (\mathbf{x}^\alpha \cdot \mathbf{t}) + \frac{1}{2} \sum_{\alpha, \beta} Q_{\alpha\beta} (\mathbf{u}^\alpha \cdot \mathbf{u}^\beta) \right. \\ & \left. - \sum_{\alpha, \beta} w_{\alpha\beta} (\mathbf{u}^\alpha \cdot \mathbf{x}^\beta) \right],\end{aligned}\quad (\text{G9})$$

where the $n \times n$ matrices $\lambda, \mathbf{Q}, \mathbf{w}$ and the n -dimensional vectors $\mathbf{m}, \mathbf{v}$ have been introduced.

2. Second contribution

The terms in the second average in Eq. (G5) can be computed from the knowledge of $\langle \det[z - \mathbf{H}^{\text{SK}}(\mathbf{x}, a)] \rangle_{\mathbf{J}}$, with z a complex number close to the real axis. In fact, the index $\mathcal{I}(\mathbf{x})$ can be obtained from the formula [75]

$$\begin{aligned}\mathcal{I}(\mathbf{x}) &= \frac{1}{2\pi i} \lim_{\epsilon \rightarrow 0} [\log \det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) - i\epsilon] \\ & - \log \det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) + i\epsilon]],\end{aligned}\quad (\text{G10})$$

which can be proved using Cauchy's residues theorem. The branch cut of the complex logarithm function is chosen so that the quantities $\log \lambda_i \pm i\epsilon$ are on the opposite side of the branch cut for any eigenvalue λ_i of $\mathbf{H}^{\text{SK}}(\mathbf{x}, a)$. For replica α , we thus obtain

$$e^{\mu \mathcal{I}(\mathbf{x}^\alpha)} = \lim_{\epsilon \rightarrow 0} \left(\frac{\det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) - i\epsilon]}{\det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) + i\epsilon]} \right)^{\mu/2\pi i}. \quad (\text{G11})$$

The absolute value of the determinant can be instead expressed as

$$\begin{aligned}|\det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)| &= \det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)] e^{\xi \pi i \mathcal{I}(\mathbf{x}^\alpha)} \\ &= \lim_{\epsilon \rightarrow 0} \det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a)] \left(\frac{\det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) - i\epsilon]}{\det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) + i\epsilon]} \right)^{\frac{\xi}{2}},\end{aligned}\quad (\text{G12})$$

where ξ can be taken to be either equal to 1 or -1 . This freedom in the choice of the sign will be exploited later

on to simplify the calculations, as done in [32]. Finally, to proceed further, we assume that $\det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a)]$, and $\det[\mathbf{H}^{\text{SK}}(\mathbf{x}, a) \pm i\epsilon]$ are independent random variables under different realizations of the disorder, i.e.,

$$\begin{aligned}\langle e^{\mu \mathcal{I}(\mathbf{x}^\alpha)} \rangle_{\mathbf{J}} &\approx \lim_{\epsilon \rightarrow 0} \left(\frac{\langle \det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) - i\epsilon] \rangle_{\mathbf{J}}}{\langle \det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) + i\epsilon] \rangle_{\mathbf{J}}} \right)^{\frac{\mu}{2\pi i}} \\ \langle |\det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a)| \rangle_{\mathbf{J}} &\approx \langle \det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) \rangle_{\mathbf{J}} \\ &\times \lim_{\epsilon \rightarrow 0} \left(\frac{\langle \det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) - i\epsilon] \rangle_{\mathbf{J}}}{\langle \det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) + i\epsilon] \rangle_{\mathbf{J}}} \right)^{\frac{\xi}{2}}.\end{aligned}\quad (\text{G13})$$

Within this approximation, we thus see that we need to compute the average $\langle \det \mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) \pm i\epsilon \rangle_{\mathbf{J}}$. We compute this next.

a. Determinant of the Hessian

We express the determinant $\det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) \pm i\epsilon]$ by means of two sets of independent Grassmann variables $\psi_i^\alpha, \bar{\psi}_i^\alpha$ [76], satisfying the anticommutation relations

$$\begin{aligned}\psi_i^\alpha \bar{\psi}_j^\alpha + \bar{\psi}_j^\alpha \psi_i^\alpha &= 0 \\ \psi_i^\alpha \psi_j^\alpha + \psi_j^\alpha \psi_i^\alpha &= 0 \\ \bar{\psi}_i^\alpha \bar{\psi}_j^\alpha + \bar{\psi}_j^\alpha \bar{\psi}_i^\alpha &= 0.\end{aligned}\quad (\text{G14})$$

We thus write

$$\begin{aligned}
\langle \det[\mathbf{H}^{\text{SK}}(\mathbf{x}^\alpha, a) \pm i\epsilon]_{\mathbf{J}} \rangle &= \int \prod_{i=1}^N d\psi_i^\alpha d\bar{\psi}_i^\alpha \left\langle \exp \left[\sum_{i,j} \bar{\psi}_i^\alpha \left(-\frac{J_{ij}}{\sqrt{N}} - \frac{J_0}{N} + (\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon) \delta_{ij} \right) \psi_j^\alpha \right] \right\rangle_{\mathbf{J}} \\
&= \int_{\psi^\alpha, \bar{\psi}^\alpha} \exp \left[\frac{1}{2N} \sum_{i \leq j} (\bar{\psi}_i^\alpha \psi_j^\alpha + \bar{\psi}_j^\alpha \psi_i^\alpha)^2 + \sum_{i,j} \bar{\psi}_i^\alpha \left(-\frac{J_0}{N} + (\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon) \delta_{ij} \right) \psi_j^\alpha \right] \\
&= \int_{\psi^\alpha, \bar{\psi}^\alpha} \exp \left[-\frac{1}{2N} \left(\sum_i \bar{\psi}_i^\alpha \psi_i^\alpha \right)^2 + \sum_{i,j} \bar{\psi}_i^\alpha \left(-\frac{J_0}{N} + (\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon) \delta_{ij} \right) \psi_j^\alpha \right] \\
&= \int_{\psi^\alpha, \bar{\psi}^\alpha} \int_{-\infty}^{\infty} \frac{dt^\alpha}{\sqrt{2\pi/N}} \exp \left[-\frac{N}{2} (t^\alpha)^2 - it^\alpha \left(\sum_i \bar{\psi}_i^\alpha \psi_i^\alpha \right) \right] \\
&\quad \times \exp \left[\sum_{i,j} \bar{\psi}_i^\alpha \left(-\frac{J_0}{N} + (\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon) \delta_{ij} \right) \psi_j^\alpha \right] \tag{G15} \\
&= \int_{-\infty}^{+\infty} \frac{idt^\alpha}{\sqrt{2\pi/N}} \exp \left[\frac{N}{2} (t^\alpha)^2 \right] \int_{\psi^\alpha, \bar{\psi}^\alpha} \exp \left[\sum_{i,j} \bar{\psi}_i^\alpha \left(-\frac{J_0}{N} + (\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon - t^\alpha) \delta_{ij} \right) \psi_j^\alpha \right] \\
&= \int_{-\infty}^{+\infty} \frac{idt^\alpha}{\sqrt{2\pi/N}} \exp \left[N \left(\frac{1}{2} (t^\alpha)^2 + \frac{1}{N} \log \det \left[(\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon - t^\alpha) \delta_{ij} - \frac{J_0}{N} \right] \right) \right] \\
&\approx \int_{-\infty}^{+\infty} dt^\alpha \exp \left[N \left(\frac{1}{2} (t^\alpha)^2 + \frac{1}{N} \log \det [(\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon - t^\alpha) \delta_{ij}] \right) \right] \\
&= \int_{-\infty}^{+\infty} dt^\alpha \exp \left[N \left(\frac{1}{2} (t^\alpha)^2 + \frac{1}{N} \sum_{i=1}^N \log (\partial_x^2 E_I(x_i^\alpha, a) \pm i\epsilon - t^\alpha) \right) \right].
\end{aligned}$$

In the second line, we computed the average over the disorder and we used the shorthand notation for the integral $\int_{\psi^\alpha, \bar{\psi}^\alpha} = \int \prod_i d\psi_i^\alpha d\bar{\psi}_i^\alpha$. In the third line, we used the anticommutation properties of the Grassmann variables to compute the square

$$\sum_{i \leq j} (\bar{\psi}_i^\alpha \psi_j^\alpha + \bar{\psi}_j^\alpha \psi_i^\alpha)^2 = - \left(\sum_i \bar{\psi}_i^\alpha \psi_i^\alpha \right)^2. \tag{G16}$$

In the fourth line, we used a Hubbard-Stratonovich transformation. In the fifth line, we changed the integration variable. In the sixth line, we expressed the integral over the Grassmann variable as the determinant of a matrix. In the seventh line, we neglected subleading contribution in N . Note that the ferromagnetic coupling is among these subleading contributions. This can be understood

from the matrix determinant lemma [77],

$$\det \left[\mathbf{A} + \frac{J_0}{N} \mathbf{t} \otimes \mathbf{t} \right] = \left(1 + \frac{J_0}{N} \mathbf{t} \cdot \mathbf{A}^{-1} \mathbf{t} \right) \det \mathbf{A}, \tag{G17}$$

where $\mathbf{t} = [1, \dots, 1]^T$ is a N -dimensional vector and $\mathbf{A}$ is a real $N \times N$ invertible matrix with entries of $O(1)$ with respect to N . A determinant of this type contributes to the exponential in Eq. (G15) through

$$\begin{aligned}
&\frac{1}{N} \log \det \left[\mathbf{A} + \frac{J_0}{N} \mathbf{t} \otimes \mathbf{t} \right] \\
&= \frac{1}{N} \log \left(1 + \frac{J_0}{N} \text{Tr}[\mathbf{A}^{-1}] \right) + \frac{1}{N} \log \det \mathbf{A} \tag{G18} \\
&= \frac{1}{N} \log \det \mathbf{A} + O(N^{-1}),
\end{aligned}$$

we thus see that the ferromagnetic alignment J_0 contributes to the integral in Eq. (G15) through a term subleading in N . From the last line of Eq. (G15), we anticipate that a saddle point evaluation of the integral is possible. Using this fact, and the approximations in Eq. (G12), we obtain

$$\begin{aligned} \langle e^{\mu \mathcal{I}(x^\alpha)} | \det \mathbf{H}^{\text{SK}}(x^\alpha, a) | \rangle_{\mathbf{J}} &\approx \lim_{\epsilon \rightarrow 0} \int dt_0^\alpha dt_+^\alpha dt_-^\alpha \exp \left[N \left((t_0^\alpha)^2 + \left(\frac{\xi}{2} + \frac{\mu}{2\pi i} \right) ((t_+^\alpha)^2 - (t_-^\alpha)^2) \right) \right. \\ &\quad \left. + \left(\frac{\xi}{2} + \frac{\mu}{2\pi i} \right) \sum_i \log \frac{\partial_x^2 E_I(x_i^\alpha, a) - i\epsilon - t_+^\alpha}{\partial_x^2 E_I(x_i^\alpha, a) + i\epsilon - t_-^\alpha} + \sum_i \log(\partial_x^2 E_I(x_i^\alpha, a) - t_0^\alpha) \right], \end{aligned} \quad (\text{G19})$$

where we neglected subleading contributions in the thermodynamic limit.

3. Saddle point evaluation of the grand-potential

We substitute Eq. (G7) and Eq. (G19) into the grandcanonical partition function of Eq. (G2). We can now evaluate the grand-potential $\Omega(\mu)$ in the thermodynamic limit $N \rightarrow \infty$ through a saddle point integral, as done for the resolvent in App. A. We obtain

$$\begin{aligned} \Omega^{\text{SK}}(\mu) &= \lim_{n \rightarrow 0} \frac{1}{n} \left[\Omega_0 - \log \lim_{\epsilon \rightarrow 0} \int \prod_{\alpha=1}^n \frac{dx^\alpha du^\alpha}{2\pi i} e^{S_\Omega[\{x^\alpha, u^\alpha\}]} \right] \\ \Omega_0 &= \sum_\alpha m_\alpha v_\alpha + \sum_{\alpha, \beta} \left[\lambda_{\alpha\beta} Q_{\alpha\beta} + \frac{1}{2} w_{\alpha\beta}^2 \right] \\ S_\Omega &= - \sum_\alpha u^\alpha \partial_x E_I(x^\alpha, a) + \sum_{\alpha, \beta} \left[\lambda_{\alpha\beta} x^\alpha x^\beta + \frac{1}{2} Q_{\alpha\beta} u^\alpha u^\beta - w_{\alpha\beta} u^\alpha x^\beta \right] + \sum_\alpha [J_0 m_\alpha u^\alpha + v_\alpha x^\alpha] \\ &\quad + \frac{1}{2} \sum_\alpha \left((t_0^\alpha)^2 + \left(\frac{\xi}{2} + \frac{\mu}{2\pi i} \right) ((t_+^\alpha)^2 - (t_-^\alpha)^2) \right) \\ &\quad + \left(\frac{\xi}{2} + \frac{\mu}{2\pi i} \right) \sum_\alpha \log \frac{\partial_x^2 E_I(x^\alpha, a) - i\epsilon - t_+^\alpha}{\partial_x^2 E_I(x^\alpha, a) + i\epsilon - t_-^\alpha} + \sum_\alpha \log(\partial_x^2 E_I(x^\alpha, a) - t_0^\alpha). \end{aligned} \quad (\text{G20})$$

The order parameters satisfy the following set of saddle point equations

$$\begin{aligned} Q_{\alpha\beta} &= \langle x^\alpha x^\beta \rangle_\Omega \\ \lambda_{\alpha\beta} &= \frac{1}{2} \langle u^\alpha u^\beta \rangle_\Omega \\ w_{\alpha\beta} &= - \langle x^\alpha u^\beta \rangle_\Omega \\ m_\alpha &= \langle x^\alpha \rangle_\Omega \\ v_\alpha &= J_0 \langle u^\alpha \rangle_\Omega \\ t_0^\alpha &= \left\langle \frac{1}{\partial_x^2 E_I(x^\alpha, a) - t_0^\alpha} \right\rangle_\Omega \\ t_\pm^\alpha &= \left\langle \frac{1}{\partial_x^2 E_I(x^\alpha) \mp i\epsilon - t_\pm^\alpha} \right\rangle_\Omega, \end{aligned} \quad (\text{G21})$$

where the average $\langle \dots \rangle_\Omega$ denotes an average over the grandcanonical mean field distribution for the replicated system of spins x^α and Lagrange multipliers u^α , defined as,

$$\langle \dots \rangle_\Omega = \frac{\int \prod_\alpha dx^\alpha du^\alpha \dots e^{S_\Omega(\{x^\alpha, u^\alpha\})}}{\int \prod_\alpha dx^\alpha du^\alpha e^{S_\Omega(\{x^\alpha, u^\alpha\})}}. \quad (\text{G22})$$

Note that the saddle point solution to the equation for t_0^α is degenerate, for if t_0^α is a solution, then its complex conjugate is a solution too. On the other hand,

the presence of the term $\mp i\epsilon$ in the equation for $t_\pm^\alpha$ lifts this degeneracy. We thus write $t_\pm^\alpha \equiv t_R^\alpha \pm it_I^\alpha$, separating the real part of $t_\pm^\alpha$ from the imaginary part, while $t_0^\alpha = t_R^\alpha \pm it_I^\alpha$. However, this degeneracy is irrelevant for the computation of the grand-potential $\beta\Omega^{\text{SK}}(\mu)$, because we have freedom on the choice of $\xi = \pm 1$, as mentioned below Eq. (G12). In fact, it is possible to show that if we choose $t_0^\alpha = t_R \pm it_I$ and $\xi = \mp 1$, then the expression of S_Ω in Eq. (G20) does not depend on the choice of the stationary solution t_0^α , and it becomes

$$\begin{aligned}
S_\Omega = & - \sum_{\alpha} u^{\alpha} \partial_x E_I(x^{\alpha}, a) + \sum_{\alpha, \beta} \left[\lambda_{\alpha\beta} x^{\alpha} x^{\beta} + \frac{1}{2} Q_{\alpha\beta} u^{\alpha} u^{\beta} - w_{\alpha\beta} u^{\alpha} x^{\beta} \right] + \sum_{\alpha} [J_0 m_{\alpha} u^{\alpha} + v_{\alpha} x^{\alpha}] \\
& + \sum_{\alpha} \left[\frac{1}{2} ((t_R^{\alpha})^2 - (t_I^{\alpha})^2) + \log [(\partial_x^2 E_I(x^{\alpha}, a) - t_R^{\alpha})^2 + (t_I^{\alpha})^2] \right] \\
& + \mu \sum_{\alpha} \left[\frac{t_R^{\alpha} t_I^{\alpha}}{\pi} + \frac{1}{2\pi i} \log \frac{\partial_x^2 E_I(x^{\alpha}, a) - i(\epsilon + t_I^{\alpha}) - t_R^{\alpha}}{\partial_x^2 E_I(x^{\alpha}, a) + i(\epsilon + t_I^{\alpha}) - t_R^{\alpha}} \right].
\end{aligned} \tag{G23}$$

The real and imaginary parts t_R^{α} and t_I^{α} satisfy the following saddle point equations, which are obtained from Eq. (G21),

$$\begin{aligned}
t_R^{\alpha} &= \lim_{\epsilon \rightarrow 0} \left\langle \frac{\partial_x^2 E_I(x^{\alpha}, a) - t_R^{\alpha}}{(\partial_x^2 E_I(x^{\alpha}, a) - t_R^{\alpha})^2 + (t_I^{\alpha} + \epsilon)^2} \right\rangle_{\Omega} \\
t_I^{\alpha} &= \lim_{\epsilon \rightarrow 0} \left\langle \frac{\epsilon + t_I^{\alpha}}{(\partial_x^2 E_I(x^{\alpha}, a) - t_R^{\alpha})^2 + (t_I^{\alpha} + \epsilon)^2} \right\rangle_{\Omega}.
\end{aligned} \tag{G24}$$

From Eq. (G21), we see that $t_R^{\alpha} \pm it_I^{\alpha}$ satisfy the same self consistent equations for the resolvent, evaluated at the origin, of a random matrix drawn from the Gaussian orthogonal ensemble, whose diagonal elements are perturbed by $\partial_x^2 E_I(x^{\alpha}, a)$, with the x^{α} following the distribution defined in the average $\langle \dots \rangle_{\Omega}$ in Eq. (G22). We now assume that $t_I^{\alpha} = 0$ which, as explained in [32], implies that the spectral density of the Hessian of the most abundant critical points of the system is zero at the origin. We checked through the numerical solution of the saddle point equation that this solution is robust when a nonzero value for the imaginary part is allowed. Within this approximation, the contribution proportional to μ to the action S_{Ω} in Eq. (G23) becomes, in the limit $\epsilon \rightarrow 0$,

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} \mu \sum_{\alpha} \left[\frac{t_R^{\alpha} t_I^{\alpha}}{\pi} + \frac{1}{2\pi i} \log \frac{\partial_x^2 E_I(x^{\alpha}, a) - i(\epsilon + t_I^{\alpha}) - t_R^{\alpha}}{\partial_x^2 E_I(x^{\alpha}, a) + i(\epsilon + t_I^{\alpha}) - t_R^{\alpha}} \right] &= \mu \sum_{\alpha} \Theta(-(\partial_x^2 E_I(x^{\alpha}, a) - t_R^{\alpha})) \\
&\equiv \mu \sum_{\alpha} \mathcal{I}_{\text{MF}}(x^{\alpha}),
\end{aligned} \tag{G25}$$

where $\Theta(x)$ is the Heaviside step function, and we have chosen the branch cut of the logarithm to lie on the positive side of the real axis. This choice has a physical justification: in this way, in the limit $\mu \rightarrow -\infty$ the energy βE_{Ω} selects for configurations x^{α} located at the minima of the energy function $\partial_x^2 E_I(x^{\alpha}, a + t_R^{\alpha})$. The last equality defines the mean field index function $\mathcal{I}_{\text{MF}}(x)$.

We now present the final result for the replica calculation of the grandcanonical partition function. We first find convenient to make a change of variable, replacing the matrices $\lambda_{\alpha\beta}$ and $w_{\alpha\beta}$ with a new pair of matrices $A_{\alpha\beta}$ and $C_{\alpha\beta}$, defined as

$$\begin{aligned}
A_{\alpha\beta} &\equiv -t_R^{\alpha} \delta_{\alpha\beta} - w_{\alpha\beta} \\
C_{\alpha\beta} &\equiv 2\lambda_{\alpha\beta}.
\end{aligned} \tag{G26}$$

With this substitution and our ansatz $t_I^{\alpha} = 0$, the expression of the grand-potential in Eq. (G20) becomes

$$\begin{aligned}
\Omega^{\text{SK}}(\mu) &= \lim_{n \rightarrow 0} \frac{1}{n} \left[\Omega_0 - \log \lim_{\epsilon \rightarrow 0} \int \prod_{\alpha=1}^n \frac{dx^{\alpha} du^{\alpha}}{2\pi i} e^{S_{\Omega}[\{x^{\alpha}, u^{\alpha}\}]} \right] \\
\Omega_0 &= \sum_{\alpha} m_{\alpha} v_{\alpha} + \frac{1}{2} \sum_{\alpha, \beta} [C_{\alpha\beta} Q_{\alpha\beta} + A_{\alpha\beta}^2] + \sum_{\alpha} A_{\alpha\alpha} t_R^{\alpha} \\
S_{\Omega} &= - \sum_{\alpha} u^{\alpha} \partial_x E_I(x^{\alpha}, a + t_R^{\alpha}) + \sum_{\alpha, \beta} \left[\frac{1}{2} C_{\alpha\beta} x^{\alpha} x^{\beta} + \frac{1}{2} Q_{\alpha\beta} u^{\alpha} u^{\beta} + A_{\alpha\beta} u^{\alpha} x^{\beta} \right] + \sum_{\alpha} [J_0 m_{\alpha} u^{\alpha} + v_{\alpha} x^{\alpha}] \\
&+ \sum_{\alpha} \log |\partial_x^2 E_I(x^{\alpha}, a + t_R^{\alpha})| + \mu \sum_{\alpha} \mathcal{I}_{\text{MF}}(x^{\alpha}).
\end{aligned} \tag{G27}$$

The order parameters satisfy the following set of saddle point equations

$$\begin{aligned}
Q_{\alpha\beta} &= \langle x^\alpha x^\beta \rangle_\Omega \\
C_{\alpha\beta} &= \langle u^\alpha u^\beta \rangle_\Omega \\
A_{\alpha\beta} &= \langle x^\alpha u^\beta \rangle_\Omega - t_R^\alpha \delta_{\alpha\beta} \\
m_\alpha &= \langle x^\alpha \rangle_\Omega \\
v_\alpha &= J_0 \langle u^\alpha \rangle_\Omega \\
t_R^\alpha &= \left\langle \frac{1}{|\partial_x^2 E_I(x^\alpha, a + t_R^\alpha)|} \right\rangle_\Omega.
\end{aligned} \tag{G28}$$

These equations are evaluated in the replica-symmetric phase in the next Section.

4. Replica symmetric ansatz

We make the following ansatz for the structure of the order parameters in Eq. (G28),

$$\begin{aligned}
A_{\alpha\beta} &= A_d \delta_{\alpha\beta} + A_o (1 - \delta_{\alpha\beta}) = \Delta A \delta_{\alpha\beta} + A_o \\
C_{\alpha\beta} &= \Delta C \delta_{\alpha\beta} + C_o \\
Q_{\alpha\beta} &= \Delta q \delta_{\alpha\beta} + q_o \\
m_\alpha &= m \\
v_\alpha &= v \\
t_R^\alpha &= t_R,
\end{aligned} \tag{G29}$$

and we compute the exponential weight

$$\begin{aligned}
&\int \prod_\alpha \frac{dx^\alpha du^\alpha}{2\pi i} e^{S_\Omega[\{x^\alpha, u^\alpha\}]} = \int_{\mathbf{x}^\alpha, \mathbf{u}^\alpha} \left(\prod_\alpha |\partial_x^2 E_I(x^\alpha, a + t_R)| \right) \\
&\times \exp \left[\frac{1}{2} \sum_\alpha [\Delta C (x^\alpha)^2 + 2\Delta A x^\alpha u^\alpha + \Delta q (u^\alpha)^2] + \sum_\alpha [x^\alpha v + J_0 m u^\alpha] \right. \\
&\left. - \sum_\alpha u^\alpha \partial_x E_I(x^\alpha, a + t_R) + \mu \sum_\alpha \mathcal{I}_{\text{MF}}(x^\alpha) + \frac{1}{2} \left[C_o \left(\sum_\alpha x^\alpha \right)^2 + 2A_o \left(\sum_\alpha x^\alpha \right) \left(\sum_\alpha u^\alpha \right) + q_o \left(\sum_\alpha u^\alpha \right)^2 \right] \right] \\
&= \int_{\mathbf{x}^\alpha, \mathbf{u}^\alpha} \int \frac{dh_1 dh_2}{2\pi} e^{-\frac{h_1^2}{2} - \frac{h_2^2}{2}} \prod_\alpha |\partial_x^2 E_I(x^\alpha, a + t_R)| \exp \left[\frac{1}{2} [\Delta C (x^\alpha)^2 + 2\Delta A x^\alpha u^\alpha + \Delta q (u^\alpha)^2] + \mu \mathcal{I}_{\text{MF}}(x^\alpha) \right. \\
&\left. + \left(v + h_1 \frac{A_o}{\sqrt{q_o}} + h_2 \sqrt{C_o - \frac{A_o^2}{q_o}} \right) x^\alpha - (\partial_x E_I(x^\alpha, a + t_R) - J_0 m - h_1 \sqrt{q_o}) u^\alpha \right] \\
&\equiv \int_{\mathbf{x}^\alpha, \mathbf{u}^\alpha} \int \frac{dh_1 dh_2}{2\pi} e^{-\frac{h_1^2}{2} - \frac{h_2^2}{2}} \prod_\alpha |\partial_x^2 E_{\Omega, \text{eff}}(x^\alpha, h_1)| \exp \left[\frac{1}{2} [\Delta C (x^\alpha)^2 + 2\Delta A x^\alpha u^\alpha + \Delta q (u^\alpha)^2] + \mu \mathcal{I}_{\text{MF}}(x^\alpha) \right. \\
&\left. + x^\alpha F(h_1, h_2) - \partial_x E_{\Omega, \text{eff}}(x^\alpha, h_1) u^\alpha \right] \\
&= \overline{\left(\int \frac{dx du}{2\pi} |\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)| \exp \left[-\frac{\Delta q}{2} \left(u - i \frac{\Delta A x - \partial_x E_{\Omega, \text{eff}}(x, h_1)}{\Delta q} \right)^2 \right. \right.} \\
&\left. \left. + \frac{1}{2} \Delta C x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) - \frac{(\Delta A x - \partial_x E_{\Omega, \text{eff}}(x, h_1))^2}{2\Delta q} \right] \right)^n} \\
&= \overline{\left(\int dx, \frac{|\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)|}{\sqrt{2\pi \Delta q}} \exp \left[\frac{1}{2} \Delta C x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) - \frac{(\Delta A x - \partial_x E_{\Omega, \text{eff}}(x, h_1))^2}{2\Delta q} \right] \right)^n},
\end{aligned} \tag{G30}$$

where in the second equality, we have used a Hubbard-Stratonovich transformation. In the third equality, we have defined an effective energy $E_{\Omega, \text{eff}}(x, h)$ and a field $F(h_1, h_2)$ as

$$\begin{aligned} E_{\Omega, \text{eff}}(x, h) &\equiv E_I(x, a + t_R) - J_0 m x - \sqrt{q_o} h x \\ F(h_1, h_2) &\equiv v + h_1 \frac{A_o}{\sqrt{q_o}} + h_2 \sqrt{C_o - \frac{A_o^2}{q_o}}. \end{aligned} \quad (\text{G31})$$

Finally in the last equality, we have rotated the integra-

tion domain of the fields u^α in the complex plane, defined the Gaussian average $\overline{\cdots} \equiv \int \frac{dh_1 dh_2}{2\pi} e^{-\frac{h_1^2}{2} - \frac{h_2^2}{2}}$ and observed that the integrals over the different replicated fields factorize. Note that within the replica symmetric ansatz, the distribution of the auxiliary field u is Gaussian, and it can be integrated out. The last line defines, up to a normalization constant, the single spin probability distribution $P_\Omega^{\text{SK}}(x, h_1, h_2)$ as

$$P_\Omega(x, h_1, h_2) \propto |\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)| \exp \left[\frac{1}{2} \Delta C x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) - \frac{(\Delta A x - \partial_x E_{\Omega, \text{eff}}(x, h_1))^2}{2 \Delta q} \right]. \quad (\text{G32})$$

Within the replica-symmetric ansatz, the saddle point equations in Eq. (G28) become

$$\begin{aligned} \Delta A &= \overline{\langle x u \rangle_\Omega} - \overline{\langle x \rangle_\Omega \langle u \rangle_\Omega} - t_R \\ A_o &= \overline{\langle x \rangle_\Omega \langle u \rangle_\Omega} \\ \Delta C &= \overline{\langle u^2 \rangle_\Omega} - \overline{\langle u \rangle_\Omega^2} \\ C_o &= \overline{\langle u \rangle_\Omega^2} \\ \Delta q &= \overline{\langle x^2 \rangle_\Omega} - \overline{\langle x \rangle_\Omega^2} \\ q_o &= \overline{\langle x \rangle_\Omega^2} \\ m &= \overline{\langle x \rangle_\Omega} \\ v &= -J_0 \overline{\langle u \rangle_\Omega}, \end{aligned} \quad (\text{G33})$$

computing the Gaussian average over the field u yields the set of equations in Eq. (70) of the main text.

a. Alternative expression of P_Ω

An alternative expression of the probability distribution $P_\Omega(x, h_1, h_2)$ can be obtained by evaluating the grand-potential average of a generic test function $f(x)$:

$$\begin{aligned} \langle f \rangle_\Omega &= \int dx P(x, h_1, h_2) f(x) \\ &\propto \int dx |\partial_x^2 E_{\Omega, \text{eff}}(x, h_1)| \exp \left[\frac{1}{2} \left(\Delta C - \frac{\Delta A^2}{\Delta q} \right) x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) \right] \\ &\times \exp \left[-\frac{(\partial_x E_{\Omega, \text{eff}})^2}{2 \Delta q} + \frac{\Delta A x}{\Delta q} \partial_x E_{\Omega, \text{eff}} \right] f(x) \\ &= \int dx \int dh_0 \delta(h_0 - \partial_x E_{\Omega, \text{eff}}) |\partial_x^2 E_{\Omega, \text{eff}}| \exp \left[\frac{1}{2} \left(\Delta C - \frac{\Delta A^2}{\Delta q} \right) x^2 + x F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(x) \right] \\ &\times \exp \left[-\frac{h_0^2}{2 \Delta q} + \frac{\Delta A x}{\Delta q} h_0 \right] f(x) \\ &= \int dh_0 e^{-\frac{h_0^2}{2 \Delta q}} \sum_{y \in \text{Crt}_x [E_{\Omega, \text{eff}}(x, h_1 + h_0)]} \exp \left[\frac{1}{2} \left(\Delta C - \frac{\Delta A^2}{\Delta q} \right) y^2 + y F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(y) + \frac{\Delta A y}{\Delta q} h_0 \right] f(y) \\ &\propto \left\langle \sum_{y \in \text{Crt } E_{\Omega, \text{eff}}(x, h_1 + h_0)} e^{\frac{1}{2} \left(\Delta C - \frac{\Delta A^2}{\Delta q} \right) y^2 + y F(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(y) + \frac{\Delta A y}{\Delta q} h_0} f(y) \right\rangle_{h_0}, \end{aligned} \quad (\text{G34})$$

where we made use of the properties of the Dirac delta, and we have defined in the last line a Gaussian average over the field h_0 , with variance Δq . We can thus give an alternative expression of the single-spin grand-potential probability density as

$$P_\Omega(x, h_1, h_2) \propto \left\langle \sum_{y \in \text{Crt}[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} e^{\frac{1}{2}(\Delta C - \frac{\Delta A^2}{\Delta q})y^2 + yF(h_1, h_2) + \mu \mathcal{I}_{\text{MF}}(y) + \frac{\Delta A y}{\Delta q} h_0} \delta(x - y) \right\rangle_{h_0} \quad (\text{G35})$$

where $E_{\Omega, \text{eff}}(x, h)$ is defined in (G31).

b. The effective energy is convex if and only if supersymmetry is preserved

If the effective grand-potential energy $E_{\Omega, \text{eff}}(x, h_1)$ is a convex function of x for all the realizations of the noise h_1 , i.e. if $a + t_R \leq 0$, then supersymmetry is preserved, namely $\Delta A = A_o = C_o = \Delta C = v = 0$. This can be seen through a self-consistent condition, by evaluating the saddle point equations around the supersymmetric point. For the parameter v , we get in fact,

$$\begin{aligned} v &= \frac{J_0}{\Delta q} \overline{\langle \partial_x E_{\Omega, \text{eff}} \rangle_\Omega} \\ &= \frac{J_0}{\Delta q} \overline{\left\langle \sum_{y \in \text{Crt}_x[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} h_0 e^{\mu \mathcal{I}_{\text{MF}}(y)} \right\rangle_{h_0}} \quad (\text{G36}) \\ &= \frac{J_0}{\Delta q} \overline{\langle h_0 \rangle_{h_0}} = 0. \end{aligned}$$

In the second equality, we used the alternative definition of the single-spin distribution P_Ω in the supersymmetric phase, given by Eq. (G35). In the third equality, we used

the fact that, since $E_{\Omega, \text{eff}}(x, h)$ is a convex function of x for all h , there is one critical point for every realization of the Gaussian field h_0 , and the mean field index evaluates to zero because of the convexity of the effective energy. In the last equality, we used the fact that the mean of h_0 is, by definition, 0. This calculation implies that the self-consistent conditions $C_o = A_o = 0$ are verified too, since these order parameters are all proportional to $\langle \partial_x E_{\Omega, \text{eff}} \rangle_\Omega$. The self-consistent condition for the parameter ΔC reads instead

$$\begin{aligned} \Delta C &= -\frac{1}{\Delta q} \left[1 - \Delta q^{-1} \overline{\langle (\partial_x E_{\Omega, \text{eff}})^2 \rangle_\Omega} \right] \\ &= -\frac{1}{\Delta q} \\ &\quad + \Delta q^{-2} \overline{\left\langle \sum_{y \in \text{Crt}_x[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} h_0^2 e^{\mu \mathcal{I}_{\text{MF}}(y)} \right\rangle_{h_0}} \quad (\text{G37}) \\ &= -\frac{1}{\Delta q} + \frac{1}{\Delta q^2} \overline{\langle h_0^2 \rangle_{h_0}} \\ &= 0, \end{aligned}$$

where we used a similar reasoning to the one adopted in Eq. (G36). Finally, we verify the self-consistent condition for ΔA in the following way

$$\begin{aligned} \Delta A &= \overline{\Delta q^{-1} \langle x \partial_x E_{\Omega, \text{eff}} \rangle_\Omega} - t_R = \overline{\left\langle \sum_{y \in \text{Crt}_x[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} \left[\frac{h_0 y}{\Delta q} - \frac{1}{\partial_x^2 E_{\Omega, \text{eff}}(y, h_1 + h_0)} \right] \right\rangle_{h_0}} \\ &= \overline{\left\langle \sum_{y \in \text{Crt}_x[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} \left[\frac{dy}{dh_0} - \frac{1}{\partial_x^2 E_{\Omega, \text{eff}}(y, h_1 + h_0)} \right] \right\rangle_{h_0}} \quad (\text{G38}) \\ &= -\overline{\left\langle \sum_{y \in \text{Crt}_x[E_{\Omega, \text{eff}}(x, h_1 + h_0)]} \frac{d}{dh_0} (\partial_x E_{\Omega, \text{eff}}(x, h_1 + h_0)|_{x=y(h_0)}) \right\rangle_{h_0}} = 0, \end{aligned}$$

where the third equality follows from integration by parts, while in the final equality we used the fact that y is a critical point of the function $E_{\Omega, \text{eff}}$. These calculations demonstrate that when $a + t_R \leq 0$, the supersymmetry

condition is self-consistently verified. Note that this is true regardless of the magnetization m being zero or not. Moreover, these self-consistent conditions hold also in the case $\Delta q \rightarrow 0$, when h_0 concentrates around the origin.

On the other hand, the self consistent condition for a supersymmetric phase is violated when the mean field potential $E_{\Omega,\text{eff}}(x, h)$ is non-convex for some value of h , i.e. when $a + t_R > 0$. Let us compute, in this case, the quantity $\Delta C + C_o$. There is a range of values of h_0 for which the number of critical points of $E_{\Omega,\text{eff}}(x, h_1 + h_0)$ is three: two minima and a saddle. The equality obtained in the supersymmetric case becomes therefore an upper bound, where only one minimum among the three critical points is counted. We thus obtain

$$\begin{aligned} \Delta C + C_o &= -\frac{1}{\Delta q} \left[1 - \frac{1}{\Delta q} \overline{\langle (\partial_x E_{\Omega,\text{eff}})^2 \rangle_\Omega} \right] \\ &> -\frac{1}{\Delta q} + \frac{1}{\Delta q} \overline{\langle h_0^2 \rangle_{h_0}} = 0, \end{aligned} \quad (\text{G39})$$

which demonstrates the violation of the supersymmetric self-consistent condition.

5. $\Omega(\mu) = 0$ in the supersymmetric phase

In the supersymmetric phase, $\Delta A = A_o = \Delta C = C_o = v = 0$. Through a calculation similar to the one performed in Sec. G we can show, starting from Eq. (G27), that

$$\begin{aligned} \Omega(\mu) &= -\overline{\log Z_{\Omega,\text{RS}}(\mathbf{h})} \\ &= -\overline{\log \left\langle \sum_{y \in \text{Crt}_x[E_{\Omega,\text{eff}}(x, h_1 + h_0)]} e^{\mu \mathcal{I}_{\text{MF}}(y)} \right\rangle_{h_0}} \\ &= -\overline{\log 1} \\ &= 0. \end{aligned} \quad (\text{G40})$$

In passing from the second to the third equality in the equation above, we exploited the fact that in the supersymmetric phase the effective energy $E_{\Omega,\text{eff}}(x, h)$ is a convex function of x for all the realizations of h , and thus it thus admits a unique critical point, as a function of x , of mean field index 0.

6. Derivation of Eq. (77)

In this Appendix we discuss how some order parameters appearing in the Kac-Rice calculation can be interpreted, in the replica symmetric phase, as susceptibilities of the grand-potential $\Omega^{\text{SK}}(\mu)$ with respect to perturbations of the single-site energy E_I described by Eq. (76). The different perturbations independently change the mean-field grand-potential action S_Ω in Eq. (G27). The perturbation proportional to ϵ_A changes S_Ω by an

amount $\Delta S_{\Omega, A}$, given by

$$\begin{aligned} \Delta S_{\Omega, A} &\equiv \epsilon_A \sum_\alpha u^\alpha x^\alpha \\ &- \sum_\alpha \log \frac{|\partial_x^2 E_I(x^\alpha, a + t_R^\alpha + \epsilon_A)|}{|\partial_x^2 E_I(x^\alpha, a + t_R^\alpha)|} \\ &+ \mu \sum_\alpha \Theta(-\partial_x^2 E_I(x^\alpha, a + t_R^\alpha + \epsilon_A)) \\ &- \mu \sum_\alpha \Theta(-\partial_x^2 E_I(x^\alpha, a + t_R^\alpha)), \end{aligned} \quad (\text{G41})$$

therefore, the derivative with respect to ϵ_A of the grand-potential becomes, within the replica symmetric ansatz,

$$\begin{aligned} \partial_{\epsilon_A} \Omega^{\text{SK}}(\mu)|_{\epsilon_A=0} &= -\lim_{n \rightarrow 0} \frac{1}{n} \sum_\alpha \left[\langle u^\alpha x^\alpha \rangle_\Omega \right. \\ &\quad \left. - \frac{1}{2} \left\langle \frac{1}{|\partial_x^2 E_I(x^\alpha, a + t_R^\alpha)|} \right\rangle_\Omega \right] \\ &= -\overline{\langle ux \rangle_\Omega} + \frac{1}{2} t_R \\ &= -A_o - \Delta A, \end{aligned} \quad (\text{G42})$$

thus proving the first relation in Eq. (77). The perturbation proportional to ϵ_C introduces a random fields on each soft spin site. It thus couples different replicas. The associated change in the mean field energy $\Delta S_{\Omega, C}$ is

$$\Delta S_{\Omega, C} = \epsilon_C \sum_{\alpha, \beta} u^\alpha u^\beta. \quad (\text{G43})$$

The derivative with respect to ϵ_C of the grand-potential becomes

$$\begin{aligned} \partial_{\epsilon_C} \Omega^{\text{SK}}(\mu)|_{\epsilon_C=0} &= -\lim_{n \rightarrow 0} \frac{1}{n} \sum_{\alpha, \beta} \langle u^\alpha u^\beta \rangle_\Omega \\ &= -\left[\overline{\langle u^2 \rangle_\Omega} - \overline{\langle u \rangle_\Omega^2} \right] \\ &= -\Delta C. \end{aligned} \quad (\text{G44})$$

This is the second relation in Eq. (77). Finally, the perturbation proportional to ϵ_v yields a change in the mean field grand-potential energy $\Delta S_{\Omega, v}$ given by

$$\Delta S_{\Omega, v} = J_0 \epsilon_v \sum_\alpha u^\alpha, \quad (\text{G45})$$

and the associated derivative of the grand-potential reads

$$\begin{aligned} \partial_{\epsilon_v} \Omega^{\text{SK}}(\mu)|_{\epsilon_v=0} &= -\lim_{n \rightarrow 0} \frac{1}{n} J_0 \sum_\alpha \langle u^\alpha \rangle_\Omega \\ &= -J_0 \overline{\langle u \rangle_\Omega} \\ &= -v, \end{aligned} \quad (\text{G46})$$

which is the last relation in Eq. (77) in the main text.

7. Direct sampling of critical points in finite size systems

The direct sampling of critical points in finite size system is performed following the algorithm of [32]. Given a system of N soft spins, we sample N_J realizations of the connectivity matrix $\mathbf{J}^{\text{SK}}$. For each realization of the connectivity matrix, we initialize the system from random initial conditions, and we find the closest critical point using Newton-conjugate gradient descent on the squared norm of the gradient of the energy, $[\nabla E(\mathbf{x}, a)]^2$, with $E(\mathbf{x}, a)$ given by Eq. (2). When sampling minima, we use Newton-conjugate gradient descent on the energy $E(\mathbf{x}, a)$ of the system. This procedure is iterated for N_{samples} times for *each* realization. A deduplication procedure is employed to avoid counting the same critical point multiple times. Pairs of critical points whose Euclidean distance is smaller than 10^{-6} are identified as the same critical point. To produce the scatter-plots in Fig. 5, we used $N = 12$, $N_J = 10$ and $N_{\text{samples}} = 5 \times 10^5$.

8. Argument for the stability of the paramagnetic solutions of $\Omega(\mu)$

The following argument rationalizes the finding that the most abundant critical point are paramagnetic in the supersymmetry-broken region, even for moderately large values of the ferromagnetic alignment J_0 . Let us consider small perturbations to the magnetization δm^α and to the associated Lagrange multiplier $i\delta v^\alpha$ around the paramagnetic, replica-symmetric, supersymmetry-breaking saddle point of the grand-potential $\Omega^{\text{SK}}(\mu)$ in Eq. (G27), where $m_\alpha = v_\alpha = 0$. The perturbation components δm_α are real, while the components $i\delta v_\alpha$ are pure imaginary numbers, since the steepest descent path for the auxiliary variables $\mathbf{v}$ across which the saddle point is evaluated is orthogonal to the real axis. The change of the grand-potential up to quadratic order in these perturbations is

$$n\delta\Omega^{\text{SK}}(\mu) = \begin{bmatrix} \delta\mathbf{m} \\ \delta\mathbf{v} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2\Omega^{\text{SK}}}{\partial m_\alpha \partial m_\beta} & i \frac{\partial^2\Omega^{\text{SK}}}{\partial m_\alpha \partial v_\beta} \\ i \frac{\partial^2\Omega^{\text{SK}}}{\partial v_\alpha \partial m_\beta} & -\frac{\partial^2\Omega^{\text{SK}}}{\partial v_\alpha \partial v_\beta} \end{bmatrix} \begin{bmatrix} \delta\mathbf{m} \\ \delta\mathbf{v} \end{bmatrix}, \quad (\text{G47})$$

where we have defined a $2n \times 2n$ matrix, made up by four $n \times n$ blocks. Their entries read

$$\begin{aligned} \frac{\partial^2\Omega^{\text{SK}}}{\partial m_\alpha \partial m_\beta} &= -J_0^2 \overline{\langle u^2 \rangle}_\Omega \delta_{\alpha\beta} = -J_0^2 \Delta C \delta_{\alpha\beta} \\ \frac{\partial^2\Omega^{\text{SK}}}{\partial m_\alpha \partial v_\beta} &= \left(1 - J_0 \overline{\langle ux \rangle}_\Omega\right) \delta_{\alpha\beta} \\ &= (1 - J_0 \Delta A + J_0 t_R) \delta_{\alpha\beta} \\ \frac{\partial^2\Omega^{\text{SK}}}{\partial v_\alpha \partial v_\beta} &= -\overline{\langle x^2 \rangle}_\Omega \delta_{\alpha\beta} = -\Delta q \delta_{\alpha\beta}, \end{aligned} \quad (\text{G48})$$

where we used the fact that in the replica-symmetric, paramagnetic phase $v = m = q_o = C_o = A_o = 0$. All $n \times n$ sub-blocks are diagonal. The paramagnetic solution

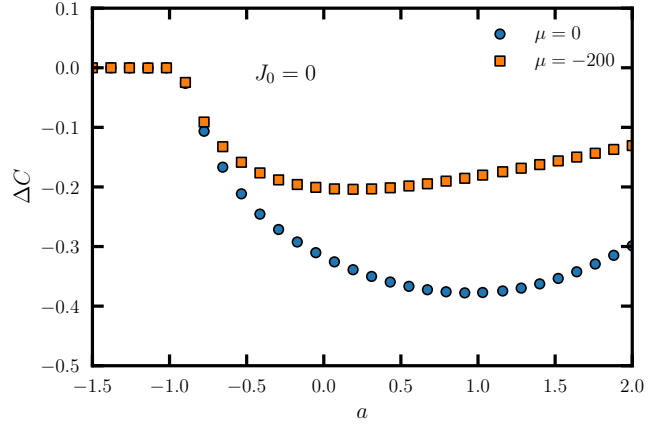


FIG. 14. **Negative values of ΔC in the paramagnetic, supersymmetry-breaking phase.** Plot of ΔC as a function of the laser gain in the absence of ferromagnetic bias, obtained from the numerical solution of Eq. (70). We use two distinct values of the chemical potential μ , corresponding to the properties of the most abundant critical points and the properties of the most abundant minima. The latter are obtained by taking a large negative value of μ . ΔC is less than zero in the supersymmetry-broken phase.

is stable if all the eigenvalues of the Hessian matrix in Eq. (G47) are positive. To check whether this condition is obeyed, let us observe that when integrating numerically the saddle point equations in Eq. (70) for $J_0 = 0$, we obtain that $\Delta C < 0$ both for $\mu = 0$ and for a large and negative value of μ , as shown in Fig. 14. We thus assume that in the paramagnetic phase, $\Delta C < 0$. Then, the top left block of the Hessian in Eq. (G47), associated to fluctuations of the magnetization is positive definite. The matrix in Eq. (G47) has therefore one eigenvalue with a positive definite real part, and thus the paramagnetic solution is stable as long as both eigenvalues are positive. The stability condition becomes thus

$$\det \begin{bmatrix} -J_0^2 \Delta C & i(1 - J_0 \Delta A + J_0 t_R) \\ i(1 - J_0 \Delta A + J_0 t_R) & \Delta q \end{bmatrix} > 0, \quad (\text{G49})$$

which yields

$$[1 - J_0(\Delta A + t_R)]^2 - J_0^2 \Delta C \Delta q > 0. \quad (\text{G50})$$

This condition is satisfied for $\Delta C < 0$. This stability analysis rationalizes our numerical observation that the most abundant critical points are paramagnetic even in presence of a net ferromagnetic coupling strength.

Appendix H: Dynamical mean field theory

In this Appendix we derive the dynamical mean field equations for the coherent Ising machine, presented in Eq. (79), using the dynamical cavity method [60]. The starting point is the gradient descent dynamics for a system of N spins given by Eq. (4), which we rewrite here

as

$$\tau \dot{x}_i = -\partial_x E_I(x_i, a(t)) + \frac{J_0}{N} \sum_j x_j + \frac{1}{\sqrt{N}} \sum_j J_{ij} x_j. \quad (\text{H1})$$

We want to describe the effective dynamics of a single spin of the system. To do so, we consider a system made by $N + 1$ spins, where an additional degree of freedom $x_0(t)$ has been added. We refer to this additional spin as the cavity variable. The new spin interacts with the other spins in the system through the ferromagnetic coupling J_0/N and through a set of symmetric, disordered couplings $J_{0j} = J_{j0}$, which are independent, identically distributed Gaussian variables of mean 0 and unit variance. The cavity variable perturbs the dynamics of the other spins in the system through a field $h_i(t)$, which reads

$$\begin{aligned} h_i(t) &\equiv \frac{J_0}{N} x_0(t) + \frac{J_{i0}}{\sqrt{N}} x_0(t) \\ &= \frac{J_{i0}}{\sqrt{N}} x_0(t) + O(N^{-1}), \end{aligned} \quad (\text{H2})$$

in the second line, we have observed that the field $h_i(t)$ is of order $O(N^{-1/2})$. Let us denote by $\tilde{x}_i(t)$ the trajectories of spins $i = 1, \dots, N$ perturbed by the cavity degree of freedom. The equation of motion of the perturbed spins is

$$\begin{aligned} \tau \dot{\tilde{x}}_i &= -\partial_x E_I(\tilde{x}_i, a(t)) + \frac{J_0}{N} \sum_{j \neq 0} \tilde{x}_j + \frac{1}{\sqrt{N}} \sum_{j \neq 0} J_{ij} \tilde{x}_j \\ &\quad + h_i(t), \end{aligned} \quad (\text{H3})$$

with $h_i(t)$ given by Eq. (H2). On the other hand, the equation of motion of the cavity variable is thus

$$\tau \dot{x}_0 = -E'_I(x_0, a(t)) + \frac{J_0}{N} \sum_{i \neq 0} \tilde{x}_i + \frac{1}{\sqrt{N}} \sum_{i \neq 0} J_{0i} \tilde{x}_i. \quad (\text{H4})$$

All the terms in Eq. (H1), (H3) and Eq. (H4) are of order $O(1)$. On the other hand we see from Eq. (H2) that $h_i(t) \sim O(N^{-1/2})$. We can thus treat the interaction between spin i and spin 0 as a small perturbation to the trajectory of spin i in the system without the cavity variable, $x_i(t)$. Using linear response theory, we thus write

$$\begin{aligned} \tilde{x}_i(t) &\approx x_i(t) + \sum_j \int_0^t d\tau \left. \frac{\delta x_i(t)}{\delta h_j(\tau)} \right|_{h_j=0} h_j(\tau) \\ &= x_i(t) + \frac{1}{\sqrt{N}} \sum_j \int_0^t d\tau \left. \frac{\delta x_i(t)}{\delta h_j(\tau)} \right|_{h_j=0} J_{j0} x_0(\tau), \end{aligned} \quad (\text{H5})$$

Where $\left. \frac{\delta x_i(t)}{\delta h_j(t')} \right|_{h_j=0}$ is the variation of the value $x_i(t)$ under a small change in the field $h_j(t')$ acting on spin j ,

evaluated in the absence of any external field. Substituting Eq. (H5) into Eq. (H4) we obtain, to leading order in N ,

$$\begin{aligned} \tau \dot{x}_0 &\approx -\partial_x E_I(x_0, a(t)) + \frac{J_0}{N} \sum_{i \neq 0} x_i(t) \\ &\quad + \frac{1}{\sqrt{N}} \sum_{i \neq 0} J_{0i} x_i \\ &\quad + \frac{1}{N} \sum_{i,j \neq 0} \int_0^t d\tau \frac{\delta x_i(t)}{\delta h_j(\tau)} J_{0j} J_{i0} x_0(\tau). \end{aligned} \quad (\text{H6})$$

This equation describes the dynamics of the cavity variable in terms of the dynamics of the unperturbed spins $x_i(t)$, which are independent from the dynamics of the cavity. The spins $x_i(t)$ can be considered as a set random variables under different realizations of the disordered couplings J_{ij} . In the thermodynamic limit, the cross-correlations among two spin x_i and x_j read

$$\begin{aligned} &\langle (x_i(t) - \langle x_i(t) \rangle_{\mathbf{J}}) (x_j(t') - \langle x_j(t') \rangle_{\mathbf{J}}) \rangle_{\mathbf{J}} \\ &= \frac{1}{N} \sum_{k,l} \langle J_{ik} J_{jl} x_i(t) x_j(t') \rangle_{\mathbf{J}} \\ &= \delta_{ij} \langle x_i(t) x_j(t') \rangle_{\mathbf{J}} + O(N^{-1}). \end{aligned} \quad (\text{H7})$$

This equation demonstrates that in the thermodynamic limit the dynamics of two different spins under different realizations of the quenched disorder are independent of each other. The term $N^{-1} J_0 \sum_{i \neq 0} x_i(t)$ in Eq. (H6) can then be written as

$$\begin{aligned} \frac{J_0}{N} \sum_i x_i(t) &= \frac{J_0}{N} \sum_i \langle x_i(t) \rangle_{\mathbf{J}} \\ &\quad + \frac{J_0}{N} \sum_i x_i(t) - \langle x_i(t) \rangle_{\mathbf{J}} \\ &= \frac{J_0}{N} \sum_i \langle x_i(t) \rangle_{\mathbf{J}} + O(N^{-1/2}) \\ &\equiv J_0 m(t) + O(N^{-1/2}). \end{aligned} \quad (\text{H8})$$

In the second equality, we applied the central limit theorem, while the third equality defines the time-dependent magnetization of the system $m(t)$. Similarly, the response term in Eq. (H6) concentrates around its average value in the thermodynamic limit:

$$\begin{aligned} &\frac{1}{N} \int_0^t d\tau \sum_{i,j} \left. \frac{\delta x_i(t)}{\delta h_j(\tau)} \right|_{h_j=0} J_{0i} J_{j0} x_0(\tau) \\ &\approx \int_0^t d\tau \left\langle \left. \frac{1}{N} \sum_i \frac{\delta x_i(t)}{\delta h_i(\tau)} \right|_{h_j=0} \right\rangle_{\mathbf{J}} x_0(\tau) \\ &\equiv \int_0^t d\tau R(t, \tau) x_0(\tau), \end{aligned} \quad (\text{H9})$$

where the second line defines the response function $R(t, \tau)$. Finally, the second term in Eq. (H6) can be

written, using the central limit theorem, as a Gaussian noise $\xi(t)$ with mean $\langle \xi(t) \rangle_{\mathbf{J}} = 0$ and variance

$$\begin{aligned} \langle \xi(t) \xi(t') \rangle_{\mathbf{J}} &= \frac{1}{N} \sum_{i,j} \langle J_{0i} J_{0j} x_i(t) x_j(t') \rangle_{\mathbf{J}} \\ &= \frac{1}{N} \sum_i \langle x_i(t) x_i(t') \rangle_{\mathbf{J}} \equiv C(t, t'), \end{aligned} \quad (\text{H10})$$

where the last equality defines the correlation function $C(t, t')$. Finally, we observe that, since the system is fully connected, the dynamics of one spin is equivalent to the dynamics of any other spin. This is true for the dynamics of the cavity variable as well. The index 0 in Eq. (H6) can then be dropped, as well as the index i in the second-to-last lines of Eqs. (H8), (H9), and (H10). We thus obtain an effective equation of motion for a spin $x(t)$ in the coherent Ising machine, namely

$$\begin{aligned} \tau \dot{x}(t) &= -\partial_x E_I(x, a) \\ &\quad + J_0 m(t) + \int_0^t d\tau R(t, \tau) x(\tau) + \xi(\tau) \\ \langle \xi(t) \xi(t') \rangle_{\mathbf{J}} &= C(t, t') = \langle x(t) x(t') \rangle_{\mathbf{J}} \\ m(t) &= \langle x(t) \rangle_{\mathbf{J}} \\ R(t, t') &= \left\langle \frac{\delta x(t)}{\delta h(t')} \right\rangle_{h=0} \Big|_{\mathbf{J}} = \left\langle \frac{\delta x(t)}{\delta \xi(t')} \right\rangle_{\mathbf{J}}. \end{aligned} \quad (\text{H11})$$

In the last line, we used the fact that the variation of $x(t)$ with respect to a variation in an external field $h(t')$ can equivalently be obtained by making an infinitesimal change in the noise at time t' . Equation (H11) is the dynamical mean field theory of the coherent Ising machine. It provides a set of self-consistent equations that can be solved to obtain the temporal correlation $C(t, t')$, magnetization $m(t)$, and response $R(t, t')$ given initial conditions on x . We show how to find such solutions next.

1. Numerical solution of the dynamical mean field theory

Equation (H11) can be solved numerically for any initial condition $x(0)$ on a discrete temporal grid, $t_i = i\Delta t$ with $i = 0, \dots, N_{\text{steps}}$, by means of an iterative self-consistent scheme [78–80]. Starting from an initial guess of the function $m(t_i)$, $C(t_i, t_j)$ and $R(t_i, t_j)$ we can generate a set of N_{traj} independent trajectories $\{x_\alpha(t_i)\}_{\alpha=1}^{N_{\text{traj}}}$ using an Euler scheme. To do this, we first generate and store $N_{\text{traj}} \times N_{\text{steps}}$ realizations of the noise, with correlations $\langle \xi_\alpha(t_i) \xi_\beta(t_j) \rangle = \delta_{\alpha\beta} C(t_i, t_j)$. We can then use the self-consistent equations for $m(t)$ and $C(t, t')$ given by Eq. (H11) to compute the new values of the correlation

function $C_{\text{new}}(t_i, t_j)$ and of the magnetization $m_{\text{new}}(t_i)$, as

$$\begin{aligned} C_{\text{new}}(t_i, t_j) &= \frac{1}{N_{\text{traj}}} \sum_{\alpha=1}^{N_{\text{traj}}} x_\alpha(t_i) x_\alpha(t_j) \\ m_{\text{new}}(t_i) &= \frac{1}{N_{\text{traj}}} \sum_{\alpha=1}^{N_{\text{traj}}} x_\alpha(t_i). \end{aligned} \quad (\text{H12})$$

To compute the response function of the system, we can use the Novikov relation [80, 81], which relates the response function to the realizations of the noise along a trajectory as

$$R(t, t') = \left\langle x(t) \int d\tau C^{-1}(t', \tau) \xi(\tau) \right\rangle_{\mathbf{J}}, \quad (\text{H13})$$

where $C^{-1}(t, t')$ is the inverse of the correlation function $C(t, t')$, defined by the identity $\int d\tau C(t, \tau) C^{-1}(\tau, t') = \delta(t - t')$. For a discrete time grid and a finite number of trajectories, Novikov relation reads

$$\begin{aligned} R_{\text{new}}(t_i, t_j) &= \frac{1}{N_{\text{traj}} \Delta t} \sum_{\alpha=1}^{N_{\text{traj}}} x_\alpha(t_i) \\ &\quad \times \sum_{k=1}^{N_{\text{steps}}} C^{-1}(t_i, t_k) \xi_\alpha(t_k). \end{aligned} \quad (\text{H14})$$

Once the new values of the order parameters C_{new} , R_{new} and m_{new} are known, we update the current values of the order parameter using a linear interpolation between the current values and the new values with parameter γ , namely

$$\begin{aligned} C(t_i, t_j) &\leftarrow \gamma C_{\text{new}}(t_i, t_j) + (1 - \gamma) C(t_i, t_j) \\ R(t_i, t_j) &\leftarrow \gamma R_{\text{new}}(t_i, t_j) + (1 - \gamma) R(t_i, t_j) \\ m(t_i) &\leftarrow \gamma m_{\text{new}}(t_i) + (1 - \gamma) m(t_i). \end{aligned} \quad (\text{H15})$$

This interpolation scheme is used to avoid large jumps as well as oscillations between updates. Once the order parameters are updated, the procedure is repeated until convergence is achieved. As a convergence criterion, we impose

$$\begin{aligned} &\frac{1}{N_{\text{steps}}} \sum_i (m(t_i) - m_{\text{new}}(t_i))^2 \\ &+ \frac{1}{N_{\text{steps}}^2} \sum_{i,j} (C(t_i, t_j) - C_{\text{new}}(t_i, t_j))^2 \leq \epsilon, \end{aligned} \quad (\text{H16})$$

with $\epsilon > 0$. The complexity of this algorithm is of order $O(N_{\text{traj}} \times N_{\text{steps}}^2)$. The results presented in the main text are obtained using $N_{\text{traj}} = 20\,000$, $T/\Delta t = 5\,000$, $\Delta t = 0.15$, $\gamma = 0.3$, $\epsilon = 10^{-4}$.

Appendix I: Wishart planted ensemble

In this Section, we compute the average of the single spin distribution at the global minima of the CIM-WPE in Eq. (98). We begin by rewriting the single spin distribution as the derivative of a modified free energy function,

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \langle \delta(x_i - x) \rangle_{\mathbf{x}, \mathbf{z}^\mu}^{\text{WPE}} &= \frac{1}{N} \left\langle \frac{\int \prod_{i=1}^N dy_i \sum_i \delta(y_i - x) e^{-\beta E^{\text{WPE}}(\mathbf{y})}}{\int \prod_i dy_i e^{-\beta E^{\text{WPE}}(\mathbf{y})}} \right\rangle_{\mathbf{z}^\mu} \\ &= \frac{1}{N} \left\langle \partial_\ell \log \int d\mathbf{y} e^{-\beta E^{\text{WPE}} + \ell \sum_i \delta(y_i - x)} \right\rangle_{\mathbf{z}^\mu} \Big|_{\ell=0} \\ &\equiv \frac{1}{N} \partial_\ell \langle \log Z_\ell^{\text{WPE}} \rangle_{\mathbf{z}^\mu} \Big|_{\ell=0}, \end{aligned} \quad (I1)$$

where the last line defines a modified partition function Z_ℓ^{WPE} . To compute the average over the disorder, we employ the replica trick described in Eq. (21), which yields

$$\begin{aligned} \frac{1}{N} \partial_\ell \langle \log Z_\ell^{\text{WPE}} \rangle_{\mathbf{z}^\mu} \Big|_{\ell=0} &= \frac{1}{N} \partial_\ell \lim_{n \rightarrow 0} \frac{1}{n} \langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} \Big|_{\ell=0} \\ &= \frac{1}{N} \partial_\ell \lim_{n \rightarrow 0} \frac{1}{n} \left\langle \int \prod_{\alpha=1}^n d\mathbf{y}^\alpha e^{-\beta \sum_\alpha E^{\text{WPE}}(\mathbf{y}^\alpha) + \ell \sum_\alpha \sum_i \delta(y_i^\alpha - x)} \right\rangle_{\mathbf{z}^\mu} \Big|_{\ell=0}. \end{aligned} \quad (I2)$$

Upon substitution of Eq. (87) and Eq. (83) into the replicated partition function in Eq. (I2) we obtain

$$\langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} = \int_{\mathbf{y}^\alpha} \left\langle \exp \left[-\frac{\beta}{2N} \sum_{\alpha=1}^n \sum_{\mu=1}^M (\mathbf{w}^\mu \cdot \mathbf{y}^\alpha)^2 + \frac{\beta}{2N} \sum_{\alpha, \mu} \sum_{i=1}^N (w_i^\mu y_i^\alpha)^2 - \sum_{\alpha, i} [\beta E_I(y_i^\alpha, a) - \ell \delta(y_i^\alpha - x)] \right] \right\rangle_{\mathbf{z}^\mu}, \quad (I3)$$

where we used the shorthand notation $\int_{\mathbf{y}^\alpha} \equiv \int \prod_{\alpha=1}^n d\mathbf{y}^\alpha$ to denote an integral over the replicated space of soft-spin systems. The second term in the exponential comes from the fact that the connectivity matrix in the Wishart planted ensemble is defined to be traceless. We use Hubbard-Stratonovich transformation in Eq. (A4) to linearize the term $(\mathbf{w}^\mu \cdot \mathbf{y}^\alpha)^2$ while introducing a set of $n \times M$ continuous degrees of freedom $\{\phi_\mu^\alpha\}$ for $\alpha = 1, \dots, n$ and $\mu = 1, \dots, M$. We obtain

$$\begin{aligned} \langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} &= \left\langle \int_{\mathbf{y}^\alpha, \phi^\alpha} \exp \left[-\sum_{\mu, \alpha} \frac{(\phi_\mu^\alpha)^2}{2} - i \sqrt{\frac{\beta}{N}} \sum_{\mu, \alpha} \phi_\mu^\alpha (\mathbf{w}^\mu \cdot \mathbf{y}^\alpha) \right. \right. \\ &\quad \left. \left. + \frac{\beta}{2N} \sum_{\alpha, \mu, i} (w_i^\mu y_i^\alpha)^2 - \sum_{\alpha, i} [\beta E_I(y_i^\alpha, a) - \ell \delta(y_i^\alpha - x)] \right] \right\rangle_{\mathbf{z}^\mu}, \end{aligned} \quad (I4)$$

where we employed a shorthand notation for the integral $\int_{\mathbf{y}^\alpha, \phi^\alpha} \equiv \int \prod_{\alpha} \frac{d\mathbf{y}^\alpha d\phi^\alpha}{\sqrt{(2\pi)^N}}$. The patterns $\mathbf{w}^\mu$ are constructed from the realizations of the quenched disorder $\mathbf{z}^\mu$, as described in Eq. (84) of the main text. Writing explicitly the average over the quenched disorder we obtain

$$\begin{aligned} \langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} &= \int \prod_{\mu=1}^M \frac{d\mathbf{z}^\mu}{(\sqrt{2\pi})^N} \int_{\mathbf{y}^\alpha, \phi^\alpha} \exp \left[-\frac{1}{2} \sum_{\mu, \alpha} (\phi_\mu^\alpha)^2 - \sum_{\alpha, i} [\beta E_I(y_i^\alpha, a) - \ell \delta(y_i^\alpha - x)] \right] \\ &\quad \times \prod_{\mu=1}^M \exp \left[-\frac{1}{2} \sum_{i,j} z_i^\mu \left(\delta_{ij} - \frac{\beta}{N} \left(\sum_{k, \alpha} \Sigma_{ik}^{1/2} \Sigma_{jk}^{1/2} (y_k^\alpha)^2 \right) \right) z_j^\mu - i \sqrt{\frac{\beta}{N}} \sum_{\alpha} \phi_\mu^\alpha \mathbf{y}^\alpha \cdot \boldsymbol{\Sigma}^{1/2} \mathbf{z}^\mu \right], \end{aligned} \quad (I5)$$

where we see that the traceless property of the connectivity matrix $\mathbf{J}^{\text{WPE}}$ has the effect of changing the correlations of the quenched disorder. The inverse of the covariance matrix of the quenched variables $\{\mathbf{z}^\mu\}$ is now given by the

matrix

$$\begin{aligned}
\delta_{ij} - \frac{\beta}{N} \sum_{k,\alpha} \Sigma_{ik}^{1/2} \Sigma_{jk}^{1/2} (y_k^\alpha)^2 &= \delta_{ij} - \frac{\beta}{N-1} \sum_{\alpha,k} \left(\delta_{ik} - \frac{1}{N} \right) \left(\delta_{jk} - \frac{1}{N} \right) (y_k^\alpha)^2 \\
&= \delta_{ij} - \frac{\beta}{N-1} \sum_{\alpha=1}^n \left(\delta_{ij} (y_i^\alpha)^2 - \frac{1}{N} ((y_j^\alpha)^2 + (y_i^\alpha)^2) + \frac{1}{N^2} \sum_k (y_k^\alpha)^2 \right) \\
&\approx \delta_{ij} \left(1 - \frac{\beta}{N} \sum_{\alpha=1}^n (y_i^\alpha)^2 \right),
\end{aligned} \tag{I6}$$

where in the last passage we have neglected all the terms of order $O(N^{-2})$ or higher. By defining the variable $\mathbf{\Gamma}^\mu$ as

$$\mathbf{\Gamma}^\mu \equiv \sqrt{\frac{\beta}{N}} \sum_{\alpha} \phi_{\mu}^{\alpha} \Sigma^{1/2} \mathbf{y}^{\alpha} \tag{I7}$$

we evaluate the Gaussian integral over the quenched disorder $\mathbf{z}^\mu$ in Eq. (I4), and we expand the argument of the exponential to the leading order in powers of N^{-1} , obtaining

$$\begin{aligned}
\prod_{\mu=1}^M \int \frac{d\mathbf{z}^\mu}{(2\pi)^{N/2}} \exp \left[-\frac{1}{2} \sum_i (z_i^\mu)^2 \left[1 - \frac{\beta}{N} \sum_{\alpha} (y_i^\alpha)^2 \right] - i \mathbf{\Gamma}^\mu \cdot \mathbf{z}^\mu \right] \\
= \exp \left[-\frac{M}{2} \sum_i \log \left(1 - \frac{\beta}{N} \sum_{\alpha} (y_i^\alpha)^2 \right) \right] \prod_{\mu} \exp \left[-\frac{1}{2} \sum_i \frac{(\Gamma_i^\mu)^2}{1 - \frac{\beta}{N} \sum_{\alpha} (y_i^\alpha)^2} \right] \\
\approx \exp \left[\frac{M\beta}{2N} \sum_{\alpha,i} (y_i^\alpha)^2 - \frac{1}{2} \mathbf{\Gamma}^\mu \cdot \mathbf{\Gamma}^\mu \right] \\
= \exp \left[\frac{\alpha\beta}{2} \sum_{\alpha,i} (y_i^\alpha)^2 - \frac{1}{2} \mathbf{\Gamma}^\mu \cdot \mathbf{\Gamma}^\mu \right].
\end{aligned} \tag{I8}$$

From this expression, we observe that the lack of a diagonal element in $\mathbf{J}^{\text{WPE}}$ contributes to a shift in the effective laser gain by an amount $\alpha = M/N$, the fraction of patterns stored in the connectivity matrix. Plugging Eq. (I8) into Eq. (I5) yields

$$\begin{aligned}
\langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} &= \int_{\mathbf{y}^\alpha, \phi^\alpha} \exp \left[-\sum_{\mu\alpha} \frac{(\phi_\mu^\alpha)^2}{2} - \beta \sum_{\alpha,i} E_I(y_i^\alpha, a + \alpha) - \frac{1}{2} \sum_{i\mu} (\Gamma_i^\mu)^2 + \ell \sum_{\alpha,i} \delta(y_i^\alpha - x) \right] \\
&= \int_{\mathbf{y}^\alpha, \phi^\alpha} \left(\prod_{\mu} \exp \left[-\sum_{\alpha} \frac{1}{2} (\phi_\mu^\alpha)^2 - \frac{\beta}{2N} \sum_{\alpha,\beta} \left(\phi_\mu^\alpha \phi_\mu^\beta \mathbf{y}^\alpha \cdot \mathbf{y}^\beta - \frac{\phi_\mu^\alpha \phi_\mu^\beta}{N} (\mathbf{y}^\alpha \cdot \mathbf{t})(\mathbf{y}^\beta \cdot \mathbf{t}) \right) \right] \right) \\
&\quad \times e^{-\beta \sum_{i\alpha} E_I(y_i^\alpha, a + \alpha) + \ell \sum_{\alpha,i} \delta(y_i^\alpha - x)},
\end{aligned} \tag{I9}$$

where in the second line we have used the definition of $\mathbf{\Gamma}^\mu$ in Eq. (I7). Integrating out the Gaussian variables ϕ_μ^α

$$\langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} = \int_{\mathbf{y}^\alpha} \exp \left[-\frac{M}{2} \text{Tr} \log(\mathbf{1}_n + \beta \mathbf{K}) - \beta \sum_{\alpha,i} E_I(y_i^\alpha, a + \alpha) + \ell \sum_{\alpha,i} \delta(y_i^\alpha - x) \right], \tag{I10}$$

where the matrix $\mathbf{1}_n$ and the matrix $\mathbf{K}$ are $n \times n$ matrices. The former is the identity matrix in replica space, while the latter is defined as

$$K_{\alpha\beta} = \frac{1}{N} \mathbf{y}^\alpha \cdot \mathbf{y}^\beta - \frac{1}{N^2} (\mathbf{y}^\alpha \cdot \mathbf{t})(\mathbf{y}^\beta \cdot \mathbf{t}). \tag{I11}$$

Now we use the identity

$$f(a) = \int dx \delta(x - a) f(a) = \int dx \delta(x - a) f(x). \tag{I12}$$

to insert a set of order parameters into the replicated partition function $\langle (Z_\ell)^n \rangle_{\mathbf{z}^\mu}$. The first one is an $n \times n$ overlap matrix $\mathbf{Q}$, while the second one is a n -dimensional magnetization vector $\mathbf{m}$. We thus obtain

$$\begin{aligned} \langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} &= \int \left(\prod_\alpha d\mathbf{y}^\alpha \right) \int \left(\prod_{\alpha \leq \beta} dQ_{\alpha\beta} \right) \delta \left(Q_{\alpha\beta} - N^{-1} \sum_i y_i^\alpha y_i^\beta \right) \left(\prod_\alpha dm_\alpha \right) \delta \left(m_\alpha - N^{-1} \sum_i t_i y_i^\alpha \right) \\ &\times \exp \left[-\frac{M}{2} \text{Tr} \log(\mathbf{1} + \beta \mathbf{K}) \right] e^{-\beta \sum_{i\alpha} E_I(y_i^\alpha, a+\alpha) + \ell \sum_{\alpha,i} \delta(y_i^\alpha - x)}. \end{aligned} \quad (\text{I13})$$

Using the Fourier representation of the Delta distribution, we now introduce a set of Lagrange multipliers, $r_{\alpha\beta}$ and $\hat{m}_\alpha$, for each entries of $\mathbf{Q}$ and $\mathbf{m}$, respectively.

$$\begin{aligned} \langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} &= \int \prod_\alpha d\mathbf{y}^\alpha \prod_{\alpha \leq \beta} \frac{idr_{\alpha\beta}}{2\pi} \prod_{\alpha \leq \beta} Q_{\alpha\beta} \prod_\alpha dm_\alpha \prod_\alpha \frac{id\hat{m}_\alpha}{2\pi} e^{-\beta \sum_{\alpha,i} E_I(y_i^\alpha, a+\alpha) + \ell \sum_{\alpha,i} \delta(y_i^\alpha - x)} \\ &\times \exp \left[-\frac{M}{2} \text{Tr} \log(\mathbf{1} + \beta \mathbf{Q} - \beta \mathbf{m} \otimes \mathbf{m}) - \sum_{\alpha\beta} \frac{1}{2} r_{\alpha\beta} (N Q_{\alpha\beta} - \sum_i y_i^\alpha y_i^\beta) - \sum_\alpha \hat{m}_\alpha (N m_\alpha - \sum_i t_i y_i^\alpha) \right] \end{aligned} \quad (\text{I14})$$

After rescaling $r_{\alpha\beta} \rightarrow \alpha\beta^2 r_{\alpha\beta}$, $\hat{m}_\gamma \rightarrow \alpha\beta \hat{m}_\gamma$, and performing more steps of algebra, we obtain

$$\begin{aligned} \langle (Z_\ell^{\text{WPE}})^n \rangle_{\mathbf{z}^\mu} &= \int d\mathbf{\Gamma} \exp N \left[-\frac{\alpha}{2} \text{Tr} \log(\mathbf{1} + \beta \mathbf{Q} - \beta \mathbf{m} \otimes \mathbf{m}) - \frac{\alpha\beta^2}{2} \sum_{\alpha\beta} r_{\alpha\beta} Q_{\alpha\beta} - \alpha\beta \sum_\alpha \hat{m}_\alpha m_\alpha + \log Z_{\text{MF},\ell}^{\text{WPE}} \right] \Big|_{\ell=0} \\ &\equiv \int d\mathbf{\Gamma} e^{NS^{\text{WPE}}[\mathbf{\Gamma}]}, \end{aligned} \quad (\text{I15})$$

where we denoted by $\mathbf{\Gamma} = \{\mathbf{r}, \mathbf{Q}, \mathbf{m}, \hat{\mathbf{m}}\}$ the set of various order parameters, and by $d\mathbf{\Gamma}$ the product of their volume factors in phase space. We also defined an action $S^{\text{WPE}}(\mathbf{\Gamma})$ and a mean field partition function $Z_{\text{MF},\ell}^{\text{WPE}}$ as

$$\begin{aligned} S^{\text{WPE}}[\mathbf{\Gamma}] &= -\frac{\alpha}{2} \text{Tr} \log(\mathbf{1} + \beta \mathbf{Q} - \beta \mathbf{m} \otimes \mathbf{m}) - \frac{\alpha\beta^2}{2} \sum_{\alpha\beta} r_{\alpha\beta} Q_{\alpha\beta} - \alpha\beta \sum_\alpha \hat{m}_\alpha m_\alpha + \log Z_{\text{MF},\ell}^{\text{WPE}}, \\ Z_{\text{MF},\ell}^{\text{WPE}} &= \int \prod_\alpha d\mathbf{y}^\alpha e^{\frac{\alpha\beta^2}{2} \sum_{\alpha\beta} r_{\alpha\beta} y^\alpha y^\beta + \alpha\beta \sum_\alpha \hat{m}_\alpha y^\alpha - \beta \sum_\alpha E_I(y^\alpha, a+\alpha) + \ell \sum_\alpha \delta(y^\alpha - x)}. \end{aligned} \quad (\text{I16})$$

Substituting Eq. (I15) into Eq. (I1) and taking the limit $N \rightarrow \infty$, we obtain

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \langle \delta(x_i - x) \rangle_{\mathbf{x}, \mathbf{z}^\mu}^{\text{WPE}} = \lim_{N \rightarrow \infty} \frac{1}{N} \partial_\ell \lim_{n \rightarrow 0} \frac{1}{n} \log \int d\mathbf{\Gamma} e^{NS^{\text{WPE}}[\mathbf{\Gamma}]}. \quad (\text{I17})$$

As in Appendix A, we ignore the subleading normalization factors in $d\mathbf{\Gamma}$. In the limit $N \rightarrow \infty$, the integral on the right hand side of Eq. (I17) is evaluated at the saddle point of the action $S^{\text{WPE}}(\mathbf{\Gamma})$. The saddle point equations are given in Eq. (103) of the main text. After the saddle point evaluation, we obtain

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \langle \delta(x_i - x) \rangle_{\mathbf{x}, \mathbf{z}^\mu}^{\text{WPE}} &= \partial_\ell \lim_{n \rightarrow 0} \frac{1}{n} \left[-\frac{\alpha}{2} \text{Tr} \log(\mathbf{1} + \beta \mathbf{Q} - \beta \mathbf{m} \otimes \mathbf{m}) \right. \\ &\quad \left. - \frac{\alpha\beta^2}{2} \sum_{\alpha\beta} r_{\alpha\beta} Q_{\alpha\beta} - \alpha\beta \sum_\alpha \hat{m}_\alpha m_\alpha + \log Z_{\text{MF},\ell}^{\text{WPE}} \right] \Big|_{\ell=0}. \end{aligned} \quad (\text{I18})$$

The derivative over ℓ eliminates all but the last term on the right hand side of the above equation, thus yielding

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \langle \delta(x_i - x) \rangle_{\mathbf{x}, \mathbf{z}^\mu}^{\text{WPE}} = \lim_{n \rightarrow 0} \frac{1}{n} \frac{1}{Z_{\text{MF}}^{\text{WPE}}} \int \prod_\alpha d\mathbf{y}^\alpha \left(\sum_\alpha \delta(y^\alpha - x) \right) e^{-\beta E_{\text{MF}}^{\text{WPE}}(\{y^\alpha\})}, \quad (\text{I19})$$

where the evaluation at $\ell = 0$ has been performed. This is Eq. (99) in the main text. In the next Section, we evaluate the expression above using the replica-symmetric ansatz.

1. Replica-symmetric ansatz

In the replica-symmetric phase, we take $Q_{\alpha\beta} = q_d\delta_{\alpha\beta} + q_o(1 - \delta_{\alpha\beta})$, $r_{\alpha\beta} = r_d\delta_{\alpha\beta} + r_o(1 - \delta_{\alpha\beta})$. We take $m_\alpha = m$, $\hat{m}_\alpha = \hat{m}$. The logarithm of the modified mean field partition function $\log Z_{\text{MF}, \ell}^{\text{WPE}}$ becomes

$$\begin{aligned}
\log Z_{\text{MF}, \ell}^{\text{WPE}} &= \log \int \prod_{\alpha} dy^{\alpha} e^{\frac{\alpha\beta^2}{2}r_o(\sum_{\alpha} y^{\alpha})^2 + \frac{\alpha\beta^2}{2}(r_d-r_o)\sum_{\alpha}(y^{\alpha})^2 + \alpha\beta\hat{m}\sum_{\alpha} y^{\alpha} - \beta\sum_{\alpha} E_I(y^{\alpha}, a+\alpha) + \ell\sum_{\alpha} \delta(y^{\alpha}-x)} \\
&= \log \overline{\prod_{\alpha} \left[\int dy^{\alpha} e^{-\beta\sqrt{\alpha r_o} h y^{\alpha} + \frac{\alpha\beta^2}{2}(r_d-r_o)(y^{\alpha})^2 + \alpha\beta\hat{m}y^{\alpha} - \beta E_I(y^{\alpha}, a+\alpha) + \ell\delta(y^{\alpha}-x)} \right]} \\
&= \log \left[\int dy^{\alpha} e^{-\beta\sqrt{\alpha r_o} h y^{\alpha} + \frac{\alpha\beta^2}{2}(r_d-r_o)(y^{\alpha})^2 + \alpha\beta\hat{m}y^{\alpha} - \beta E_I(y^{\alpha}, a+\alpha) + \ell\delta(y^{\alpha}-x)} \right]^n \\
&= n \log \left[\int dy^{\alpha} e^{-\beta\sqrt{\alpha r_o} h y^{\alpha} + \frac{\alpha\beta^2}{2}(r_d-r_o)(y^{\alpha})^2 + \alpha\beta\hat{m}y^{\alpha} - \beta E_I(y^{\alpha}, a+\alpha) + \ell\delta(y^{\alpha}-x)} \right] + O(n) \\
&\equiv n \log \int dy e^{-\beta E_{\text{eff}}^{\text{WPE}}(x, h) + \ell\delta(y-x)} + O(n).
\end{aligned} \tag{I20}$$

In the second line, we performed a Hubbard-Stratonovich transformation to linearize the squared sum over the replicated spins y^{α} , thus introducing a Gaussian random field h of mean zero and unit variance. The overline denotes an average over this Gaussian distribution. In the third line, we used the fact that integral over the different replicas factorizes into the product of n identical integrals. In the fourth line, we expanded to the lowest order in n . The last line defined an effective single-site mean field energy $E_{\text{eff}}^{\text{WPE}}(x, h)$, which is given in Eq. (107) of the main text. Through these manipulations, Eq. (I19) in the replica-symmetric phase becomes

$$\begin{aligned}
P_{\text{MF}}^{\text{WPE}}(x) &= e^{-\beta E_{\text{eff}}^{\text{WPE}}(x, h)} [Z_{\text{RS}}^{\text{WPE}}]^{-1} \\
&\equiv P_{\text{RS}}^{\text{WPE}}(x),
\end{aligned} \tag{I21}$$

which is Eq. (106) in the main text. Following relatively similar manipulations to the ones performed above and in Appendices A and C, we obtain the saddle point equations of Eq. (103) in the main text.

2. Stability of the replica symmetric solution

To investigate the stability of the replica-symmetric phase, we study the Hessian of the action S^{WPE} with respect to perturbations in the space of overlaps $Q_{\alpha\beta}$ and Lagrange multipliers $r_{\alpha\beta}$. The Hessian matrix has the following block structure,

$$\begin{bmatrix} \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\gamma\delta}} & \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial r_{\gamma\delta}} \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial Q_{\gamma\delta}} & \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\gamma\delta}} \end{bmatrix}. \tag{I22}$$

Let us observe first that the top right and bottom left blocks are diagonal, because of the term proportional to $Q_{\alpha\beta}r_{\alpha\beta}$ in the action S^{WPE} in Eq. (I16). The structure of the Hessian matrix is then identical to the one that arises when studying the stability of the replica symmetric phase in the Hopfield or anti-Hopfield models [62, 82]. The condition for the stability of the convex phase derived in those cases applies also here, and it reads

$$1 - \frac{1}{\alpha\beta^2} \left[\frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\alpha\beta}} - 2 \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\alpha\gamma}} + \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\gamma\delta}} \right] \left[\frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\alpha\beta}} - 2 \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\alpha\gamma}} + \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\gamma\delta}} \right] < 0. \tag{I23}$$

We compute first the terms involving derivatives with respect to $Q_{\alpha\beta}$, which depend only on the logarithmic term in the action S^{WPE} in Eq. (I16). To derive this, we first use the following identity, valid for any continuous function f and any symmetric matrix $\mathbf{A}$,

$$\begin{aligned}
\partial_{A_{\alpha\beta}} \text{Tr } f(\mathbf{A}) &= f'(\mathbf{A})_{\alpha\beta} + f'(\mathbf{A})_{\beta\alpha} \\
&\quad - \delta_{\alpha\beta} f'(\mathbf{A})_{\alpha\beta}.
\end{aligned} \tag{I24}$$

In our case, $f(\mathbf{A}) = \log(\mathbf{A} + \mathbf{B})$. The first derivative of f is

$$\begin{aligned}
f'(\mathbf{A})_{\alpha\beta} &= (\mathbf{A} + \mathbf{B})_{\alpha\beta}^{-1} + (\mathbf{A} + \mathbf{B})_{\beta\alpha}^{-1} \\
&\quad - \delta_{\alpha\beta} (\mathbf{A} + \mathbf{B})_{\alpha\beta}^{-1}.
\end{aligned} \tag{I25}$$

Now we let $\mathbf{K} \equiv (\mathbf{A} + \mathbf{B})^{-1}$. We seek $\partial_{A_{\gamma\delta}} K_{\alpha\beta}$. Using

a matrix calculus identity, we get

$$\partial_{A_{\gamma\delta}} K_{\alpha\beta} = - \sum_{\rho\sigma} K_{\alpha\rho} \frac{\partial A_{\rho\sigma}}{\partial A_{\gamma\delta}} K_{\sigma\beta} \quad (\text{I26})$$

Because A is symmetric,

$$\frac{\partial A_{\rho\sigma}}{\partial A_{\gamma\delta}} = \begin{cases} \mathbb{I}_{\rho\gamma}\mathbb{I}_{\sigma\delta} + \mathbb{I}_{\rho\delta}\mathbb{I}_{\sigma\gamma}, & \gamma \neq \delta \\ \mathbb{I}_{\rho\gamma}\mathbb{I}_{\rho\gamma}, & \gamma = \delta, \end{cases} \quad (\text{I27})$$

where we have used the symbol $\mathbb{I}$ to represent an indicator function (i.e. a Kronecker delta) to avoid confusion with the δ representing an index.

We thus obtain

$$\begin{aligned} \frac{\partial^2 \text{Tr} f(A)}{\partial A_{\alpha\beta} \partial A_{\gamma\delta}} &= \begin{cases} 2\partial_{A_{\gamma\delta}} K_{\alpha\beta}, & \alpha \neq \beta \\ \partial_{A_{\gamma\delta}} K_{\alpha\alpha}, & \alpha = \beta \end{cases} \\ &= \begin{cases} -2(K_{\alpha\gamma}K_{\delta\beta} + K_{\alpha\delta}K_{\beta\gamma}), & \alpha \neq \beta, \gamma \neq \delta \\ -2K_{\alpha\gamma}K_{\gamma\beta}, & \alpha \neq \beta, \gamma = \delta \\ -(K_{\alpha\gamma}K_{\delta\alpha} + K_{\alpha\delta}K_{\gamma\alpha}), & \alpha = \beta, \gamma \neq \delta \\ -K_{\alpha\gamma}K_{\alpha\gamma}, & \alpha = \beta, \gamma = \delta. \end{cases} \end{aligned} \quad (\text{I28})$$

In our case, $\mathbf{A} + \mathbf{B} = \mathbf{1}_n + \beta\mathbf{Q} - \beta\mathbf{m} \otimes \mathbf{m}$. Therefore in the replica symmetric phase, $\mathbf{K}$ becomes

$$\mathbf{K} = ((1 + \Delta\tilde{q})\mathbf{1}_n + \beta(q_o - m^2)\mathbf{t}_n \otimes \mathbf{t}_n)^{-1}, \quad (\text{I29})$$

where $\mathbf{t}_n \equiv [1, \dots, 1]^T$ is a n -dimensional vector with all entries equal to 1. Using the Sherman-Morrison formula [83], we obtain

$$\begin{aligned} \mathbf{K} &= \frac{1}{1 + \Delta\tilde{q}}\mathbf{1}_n - \frac{\beta(q_o - m^2)}{(1 + \Delta\tilde{q})^2}\mathbf{t}_n \otimes \mathbf{t}_n \\ &= -\Delta\tilde{r}\mathbf{1}_n - \beta r_o \mathbf{t}_n \otimes \mathbf{t}_n, \end{aligned} \quad (\text{I30})$$

where in the second line we used the replica-symmetric saddle point equations given by Eq. (I08). We can now compute all the matrix elements of the top left block of the Hessian of the CIM-WPE action in Eq. (I22), obtaining,

$$\begin{aligned} \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\alpha} \partial Q_{\alpha\alpha}} &= \frac{\alpha\beta^2}{2} K_{\alpha\alpha}^2 \\ \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\alpha} \partial Q_{\alpha\gamma}} &= \alpha\beta^2 K_{\alpha\alpha} K_{\gamma\alpha} \\ \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\alpha} \partial Q_{\gamma\delta}} &= \frac{\alpha\beta^2}{2} (K_{\alpha\gamma}K_{\delta\alpha} + K_{\alpha\delta}K_{\gamma\alpha}) \\ \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\alpha} \partial Q_{\gamma\gamma}} &= \frac{\alpha\beta^2}{2} K_{\alpha\gamma}^2 \\ \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\alpha\beta}} &= \alpha\beta^2 (K_{\alpha\alpha}^2 + K_{\alpha\beta}^2) \\ \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\alpha\gamma}} &= \alpha\beta^2 (K_{\alpha\alpha}K_{\beta\gamma} + K_{\alpha\gamma}K_{\alpha\beta}) \\ \frac{\partial^2 S^{\text{WPE}}}{\partial Q_{\alpha\beta} \partial Q_{\gamma\delta}} &= \alpha\beta^2 (K_{\alpha\gamma}K_{\delta\beta} + K_{\alpha\delta}K_{\beta\gamma}) \end{aligned} \quad (\text{I31})$$

The elements of $\mathbf{K}$ are, in the replica symmetric phase,

$$\begin{aligned} K_{\alpha\alpha} &= \frac{1}{1 + \Delta\tilde{q}} - \frac{\beta(q_o - m^2)}{(1 + \Delta\tilde{q})^2} \\ &= -\Delta\tilde{r} - \beta r_o \\ K_{\alpha\beta} &= -\frac{\beta(q_o - m^2)}{(1 + \Delta\tilde{q})^2} \\ &= -\beta r_o, \end{aligned} \quad (\text{I32})$$

with $\alpha \neq \beta$.

Now we turn to the computation of derivatives with respect to the Lagrange multipliers $r_{\alpha\beta}$, which involve the term $-\log Z_{\text{MF}}^{\text{WPE}}$ in Eq. (I16). Using the fact that $r_{\alpha\beta}$ is a symmetric matrix, we obtain

$$\begin{aligned} -\partial_{r_{\alpha\beta}} \partial_{r_{\gamma\delta}} \log Z &= -\alpha^2 \beta^4 \langle x^\alpha x^\beta x^\gamma x^\delta \rangle_{\text{MF}}^{\text{WPE}} \\ &\quad + \alpha^2 \beta^4 \langle x^\alpha x^\beta \rangle_{\text{MF}}^{\text{WPE}} \langle x^\gamma x^\delta \rangle_{\text{MF}}^{\text{WPE}}. \end{aligned} \quad (\text{I33})$$

In the replica-symmetric phase, the different types of matrix elements become

$$\begin{aligned} \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\alpha} \partial r_{\alpha\alpha}} &= \alpha^2 \beta^4 \left[\overline{\langle x^4 \rangle_{\text{RS}}^{\text{WPE}}} - \overline{\langle x^2 \rangle_{\text{RS}}^{\text{WPE}}}^2 \right] \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\alpha} \partial r_{\alpha\gamma}} &= \alpha^2 \beta^4 \left[\overline{\langle x^3 \rangle_{\text{RS}}^{\text{WPE}} \langle x \rangle_{\text{RS}}^{\text{WPE}}} - \overline{\langle x^2 \rangle_{\text{RS}}^{\text{WPE}}} \overline{\langle x \rangle_{\text{RS}}^{\text{WPE}}}^2 \right] \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\alpha} \partial r_{\gamma\delta}} &= \alpha^2 \beta^4 \left[\overline{\langle x^2 \rangle_{\text{RS}}^{\text{WPE}} (\langle x \rangle_{\text{RS}}^{\text{WPE}})^2} \right. \\ &\quad \left. - \overline{\langle x^2 \rangle_{\text{RS}}^{\text{WPE}}} \overline{(\langle x \rangle_{\text{RS}}^{\text{WPE}})^2} \right] = 0 \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\alpha} \partial r_{\gamma\gamma}} &= \alpha^2 \beta^4 \left[\overline{(\langle x^2 \rangle_{\text{RS}}^{\text{WPE}})^2} - \overline{\langle x^2 \rangle_{\text{RS}}^{\text{WPE}}}^2 \right] = 0 \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\alpha\beta}} &= \alpha^2 \beta^4 \left[\overline{(\langle x^2 \rangle_{\text{RS}}^{\text{WPE}})^2} - \overline{(\langle x \rangle_{\text{RS}}^{\text{WPE}})^2}^2 \right] \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\alpha\gamma}} &= \alpha^2 \beta^4 \left[\overline{\langle x^2 \rangle_{\text{RS}}^{\text{WPE}} (\langle x \rangle_{\text{RS}}^{\text{WPE}})^2} - \overline{(\langle x \rangle_{\text{RS}}^{\text{WPE}})^2}^2 \right] \\ \frac{\partial^2 S^{\text{WPE}}}{\partial r_{\alpha\beta} \partial r_{\gamma\delta}} &= \alpha^2 \beta^4 \left[\overline{(\langle x \rangle_{\text{RS}}^{\text{WPE}})^4} - \overline{(\langle x \rangle_{\text{RS}}^{\text{WPE}})^2}^2 \right] = 0, \end{aligned} \quad (\text{I34})$$

where $\langle \dots \rangle_{\text{RS}}^{\text{WPE}}$ is an average over the Boltzmann distribution of the effective mean field energy $E_{\text{eff}}^{\text{WPE}}(x, h)$, for a given realization of the quenched field h , namely,

$$\langle \dots \rangle_{\text{RS}}^{\text{WPE}} \equiv \frac{\int dx \dots e^{-\beta E_{\text{eff}}^{\text{WPE}}(x, h)}}{\int dx e^{-\beta E_{\text{eff}}^{\text{WPE}}(x, h)}}, \quad (\text{I35})$$

and we recall that the overline denotes an average over the realizations of the quenched field h , which follows a Gaussian distribution of zero mean and unit variance. Plugging the relevant entries of Eq. (I31) and Eq. (I34) into the stability condition in Eq. (I23), and evaluating the resulting expression in the low temperature limit $\beta \rightarrow$

∞ , we obtain the stability condition

$$\begin{aligned}\tilde{\Lambda}_R^{\text{WPE}} &\equiv -1 + \alpha \Delta \tilde{r}^2 \beta^2 \left[\left\langle x^2 - (\langle x \rangle_{\text{RS}}^{\text{WPE}})^2 \right\rangle_{\text{RS}}^{\text{WPE}} \right]^2 \\ &= -1 + \alpha \Delta \tilde{r}^2 \left[\partial_x^2 E_{\text{eff}}^{\text{WPE}}(x, h) |_{x=x^*(h)} \right]^{-2} < 0,\end{aligned}\quad (\text{I36})$$

where $x^*(h)$ is the global minimum of the effective energy $E_{\text{eff}}^{\text{WPE}}(x, h)$ for a fixed realization of h . This is the stability condition given by Eq. (109) of the main text.

-
- [1] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization: algorithms and complexity* (Courier Corporation, 1998).
 - [2] T. Wang and J. Roychowdhury, Oim: Oscillator-based ising machines for solving combinatorial optimisation problems, in *International Conference on Unconventional Computation and Natural Computation* (Springer, 2019) pp. 232–256.
 - [3] F. Molnár, S. R. Kharel, X. S. Hu, and Z. Toroczkai, Accelerating a continuous-time analog sat solver using gpus, *Computer Physics Communications* **256**, 107469 (2020).
 - [4] N. Mohseni, P. L. McMahon, and T. Byrnes, Ising machines as hardware solvers of combinatorial optimization problems, *Nature Reviews Physics* **4**, 363 (2022).
 - [5] A. Mallick, M. K. Bashar, D. S. Truesdell, B. H. Calhoun, S. Joshi, and N. Shukla, Using synchronized oscillators to compute the maximum independent set, *Nature communications* **11**, 4689 (2020).
 - [6] Z. Wang, A. Marandi, K. Wen, R. L. Byer, and Y. Yamamoto, Coherent ising machine based on degenerate optical parametric oscillators, *Physical Review A* **88**, 063853 (2013).
 - [7] P. L. McMahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara, *et al.*, A fully programmable 100-spin coherent ising machine with all-to-all connections, *Science* **354**, 614 (2016).
 - [8] T. Inagaki, Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi, P. L. McMahon, T. Umeki, K. Enbutsu, *et al.*, A coherent ising machine for 2000-node optimization problems, *Science* **354**, 603 (2016).
 - [9] T. Honjo, T. Sonobe, K. Inaba, T. Inagaki, T. Ikuta, Y. Yamada, T. Kazama, K. Enbutsu, T. Umeki, R. Kasahara, *et al.*, 100,000-spin coherent ising machine, *Science advances* **7**, eabh0952 (2021).
 - [10] Y. Yamamoto, T. Leleu, S. Ganguli, and H. Mabuchi, Coherent ising machines—quantum optics and neural network perspectives, *Applied Physics Letters* **117** (2020).
 - [11] A. Lucas, Ising formulations of many np problems, *Frontiers in physics* **2**, 5 (2014).
 - [12] S. Reifenstein, S. Kako, F. Khoyratee, T. Leleu, and Y. Yamamoto, Coherent ising machines with optical error correction circuits, *Advanced Quantum Technologies* **4**, 2100077 (2021).
 - [13] B. Lu, C.-R. Fan, L. Liu, K. Wen, and C. Wang, Speed-up coherent ising machine with a spiking neural network, *Optics Express* **31**, 3676 (2023).
 - [14] M. Syed and N. G. Berloff, Physics-enhanced bifurcation optimisers: all you need is a canonical complex network, *IEEE Journal of Selected Topics in Quantum Electronics* **29**, 1 (2023).
 - [15] S. H. G. Mastiyage Don, Y. Inui, S. Kako, Y. Yamamoto, and T. Aonishi, Mean-field coherent ising machines with artificial zeeman terms, *Journal of Applied Physics* **134** (2023).
 - [16] Z.-Y. Zhou, C. Gneiting, J. You, and F. Nori, Frustration elimination and excited state search in coherent ising machines, *Physical Review Letters* **134**, 090401 (2025).
 - [17] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, Vol. 9 (World Scientific Publishing Company, 1987).
 - [18] H. Nishimori, *Statistical physics of spin glasses and information processing: an introduction*, 111 (Clarendon Press, 2001).
 - [19] G. Parisi, P. Urbani, and F. Zamponi, *Theory of simple glasses: exact solutions in infinite dimensions* (Cambridge University Press, 2020).
 - [20] P. Charbonneau, E. Marinari, G. Parisi, F. Ricci-terseghi, G. Sicuro, F. Zamponi, and M. Mezard, *Spin glass theory and far beyond: replica symmetry breaking after 40 years* (World Scientific, 2023).
 - [21] N. Sourlas, Spin-glass models as error-correcting codes, *Nature* **339**, 693 (1989).
 - [22] N. Sourlas, Spin glasses, error-correcting codes and finite-temperature decoding, *Europhysics Letters* **25**, 159 (1994).
 - [23] M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, 2009).
 - [24] S. Sarao Mannelli, G. Biroli, C. Cammarota, F. Krzakala, and L. Zdeborová, Who is afraid of big bad minima? analysis of gradient-flow in spiked matrix-tensor models, *Advances in neural information processing systems* **32** (2019).
 - [25] E. Agliari, A. Barra, P. Sollich, and L. Zdeborová, Machine learning and statistical physics: theory, inspiration, application, *J. Phys. A* (2020).
 - [26] Y. Bahri, J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli, Statistical mechanics of deep learning, *Annual review of condensed matter physics* **11**, 501 (2020).
 - [27] R. M. May, Will a large complex system be stable?, *Nature* **238**, 413 (1972).
 - [28] S. Diederich and M. Oppen, Replicators with random interactions: A solvable model, *Physical Review A* **39**, 4333 (1989).
 - [29] B. Derrida and L. Peliti, Evolution in a flat fitness landscape, *Bulletin of mathematical biology* **53**, 355 (1991).
 - [30] G. Bunin, Ecological communities with lotka-volterra dynamics, *Physical Review E* **95**, 042414 (2017).
 - [31] A. Agarwala and D. S. Fisher, Adaptive walks on high-dimensional fitness landscapes and seascapes with distance-dependent statistics, *Theoretical population bi-*

- ology **130**, 13 (2019).
- [32] A. Yamamura, H. Mabuchi, and S. Ganguli, Geometric Landscape Annealing as an Optimization Principle Underlying the Coherent Ising Machine, *Physical Review X* **14**, 031054 (2024).
 - [33] D. Sherrington and S. Kirkpatrick, Solvable model of a spin-glass, *Physical review letters* **35**, 1792 (1975).
 - [34] F. Hamze, J. Raymond, C. A. Pattison, K. Biswas, and H. G. Katzgraber, Wishart planted ensemble: A tunably rugged pairwise ising model with a first-order phase transition, *Physical Review E* **101**, 052102 (2020).
 - [35] M. Potters and J.-P. Bouchaud, *A first course in random matrix theory: for physicists, engineers and data scientists* (Cambridge University Press, 2020).
 - [36] G. Toulouse, On the mean field theory of mixed spin glass-ferromagnetic phases, *Journal de Physique Lettres* **41**, 447 (1980).
 - [37] G. Parisi, A sequence of approximated solutions to the sk model for spin glasses, *Journal of Physics A: Mathematical and General* **13**, L115 (1980).
 - [38] M. Talagrand, The parisi formula, *Annals of mathematics* , 221 (2006).
 - [39] A. Montanari, Optimization of the sherrington-kirkpatrick hamiltonian, *SIAM Journal on Computing* **54**, FOCS19 (2025).
 - [40] L. F. Cugliandolo and J. Kurchan, Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model, *Physical Review Letters* **71**, 173 (1993).
 - [41] E. P. Wigner, Random matrices in physics, *SIAM review* **9**, 1 (1967).
 - [42] J. Baik, G. B. Arous, and S. P        , Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices, *The Annals of Probability* **33**, 1643 (2005).
 - [43] S. F. Edwards and R. C. Jones, The eigenvalue spectrum of a large symmetric random matrix, *Journal of Physics A: Mathematical and General* **9**, 1595 (1976).
 - [44] P. Urbani, Field theory for zero temperature soft anharmonic spin glasses in a field, *Journal of Physics A: Mathematical and Theoretical* **55**, 335002 (2022).
 - [45] L. A. Pastur, On the spectrum of random matrices, *Teoreticheskaya i Matematicheskaya Fizika* **10**, 102 (1972).
 - [46] T. Tao, *Topics in random matrix theory*, Vol. 132 (American Mathematical Soc., 2012).
 - [47] E. Bouchbinder, E. Lerner, C. Rainone, P. Urbani, and F. Zamponi, Low-frequency vibrational spectrum of mean-field disordered systems, *Physical Review B* **103**, 174202 (2021).
 - [48] G. Parisi, Infinite number of order parameters for spin-glasses, *Physical Review Letters* **43**, 1754 (1979).
 - [49] M. J. Schmidt and R. Oppermann, Method for replica symmetry breaking at and near $t = 0$ with application to the sherrington-kirkpatrick model, *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* **77**, 061104 (2008).
 - [50] V. Ros, G. Ben Arous, G. Biroli, and C. Cammarota, Complex energy landscapes in spiked-tensor and simple glassy models: Ruggedness, arrangements of local minima, and phase transitions, *Physical Review X* **9**, 011003 (2019).
 - [51] V. Ros and Y. V. Fyodorov, The high-d landscapes paradigm: spin-glasses, and beyond, *arXiv preprint arXiv:2209.07975* (2022).
 - [52] J. Kent-Dobias and J. Kurchan, How to count in hierarchical landscapes: A full solution to mean-field complexity, *Physical Review E* **107**, 064111 (2023).
 - [53] V. Ros, High-dimensional random landscapes: from typical to large deviations, *arXiv preprint arXiv:2502.14084* (2025).
 - [54] J.-M. Aza  s and M. Wschebor, *Level sets and extrema of random processes and fields* (John Wiley & Sons, 2009).
 - [55] A. Annibale, A. Cavagna, I. Giardina, G. Parisi, and E. Trevisan, The role of the beccchi-rouet-stora-tyutin supersymmetry in the calculation of the complexity for the sherrington-kirkpatrick model, *Journal of Physics A: Mathematical and General* **36**, 10937 (2003).
 - [56] A. Annibale, A. Cavagna, I. Giardina, and G. Parisi, Supersymmetric complexity in the sherrington-kirkpatrick model, *Physical Review E* **68**, 061103 (2003).
 - [57] A. Annibale, G. Gualdi, and A. Cavagna, Coexistence of supersymmetric and supersymmetry-breaking states in spherical spin-glasses, *Journal of Physics A: Mathematical and General* **37**, 11311 (2004).
 - [58] A. Cavagna, I. Giardina, and G. Parisi, Cavity method for supersymmetry-breaking spin glasses, *Physical Review B—Condensed Matter and Materials Physics* **71**, 024422 (2005).
 - [59] E. S  li and D. F. Mayers, *An introduction to numerical analysis* (Cambridge university press, 2003).
 - [60] E. Agoritsas, G. Biroli, P. Urbani, and F. Zamponi, Out-of-equilibrium dynamical mean-field equations for the perceptron model, *Journal of Physics A: Mathematical and Theoretical* **51**, 085002 (2018).
 - [61] K. Hukushima and K. Nemoto, Exchange monte carlo method and application to spin glass simulations, *Journal of the Physical Society of Japan* **65**, 1604 (1996).
 - [62] K. Nokura, Spin glass states of the anti-hopfield model, *Journal of Physics A: Mathematical and General* **31**, 7447 (1998).
 - [63] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities., *Proceedings of the national academy of sciences* **79**, 2554 (1982).
 - [64] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Spin-glass models of neural networks, *Physical Review A* **32**, 1007 (1985).
 - [65] G. Folena and P. Urbani, Marginal stability of soft anharmonic mean field spin glasses, *Journal of Statistical Mechanics: Theory and Experiment* **2022**, 053301 (2022).
 - [66] T. Leleu, Y. Yamamoto, S. Utsunomiya, and K. Aihara, Combinatorial optimization using dynamical phase transitions in driven-dissipative systems, *Physical Review E* **95**, 022118 (2017).
 - [67] T. Leleu, Y. Yamamoto, P. L. McMahon, and K. Aihara, Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity, *Physical review letters* **122**, 040607 (2019).
 - [68] F. Ghimenti, L. Berthier, G. Szamel, and F. van Wijland, Sampling Efficiency of Transverse Forces in Dense Liquids, *Physical Review Letters* **131**, 257101 (2023).
 - [69] W. Xu, Z. Shen, Y. Tang, and P. Zhang, Combinatorial optimization with dual mean-field dynamics, *Communications in Theoretical Physics* (2025).
 - [70] P. Urbani, Disordered high-dimensional optimal control, *Journal of Physics A: Mathematical and Theoretical* **54**, 324001 (2021).
 - [71] S. Zhou, K. Wong, J. Wang, D. S. W. Hui, D. Ebler, and J. Sun, Phase analysis of ising machines and their implications on optimization, *arXiv preprint arXiv:2507.08533*

- (2025).
- [72] S. Franz, G. Parisi, P. Urbani, and F. Zamponi, Universal spectrum of normal modes in low-temperature glasses, *Proceedings of the National Academy of Sciences* **112**, 14539 (2015).
 - [73] D. Sherrington, Stability of the sherrington-kirkpatrick solution of a spin glass model: a reply, *Journal of Physics A: Mathematical and General* **11**, L185 (1978).
 - [74] C. De Dominicis and I. Giardina, *Random fields and spin glasses: a field theory approach* (Cambridge University Press, 2006).
 - [75] A. Cavagna, J. P. Garrahan, and I. Giardina, Index distribution of random matrices with an application to disordered systems, *Physical Review B* **61**, 3960 (2000).
 - [76] J. Zinn-Justin, *Quantum field theory and critical phenomena*, Vol. 171 (Oxford university press, 2021).
 - [77] J. Ding and A. Zhou, Eigenvalues of rank-one updated matrices with some applications, *Applied Mathematics Letters* **20**, 1223 (2007).
 - [78] H. Eissfeller and M. Opper, New method for studying the dynamics of disordered spin systems without finite-size effects, *Physical review letters* **68**, 2094 (1992).
 - [79] H. Eissfeller and M. Opper, Mean-field monte carlo approach to the sherrington-kirkpatrick model with asymmetric couplings, *Physical Review E* **50**, 709 (1994).
 - [80] F. Roy, G. Biroli, G. Bunin, and C. Cammarota, Numerical implementation of dynamical mean field theory for disordered systems: Application to the Lotka–Volterra model of ecosystems, *Journal of Physics A: Mathematical and Theoretical* **52**, 484001 (2019).
 - [81] E. A. Novikov, Functionals and the random-force method in turbulence theory, *Sov. Phys. JETP* **20**, 1290 (1965).
 - [82] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Statistical mechanics of neural networks near saturation, *Annals of physics* **173**, 30 (1987).
 - [83] M. S. Bartlett, An inverse matrix adjustment arising in discriminant analysis, *The Annals of Mathematical Statistics* **22**, 107 (1951).