Enhanced Associative Memory, Classification, and Learning with Active Dynamics

Agnish Kumar Behera,[1] Madan Rao,[2] Srikanth Sastry,[3] and Suriyanarayanan Vaikuntanathan[1,4,*]

[1]*Department of Chemistry, University of Chicago, Chicago, Illinois 60637, USA*

[2]*Simons Centre for the Study of Living Machines, National Centre for Biological Sciences-Tata Institute of Fundamental Research, Bangalore, India*

[3]*Theoretical Sciences Unit and School of Advanced Materials, Jawaharlal Nehru Centre For Advanced Scientific Research, Jakkur Campus, 560064 Bengaluru, India*

[4]*The James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA*

(Received 11 July 2023; revised 26 September 2023; accepted 9 October 2023; published 6 December 2023)

Motivated by advances in the field of active matter where nonequilibrium forcing has been shown to activate new assembly pathways, here we study how nonequilibrium driving in prototypical memory formation models can affect their information processing capabilities. Our results reveal that activity can provide a new and surprisingly general way to dramatically improve the memory and information processing performance of the memory-forming systems without the need for additional interactions or changes in connectivity. Nonequilibrium dynamics can allow these systems to have memory capacity, assembly or pattern recognition properties, and learning ability, in excess of their corresponding equilibrium counterparts. Our results demonstrate the generality of the enhancement of memory capacity arising from nonequilibrium, active dynamics when compared to noise sources characteristic of equilibrium dynamics. These results are of significance to a variety of processes that take place under nonequilibrium dynamics, and involve information storage and retrieval, as well as *in silico* learning and memory-forming systems for which nonequilibrium dynamics may provide an approach for modulating memory formation.

DOI: 10.1103/PhysRevX.13.041043

Subject Areas: Complex Systems, Computational Physics, Statistical Physics

I. INTRODUCTION

Biological systems ranging from neuronal circuits in multicellular organisms, to biological circuits responsible for immune memory, display a remarkable array of information storage and retrieval dynamics [1–5]. The olfactory system is involved in storing information about a wide array of smells and retrieving them accurately even with a low signal-to-noise ratio arising from mixed odors [6,7]. The immune system is responsible for storing memory from a previous infection through B-cell populations and using them to elicit responses during new infections [8]. Biological assemblies, for example, those connected to the cytoskeleton, have been shown to possess information processing abilities that allow them to respond in a desired manner to external stimuli [9,10]. All these processes involve some form of information storage across large timescales as well as retrieval of the same, which is inherently dynamic in nature. The foundational paradigms used to understand information storage and retrieval, however, are arguably rooted in the principles of equilibrium statistical mechanics. These principles provide a prescription to understand how interactions between constituent particles (in the case of molecular recognition or assembly problems) or connectivity (in the case of neural network computation problems) determines an overall energetic or free-energetic landscape in the space of a relevant order parameter. In turn, the characteristics of such landscapes are employed to rationalize information storage and retrieval processes. Here, motivated in part by advances in the field of active matter [11–16] where nonequilibrium forcing has been shown to activate new assembly and organization pathways, we study how modulating the dynamics of systems through nonequilibrium activity can affect their information processing capabilities.

We investigate prototypical models of associative memory [17,18], pattern recognition and assembly [19], and neural networks [20]. Our results reveal that activity can provide a new and surprisingly general way to dramatically improve memory and information processing performance. In particular, nonequilibrium dynamics can allow these systems to have memory capacity, assembly or pattern recognition properties, and learning ability, in excess of their corresponding equilibrium counterparts.

First, we motivate our approach using an adaptation of a particle moving in a double-well potential. Results obtained

*svaikunt@uchicago.edu

from this minimal model show how strategies based on modifying the dynamics with active terms may result broadly in improved memory storage and recall. Next, we demonstrate our results using a version of the Hopfield model [18], a paradigmatic model for associative memory [21–23]. By studying a version of the Hopfield model with nonequilibrium dynamics, we show how the associative memory characteristics of a system may be enhanced far beyond the bounds placed by equilibrium dynamics. Our model of nonequilibrium dynamics is motivated by choices commonly made in the field of active matter [24,25].

The observed connection between nonequilibrium forcing and the promotion of specific ordered states in active matter problems and the general nature of our analytical results suggests that our conclusions may apply more broadly. Exploring this possibility, we demonstrate how a model elastic material endowed with associative memory-like properties [19] can be made to store configurations far in excess of what is allowed under equilibrium dynamics. Insight gleaned from studies of this elastic material may be directly relevant for the design of various bioinspired materials and molecular recognition processes. It is also relevant for understanding how desired structures may be self-assembled even in systems with highly promiscuous interactions. Our work shows how nonequilibrium activity can provide a new route to control the properties of such systems [19].

As a final example we show how nonequilibrium activity may also help improve the performance of neural networks as illustrated in Fig. 1. We do so by considering a simple yet prototypical "phase retrieval" problem in which a neural network attempts to reconstruct a signal from a set of measurements [20]. The effectiveness of this network may

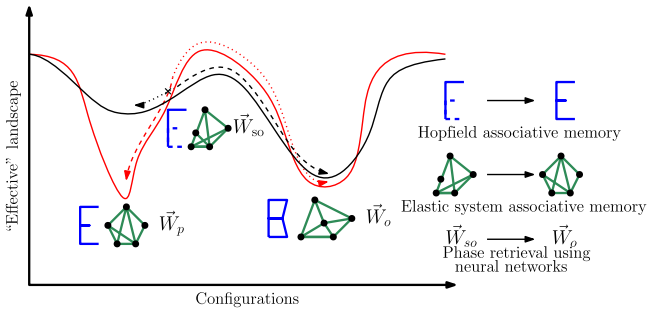

FIG. 1. Activity can be used to improve the information processing properties of a large class of systems, including an associative memory exhibiting spin-based Hopfield network which stores patterns, an elastic material with promiscuous interactions which can store certain configurations, and finally neural network used for the prototypical phase retrieval problem. In all the three cases, our work shows how nonequilibrium activity can (in some cases dramatically) improve the desired information processing capacity of the system. Viewed in the commonly used landscape caricature, our work suggests that activity (red) can generically improve the stability of the desired states or outcomes while suppressing undesirable outcomes.

be measured by a loss function that is minimized when the signal is retrieved with complete accuracy. The landscape of the loss function (as the weights of the neural network are tuned) can, however, be highly nonconvex leading to imperfect signal recovery. Intuition from the above described examples suggests that activity may provide a route to deepen (or make more convex) landscape around desired memories or self-assembled structures. Here we show that adding nonequilibrium activity to the dynamical equations of motion dictating how the weights of the neural network change can dramatically improve the performance of the network. Like in the previous examples, nonequilibrium activity provides a new route to modulate the "loss" landscape of our model neural network and improve its performance without the need for additional layers or connectivity. Our results are consistent with recent work in Ref. [20] where the accuracy of the phase retrieval algorithm was improved by introducing a persistent mini-batch sampling procedure. Taken together, our work opens up general strategies for enhancing memory capacity and classification properties in nonequilibrium materials.

II. PRIMER: INSIGHTS FROM A MINIMAL MODEL

In order to elucidate how activity may provide a general route to modulating memories, we first discuss a highly simplified minimal model of a single particle in a double-well potential, $V(x) = -(a/2)x^2 + (b/4)x^4$, with over-damped Brownian motion dynamics,

$$\dot{x} = -\frac{\partial V}{\partial x} + \eta(t), \tag{1}$$

where $\eta(t)$ is a white noise source with zero mean and variance $\langle \eta(t)\eta(t')\rangle = 2T\delta(t-t')$, and T is a temperature scale. The kinetics of escape rates across the energy barrier follow an Arrhenius-like form, with barrier crossing rates that are exponentially suppressed as the barrier height is increased or the temperature of the system is decreased. Replacing the white noise source with an exponentially correlated noise source, $\xi(t)$ with zero mean and with correlations $\langle \xi(t)\xi(t')\rangle = T/\tau \exp(-|t-t'|/\tau)$ where τ is a timescale, breaks the condition of detailed balance and drives the system away from equilibrium. Indeed in this case, fluctuations in the exponentially correlated noise source are not dissipated through a corresponding friction source. The effect of such nonequilibrium activity on barrier crossing times can be inferred following the works of Ref. [26] in the limit $\tau \ll 1$. Adapting the calculations in Ref. [26], we show in Appendix A how nonequilibrium activity can lead to an increase in the escape time of a system over the barriers (Fig. 2). Specifically, as we describe in Appendix A, the ratio of the average escape times in and out of equilibrium, τ_p and τ_a, respectively, satisfies

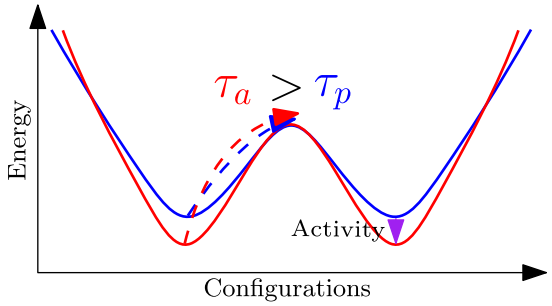

FIG. 2. A minimal landscape model for improved associative memory recall due to activity. The addition of activity can modulate the landscape in specific ways leading to longer residence times in portions of the landscape corresponding to associative memory. Here τ_a and τ_p denote the average escape times in the active and passive case, respectively.

$$\ln\left[\frac{\tau_a}{\tau_p}\right] = \frac{1}{2}\tau a. \tag{2}$$

We note that, in general, the effect of activity on barrier crossing times cannot be simply described by an effective potential or temperature and that it depends on the details of the nonequilibrium forcing and the underlying potential [26,27].

This minimal calculation motivates us to consider whether similar phenomena might happen for highly disordered systems when subjected to detailed balance breaking colored noise. In the context of associative memory and neural network systems, where memories and solutions are stored in attractor basins, such effects might result in improved memory storage and information processing capacity.

III. SPHERICAL HOPFIELD MODEL WITH NONEQUILIBRIUM DYNAMICS

The Hopfield model is an interacting spin system with a Hamiltonian which is fully connected; i.e., every spin is connected to every other spin. We work with a version of the Hopfield model [18] where the spins are continuous and obey the constraint $\sum_{i=1}^{N}\sigma_i^2 = N$, where σ_i denotes the value of the ith spin and N is the total number of spins in the system:

$$\mathcal{H}_0 = -\frac{1}{2}\sum_{i\neq j} J_{ij}\sigma_i\sigma_j - \frac{u_0}{4}\sum_{ijkl} J_{ijkl}\sigma_i\sigma_j\sigma_k\sigma_l, \tag{3}$$

$$J_{ij} = \frac{1}{N}\sum_{\mu}\xi_i^\mu\xi_j^\mu, \qquad J_{ijkl} = \frac{1}{N^3}\sum_{\mu}\xi_i^\mu\xi_j^\mu\xi_k^\mu\xi_l^\mu. \tag{4}$$

Henceforth, repeated indices would mean a summation over that index. The pattern variables are denoted as ξ_i^μ, where μ denotes the pattern index and i denotes the site index. The N components of the patterns are drawn from independent identically distributed (IID) normal distributions, $\xi_i^\mu \sim \mathcal{N}(0,1)$. The pattern loading of the system is denoted by α, which is the ratio of the number of stored patterns (P) to the total number of spins (N) in the system. In the model, the coupling strengths between spins J_{ij} and J_{ijkl} depend on the patterns through the Hebbian rule [28]. Quartic terms are included in the Hamiltonian following Ref. [18], where it was demonstrated that such higher-order terms are a necessary requirement for associative memory-like properties in a system with continuous spins. The spins evolve according to the following equations of motion,

$$\partial_t\sigma_i = -\mu(t)\sigma_i(t) - \frac{\delta\mathcal{H}_0(\sigma)}{\delta\sigma_i(t)} + \eta_i(t). \tag{5}$$

Here, $\mu(t)$ is the Lagrange multiplier which ensures the normalization of the spins and $\delta\mathcal{H}_0/\delta\sigma$ is the relaxational term. Finally, $\eta(t)$ models the effect of various thermal and athermal fluctuations,

$$\eta_i(t) = \eta_{w,i}(t) + \eta_{a,i}(t), \tag{6}$$

$$\langle\eta_{w,i}(t)\rangle = 0 = \langle\eta_{a,i}(t)\rangle \quad \forall\ i,t, \tag{7}$$

$$\langle\eta_{w,i}(t)\eta_{w,j}(t')\rangle = 2T_p\delta_{ij}\delta(t-t'), \tag{8}$$

$$\langle\eta_{a,i}(t)\eta_{a,j}(t')\rangle = \frac{T_a}{\tau}\delta_{ij}\exp\left(-\frac{|t-t'|}{\tau}\right), \tag{9}$$

where the thermal fluctuations (thermal noise) are modeled using a δ function correlated white noise $\vec{\eta_w}$, and the exponentially correlated $\vec{\eta_a}$ is a so-called *active* or a colored noise source. As mentioned earlier, the addition of colored noise into the system without a corresponding change in dissipation breaks detailed balance as has been demonstrated in Refs. [11,29,30]. Thus, even in the limit $\tau \ll 1$, the system will still be forced away from equilibrium. In what follows, we report results from our numerical and analytical calculations in terms of two parameters, namely, the passive fraction, $f \equiv T_p/(T_p+T_a)$, and the so-called effective temperature, $T_{\text{eff}} \equiv T_p + T_a$ [24,25]. Note that such an effective temperature prescription does not hold in general—indeed, the effective temperature T_{eff} is most appropriate in the low persistence time limit [24,25]—T_{eff} provides a convenient way to characterize the strength of the nonequilibrium forcing [31].

$$\lim_{\tau\to 0}\langle\eta_{a,i}(t)\eta_{a,j}(t')\rangle = 2T_a\delta_{ij}\delta(t-t'). \tag{10}$$

We perform numerical simulations in which a system with $N = 200$ spins is evolved forward in time using Eq. (5). In these numerical simulations, we probe the ability of the system to retrieve a stored pattern by initializing the spin system in configurations close to those

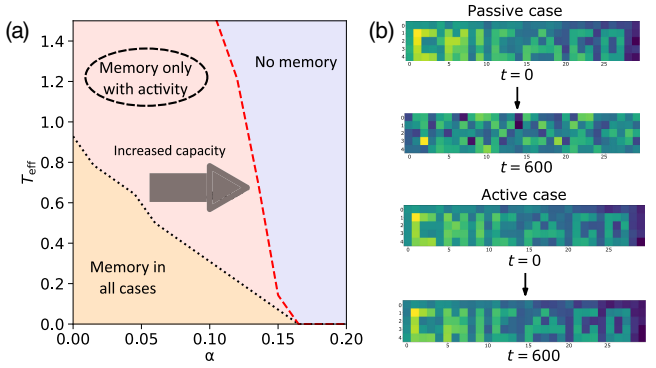

FIG. 3. Memory retrieval in an associative memory model with equilibrium and nonequilibrium dynamics. (a) A phase diagram demarcating regions exhibiting associative memory. The orange region represents the parameter space where retrieval is possible in both the passive (equilibrium) and active (out-of-equilibrium) case, the red region is where the active case shows retrieval whereas the passive system does not, and in the blue region memory retrieval is lost altogether. The phase boundaries were obtained using a mean field technique described in Appendix D. Parameters used are $\tau = 5$ and the passive temperature $T_p = 0$ for active simulations. (b) As an example, we show the retrieval dynamics of a pattern stored at $\alpha = 0.1$, $T_{\text{eff}} = 0.4$ in the phase diagram. For illustrative purposes, the pattern is arranged such that it spells out Chicago. Simulations with equilibrium dynamics, when initialized in the vicinity of this pattern, fail to retrieve it. On the other hand, this pattern is successfully retrieved with nonequilibrium dynamics at the same effective temperature. The "Chicago" pattern simulation was performed with $N = 150$ spins and 10 patterns encoded in the interactions.

corresponding to the stored memory states. We choose an initial overlap of 0.9–0.93. This is in accordance with Ref. [17] where similar error limits are considered. Retrieval is considered successful if the dynamics are able to recover the full stored pattern as it reaches its steady state. Quantitatively, the retrieval ability is measured by tracking the steady state value of the overlap of the final spin configuration of the system with the pattern it was initialized near [17,21]. For a particular pattern μ, this overlap can be measured as $m^\mu \equiv (1/N)\sum_i \xi_i^\mu \sigma_i$. Following Refs. [18,32] we look at the condensation of a single pattern [33]. Figure 3 describes results from numerical simulations using Eq. (5).

The phase portrait in Fig. 3(a) shows the regimes under which the different dynamics are able to successfully retrieve the stored patterns. To construct this phase portrait, multiple numerical simulations were performed at various values of α and T_{eff} for both equilibrium and nonequilibrium dynamics. Here, equilibrium dynamics implies T_a is set to zero and $T_{\text{eff}} = T_p$ while nonequilibrium dynamics implies T_p is set to zero and $T_{\text{eff}} = T_a$. The phase boundaries themselves were obtained using a mean field technique described in Sec. IV B. As can be clearly seen in Fig. 3(a), the ability of the spin system to retrieve patterns is markedly increased due to nonequilibrium driving. Specifically, the orange shaded region demarcates the parameter combinations under which associative memory or memory retrieval is possible under equilibrium dynamics [21]. Our nonequilibrium simulations demonstrate memory retrieval in the red shaded region in addition to the orange shaded region. The nonequilibrium dynamics chosen for the simulations in Fig. 3(b) were performed with $T_{\text{eff}} = 0.4, \tau = 5$. Qualitatively similar results can be obtained for other choices of the nonequilibrium parameters.

In the subsequent sections, we explore the theoretical basis of this improved associative memory due to nonequilibrium dynamics. First, in Sec. IV A we perform a perturbative analysis in the limit of small persistence time τ. In this limit the nonequilibrium distribution function can be approximated using Boltzmann statistics with an effective Hamiltonian and an effective temperature. Our calculations show how the effective Hamiltonian supports enhanced interactions between spins as well as new higher-order interactions at first order in τ. A replica calculation reveals that as a consequence of these enhanced interactions—these emerge due to the nonequilibrium forcing—the spin system possesses enhanced associative memory recall. Then in Sec. IV B we perform a Martin-Siggia-Rose (MSR) calculation which describes our system in a mean field limit and provides an analytically tractable route to quantify how the robustness of pattern retrieval increases away from equilibrium.

IV. RATIONALIZING IMPROVED ASSOCIATIVE MEMORY UNDER NONEQUILIBRIUM DYNAMICS

The equilibrium Hopfield model can be solved analytically using the replica method [21]. Since our model is out of equilibrium, a direct application of the replica method is not possible. Through unified active noise approximation (UCNA) [34] and following recent work by Ref. [24], we show that our active system can be described using an *effective* Hamiltonian and a new *effective* temperature. We then use the standard replica technique with this effective Hamiltonian to show how the addition of activity enhances associative memory recall. In Sec. IV B we derive an exact mean field set of evolution equations for our active system using the Martin-Siggia-Rose generating functional formalism [18,32,35] and further illustrate how memory recall is improved by the introduction of activity.

A. Effective interactions due to nonequilibrium forcing provide a mechanism for improved associative memory recall

UCNA suggests [11,34] suggests that at small τ our nonequilibrium system can be described by an effective Hamiltonian with an effective temperature. As outlined in Appendix B, we show that at first order in τ, the

perturbed Hamiltonian and the effective temperature ($T_{\rm eff}$) are given by

$$\mathcal{H} = \mathcal{H}_0 + \frac{\tau T_a}{T_{\rm eff}}\left(\frac{1}{2}|\nabla_\sigma \mathcal{H}_0|^2 - T_{\rm eff}\nabla_\sigma^2 \mathcal{H}_0\right), \quad (11)$$

$$T_{\rm eff} = T_p + T_a. \quad (12)$$

Substituting the Hopfield Hamiltonian, Eq. (3), into Eq. (11) yields

$$\mathcal{H} = -\frac{\tilde{v}}{2}J_{ij}\sigma_i\sigma_j - \frac{\tilde{u}}{4}J_{ijkl}\sigma_i\sigma_j\sigma_k\sigma_l + \frac{k}{6}J_{ijklmn}\sigma_i\sigma_j\sigma_k\sigma_l\sigma_m\sigma_n + O(1/N)\text{ terms}, \quad (13)$$

$$\tilde{v} \equiv 1 + \frac{\tau T_a(2\tilde{\mu} - s + 2\alpha)}{T_{\rm eff}},$$
$$\tilde{u} \equiv u\left[1 + \frac{2\tau T_a(2\tilde{\mu} - 2s + \alpha)}{T_{\rm eff}}\right], \quad (14)$$

$$k = \frac{3\tau T_a u^2 s}{T_{\rm eff}},\quad s = \langle(\xi_i^\mu)^2\rangle,\quad J_{ijklmn} = \frac{\xi_i^\mu\xi_j^\mu\xi_k^\mu\xi_l^\mu\xi_m^\mu\xi_n^\mu}{N^5}, \quad (15)$$

where, as outlined in Appendix B, $\tilde{\mu}$ is approximated as the average value of the Lagrange multiplier, $\tilde{\mu} \approx \langle\mu(t)\rangle$. Numerically we find that for values of $\tilde{\mu}$ close to the observed average in the simulations, $\langle\mu(t)\rangle \approx 2$, the results from the replica approach (described below) lead to qualitatively similar results to those obtained from explicit numerical calculations [36]. We see that activity enhances the strength of the quadratic and quartic terms. It also generates higher-order coupling between spins; for instance, at first order it gives rise to a sextic term. The enhanced strength of the quadratic and quartic terms in Eq. (11) may potentially offer a route to stabilize associative memory recall by modifying the effective free-energy landscape supporting pattern retrieval [Fig. 3(a)]. We now confirm this picture by performing a replica calculation on this effective Hamiltonian. Our calculation identifies regimes where the nonequilibrium model can exhibit associative memory whereas an equivalent equilibrium model fails to have associative memory properties. The details of the calculation are provided in Appendix C and follow standard calculations performed on the equilibrium Hopfield model [22].

The replica calculation provides an estimate of the order parameter m that characterizes the degree of polarization of the system toward one of the stored patterns:

$$m^\mu = \frac{1}{N}\xi_i^\mu\sigma_i. \quad (16)$$

A nonzero value of this order parameter, along with conditions on an additional order parameter that characterizes the spin-glass nature of the system, identify regimes in which memory retrieval is possible. In Fig. 4 we plot the results from our replica calculation and demarcate regimes in which memory retrieval is possible both in and out of equilibrium. The specific calculation in Fig. 4 was performed with the passive fraction, f held a constant at $f = 0.2$ and $\tau = 0.05$ and shows that nonequilibrium dynamics permit memory retrieval even in regions with higher effective temperatures. While the presence of sixth-order terms in Eq. (13) might lead to the formation of spurious energy minimas, our replica calculation reveals that the system can robustly retrieve the original patterns stored in it.

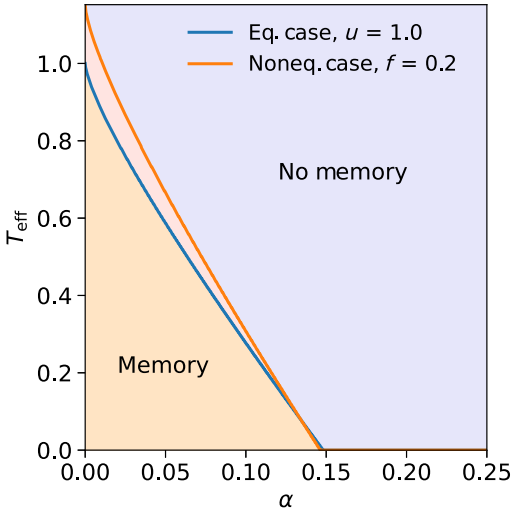

FIG. 4. Phase diagram of the spherical Hopfield system using replica calculation in equilibrium and nonequilibrium regimes. The region to the left of the lines labeled as "Memory" represents the retrieval phase. With a passive fraction, $f = 0.2$, $\tau = 0.05$, $s = 1.0$, and Lagrange multiplier $\mu = 2.4$, and taking into account the first-order corrections, the phase diagram shows an enhancement in memory; i.e., the overlap parameter m is nonzero over a larger region. Note that the replica calculation can be performed only in the small τ limit as explained in the main text. Consequently, the enlargement of the region with associative memory is relatively low in this theoretical phase diagram. More substantial improvement in the associative memory properties occurs in regimes with larger τ.

While this is only an approximate result as the perturbation is valid only for small τ and we ignore the higher-order corrections to the effective Hamiltonian, our results nonetheless show how the strengthening of stabilizing interactions in the effective Hamiltonian due to nonequilibrium activity results in improved associative memory recall.

B. Mean field approach using the Martin-Siggia-Rose generating functional Lagrangian to explain improvement in nonequilibrium associative memory recall

The results of the previous section analytically show how associative memory can be improved in the low persistence time limit. In order to probe the effects of

activity in other limits, we use the Martin-Siggia-Rose generating functional approach to write down and study mean field coarse-grained equations for the evolution of the order parameter $m(t)$.

In the MSR approach, the statistics of the evolution path of the system are captured by the disorder averaged generating functional [Eq. (17)]:

$$\overline{Z[\psi,\theta]} = \int D\sigma D\eta P(\eta) \overline{\exp\left[i\sum_{i=1}^{N}\int dt \psi_i(t)\sigma_i(t)\right]} \times \overline{\prod_{i=1}^{N}\delta\left(\partial_t\sigma_i(t)+\mu(t)\sigma_i(t)+\frac{\delta\mathcal{H}(\sigma)}{\delta\sigma_i(t)}-\theta_i(t)-\eta_i(t)\right)}. \tag{17}$$

Here $\theta_i(t)$ is the external field on a spin which is eventually set to 0. When we perform the disorder average over Z in Eq. (17) we decouple the spins but couple different times. This is similar to the coupling of different replicas in the replica approach. The entire procedure is described in Appendix D. We obtain an equation of motion for the decoupled spins following the procedure. Using the decoupled equation of motion for single spins, Eq. (D35), we can easily write down the equations of motion for the macroscopic variable m [Eq. (D46)].

These equations can be simulated numerically and memory retrieval phase diagrams can be constructed based on the steady state values of m. In Fig. 5(a) we plot the evolution of m as a function of time, for various values of the passive fraction $f = T_p/T_{\text{eff}}$ (at fixed α, T_{eff}, and τ). These trends clearly demonstrate how nonequilibrium forcing improves memory recall. For a fixed T_{eff}, memory can be markedly improved if the fluctuations are mainly due to detailed balance violating noise.

To more comprehensively characterize the improved memory recall due to nonequilibrium dynamics in our system, we use the MSR mean field framework to obtain steady state values of the memory order parameter m at various values of T_{eff}, f, α, and τ in Fig. 5(b). We observe that at constant $T_{\text{eff}} = 1.0$ and $f = 0$, the capacity of the system increases with increasing the persistence time as shown in Fig. 5(b). Note that at $T_{\text{eff}} = 1.0$ at equilibrium (i.e., $\tau = 0$), this Hopfield model has no retrieval capacity [inset in Fig. 5(b)]. We have also provided a comparison between the full numerical simulations by evolving Eqs. (5) for a system with $N = 200$ spins and the MSR equations in Fig. 11(b). Finally, analogous to the replica calculation enhancement in Fig. 4, we also computed phase boundaries between regimes with memory and regimes without memory in both active and passive cases. The resultant $T_{\text{eff}} - \alpha$ phase boundaries from MSR calculation is presented in Fig. 3(a). This also shows how activity can vastly increase the conditions under which memory retrieval is possible. Together, these results further show how increasing the nonequilibrium forcing may enhance the associative memory recall of a system.

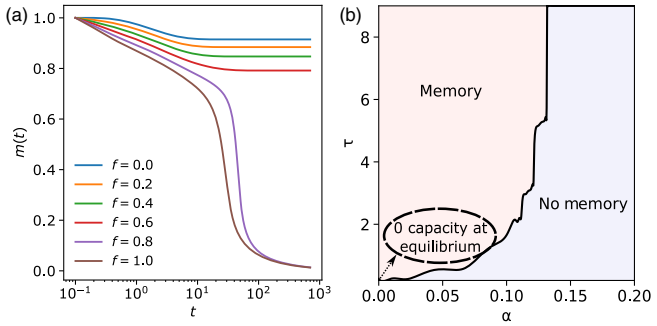

FIG. 5. Enhanced retrieval region from MSR calculations. (a) Evolution of m as a function of t from the MSR formalism [Eqs. (D46), (D48), (D47)]. Here, $\tau = 5$ and $T_{\text{eff}} = 1.0$ for the simulations. (b) Retrieval is improved when the amount of activity is increased by tuning the persistence time τ. At $\tau = 0$ and $T_{\text{eff}} = 1$, the system is effectively at equilibrium, and as per the equilibrium phase diagram of our spherical Hopfield model, the system has no memory retrieval at this temperature (inset). When $\tau \neq 0$, the system is out of equilibrium and we start observing memory phase. The parameter f was held fixed at $f = 0$ for this phase plot.

V. ELASTIC MATERIALS WITH MEMORY: ENHANCED ASSOCIATIVE MEMORY RECALL DUE TO NONEQUILIBRIUM ACTIVITY

In Ref. [19] it was shown that nonlinear spring networks can be engineered to posses associative memory properties. Specifically, the interactions between a system of N particles can be programmed such that specific desired spatial configurations sit at metastable minima in the energy landscape. Akin to the Hopfield model, these configurations can be recovered if the system of particles is initialized in its vicinity. The learning rule used to encode the memory configurations has features similar to the Hebb rule used in the Hopfield case with the important distinction being that the interactions are all local. Specifically, the process of encoding memory proceeds in the following way. The particles are arranged so that their positions overlap with the first target two-dimensional configuration to be stored as memory, $\{(\vec{x}^{\mu},\vec{y}^{\mu})\}$, where the label μ identifies the pattern. Nonlinear springs are then "grown" between neighboring nodes such that their rest length is equal to the distance between the nodes in the pattern. Mathematically, this can be represented as

$$k_{ij}^{\mu} = k_0\Theta(R_c - |\vec{r}_i^{\mu} - \vec{r}_j^{\mu}|), \tag{18}$$

$$l_{ij}^{\mu} = |\vec{r}_i^{\mu} - \vec{r}_j^{\mu}|, \tag{19}$$

where R_c is the cutoff for two nodes to be considered neighbors, Θ is the Heaviside function, $\vec{r}_i^{\,\mu}$ is the position of particle "i" in pattern μ, $k_{\vec{r}_i^{\,\mu},\vec{r}_j^{\,\mu}}$ and $l_{\vec{r}_i^{\,\mu},\vec{r}_j^{\,\mu}}$ are the spring constant and rest length of the spring connecting particle "i" with particle "j" in pattern μ. The energy due to the nonlinear springs is given by

$$\mathcal{H} = \frac{1}{2}\sum_{\mu,i,j} k_{ij}^{\mu} \frac{(s_{ij}^{\mu})^2}{[\sigma^2 + (s_{ij}^{\mu})^2]^{1-(1/2)\xi}}, \tag{20}$$

$$s_{ij}^{\mu} = ||\vec{r}_i - \vec{r}_j| - l_{ij}^{\mu}|, \tag{21}$$

where $\xi \neq 2$ sets the extent of nonlinearity and σ is a length scale. Analogous to the Hopfield model, this network also has a critical number of patterns that the system can remember after which its memory capacity degrades rapidly.

Here, we investigate if activity can provide a route to improve the memory capacity of this system. The system is evolved according to the following equations of motion:

$$\partial_t r_i = -\frac{\partial \mathcal{H}}{\partial r_i} + \eta_i. \tag{22}$$

Here the noise η has the same statistics as Eqs. (6)–(9). Following Ref. [19] all the configurations considered were two dimensional. The simulations were carried out in a two-dimensional box with linear dimension L. The initial state is a stored pattern configuration, i.e., $\vec{r}_0 = \vec{r}^{\mu}$. We calculate the average local displacement (d_L^{μ}) of a given state from the specific pattern configuration it was initialized near, in the following fashion. For every node in the system, we calculate the average displacement of all other nodes connected to it in the specific pattern. Then we average over all these local displacements for all the nodes. Mathematically, $d_L^{\mu} \equiv (1/N_S^{\mu})\sum_{ij} s_{ij}^{\mu}$, where N_S^{μ} is the number of springs in pattern μ. The parameter σ was used to specify our numerical memory retrieval criteria. Specifically, if $d_L^{\mu} \leq \sigma$, we say that the system can retrieve the patterns. For this system Ref. [19] discusses how the effectiveness of memory retrieval is determined by the radius of attraction around the pattern configurations, i.e., the maximum displacement from the minima for which the system still relaxes to the desired configuration. Reference [19] shows that the radius of attraction around the patterns is comparable to the length scale set by σ. We choose σ to be 1% the box length following the choices made in Ref. [19]. As outlined in Ref. [19], σ and ξ can also set a so-called threshold energy scale for the stability of stored patterns. Hence, modulating the value of σ may also potentially affect the maximum effective temperature at which memory storage is possible. In the following we simply focus on numerical simulations with the aforementioned value of σ and leave a more comprehensive exploration to future work [37].

This system is not analytically tractable due to the presence of the nonlinearities in springs. Thus we resorted to using numerical methods for obtaining a phase diagram. Specifically, we used a fixed-time step second-order Runge-Kutta integrator for numerically evolving the system using Eq. (22). The zero temperature capacity for the set of parameters mentioned in Fig. 6 is $\alpha \approx 0.15$. This capacity decreases as the temperature is increased. Similar to the spherical Hopfield model, we find that the capacity of the system to store patterns—at the same effective temperature—is increased if the passive noise source is replaced with an active noise source (Fig. 6).

Following the construction in Ref. [19], we also constructed a test to study the classification ability of the elastic network under passive and active dynamics. For this, we used data from the MNIST dataset, a set of 60 000 training and 10 000 test images of handwritten digits. The training dataset is transformed in the following fashion: Each 28×28 pixel image is first truncated to a 20×20 image with 400 pixels. Then for every label, the average over all its training images is taken. The grayscale image is then interpreted as a configuration of 400 nodes in a 1D box of length $L = 10$ with the node displacements given as the pixel values scaled between 0 and 10. Thus we have a set of configurations, one for every label, which we call "pattern" in accordance with the terminology being used for elastic networks. Using the above described procedure, we store these patterns into the 1D network of nodes with a critical radius of $R_c = L$.

For this work, we store the configurations from training dataset corresponding to labels "0" and "1." After this we initialize the system with one of the test images corresponding to labels 0 or 1, truncated and transformed as before. In Fig. 7 we characterize the classification ability by initiating the system at a test configuration corresponding to the 0 label with the dynamics prescribed by Eq. (22). When evolved with passive dynamics ($\tau = 0$), a test pattern diverges away from the stored 0 label state (Fig. 7). Under active dynamics (at the same effective temperature), however, the designed attractor around the 0 state is more stably accessed (Fig. 7). In effect, the active system manages to classify the test pattern as a 0 label correctly while the passive variant does not.

VI. IMPROVING SIGNAL RECOVERY AND PERFORMANCE OF A PHASE RETRIEVAL NEURAL NETWORK USING ACTIVE DYNAMICS

In this section, we show that the addition of active dynamics may also improve the performance of a model neural network. We illustrate this by considering the prototypical problem of phase retrieval where the task is one of recovering a signal from a set of measurements. This is achieved using a single-layer perceptron network.

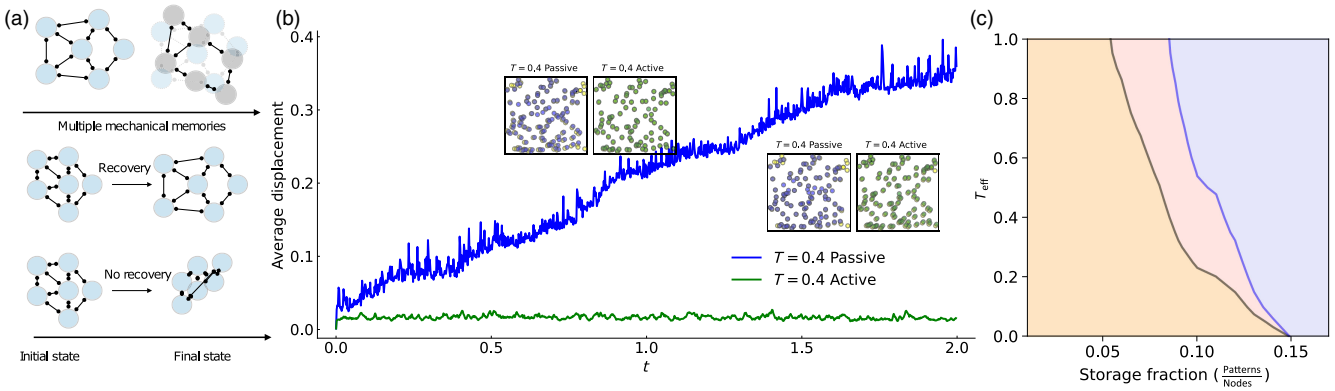

FIG. 6. Enhancement of associative memory recall due to nonequilibrium activity in elastic materials. (a) Memory storage and retrieval in elastic associative memory system. (b) Characteristic retrieval dynamics for patterns in the active and passive cases. The system configuration is initialized near one of the patterns stored in the system. Two simulations are then performed, one in which the system is evolved at $T_{\text{eff}} = 0.04$ with passive noise ($T_p = 0.4$, $T_a = 0$) and another in which it is evolved at the same effective temperature but with active noise ($T_p = 0$, $T_a = 0.4$). Shown is a single trial for a specific pattern. The y axis denotes the absolute value of average displacement of the nodes. The insets show the configurations at two time instants, $t = 1.0$ and $t = 2.0$ for the active and passive cases. The yellow circles represent the stored patterns. The displacement order parameter tracks the loss of overlap between the stored patterns and the instantaneous configuration of the system. The active dynamics are able to stabilize the system close to the stored configuration, whereas in the passive case about 10 nodes suffer large displacements leading to no recovery. (c) A phase portrait describing memory retrieval in the elastic associative memory system. The yellow region denotes recovery in both passive and active cases. The light red region is where active case alone shows recovery. This phase diagram was built by averaging over 10 different systems with 100 nodes (particles), $\xi = 0.4$, box length $(L) = 10.0$, and $\sigma = 0.1$. Four trials were run per pattern at a every given temperature. Data obtained from numerical simulations were used to generate the smoothened phase boundaries using a 2D Gaussian interpolation.

Specifically, we consider, M IID Gaussian measurements of dimension N, $\vec{\xi}^{\mu}$ for $\mu \in 1 \ldots M$ from which we want to recover an N-dimensional weight vector $\vec{w}$ such that it provides the best approximation for the labels $y^{\mu} = |(1/\sqrt{N})\vec{\xi}^{\mu} \cdot \vec{w}_0|$. Here $\vec{w}_0$ is the target weight or "signal" that corresponds to the right measurements. In the typical implementation of the phase retrieval problem, the weights $\vec{w}$ are evolved such that a loss function given by $\mathcal{L} = \sum_{\mu=1}^{M} \mathcal{L}^{\mu} = \frac{1}{4}\sum_{\mu=1}^{M}[(\hat{y}^{\mu})^2 - (y^{\mu})^2]^2$, where $\hat{y}^{\mu} = |(1/\sqrt{N})\vec{\xi}^{\mu} \cdot \vec{w}|$, is minimized.

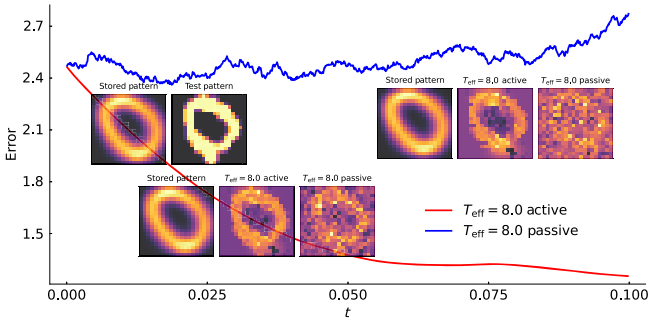

FIG. 7. Enhancement of MNIST classification due to nonequilibrium activity in elastic materials. Unlike standard MNIST classification which takes place using a deterministic scheme with randomness arising only from minibatch sampling, here we consider a case where we have an additional randomness arising due to noise. After the training images have been averaged and stored into elastic networks, the test images are evolved either with passive white noise or active correlated noise. At higher temperatures, with active noise, the test image reduces its "distance" from the stored image or "error" thus improving the parameter used for classification. The parameters used for the simulation are $T_{\text{eff}} = 8.0$, $\tau = 0.05$, $\xi = 0.5$.

Dynamical equations with gradient descent are typically used to evolve the weights to their optimal values. One of the most popular implementations is the stochastic gradient descent (SGD) algorithm, where a minibatch is sampled from the available data and a gradient descent is performed,

$$w_i(t+\eta) - w_i(t) = -\frac{1}{b}\eta\left[\sum_{\mu=1}^{M} s_{\mu}(t)\partial_{w_i}\mathcal{L} + brw_i(t)\right], \tag{23}$$

where η is the learning rate, r is a ridge regularizer, $P[s_{\mu}(t) = 1] = b$ denotes the probability associated with including a certain point from the data in the minibatch, and s_{μ} is an indicator for the examples which are a part of the minibatch. The weights $\vec{w}$ hence evolve on a landscape dictated by the loss function, $\mathcal{L}$. This landscape (i.e., the dependency of the loss function on the weights $\vec{w}$) can be highly nonconvex leading to instances where the SGD algorithm gets trapped in regimes far from the globally optimal point. This leads to errors in recovery and degrades the performance of the network.

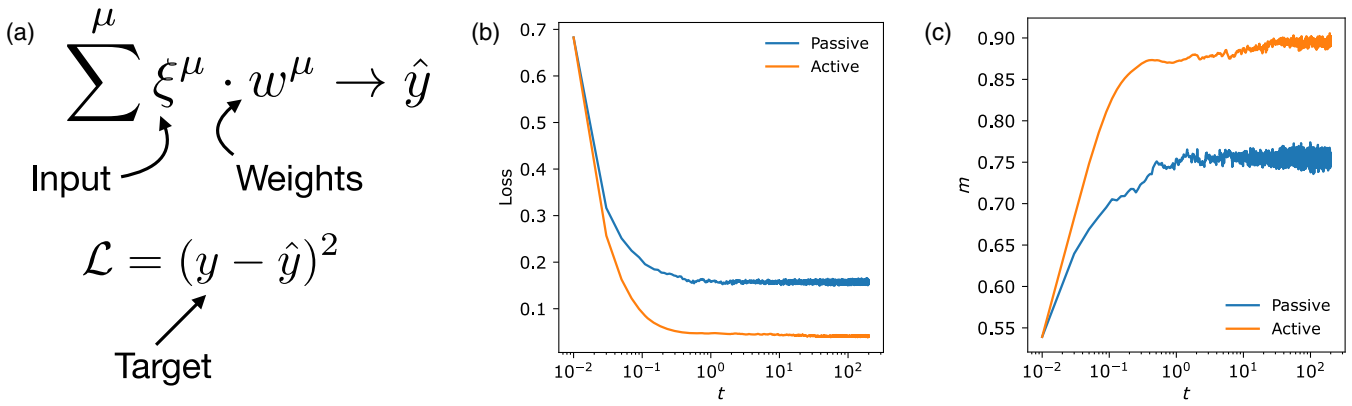

FIG. 8. (a) A schematic outline of the phase retrieval problem. Here we attempt to construct the weight vector $\vec{\hat{w}}$ which is the best approximation of the true weight $\vec{w}_0$ which provides labels to a set of measurements $\vec{\xi}^{\mu}$. Once initialized, we let the system evolve through Eq. (24) and we track the loss $\mathcal{L}$ and the "magnetization" $m = (1/N)\vec{\hat{w}} \cdot \vec{w}_0$. This is shown in (b) and (c). The active case outperforms the passive case. All the numerical simulations were carried out with $N = 100$, $\alpha = 3$. Parameters for (b) and (c) are $T_{\text{eff}} = 1.0$, $\tau = 1.0$, $m_0 = 0.5$.

Inspired by recent work by Mignacco and co-workers [20,38] where they reported that adding persistence associated with resampling data, i.e., data sampled in one minibatch is retained with a certain probability for the subsequent minibatch, can provide a route to improve the performance of the phase retrieval problem, and following the results from the previous sections, we investigated if the performance of such systems could be improved by modulating the dynamics. To make progress, we move away from the SGD equations of motion in Eq. (23) and consider instead a simpler Langevin-equation-like motion for the evolution of the weights:

$$\partial_t w_i(t) = -\partial_{w_i}\mathcal{L} - r w_i(t) + \chi(t). \tag{24}$$

Similar to the previous examples, we compared the performance of networks with weights evolving according to this equation for both passive and active (or persistent) choices of noise $\chi(t)$. Here r is the Lagrange multiplier that ensures the normalization of weights to 1. In our case we have also tried a simple ridge regularizer. Both approaches give similar results. In order to find $\vec{\hat{w}}$, we first perform a warm initialization of $\vec{\hat{w}}$, i.e., $\vec{\hat{w}}_{\text{init}} = c(m_0\vec{w}_0 + \vec{\zeta})$, where $\vec{w}_0$ is the true weight vector, m_0 denotes how close to the true weight we want the initial guess, $\vec{\zeta}$ is a vector of random numbers drawn from IID standard normal distribution, and c is just a normalization constant such that $\vec{\hat{w}} \cdot \vec{\hat{w}} = N$, where N is the dimension of each measurement $\vec{\xi}^{\mu}$. We then evolve the system and track the loss $\mathcal{L}$ and the "magnetization," $m = (1/N)\vec{\hat{w}} \cdot \vec{w}_0$.

One important distinction between the previous two models and this model is the fact that previously we were navigating the energy landscape by adding activity to the "physical" degrees of freedom, i.e., spins in the Hopfield

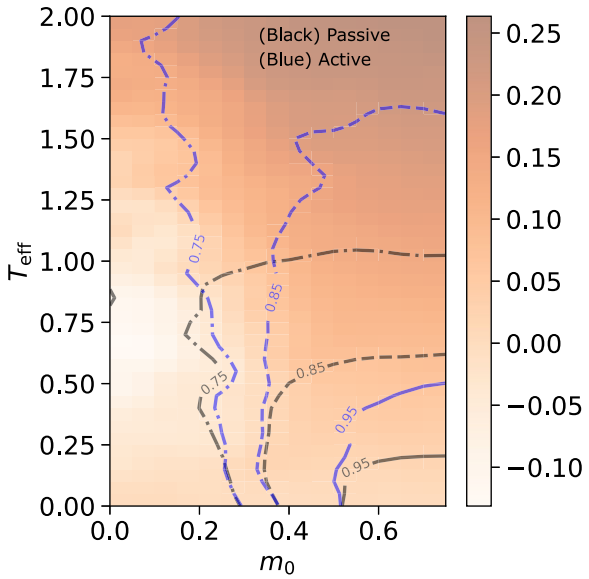

FIG. 9. Effectiveness of phase retrieval under passive and active dynamics. This figure shows the difference between the final overlaps of the predicted signal with the true signal in the active and the passive case. The active case performs better than the passive case in most regions, and thus $m_{\text{active}} - m_{\text{passive}}$ is positive. The blue contours correspond to values of overlap in the active case and the black ones correspond to the passive case. Note that even if we start with a small warm initialization (e.g., $m_0 = 0.2$), the active case can reach overlaps up to 0.75 whereas the passive case cannot. Indeed, the "0.75" active contour is present at $m_0 = 0.2$ even for higher temperatures. For a particular value of T_{eff} and m_0, the overlap was averaged over 20 different systems with 10 trials run for every system. The active and passive simulations are run for the same amount of time, i.e., from $t = 0$ to 200 with $dt = 0.001$ using spherical normalization. The simulation is stopped at $t = 200$ since we observed no significant changes in the overlaps thereafter. An average is taken over the last 50 time steps.

model and the particle positions in the elastic material, whereas here we will navigate the loss landscape by adding activity to the "learning" degrees of freedom.

As shown in Fig. 8, there is a marked improvement in the performance of the phase retrieval network. The signal recovery is markedly higher and the loss is lowered. In Fig. 9, we compare the performances of the active and passive phase retrieval dynamics for various different initial conditions. In general, for extremely small values of m_0 the system gets lost and is unable to recover the signal in both the active and passive cases. For slightly higher m_0 ($m_0 \approx 0.2$), however, active case starts to outperforms the zero temperature as well as passive case. This trend is maintained for larger m_0. On the whole, active dynamics provide a more effective strategy to find close to optimal solutions for the phase retrieval problem. Following intuition from the previous examples, we expect this improved performance to be due to activity deepening, or making more convex, the loss landscape in the vicinity of the optimal solution. While we have focused on a highly simplified and idealized learning problem, our collection of results suggests that similar advances may be possible in more complex deep neural networks.

VII. DISCUSSION AND CONCLUSIONS

Our work here shows that the storage capacity of a system which uses a Hopfield-like strategy to store memory as well as that of a system with promiscuous interactions can be increased significantly compared to equilibrium dynamics through the introduction of activity into the system. It also provides a way to improve classification and phase retrieval at the same "effective" temperature through the introduction of activity. While we have explored the improved associative memory capacity using specific numerical and analytical tools, it may become possible to develop broader principles for improved associative memory [39]. While these qualitative results suggest that nonequilibrium activity may provide a general route to enhance the information processing abilities of a material, we note that there are important caveats. Associative memory recall away from equilibrium is dictated by a balance between two competing factors. On the one hand, as our theoretical analysis suggests, nonequilibrium forcing may generate deeper minima in the effective landscape. On the other hand, nonequilibrium forcing also has the potential to generate many spurious minima and hasten the transition to a glassy regime where associative memory properties are lost. Future work will explore these trade-offs more comprehensively. Note also that our findings in the previous two sections (spring based associative memory and phase retrieval) are mainly based on numerics and we have explored limited ranges of various parameters such as τ. Future work will also explore these systems more rigorously.

The stored memory patterns or configurations in the associative memory models considered in this work are thought to be point attractors of the dynamics of the model. Associative memory phenomena can also be achieved with stored memory patterns resembling continuous attractors. The self-assembly of desired patterns from building blocks with multifarious interactions [40] and associative memory models of spatial learning in the place cells of the hippocampus provide two illustrative examples of this class [41]. We expect that the memory capacity of such models can also be increased using active noise sources. Further, in this class of systems and materials, it may be possible to robustly achieve regimes where the addition of activity leads to an increase in capacity beyond the capacity of the zero temperature passive system. Indeed, in the context of associative memory models of spatial learning in place cells, Ref. [41] shows how the capacity may be already improved in the presence of thermal noise. Active noise sources can potentially lead to a further improvement in the memory capacity. We have also restricted our attention to models with two body connectivity. p-spin associative memory models [42], with $p \gg 2$, should also show phenomenology similar to that discovered here when driven by active sources.

Memory in physical systems can occur in a multitude of forms [43]. There is also precedent from driven systems with transient memories wherein noise can lead to better memory formation [44,45]. In such systems, memory of the driving amplitudes is encoded in the self-organization of particles into configurations that eliminate collisions within a range determined by the training amplitudes, and thus memory of all but the largest amplitude is transient. The presence of noise enables retention of memory of additional, smaller, amplitudes of training. The associative memory systems we explore here are different as the memories are encoded as point attractors in a free-energy landscape. Future work can explore how different noise sources can help transient memory as well as other systems.

Our work might also have broader implications for the design of artificial neural networks. Indeed, as discussed above, recent work in Ref. [20] has shown that learning tasks using stochastic gradient descent with a certain persistence time for minibatch sampling can lead to better performance. It may be possible to connect the persistence time in such methods to the persistence time introduced in the context of nonequilibrium activity here. Such a connection might help us understand how the performance of deep neural networks can be improved [38,46].

Finally, in the limit of small persistence time, it may become possible to express the change in the effective energetic landscape in terms of the rate of work done by the active forces. Similar insights have proven useful in the context of active matter systems [25,47] to establish connections between dissipation and assembly or organization. Such connections may suggest how nonequilibrium

forcing might provide a general mechanism to enhance memory recall [48,49].

ACKNOWLEDGMENTS

Theory development, computational modeling, and conception was supported by funds from DOE BES Grant No. DE-SC0019765 to S. V. A. K. B. acknowledges support from a fellowship from the Department of Chemistry at the University of Chicago. Earlier parts of the computational modeling and analytical calculations were supported by funds from ARO under Grant No. W911NF- 720 14-1-0403 to A. K. B. M. R. acknowledges support from the Department of Atomic Energy (India), under Project No. RTI4006, and the Simons Foundation (Grant No. 287975). M. R. and S. S. (JBR/2020/000015) acknowledge SERB, DST (India) for JC Bose Fellowships. We gratefully acknowledge the helpful discussions we had with Thomas Witten, Sidney Nagel, Arvind Murugan, Francesca Mignacco, Pierfrancesco Urbani, and Giampaolo Folena.

APPENDIX A: INSIGHTS FROM A MINIMAL MODEL

Consider a particle in 1D executing dynamics on a double-well potential given by $V(x) = -(a/2)x^2 + (b/4)x^4$, in the presence of active noise,

$$\dot{x} = f(x) + \xi(t), \tag{A1}$$

$$\langle \xi(t) \rangle = 0, \tag{A2}$$

$$\langle \xi(t)\xi(t') \rangle = \frac{T}{\tau} \exp\left(-\frac{|t-t'|}{\tau}\right), \tag{A3}$$

$$f(x) = -\frac{\partial V}{\partial x} = ax - bx^3. \tag{A4}$$

Note that we are implicitly working in units where the friction constant associated with the overdamped dynamics has been set to unity. This choice implicitly sets the timescales in the equations that follow. Our calculations below are based on Ref. [26]. This is a special case of the more general case discussed in Ref. [26]. In case the white noise is present along with colored noise, the temperature gets replaced by an effective temperature. Under these conditions the average escape time of the particle from a well is given as

$$\tau_a = [|(1-\tau f')|_{x_b}|(1-\tau f')|_{x_t}]^{1/2} \exp\left(\frac{\Delta\Phi(\tau)}{T}\right). \tag{A5}$$

$$\Delta\Phi(\tau) = -\int_{x_b}^{x^t} f(y)[1-\tau f(y)']dy, \tag{A6}$$

where x_b is the bottom of the well and x_t is the top of the well. Using the identity $1 - x \approx e^{-x}$ for small x in the above expression for when $\tau \ll 1$, we obtain

$$\tau_a = \tau_0 \exp\left(\frac{\Phi_{\text{eff}}(\tau)}{T}\right), \tag{A7}$$

$$\Phi_{\text{eff}} \equiv \{V(x) + \tau_a[\partial_x V(x)]^2\}|_{x_b}^{x_t} + \frac{1}{2}\tau T[\partial_x^2 V(x)|_{x_b} + \partial_x^2 V(x)|_{x_t}]. \tag{A8}$$

Note that $\partial_x V(x)|_{x_b} = 0$ and $\partial_x V(x)|_{x_t} = 0$ since the force is zero at the top and bottom of the well. We can now compare the average escape times in the equilibrium $\tau = 0$ and the nonequilibrium $\tau \neq 0$ limits, τ_p and τ_a respectively,

$$\ln\left[\frac{\tau_a}{\tau_p}\right] = \frac{1}{2}\tau a. \tag{A9}$$

APPENDIX B: EFFECTIVE HAMILTONIAN IN PRESENCE OF ACTIVE NOISE

The equations of motion for the spins and the active field are given by

$$\Gamma_0^{-1}\frac{\partial \sigma_i}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \sigma_i} + \chi_i(t) + \xi_i(t), \tag{B1}$$

$$\tau\frac{\partial \xi_i}{\partial t} = -\xi_i + \eta_i(t), \tag{B2}$$

$$\langle \chi_i(t) \rangle = 0 = \langle \eta_i(t) \rangle, \qquad \langle \chi_i(t)\chi_j(t') \rangle = 2T\delta_{ij}\delta(t-t'), \qquad \langle \eta_i(t)\eta_j(t') \rangle = 2T_a\delta_{ij}\delta(t-t'). \tag{B3}$$

Here, Γ_0^{-1} sets the microscopic processing time and for simplicity it is taken to be 1. $\chi(t)$ is the δ correlated white noise, $\xi(t)$ is the active colored noise, and τ is the persistence time. Thus the Fokker-Planck equation is given by

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \sigma_i}\left(h_i\rho - \xi_i\rho + T\frac{\partial \rho}{\partial \sigma_i}\right) + \frac{1}{\tau}\frac{\partial}{\partial \xi_i}\left(\xi_i\rho + \frac{T_a}{\tau}\frac{\partial \rho}{\partial \xi_i}\right), \tag{B4}$$

$$h_i = \frac{\delta \mathcal{H}}{\delta \sigma_i}. \tag{B5}$$

It is important to note that here the probability distribution is a function of both the spin degrees and the active degrees of freedom as well. In order to recover the marginal distribution with respect to only the spins, we need to integrate out the active degrees of freedom. We reproduce the procedure outlined in Ref. [24] and extend it for the general case. First let us define the quantity R:

$$R_{(n)i_1 i_2 \ldots i_n} = \int \boldsymbol{D\xi}\, \xi_{i_1}\xi_{i_2}\ldots\xi_{i_n}\rho. \tag{B6}$$

The properties of R are derived through integration by parts and by setting the boundary term to 0. Now from the Fokker-Planck equation in (B4) we can write down a hierarchy of equations for $R_{(k)}$, where the equation for R_k depends on $R_{(k+1)}$. We can truncate the hierarchy at any k of our choosing.

$$\partial_i J_{(1)i} - \partial_i R_{(1)i} = 0, \tag{B7}$$

$$\partial_i J_{(2)ij} - \partial_i R_{(2)ij} - \frac{R_{(1)j}}{\tau} = 0, \tag{B8}$$

$$\partial_i J_{(3)ijk} - \partial_i R_{(3)ijk} - \frac{2R_{(2)jk}}{\tau} + \frac{2T_a}{\tau^2} S_{(2)jk} = 0, \tag{B9}$$

where

$$\delta_{jk} R_{(0)} \equiv S_{(2)jk}, \quad \delta_{ij} R_{(1)k} + \delta_{ik} R_{(1)j} + \delta_{jk} R_{(1)i} \equiv S_{(3)ijk},$$
$$h_{i_1} R_{(n-1)i_2 \ldots i_n} + T\partial_{i_1} R_{(n-1)i_2 \ldots i_n} \equiv J_{(n)i_1 \ldots i_n}. \tag{B10}$$

As a generalization, $S_{(n)i_1 \ldots i_n}$ is defined as the combinatorial sum of $\binom{n}{2}$ terms of the form $\delta_{ij} R_{(n-2)i_1 \ldots i_{n-2}}$.

Now to ease the readability, we introduce a few more definitions. Indices implicitly exist for $R_{(n)}$ and $S_{(n+1)}$ for $n \geq 1$.

$$R_{(n)i_1 i_2 \ldots i_n} \equiv R_{(n)}, \qquad S_{(n)i_1 i_2 \ldots i_n} \equiv S_{(n)}. \tag{B11}$$

Continuing the recursion from Eq. (B9), we obtain

$$\partial_i J_{(n)} - \partial_i R_{(n)} - \frac{(n-1)R_{(n-1)}}{\tau} + \frac{2T_a}{\tau^2} S_{(n-1)} = 0. \tag{B12}$$

For example, we want to truncate at $R_{(2)}$, we compare the terms at $O(\tau)$ in Eq. (B9), i.e., $R_{(2)} = (2T_a/\tau)S_{(2)}$, and substitute it in the previous equation and continue this up to Eq. (B7). Generalizing this, after some algebra, we can write the following:

$$\partial_i[(h_i + T_{\text{eff}})R_0] = 2T_a \sum_{n=1}^{\infty} \frac{(-1)^{n+1}\tau^n}{(n+2)!} \partial^{n+2} S_{n+2} + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}\tau^n}{n!} \partial^{n+1} J_{n+1}, \tag{B13}$$

$$T_{\text{eff}} = T + T_a, \tag{B14}$$

$$\partial^n \equiv \partial_{i_1}\partial_{i_2} \ldots \partial_{i_n}, \qquad i_1 \neq i_2 \cdots \neq i_n, \tag{B15}$$

$$\partial_i \partial_i \equiv \partial_i^2. \tag{B16}$$

Now we can simplify this expression further:

$$\partial^n S_n \equiv \partial_{i_1}\partial_{i_2} \ldots \partial_{i_n} S_{n i_1 i_2 .. i_n} \tag{B17}$$

$$= \binom{n}{2} \partial_{i_1}^2 \partial_{i_3}\partial_{i_4} \ldots \partial_{i_n} R_{(n-2)i_3 i_4 \ldots i_n}$$

$$= \binom{n}{2} \partial_i^2 \partial^{n-2} R_{(n-2)}, \tag{B18}$$

$$\partial^n J_n \equiv \partial_{i_1}\partial_{i_2} \ldots \partial_{i_n} J_{n i_1 i_2 .. i_n} \tag{B19}$$

$$= \partial_{i_n}\partial_{i_{n-1}} \ldots \partial_{i_2}\partial_{i_1} h_{i_1} R_{(n-1)i_2 i_3 \ldots i_n} + T\partial_{i_1} R_{(n-1)i_2 i_3 \ldots i_n} \tag{B20}$$

$$= \partial^{n-1}\partial_i (h_i R_n) + T\partial^{n-1}\partial_i^2 R_{(n-1).} \tag{B21}$$

Substituting these results in Eq. (B13) and after some algebra, we obtain

$$\partial_i[(h_i + T_{\text{eff}}\partial_i)R_0] = -\sum_{n=1}^{\infty} \frac{(-\tau)^n}{n!} [T_{\text{eff}} \partial_i^2 \partial^n R_n + \partial^n \partial_i (h_i R_n)]. \tag{B22}$$

Using this relation we can express R_0 as $R_0 = R_0^0 + \tau R_0^1 + \tau^2 R_0^2 + \cdots$, where R_0^1 is the first-order correction in τ to R_0, R_0^2 the second-order correction, and so on. Now we will use the notation R_n^k to denote the kth-order correction to R_n. At the zeroth-order correction for R_0, we have

$$\partial_i[(h_i + T_{\text{eff}}\partial_i)R_0^0] = 0 \tag{B23}$$

$$\Rightarrow R_0^0 = \exp\left(-\frac{H}{T_{\text{eff}}}\right). \tag{B24}$$

At first-order correction, we have

$$\partial_i[(h_i + T_{\text{eff}}\partial_i)R_0^1] = [T_{\text{eff}}\partial_i^2 \partial_j R_{1j}^0 + \partial_j \partial_i (h_i R_{1j})] - \frac{\tau}{2}[T_{\text{eff}}\partial_i^2 \partial_j \partial_k R_{2jk}^0 + \partial_j \partial_k \partial_i (h_i R_{2jk}^{-1})]. \tag{B25}$$

From Eq. (B9) we have $R_{2jk}^{-1} = (T_a/\tau)S_{2jk} = (T_a/\tau)\delta_{jk} R_0^0$, and from Eq. (B8) we have $R_{1j}^0 = -T_a \partial_k S_{2kj} = -T_a \partial_j R_0^0$. Substituting these in the equation above and after some algebra, we obtain

$$\partial_i\left[\exp\left(\frac{H}{T_{\text{eff}}}\right)R_0^1\right] = \frac{T_a}{T_{\text{eff}}}\left(h_{ijj} - \frac{1}{T_{\text{eff}}}h_{ij}h_j\right) = \frac{T_a}{T_{\text{eff}}}\partial_i\left(h_{jj} - \frac{1}{2T_{\text{eff}}}|h_j|^2\right). \quad \text{(B26)}$$

Thus we can now express R_0 including the first correction term and from there we can compute the effective Hamiltonian.

$$R_0 = R_0^0\left[1 + \frac{\tau T_a}{T_{\text{eff}}}\left(h_{jj} - \frac{1}{2T_{\text{eff}}}|h_j|^2\right)\right] + O(\tau^2), \quad \text{(B27)}$$

$$R_0 = \exp\left(-\frac{H}{T_{\text{eff}}}\right)\exp\left[\frac{\tau T_a}{T_{\text{eff}}}\left(h_{jj} - \frac{1}{2T_{\text{eff}}}|h_j|^2\right)\right], \quad \text{(B28)}$$

$$R_0 = \exp\left(-\frac{H_{\text{eff}}}{T_{\text{eff}}}\right),$$
$$H_{\text{eff}} = H - \tau T_a\left(h_{jj} - \frac{1}{2T_{\text{eff}}}|h_j|^2\right). \quad \text{(B29)}$$

This is the same expression as derived in Ref. [24].

Now we turn back to the Hopfield Hamiltonian, and using this expression for the effective Hamiltonian, write down the new terms which arise due to activity. The effective Hamiltonian procedure cannot be readily carried out in the presence of a Lagrange multiplier. Hence, we include in the Hamiltonian a term that simulates the presence of a Lagrange multiplier:

$$\mathcal{H}_0 = \mathcal{H}_0 + \frac{\kappa}{4}(\sigma_i^2 - N)^2. \quad \text{(B30)}$$

In practice, setting $\kappa \gg 1$ will impose a spherical constraint. Simulations or calculations carried out with this extra term in the Hamiltonian should mimic the presence of a Lagrange multiplier. We now carry out the procedure detailed above for this Hamiltonian and find the following effective Hamiltonian. We first consider the term h_{ii}:

$$h_{ii} = \sum_{i=1}^{N}\kappa\partial_i\left[\left(\sum_j \sigma_j^2 - N\right)\sigma_i\right] - \frac{3u}{N^3}\sum_{\mu,i,k,l}\xi_i^\mu\xi_i^\mu\xi_k^\mu\xi_l^\mu\sigma_k\sigma_l. \quad \text{(B31)}$$

We begin by considering the first term. Recall that the force on each spin due to the extra field we have added is $f_c = -\kappa(\sum_j \sigma_j^2 - N)\sigma_i$. This has a form close to the force that would have been exerted by a Lagrange multiplier $\mu(t)$, $f_l = -\mu(t)\sigma_i$. Numerically, we find that $\langle\mu(t)\rangle \equiv \tilde{\mu} \approx 2$ in regimes with recovery. Given this, in the limit $\kappa \gg 1$ where our constrain force is expected to mimic a Lagrange multiplier, we can reasonably expect $\kappa(\sum_j \sigma_j^2 - N) \approx \tilde{\mu} \approx 2$. Numerical simulations bear out this expectation. In the replica calculations that will follow—performed in the limit of $\kappa \gg 1$—we will replace $\kappa(\sum_j \sigma_j^2 - N)$ with $\tilde{\mu} \approx 2$. With this replacement, the first term in the above equation is simply a constant and can be ignored.

$$h_{ii} = rN - \frac{3u}{N}\sum_\mu\sum_{i=1}^{N}(\xi_i^\mu)^2(m^\mu)^2, \quad \text{(B32)}$$

where $r \equiv \sum_j \sigma_j^2 - N$. The second term is substantially simplified in the limit where only one of the pattern is condensed, i.e., $m^\mu = m^\mu\delta_{\mu,\nu}$, where ν is the condensed pattern. Putting this back, we see that it is O(1), thus not extensive.

Now we look at $|h_i|^2$. After some algebra and taking into consideration that only one pattern has condensed, we obtain

$$|h_i|^2 = (r\sigma_i - J_{ij}\sigma_j + J_{ii}\sigma_i - J_{ijkl}\sigma_j\sigma_k\sigma_l)^2 \quad \text{(B33)}$$

$$= (m^\nu)^2\left[\sum_{j=1}^{N}(\xi_j^\mu)^2 - 2rN - 2\alpha\right] + 2u(m^\nu)^4\left[\sum_{i=1}^{N}(\xi^\nu)^2 - rN - \alpha\right] + u^2(m^\nu)^6\sum_{i=1}^{N}(\xi_i^\nu)^2; \quad \text{(B34)}$$

here $r \equiv \kappa(\sum_j \sigma_j^2 - N) \approx \tilde{\mu}$ following the reasoning given above. Note that the pattern variables ξ_i^μ are IID normal random variables. Thus $\sum_{i=1}^{N}(\xi_i^\nu)^2$ form a χ-squared distribution with $N \to \infty$ degrees of freedom which essentially becomes a Gaussian distribution with mean N and variance $2N$. Thus $\sum_{i=1}^{N}(\xi_i^\nu)^2 \sim O(N)$. With these approximations, and simplifications, we recover the effective Hamiltonian written down in the main text.

APPENDIX C: REPLICA CALCULATION

In the following calculations we follow Refs. [32,50]. As detailed in the previous appendix, the effective Hamiltonian calculation for our Hopfield-like network in the presence of active noise leads to the following energy function which dictates the evolution of the spins:

$$\mathcal{H}(\sigma) = \frac{1}{2}\tilde{\mu}\sigma_i^2 - \frac{v}{2}J_{ij}\sigma_i\sigma_j - \frac{u}{4}J_{ijkl}\sigma_i\sigma_j\sigma_k\sigma_l + \frac{k}{6}J_{ijklmn}\sigma_i\sigma_j\sigma_k\sigma_l\sigma_m\sigma_n + O(1/N) \text{ terms}, \quad \text{(C1)}$$

$$v = 1 + \frac{\tau T_a}{T_{\text{eff}}}(2\tilde{\mu} - 1), \qquad u = u\left[1 + \frac{2\tau T_a}{T_{\text{eff}}}(2\tilde{\mu} - 1)\right], \quad \text{(C2)}$$

$$k = \frac{3\tau T_a}{T_{\text{eff}}} u^2, \qquad J_{ijklmn} = \frac{1}{N^5} \xi_i^\mu \xi_j^\mu \xi_k^\mu \xi_l^\mu \xi_m^\mu \xi_n^\mu, \tag{C3}$$

where $\tilde{\mu}$ is related to the average force exerted by the constraint potential as explained in the previous appendix. From numerical calculations with $\kappa = 0.25$ and $N = 200$ with 10 stored patterns, we find $\tilde{\mu} \approx 2.4$. For other values of number of stored patterns, the value of μ ranges from 1.9 to 2.5.

The partition function is given by

$$Z(\beta) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} d\sigma_i\right] \delta\left(\sum_{i=1}^{N} \sigma_i^2 - N\right) \exp[-\beta \mathcal{H}(\sigma)]. \tag{C4}$$

We use $\overline{A}$ to denote the quenched average of macroscopic variable A. Now using the replica trick,

$$\beta N f = -\overline{\ln Z(\beta)} = -\lim_{n\to 0} \frac{1}{n} [\overline{Z^n(\beta) - 1}] = \lim_{n\to 0} \frac{1}{n} \ln \overline{Z^n}, \tag{C5}$$

$$Z^n(\beta) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^{N} \prod_{\gamma=1}^{n} d\sigma_i^\gamma\right] \prod_{\gamma=1}^{n} \delta\left(\sum_{i=1}^{N} (\sigma_i^\gamma)^2 - N\right) \times \exp\left[-\beta \sum_{\gamma=1}^{n} \mathcal{H}(\sigma_i^\gamma)\right] \tag{C6}$$

where γ is the replica index and goes from $1, \ldots, n$. Henceforth, repeated indices imply summation. Only in certain cases will the summation be explicitly denoted.

$$\begin{aligned} \overline{Z^n(\beta)} = \int D\sigma D\xi \exp\Bigg[\beta\Bigg(&\frac{v}{2N} \xi_i^\mu \xi_j^\mu \sigma_i^\gamma \sigma_j^\gamma \\ &+ \frac{u}{4N^3} \xi_i^\mu \xi_j^\mu \xi_k^\mu \xi_l^\mu \sigma_i^\gamma \sigma_j^\gamma \sigma_k^\gamma \sigma_l^\gamma \\ &- \frac{k}{6N^5} \xi_i^\mu \xi_j^\mu \xi_k^\mu \xi_l^\mu \xi_m^\mu \xi_n^\mu \sigma_i^\gamma \sigma_j^\gamma \sigma_k^\gamma \sigma_l^\gamma \sigma_m^\gamma \sigma_n^\gamma \\ &- \frac{v}{2N} \sum_{\mu,i,\gamma} (\xi_i^\mu)^2 (\sigma_i^\gamma)^2\Bigg)\Bigg] \\ &\times \prod_{\gamma=1}^{n} \delta((\sigma_i^\gamma)^2 - N) \exp\left[-\frac{(\xi_i^\mu)^2}{2}\right]. \end{aligned} \tag{C7}$$

Now we introduce the overlap parameter $m^{\mu\gamma} = (1/N)\xi_i^\mu \sigma_i^\gamma$. For simplicity, we assume that only one pattern, pattern no. 1, is condensed, i.e., $m^\mu \sim O(1)$ for $\mu = 1$ and $O(1/N)$ for $\mu \neq 1$. Using this we can separate the partition function into a condensed and a noncondensed part. We introduce this macroscopic variable through a δ function and express the δ function as an exponential. Henceforth it is implicitly assumed that m denotes the overlap with only pattern 1. This yields

$$\int Dm \delta(N m^\gamma - \xi_i^1 \sigma_i^\gamma) = 1, \tag{C8}$$

$$\int Dm D\tilde{m} \exp[i\tilde{m}^\gamma (N m^\gamma - \xi_i^1 \sigma_i^\gamma)] = 1, \tag{C9}$$

$$\overline{Z^n(\beta)} = \int D\sigma D\xi Dm D\tilde{m} \exp(U) \prod \delta[(\sigma^\gamma)^2 - N], \tag{C10}$$

$$\begin{aligned} U = {} & N\beta[v(m^\gamma)^2 + u(m^\gamma)^4 - k(m^\gamma)^6] \\ & + \frac{\beta v}{2N} \sum_{\mu\neq 1,i,j,\gamma} \xi_i^\mu \xi_j^\mu \sigma_i^\gamma \sigma_j^\gamma - \frac{(\xi^\mu)^2}{2} \\ & + i\tilde{m}^\gamma (N m^\gamma - \xi_i^1 \sigma_i) + O(1). \end{aligned} \tag{C11}$$

We ignore the higher-order contributions of $\sum_{\mu\neq 1 i_1..i_p} \xi_{i_1}^\mu .. \xi_{i_p}^\mu \sigma_{i_1}^\gamma .. \sigma_{i_p}^\gamma$ for $p > 2$ as they scale as $O(N^{1-p/2})$ [42].

Now we first carry out integration over the quenched disorder ξ. The relevant terms in U are

$$\begin{aligned} & -\frac{1}{2} \sum_{\mu\neq 1,i,j} \xi_i^\gamma \left(\delta_{ij} - \frac{\beta v}{N} \sum_\gamma \sigma_i^\gamma \sigma_j^\gamma\right) \xi_j^\mu - \sum_i \frac{(\xi_i^1)^2}{2} \\ & - i \sum_\gamma \tilde{m}^\gamma \sum_i \xi_i^1 \sigma_i^\gamma. \end{aligned} \tag{C12}$$

The $\mu = 1$ (here μ refers to the pattern index) integral yields

$$\begin{aligned} & \exp\left[-\frac{1}{2} \sum_{\gamma,\kappa} \tilde{m}^\gamma \left(\sum_i \sigma_i^\gamma \sigma_i^\kappa\right) \tilde{m}^\kappa\right] \\ & = \int DQ \delta\left(N Q^{\gamma\kappa} - \sum_i \sigma_i^\gamma \sigma_i^\kappa\right) \exp\left[-\frac{N}{2} \sum_{\gamma,\kappa} \tilde{m}^\gamma Q^{\gamma\kappa} \tilde{m}^\kappa\right], \end{aligned} \tag{C13}$$

$$Q^{\gamma\kappa} = \delta^{\gamma\kappa} + (1 - \delta^{\gamma\kappa}) q^{\gamma\kappa}, \tag{C14}$$

$$q^{\gamma\kappa} = \frac{1}{N} \sum_i \sigma_i^\gamma \sigma_i^\kappa \quad \forall \ \gamma \neq \kappa. \tag{C15}$$

Thus the correlation of spins between different replicas arises naturally just as in other spin-glass systems. The $\mu \neq 1$ integral yields

$$\prod_{\mu=2}^{\alpha N} \left[\det\left(\delta_{ij} - \frac{\beta v}{N} \sum_\gamma \sigma_i^\gamma \sigma_j^\gamma\right)\right]^{-1/2} = [\det(\mathbb{1} - \beta v \mathbf{Q})]^{-\alpha N/2}. \tag{C16}$$

This step can be carried out through the Weinstein-Aronsazn theorem. After all this, we have

$$\overline{Z^n(\beta)} = \int D\sigma \int D\tilde{m} \int_{-\infty}^{\infty} Dm \int_{-\infty}^{\infty} Dq \int_{-i\infty}^{i\infty} D\tilde{\lambda} \int_{-i\infty}^{i\infty} D\hat{q} \times [\det(\mathbb{1} - \beta v \mathbf{Q})]^{-\alpha N/2} e^{U}. \tag{C17}$$

$$U = N\beta\left[\frac{v(m^\gamma)^2}{2} + \frac{u(m^\gamma)^4}{4} + \frac{k(m^\gamma)^6}{6}\right] - \frac{N}{2}\tilde{m}^\gamma Q^{\gamma\kappa}\tilde{m}^\kappa + iNm^\gamma\tilde{m}^\gamma + \sum_{\gamma,\kappa,\gamma\neq\kappa}\tilde{q}^{\gamma\kappa}(Nq^{\gamma\kappa} - \sigma_i^\gamma\sigma_i^\kappa) + \sum_\gamma \tilde{\lambda}^\gamma[N - (\sigma_i^\gamma)^2]. \tag{C18}$$

Here we have expressed the δ functions for the constraints and that for q as exponentials. Now we integrate over $\tilde{m}$. The relevant terms are

$$-\frac{N}{2}\tilde{m}^\gamma Q^{\gamma\kappa}\tilde{m}^\kappa + iNm^\gamma\tilde{m}^\gamma. \tag{C19}$$

Integration yields

$$\det(Q)^{-1/2}\exp\left(-\frac{1}{2}m^\gamma Q^{\gamma\kappa}m^\kappa\right). \tag{C20}$$

Finally, we carry out integration over the spin variables σ. The relevant terms are

$$-\sum_{\gamma,\kappa,i}\sigma_i^\gamma\tilde{Q}^{\gamma\kappa}\sigma_i^\kappa \tag{C21}$$

$$\tilde{Q}^{\gamma\kappa} = \tilde{\lambda}^\gamma\delta^{\gamma\kappa} + (1-\delta^{\gamma\kappa})\tilde{q}^{\gamma\kappa}. \tag{C22}$$

Integration yields

$$[\det\tilde{Q}]^{-N/2}. \tag{C23}$$

Combining everything, we have

$$\beta f = \lim_{N\to\infty}\lim_{n\to 0}\frac{-1}{n}\ln\frac{1}{N}\left[\int_{-\infty}^{\infty} Dm\int_{-\infty}^{\infty} Dq\int_{-i\infty}^{i\infty} D\tilde{\lambda}\int_{-i\infty}^{i\infty} D\hat{q}e^{-Ng}\right], \tag{C24}$$

$$Dm = \prod_{\gamma=1}^{n} dm_\gamma, \quad Dq = \prod_{\gamma\neq\kappa} dq_{\gamma\kappa}, \quad D\tilde{\lambda} = \prod_{\gamma=1}^{n} d\tilde{\lambda}_\gamma, \quad D\hat{q} = \prod_{\gamma\neq\kappa} d\hat{q}_{\gamma\neq\kappa},$$

$$g = -\frac{\beta v}{2}\sum_{\gamma=1}^{n} m_\gamma^2 - \frac{\beta u}{4}\sum_{\gamma=1}^{n} m_\gamma^4 + \frac{\beta k}{6}\sum_{\gamma=1}^{n} m_\gamma^6 + \frac{1}{2}\sum_{\gamma,\kappa=1}^{n} m_\alpha q_{\gamma\kappa}^{-1} m_\kappa + \frac{\alpha}{2}\ln\det(\mathbb{1} - \beta\mathbf{Q}) + \frac{1}{2}\ln\det\tilde{\mathbf{Q}} - \frac{1}{2}\sum_{\gamma,\kappa=1}^{n}\tilde{Q}_{\gamma\kappa}Q_{\gamma,\kappa}. \tag{C25}$$

This integral can be evaluated through the saddle point method. First, finding the saddle point with respect to $\tilde{Q}$ gives us

$$\tilde{Q}_{ij}^{-1} = Q_{ij} \Rightarrow \tilde{\mathbf{Q}} = \mathbf{Q}^{-1}. \tag{C26}$$

Extremizing with respect to the other variables will be done later. Now we need to assume a form for the matrix $\mathbf{Q}$ and for m^γ. Let us assume that Q is replica symmetric (RS), i.e., $Q^{\gamma\kappa} = q + (1-q)\delta^{\gamma\kappa}$ and $m^\gamma = m \;\; \forall \;\; \gamma$. We will find the free energy for this form of Q and check for its stability against RS-breaking fluctuations. As it turns out, for our purposes, the RS form is stable for the retrieval and the paramagnetic phase. It becomes unstable only at extremely low temperatures.

$$Q_{\gamma\kappa} = q + (1-q)\delta_{\gamma\kappa}, \tag{C27}$$

$$Q_{\gamma\kappa}^{-1} = A\delta_{\gamma\kappa} + B, \qquad A = \frac{1}{1-q},$$

$$B = \frac{q}{(1-q)[1+(n-1)q]}, \tag{C28}$$

$$\lim_{n\to 0} B = -\frac{q}{(1-q)^2}, \tag{C29}$$

$$\det(Q) = \left(1 + n\frac{q}{1-q}\right)\det[(1-q)\mathbb{1}], \tag{C30}$$

$$\ln\det(Q) = \ln\left[1 + n\frac{q}{1-q}\right] + n\ln(1-q), \tag{C31}$$

$$\lim_{n\to 0}\frac{1}{n}\ln\det(Q) = \frac{q}{1-q} + \ln(1-q). \tag{C32}$$

Similarly,

$$\lim_{n\to 0}\frac{1}{n}\ln\det(1-\beta vQ) = \frac{-\beta qv}{1-\beta v(1-q)} + \ln[1-\beta v(1-q)]. \tag{C33}$$

All this was found using the matrix determinant lemma. Putting all this back in the expression for f, we finally obtain

$$\beta f = \mathrm{extr}\left(-\frac{\beta v}{2}m^2 - \frac{\beta u}{4}m^4 + \frac{\beta k}{6}m^6 + \frac{\alpha}{2}\left[\ln[1-\beta v(1-q)] - \frac{\beta vq}{1-\beta v(1-q)}\right] - \frac{1}{2}\left[\ln(1-q) + \frac{q-m^2}{1-q}\right]\right). \tag{C34}$$

The saddle point equations for variation across m and q are thus given by

$$\frac{\partial f}{\partial m} = 0 \Rightarrow m\left[1 + um^2 - km^4 - \frac{1}{\chi}\right] = 0, \tag{C35}$$

$$\frac{\partial f}{\partial q} = 0 \Rightarrow \frac{\alpha q}{(1-\chi)^2} = \frac{q-m^2}{\chi^2}, \tag{C36}$$

$$\chi = \beta(1-q). \tag{C37}$$

We can investigate the stability of the RS solution by adding RS-breaking fluctuations to Q. We denote the RS broken matrix as Q^B:

$$Q^B_{ij} = Q_{ij} + \eta_{ij}, \qquad \eta_{ij} = \eta_{ji}, \qquad \eta_{ii} = 0. \tag{C38}$$

$$(Q^B)^{-1} = [Q+\eta]^{-1} = Q^{-1} - Q^{-1}\eta Q^{-1} + Q^{-1}\eta Q^{-1}\eta Q^{-1}. \tag{C39}$$

We ignore terms linear in η as the first-order variation is set to 0 while taking extrema.

$$f(m,\tilde{Q}^B,Q^B) - f(m,\tilde{Q},Q) = T_1 + T_2 + T_3 + T_4, \tag{C40}$$

$$T_1 = \sum_{\gamma,\kappa} m^\gamma[(Q^B)^{\gamma\kappa} - Q^{\gamma\kappa}]m^\kappa, \tag{C41}$$

$$T_2 = \frac{\alpha}{2}\ln\frac{\det[\mathbb{1}-\beta(Q+\eta)]}{\det(\mathbb{1}-\beta Q)}, \tag{C42}$$

$$T_3 = \frac{1}{2}\ln\frac{\det Q^B}{\det Q}, \tag{C43}$$

$$T_4 = \frac{1}{2}\sum_{\gamma\kappa}[(Q_{\gamma\kappa}+\eta_{\gamma\kappa})\tilde{Q}_{\gamma\kappa} - Q_{\gamma\kappa}\tilde{Q}_{\gamma\kappa}]. \tag{C44}$$

After some algebra, it can be easily shown that

$$T_1 = m^2(a+bn)^2\left[a\sum_{ikz}\eta_{ik}\eta_{kz} + b\left(\sum_{ik}\eta_{1k}\right)^2\right], \tag{C45}$$

$$T_2 = -\frac{\alpha}{4}\left[c^2\sum_{ij}\eta_{ij}^2 + 2cd\sum_{ijk}\eta_{ij}\eta_{jk} + d^2\left(\sum_{ij}\eta_{ij}\right)^2\right], \tag{C46}$$

$$T_3 = \frac{1}{4}\left[a^2\sum_{ik}\eta_{ik}^2 + 2ab\sum_{ijk}\eta_{ij}\eta_{jk} + b^2\left(\sum_{ij}\eta_{ij}\right)^2\right], \tag{C47}$$

$$T_4 = 0, \tag{C48}$$

$$a = \frac{1}{1-q}, \qquad b = -\frac{q}{(1-q)^2}, \qquad c = \frac{\beta}{1-\beta(1-q)}, \qquad d = \frac{\beta^2 q}{[1-\beta(1-q)]^2}, \tag{C49}$$

$$f(m,\tilde{Q}^B,Q^B) - f(m,\tilde{Q},Q) = A\sum_{ij}\eta_{ij}^2 + B\sum_{ijk}\eta_{ij}\eta_{jk} + D\left(\sum_{ij}\eta_{ij}\right)^2, \tag{C50}$$

$$A = -\frac{1}{4}(\alpha c^2 - a^2), \tag{C51}$$

$$B = -\frac{\alpha}{2}cd + \frac{1}{2}ab + am^2(a+bn)^2, \tag{C52}$$

$$D = -\frac{1}{4}(\alpha d^2 - b^2) + bm^2(a+bn)^2. \tag{C53}$$

In order for the solution to be stable, we need the eigenvalues of the quadratic form Eq. (C50) to be positive. The eigenvalues of the quadratic form are eigenvalues of the equation

$$A\eta_{ij} + \frac{B}{2}\sum_k(\eta_{ik}+\eta_{kj}) + D\left(\sum_{kl}\eta_{kl}\right) = \Lambda\eta_{ij}. \tag{C54}$$

The eigenvalues for this equation are

$$\Lambda_1 = A, \tag{C55}$$

$$\Lambda_2 = \Lambda_1 + (n-2)\frac{B}{2} = A - B + \frac{nB}{2}, \tag{C56}$$

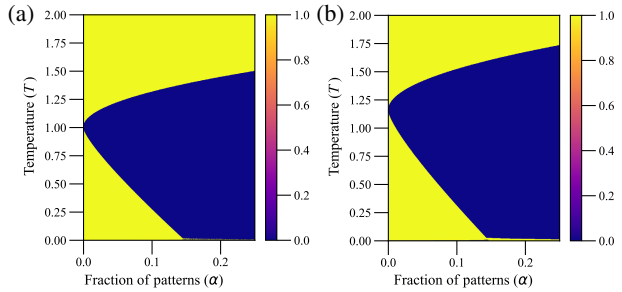

FIG. 10. Region of stability of solutions. The yellow region denotes regions where the RS solution is stable. (a) Region of stability for the equilibrium case. (b) Region of stability for passive fraction, $f = 0.2$, $\tau = 0.05$, $s = 1.0$.

$$\Lambda_3 = \Lambda_2 + \frac{nB}{2} + n(n-1)D = A - B + n(B-D) + n^2 D. \tag{C57}$$

Following Ref. [23], the relevant eigenvalue is $\Lambda_1 = A$. $A > 0$ gives us the region of stable solutions as shown in Fig. 10.

APPENDIX D: MARTIN-SIGGIA-ROSE APPROACH TO DYNAMICS IN PRESENCE OF ACTIVE NOISE

For working out the decoupled dynamics of the spins in the spherical Hopfield model, we shall be following Refs. [18,32,35,51]. The Langevin equation for the spins is given by

$$\partial_t \sigma_i = -\mu(t)\sigma_i(t) + \theta_i - \frac{\delta \mathcal{H}(\sigma)}{\delta \sigma_i(t)} + \eta_i(t), \tag{D1}$$

where μ is the Lagrange multiplier that ensures that the spins obey the spherical constraint, θ_i is the external field at each site, $\mathcal{H}(\sigma)$ is the Hamiltonian, and $\eta_i(t)$ is the noise in the system. The noise includes both white noise and active noise.

$$\eta_i(t) = \eta_{w,i}(t) + \eta_{a,i}(t), \tag{D2}$$

$$\langle \eta_{w,i}(t) \rangle = 0 = \langle \eta_{a,i}(t) \rangle \quad \forall \ \ i, t, \tag{D3}$$

$$\begin{aligned} \langle \eta_{w,i}(t)\eta_{w,j}(t') \rangle &= 2T_p \delta_{ij}\delta(t-t'), \\ \langle \eta_{a,i}(t)\eta_{a,j}(t') \rangle &= \frac{T_a}{\tau}\delta_{ij} \exp\left(-\frac{|t-t'|}{\tau}\right), \\ \langle \eta_{w,i}(t)\eta_{a,j}(t') \rangle &= 0 \end{aligned} \tag{D4}$$

$$\begin{aligned} \Rightarrow \langle \eta_i(t)\eta_j(t') \rangle &= 2T_p \delta_{ij}\delta(t-t') + \frac{T_a}{\tau}\delta_{ij} \exp\left(-\frac{|t-t'|}{\tau}\right) \\ &= D(t,t'), \end{aligned} \tag{D5}$$

where we have labeled the entire function as $D(t,t')$. Now we can write the probability of noise η as

$$P(\eta) \sim \exp\left(-\frac{1}{2}\int dt dt' \eta(t) D^{-1}(t,t')\eta(t')\right). \tag{D6}$$

Now we write the generating functional for the system as

$$Z[\boldsymbol{\psi}] = \int D\sigma D\eta P(\eta) \exp\left[i\sum_{i=1}^{N}\int dt \psi_i(t)\sigma_i(t)\right] \prod_{i=1}^{N} \delta\left(\partial_t \sigma_i(t) + \mu(t)\sigma_i(t) + \frac{\delta \mathcal{H}(\boldsymbol{\sigma})}{\delta \sigma_i(t)} - \eta_i(t) - \theta_i(t)\right). \tag{D7}$$

The various physical quantities, such as overlap with a pattern (m), correlations between spins (C), and the response of the spins to an external field (G), can be obtained from differentiating Z with respect to the conjugate fields ψ and θ.

$$m^{\mu}(t) = \sum_{i=1}^{N} \xi_i^{\mu} \langle \sigma_i(t) \rangle = -i\xi_i^{\mu} \lim_{\psi \to 0, \theta \to 0} \frac{\delta Z[\psi,\theta]}{\delta \psi_i(t)}, \tag{D8}$$

$$C_{ij}(t,t') = \langle \sigma_i(t)\sigma_j(t') \rangle = \lim_{\psi \to 0, \theta \to 0} \frac{\delta^2 Z[\psi,\theta]}{\delta \psi_i(t)\delta \psi_j(t')}. \tag{D9}$$

$$G_{ij}(t,t') = \frac{\delta \langle \sigma_i(t) \rangle}{\delta \theta_j(t')} = -i \lim_{\psi \to 0, \theta \to 0} \frac{\delta^2 Z[\psi,\theta]}{\delta \psi_i(t)\delta \theta_j(t')}. \tag{D10}$$

We now express the δ functions as exponentials and integrate out the noise using the definition of probability given in Eq. (D6). This yields

$$\begin{aligned} Z[\psi,\theta] = \int D[\sigma,\hat{\sigma}] \exp\Bigg[&\sum_{i=1}^{N}\int dt \psi_i(t)\sigma_i(t) \\ &+ \sum_{i=1}^{N}\int dt \theta_i(t)\tilde{\sigma}_i(t) + A[\sigma,\hat{\sigma}]\Bigg], \end{aligned} \tag{D11}$$

$$\begin{aligned} A[\sigma,\hat{\sigma}] = &-\frac{1}{2}\sum_{i=1}^{N}\int dt dt' \hat{\sigma}_i(t) D(t,t') \hat{\sigma}_i(t') \\ &+ i\sum_{i=1}^{N}\int dt \hat{\sigma}_i(t)(L_{0,i} + L_{\xi,i}), \end{aligned} \tag{D12}$$

$$L_{0,i} = \partial_t \sigma_i(t) + \mu(t)\sigma_i(t), \tag{D13}$$

$$L_{\xi,i} = \frac{\delta \mathcal{H}(\sigma)}{\delta \sigma_i} \tag{D14}$$

$$= -\sum_j J_{ij}\sigma_j - \sum_{j,k,l} J_{ijkl}\sigma_j\sigma_k\sigma_l \tag{D15}$$

$$= -\frac{1}{N}\sum_{\mu,j,j\neq i} \xi_i^\mu \xi_j^\mu \sigma_j - \frac{u_0}{N^3}\sum_{\mu,j,k,l} \xi_i^\mu \xi_j^\mu \xi_k^\mu \xi_l^\mu \sigma_j\sigma_k\sigma_l. \tag{D16}$$

Henceforth we will omit θ because it does not affect further calculations and just leads to cumbersome notation. It is understood that θ will always be to taken to 0 eventually. Now we need to take the average of Z over the quenched disorder ξ. Henceforth, $\overline{A}$ will denote the quenched average of A over the pattern variables.

$$\bar{A} = \int D\xi \exp\left(-\frac{\xi^2}{2}\right) A(\xi), \tag{D17}$$

$$\overline{Z[\psi]} = \overline{\int D[\sigma,\hat{\sigma}] \exp\left[\sum_{i=1}^N \int dt \psi_i(t)\sigma_i(t) + A[\sigma,\hat{\sigma}]\right]}. \tag{D18}$$

But we observe that the only part in the entire expression that depends on the pattern variables is the part associated with L_{ξ}. Thus the quenched average needs to be taken only over L_{ξ}. Henceforth, the integrals over time will be implicitly assumed and Einstein summation convention will be used.

We assume that only a single pattern, the first pattern ($\mu = 1$), is condensed, i.e., the overlap of the spin state with pattern 1 is O(1) at long times and large N. For other patterns it is $O(1/N)$ and decays to 0 for large N. We define a few macroscopic variables:

$$m(t) = \frac{1}{N}\xi_i^1\sigma_i(t), \qquad \delta_m = \delta[Nm(t) - \xi_i^1\sigma_i(t)], \tag{D19}$$

$$w(t) = \frac{1}{N}\xi_i^1\hat{\sigma}_i(t), \qquad \delta_w = \delta[Nw(t) - \xi_i^1\hat{\sigma}_1(t)], \tag{D20}$$

$$\mathcal{J} = \overline{\exp\left[i\hat{\sigma}_i \cdot L_{\xi,i}\right]} = \overline{\exp\left[-i\hat{\sigma}_i \cdot \left(\frac{1}{N}\xi_i^\mu \xi_j^\mu \sigma_j + \frac{u_0}{N^3}\xi_i^\mu \xi_j^\mu \xi_k^\mu \xi_l^\mu \sigma_j\sigma_k\sigma_l - \frac{1}{N}(\xi_i^\mu)^2\sigma_i\right)\right]} \tag{D21}$$

$$= \overline{\exp\left[-iNw(t)(m(t) + u_0 m(t)^3) - i\frac{1}{\sqrt{N}}\xi_i^\mu\hat{\sigma}_i\frac{1}{\sqrt{N}}\xi_j^\mu\sigma_j + i\frac{1}{N}(\xi_i^\mu)^2\hat{\sigma}_i\sigma_i + O(1/N)\right]}\delta_m\delta_w. \tag{D22}$$

We define a few two-time quantities (which will emerge naturally further down the calculation):

$$K(t,t') = \frac{1}{N}\sigma_i(t)\hat{\sigma}_i(t'), \qquad q(t,t') = \frac{1}{N}\sigma_i(t)\sigma_i(t'), \qquad Q(t,t') = \frac{1}{N}\hat{\sigma}_i(t)\hat{\sigma}_i(t'), \tag{D23}$$

$$\mathcal{Z}[\psi] = \int D[\sigma,\hat{\sigma}]DKD\hat{K}DmD\hat{m}DwD\hat{w}DQD\hat{Q}DqD\hat{q} \exp\left[-\frac{1}{2}\sum_{i=1}^N \int dtdt'\hat{\sigma}_i(t)D(t,t')\sigma_i(t') + i\sum_{i=1}^N \int dt\hat{\sigma}_i(t)L_{0,i}\right]$$
$$\times \exp[N[\Psi(\mathbf{m},\hat{\mathbf{m}},\mathbf{w},\hat{\mathbf{w}},\mathbf{q},\hat{\mathbf{q}},\mathbf{Q},\hat{\mathbf{Q}},\mathbf{K},\hat{\mathbf{K}}) + \Phi(\mathbf{m},\mathbf{w},\mathbf{q},\mathbf{Q},\mathbf{K}) + \Theta(\hat{\mathbf{m}},\hat{\mathbf{w}},\hat{\mathbf{q}},\hat{\mathbf{Q}},\hat{\mathbf{K}})]], \tag{D24}$$

where Ψ, Φ, and Θ are defined as follows:

$$\Psi = i[m(t)\hat{m}(t) + w(t)\hat{w}(t) + q(t,t')\hat{q}(t,t') + Q(t,t')\hat{Q}(t,t') + K(t,t')\hat{K}(t,t')], \tag{D25}$$

$$\Phi = -iw(t)[m(t) + u_0 m(t)^3] - \frac{\alpha}{2}\ln\{\det[1 - 2i\delta_{t,t'}K(t,t')]\}$$
$$+ \alpha \ln\left(\int D\hat{x}D\hat{y} \exp\left[-\frac{1}{2}\int dtdt'[\hat{x}(t)q(t,t')\hat{x}(t') + 2\hat{x}(t)K(t',t)\hat{y}(t') + \hat{y}(t)Q(t,t')\hat{y}(t') - 2i\hat{x}(t)\delta(t-t')\hat{y}(t')]\right]\right), \tag{D26}$$

$$\Theta = -i\frac{1}{N}[\hat{m}(t)\xi_i^1\sigma_i(t) + \hat{w}(t)\xi_i^1\hat{\sigma}_i(t) + \hat{q}(t,t')\sigma_i(t)\sigma_i(t') - i\hat{Q}(t,t')\hat{\sigma}_i(t)\hat{\sigma}_i(t') + \hat{K}(t,t')\sigma_i(t)\hat{\sigma}_i(t') - \psi_i(t)\sigma_i(t)]. \tag{D27}$$

Equation (D24) is of the form where we can use the saddle point approximation to calculate the integral. For the saddle point equations we set $\partial_v(\Psi+\Phi+\Theta)=0$, where v is one of the macroscopic variables, $m,\hat{m},\ldots,K,\hat{K}$. This yields

$$w(t)=\hat{m}(t)=0, \qquad \hat{w}(t)=[m(t)+u_0 m(t)^3], \tag{D28}$$

$$K(t,t')=iG(t',t), \qquad G(t',t)=\lim_{\psi\to 0}\langle\sigma(t)\hat{\sigma}(t')\rangle_*, \tag{D29}$$

$$q(t,t')=-\frac{1}{2}\alpha i\frac{\int d\hat{x}d\hat{y}\,\hat{x}(t)\hat{x}(t')\exp(-\frac{1}{2}[\hat{x}q\hat{x}+2\hat{y}K\hat{x}-2i\hat{x}\,\hat{y}+\hat{y}Q\hat{y}])}{\int d\hat{x}d\hat{y}\exp(-\frac{1}{2}[\hat{x}q\hat{x}+2\hat{y}K\hat{x}-2i\hat{x}\,\hat{y}+\hat{y}Q\hat{y}])}=0, \tag{D30}$$

$$Q(t,t')=-\frac{1}{2}\alpha i\frac{\int d\hat{x}d\hat{y}\,\hat{y}(t)\hat{y}(t')\exp(-\frac{1}{2}[\hat{x}q\hat{x}+2\hat{y}K\hat{x}-2i\hat{x}\,\hat{y}+\hat{y}Q\hat{y}])}{\int d\hat{x}d\hat{y}\exp(-\frac{1}{2}[\hat{x}q\hat{x}+2\hat{y}K\hat{x}-2i\hat{x}\,\hat{y}+\hat{y}Q\hat{y}])}=-\frac{1}{2}\alpha i[(1-G)^{-1}C(1-G^\dagger)^{-1}](t,t'), \tag{D31}$$

$$K(t,t')=-\alpha i\frac{\int d\hat{x}d\hat{y}\,\hat{y}(t)\hat{x}(t')\exp(-\frac{1}{2}[\hat{x}q\hat{x}+2\hat{y}K\hat{x}-2i\hat{x}\,\hat{y}+\hat{y}Q\hat{y}])}{\int d\hat{x}d\hat{y}\exp(-\frac{1}{2}[\hat{x}q\hat{x}+2\hat{y}K\hat{x}-2i\hat{x}\,\hat{y}+\hat{y}Q\hat{y}])}-\alpha[\mathbb{1}-\delta_{t,t'}K(t,t')]^{-1}\delta(t,t') \tag{D32}$$

$$=\alpha(1-G)^{-1}(t,t')-\alpha\mathbb{1}=\alpha G(1-G)^{-1}. \tag{D33}$$

Substituting these results in Eq. (D24), we obtain the final result,

$$Z[\psi]=\mathbb{K}\int D[\sigma,\hat{\sigma}]\exp\left[-\frac{1}{2}\hat{\sigma}_i\{D+\alpha[(1-G)^{-1}C(1-G^\dagger)^{-1}]\}\hat{\sigma}_i+i\hat{\sigma}_i\{\partial_t\sigma_i+\mu\sigma_i-[m+u_0m^3]\xi_i^1-\alpha G(\mathbb{1}-G)^{-1}\sigma_i\}\right], \tag{D34}$$

where $\mathbb{K}$ is a constant. From here we can write the effective equation of motion of a single spin which is decoupled from all other spins as

$$\partial_t\sigma_i(t)=-\mu(t)\sigma_i(t)+[m(t)+u_0m(t)^3]\xi_i^1+\int dt'\alpha G(\mathbb{1}-G)^{-1}(t,t')\sigma_i(t')+\chi_i(t), \tag{D35}$$

$$\langle\chi_i(t)\chi_j(t')\rangle=\delta_{ij}D(t,t')+\delta_{ij}\alpha[(1-G)^{-1}C(1-G^\dagger)^{-1}](t,t'), \tag{D36}$$

$$D(t,t')=2T_p\delta(t-t')+\frac{T_a}{\tau}\exp\left(-\frac{|t-t'|}{\tau}\right), \tag{D37}$$

$$C(t,t')=\langle\sigma(t)\sigma(t')\rangle_*, \tag{D38}$$

$$G(t,t')=\frac{\partial\langle\sigma(t)\rangle}{\partial\theta(t')}. \tag{D39}$$

Using this we can write the equations for m, C, and G as follows:

$$\left(\frac{\partial}{\partial t}+\mu(t)\right)m(t)=[m(t)+um(t)^3]+\int_{-\infty}^{t}dt'R(t,t')m(t'), \tag{D40}$$

$$\left(\frac{\partial}{\partial t}+\mu(t)\right)G(t,t')=\delta(t-t')+\alpha\int_{t'}^{t}dt_1R(t,t_1)G(t_1,t'), \tag{D41}$$

$$\left(\frac{\partial}{\partial t}+\mu(t)\right)C(t,t')=[m(t)+um(t)^3]m(t')+\alpha\int_{-\infty}^{t}dt_1R(t,t_1)C(t',t_1)+2T_pG(t',t)+\alpha\int_{-\infty}^{t'}dt_1\left[S(t,t_1)G(t',t_1)+\frac{T_a}{\tau}\exp\left(-\frac{|t-t_1|}{\tau}\right)G(t',t_1)\right], \tag{D42}$$

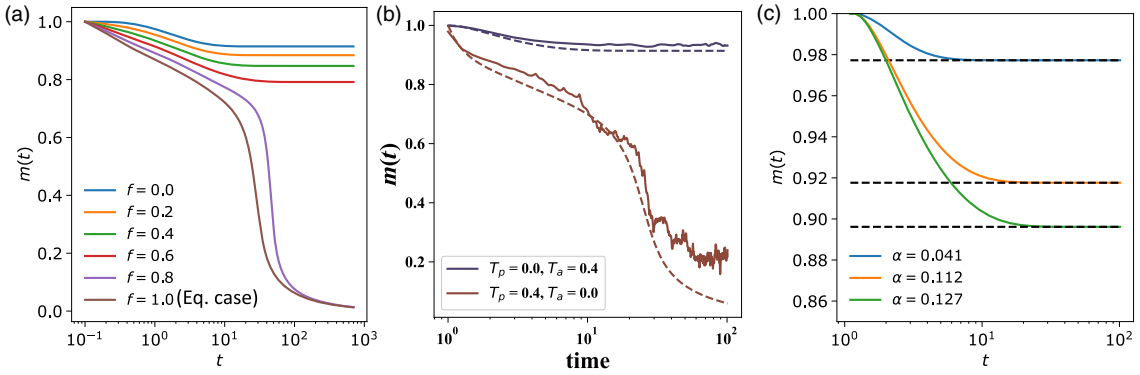

FIG. 11. (a) The simulation in this case has been run for $t = 800$. (b) Comparison between numerical simulation and MSR simulation for $\alpha = 0.1$, $T_{\text{eff}} = 0.4$. (c) Comparison between the results in Fig. 2 of Ref. [18] and the integrator we used for MSR simulations.

$$R(t,t') = [(1-G)^{-1}G](t,t'), \tag{D43}$$

$$S(t,t') = (1-G)^{-1}C(1-G^{\dagger})^{-1}(t,t'). \tag{D44}$$

In case the spins follow a non-Markovian Langevin equation due to the presence of a drag kernel, the equations of motion for the spins are given by

$$\partial_t\sigma_i + \int_{-\infty}^{t} ds\gamma_d(t-s)\partial_s\sigma(s) = -\mu(t)\sigma_i(t) + \theta_i - \frac{\delta\mathcal{H}(\sigma)}{\delta\sigma_i(t)} + \eta_i(t), \tag{D45}$$

where $\gamma_d(t)$ is the drag kernel. When the kernel satisfies $T_p\gamma_d(t-s) = \langle\eta_a(t)\eta_a(s)\rangle$, then fluctuation-dissipation theorem is satisfied and the system is at equilibrium at an effective temperature of T_p. Any deviation from this results in nonequilibrium behavior. In this case, the equations for m, C, and G get modified to

$$\left(\frac{\partial}{\partial t} + \mu'(t)\right)m(t) = [m(t) + um(t)^3] + \int_{-\infty}^{t} dt'[\alpha R(t,t') + \partial_{t'}\gamma_d(t-t')]m(t'), \tag{D46}$$

$$\left(\frac{\partial}{\partial t} + \mu'(t)\right)G(t,t') = \delta(t-t') + \int_{t'}^{t} dt_1[\alpha R(t,t_1) + \partial_{t'}\gamma_d(t-t')]G(t_1,t'), \tag{D47}$$

$$\left(\frac{\partial}{\partial t} + \mu' s(t)\right)C(t,t') = [m(t) + um(t)^3]m(t') + \int_{-\infty}^{t} dt_1[\alpha R(t,t_1) + \partial_{t'}\gamma_d(t-t')]C(t',t_1) + 2T_pG(t',t) + \alpha\int_{-\infty}^{t'} dt_1\left[S(t,t_1)G(t',t_1) + \frac{T_a}{\tau}\exp\left(-\frac{|t-t_1|}{\tau}\right)G(t',t_1)\right]. \tag{D48}$$

The numerical integrations are carried out using a simple explicit nonadaptive time stepping method. The $C(t,t')$ matrix is not time-translation invariant and is initialized as a 1×1 matrix with entry 1. The $G(t,t')$ matrix is initialized as a 1×1 matrix with entry 0 and $m(t)$ is initialized as 0.95. The time step size is chosen to be 0.1. At every step, inverses and matrix products are computed and the equations are propagated forward. At each step, the side length of C and G matrices increase by 1. We carry this procedure for a long time till m is constant for > 100 time steps. We stop our procedure here. We have verified our procedure for the equilibrium case where exact results exist in the literature [18,52].

[1] Bruno Bontempi, Catherine Laurent-Demir, Claude Destrade, and Robert Jaffard, *Time-dependent reorganization of brain circuitry underlying long-term memory storage*, Nature (London) **400**, 671 (1999).
[2] Larry R. Squire, *Memory and Brain* (Oxford University Press, New York, 1987).
[3] P. Peretto, *Collective properties of neural networks: A statistical physics approach*, Biol. Cybern. **50**, 51 (1984).
[4] W. A. Little, *The Existence of Persistent States in the Brain* (Springer, Boston, 1996), pp. 145–164.
[5] John P. Barton, Mehran Kardar, and Arup K. Chakraborty, *Scaling laws describe memories of host-pathogen riposte in the HIV population*, Proc. Natl. Acad. Sci. U.S.A. **112**, 1965 (2015).

[6] Thomas A. Cleland and Christiane Linster, *Computation in the olfactory system*, Chem. Senses **30**, 801 (2005).

[7] Richard J. Stevenson, *Olfactory perception, cognition, and dysfunction in humans*, WIREs Cognit. Sci. **4**, 273 (2013).

[8] Rolf M. Zinkernagel, Martin F. Bachmann, Thomas M. Kündig, Stephan Oehen, Hanspeter Pirchet, and Hans Hengartner, *On immunological memory*, Annu. Rev. Immunol. **14**, 333 (1996).

[9] Sumit Sinha, Xin Li, Rajsekhar Das, and D. Thirumalai, *Mechanical feedback controls the emergence of dynamical memory in growing tissue monolayers*, J. Chem. Phys. **156**, 245101 (2022).

[10] Shaobo Zhang, Xiang Teng, Yusuke Toyama, and Timothy E. Saunders, *Periodic oscillations of myosin-II mechanically proofread cell-cell connections to ensure robust formation of the cardiac vessel*, Curr. Biol. **30**, 3364 (2020).

[11] Étienne Fodor, Cesare Nardini, Michael E. Cates, Julien Tailleur, Paolo Visco, and Frédéric van Wijland, *How far from equilibrium is active matter?*, Phys. Rev. Lett. **117**, 038103 (2016).

[12] Gerhard Gompper, Roland G Winkler, Thomas Speck, Alexandre Solon, Cesare Nardini, Fernando Peruani, Hartmut Löwen, Ramin Golestanian, U Benjamin Kaupp, Luis Alvarez *et al.*, *The 2020 motile active matter roadmap*, J. Phys. Condens. Matter **32**, 193001 (2020).

[13] M. Cristina Marchetti, Jean-François Joanny, Sriram Ramaswamy, Tanniemola B. Liverpool, Jacques Prost, Madan Rao, and R Aditi Simha, *Hydrodynamics of soft active matter*, Rev. Mod. Phys. **85**, 1143 (2013).

[14] Étienne Fodor and M. Cristina Marchetti, *The statistical physics of active matter: From self-catalytic colloids to living cells*, Physica (Amsterdam) **504A**, 106 (2018).

[15] Gabriel S. Redner, Michael F. Hagan, and Aparna Baskaran, *Structure and dynamics of a phase-separating active colloidal fluid*, Phys. Rev. Lett. **110**, 055701 (2013).

[16] Alexandre P. Solon, Joakim Stenhammar, Michael E. Cates, Yariv Kafri, and Julien Tailleur, *Generalized thermodynamics of phase equilibria in scalar active matter*, Phys. Rev. E **97**, 020602 (2018).

[17] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).

[18] D. Bollé, Th. M. Nieuwenhuizen, I. Pérez Castillo, and T. Verbeiren, *A spherical Hopfield model*, J. Phys. A **36**, 10269 (2003).

[19] Menachem Stern, Matthew B. Pinson, and Arvind Murugan, *Continual learning of multiple memories in mechanical networks*, Phys. Rev. X **10**, 031044 (2020).

[20] Francesca Mignacco, Pierfrancesco Urbani, and Lenka Zdeborová, *Stochasticity helps to navigate rough landscapes: Comparing gradient-descent-based algorithms in the phase retrieval problem*, Mach. Learn. Sci. Technol. **2**, 035029 (2021).

[21] Daniel J. Amit, Hanoch Gutfreund, and H. Sompolinsky, *Storing infinite numbers of patterns in a spin-glass model of neural networks*, Phys. Rev. Lett. **55**, 1530 (1985).

[22] Daniel J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, Cambridge, England, 1989).

[23] A. Crisanti and H.-J. Sommers, The spherical p-spin interaction spin glass model: The statics Z. Phys. B **87**, 341 (1992).

[24] Ananyo Maitra and Raphael Voituriez, *Enhanced orientational ordering induced by an active yet isotropic bath*, Phys. Rev. Lett. **124**, 048003 (2020).

[25] Étienne Fodor, Takahiro Nemoto, and Suriyanarayanan Vaikuntanathan, *Dissipation controls transport and phase transitions in active fluids: Mobility, diffusion and biased ensembles*, New J. Phys. **22**, 013052 (2020).

[26] Peter Hänggi, *Escape over fluctuating barriers driven by colored noise*, Chem. Phys. **180**, 157 (1994).

[27] E. Woillez, Y. Zhao, Y. Kafri, V. Lecomte, and J. Tailleur, *Activated escape of a self-propelled particle from a metastable state*, Phys. Rev. Lett. **122**, 258001 (2019).

[28] G. L. Shaw, *Donald Hebb: The organization of behavior*, in *Brain Theory*, edited by Günther Palm and Ad Aertsen (Springer, Berlin, 1986), pp. 231–233.

[29] L. L. Bonilla, *Active Ornstein-Uhlenbeck particles*, Phys. Rev. E **100**, 022601 (2019).

[30] R Kubo, *The fluctuation-dissipation theorem*, Rep. Prog. Phys. **29**, 255 (1966).

[31] In cases where the configurational degree of freedom of a single noninteracting particle is driven by an active noise source, T_{eff} provides an accurate measure of the diffusion constant just like a native temperature would.

[32] A. C. C. Coolen, *Statistical mechanics of recurrent neural networks II—Dynamics*, in *Handbook of Biological Physics* Vol. 4, edited by F. Moss and S. Gielen (North-Holland, Amsterdam, 2001), Chap. 15, pp. 619–684.

[33] In the steady state, if the system is able to recover a pattern, the overlaps $m^{\mu} \sim O(1)$ for the pattern μ it was initialized near and the overlap with other patterns go as $1/\sqrt{N}$ [22]. In principle, one can consider retrieval of a finite number $\sim O(1)$ of patterns and perform similar calculations. We have chosen to work with the specific case in which one pattern is retrieved for simplicity.

[34] Peter Jung and Peter Hänggi, *Dynamical systems: A unified colored-noise approximation*, Phys. Rev. A **35**, 4464 (1987).

[35] P. C. Martin, E. D. Siggia, and H. A. Rose, *Statistical dynamics of classical systems*, Phys. Rev. A **8**, 423 (1973).

[36] Modifying s may also change effectiveness of recovery, but we leave a full replica analysis of these effects to a subsequent paper.

[37] For highly nonlinear springs $\xi < 1$, larger σ leads to a larger basin of recovery. However, if σ becomes too large, basins of recovery around different competing configurational memory states start interfering with each other, leading to loss in capacity.

[38] Francesca Mignacco and Pierfrancesco Urbani, *The effective noise of stochastic gradient descent*, J. Stat. Mech. (2022) 083405.

[39] Marco Benedetti, Enrico Ventura, Enzo Marinari, Giancarlo Ruocco, and Francesco Zamponi, *Supervised perceptron learning vs unsupervised Hebbian unlearning: Approaching optimal memory retrieval in Hopfield-like networks*, J. Chem. Phys. **156**, 104107 (2022).

[40] Weishun Zhong, David J Schwab, and Arvind Murugan, *Associative pattern recognition through macro-molecular self-assembly*, J. Stat. Phys. **167**, 806 (2017).

[41] R. Monasson and S. Rosay, *Transitions between spatial attractors in place-cell models*, Phys. Rev. Lett. **115**, 098101 (2015).

[42] E. Gardner, *Multiconnected neural network models*, J. Phys. A **20**, 3453 (1987).

[43] Nathan C. Keim, Joseph D. Paulsen, Zorana Zeravcic, Srikanth Sastry, and Sidney R. Nagel, *Memory formation in matter*, Rev. Mod. Phys. **91**, 035002 (2019).

[44] Laurent Corté, P. M. Chaikin, J. P. Gollub, and D. J. Pine, *Random organization in periodically driven systems*, Nat. Phys. **4**, 420 (2008).

[45] Nathan C. Keim and Sidney R. Nagel, *Generic transient memory formation in disordered systems with noise*, Phys. Rev. Lett. **107**, 010603 (2011).

[46] Elisabeth Agoritsas, Giulio Biroli, Pierfrancesco Urbani, and Francesco Zamponi, *Out-of-equilibrium dynamical mean-field equations for the perceptron model*, J. Phys. A **51**, 085002 (2018).

[47] Laura Tociu, Étienne Fodor, Takahiro Nemoto, and Suriyanarayanan Vaikuntanathan, *How dissipation constrains fluctuations in nonequilibrium liquids: Diffusion, structure, and biased interactions*, Phys. Rev. X **9**, 041026 (2019).

[48] Gili Bisker and Jeremy L. England, *Nonequilibrium associative retrieval of multiple stored self-assembly targets*, Proc. Natl. Acad. Sci. U.S.A. **115**, E10531 (2018).

[49] Arvind Murugan, Zorana Zeravcic, Michael P. Brenner, and Stanislas Leibler, *Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures*, Proc. Natl. Acad. Sci. U.S.A. **112**, 54 (2015).

[50] C. De Dominicis, *Dynamics as a substitute for replicas in systems with quenched random impurities*, Phys. Rev. B **18**, 4913 (1978).

[51] A. Crisanti, H. Horner, and H. Sommers, *The spherical p-spin interaction spin-glass model: The dynamics*, Z. Phys. B **92**, 257 (1993).

[52] Stefano Sarao Mannelli, Giulio Biroli, Chiara Cammarota, Florent Krzakala, Pierfrancesco Urbani, and Lenka Zdeborová, *Marvels and pitfalls of the Langevin algorithm in noisy high-dimensional inference*, Phys. Rev. X **10**, 011057 (2020).