Noise Effects on Memory Recall in Dense Hopfield Networks

Arun Kamath
Massachusetts Institute of Technology

(Dated: May 15, 2026)

In this study, generating functional analysis (GFA) is employed to derive an asymptotically exact analysis of the trajectory of a dense Hopfield network with noisy updates. The effect of thermal and cross-talk noise on the convergence time and the basin of attraction is presented through theoretical and numerical simulations. Future research directions are outlined.

I. INTRODUCTION

Hopfield networks generated a flurry of research in the latter 20th century because of their application to neural computation, and more specifically their implementation of associative memory [1]. The ability to reconstruct a memory from an incomplete input is a key feature of biological memory, and Hopfield networks represented a promising first step in understanding such a phenomenon. In Hopfield's implementation, patterns of binary vectors $\boldsymbol{\xi^\mu} \in \{\pm 1\}^N$ can be stored in a network:

$$J_{ij} = \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu \tag{1}$$

and repeated updates, whether one index at a time (asynchronous), or all at once (synchornous), transform input an state $\boldsymbol{S}$ to its nearest stored pattern $\boldsymbol{\xi^\mu}$. It can also be shown that repeated updates descend an associated energy function(i.e. a Lyapunov function), which contains local minima at the stored patterns. The energy function was originally written in terms of the overlaps m_μ as

$$E = -\sum_{\mu, i<j} \xi_i^\mu \xi_j^\mu S_i S_j \tag{2}$$

$$= -\frac{1}{2}\sum_{\mu} m_\mu^2 \tag{3}$$

$$m_\mu = \frac{1}{N}\sum_i \xi_i^\mu S_i \tag{4}$$

with the second equality in the energy function being valid up to an additive constant. While there will be local minima at each stored pattern $\boldsymbol{\xi^u}$ when few patterns are stored in the network, the quadratic scaling of the energy function is too weak to suppress the noise (so called cross-talk) generated by many patterns, and thus the limit on the number of storable patterns scales as αN, with $\alpha \approx 0.14$ when the patterns are randomly chosen. This linear scaling limitation cannot be mitigated by having a hierarchical correlation of stored patterns, necessitating the need for higher order interactions between spins [2, 3].

II. DENSE ASSOCIATIVE MEMORY

Dense associative memory is a rediscovery of an n-body interaction Hopfield model, wherein higher order interactions lead to a sharper energy function near the stored patterns[4, 5]. Higher order interactions are represented as an energy of the form $E \propto -\sum_\mu m_\mu^p$, where $p = 2$ corresponds to the classical Hopfield network, while $p > 2$ corresponds to higher order, "dense", networks. Various authors have further shown that this can lead to an increase in storage capacity, of the order $\alpha_p' N^{p-1}$ [6, 7]. This result, along with with connections to generative models and attention mechanisms, has renewed interest in Hopfield-like networks, and more generally the use of statistical physics to analyze the limits of modern neural networks [8, 9, 10]. An important property of an associative memory system is the radius of the *basin of attraction*, in other words, the minimum level of corruption in $m_\mu(t = 0)$ for which a target pattern can still be recalled. Ideally, a robust network will be able to recall a pattern from a small initial overlap. In addition, minimizing the number of time steps required to reach the stored pattern will be of interest. However, the presence of thermal noise, whether it be biological or electrical, will be present in implementations of the these networks, and it may affect both the convergence time and the radius of the basin of attraction. In this work, theory and simulation are used to study the effect of thermal noise on these two metrics.

III. GENERATING FUNCTIONAL

In the following work, we build upon the methods outlined in Mimura et al [11]. We consider a synchronous update rule where every spin index is updated synchronously via Glauber dynamcis as follows:

$$h_i^{(t)} = \sum_\mu p(m_\mu^{(t)} + \theta_i^{(t)})^{p-1} \xi_i^\mu \tag{5}$$

$$\Pr(S_i^{(t+1)} = 1 | h_i^{(t)}) = \frac{1}{1 + \exp(-2\beta h_i^{(t)})} \tag{6}$$

Here, θ_i is an external field (or threshold for activation) which can be taken to zero after relevant derivatives are taken of the generating functional. In the limit $\beta \to$

∞, the update becomes a sgn function, and we recover the update rule commonly studied in the literature. To analyse the trajectory of a system, we follow the procedure of Mimura et al to develop a self-consistent theory for $m_\mu(t)$. To begin, we define the path probability, which is the probability of a given trajectory of spin states consistent with the updating rule of eqn(6). In terms of the local fields, the path probability can be expressed as a Markovian process as

$$Pr[\boldsymbol{S}^{(t+1)}|\boldsymbol{S}^{(t)}] = \prod_{i=1}^{N} \frac{1}{1+\exp(-2\beta h_i^{(t)} S_i^{(t)})} \tag{7}$$

$$Pr[\boldsymbol{S}^{(0)},...,\boldsymbol{S}^{(T)}] = Pr(\boldsymbol{S}^{(0)}) \prod_{t=0}^{T} Pr[\boldsymbol{S}^{(t+1)}|\boldsymbol{S}^{(t)}] \tag{8}$$

Then we write the generating functional, where we have introduced $\boldsymbol{\psi} = (\psi^1,...,\psi^N)$ as a generating variable.

$$\bar{Z}[\boldsymbol{\psi}] = \mathbb{E}_{\boldsymbol{\xi^1},...,\boldsymbol{\xi^M}} \sum_{paths} \Pr(path) \exp(-i \sum_{t=0}^{T} \boldsymbol{S}^{(t)} \cdot \boldsymbol{\psi}^{(t)}) \tag{9}$$

Using a Dirac delta function to enforce the definition of the local field, and transforming into Fourier space, the generating functional can be written in the following way, where the contributions of spins at each time are decoupled.

$$\begin{aligned}
\bar{Z}[\boldsymbol{\psi}] = & \sum_{\boldsymbol{S}^{(0)},\cdots,\boldsymbol{S}^{(T)}} \int_{\mathbb{R}^T} d\boldsymbol{h} \delta\hat{\boldsymbol{h}} \qquad (10)\\
& \times Pr[\boldsymbol{S}^{(0)}] \Big(\prod_{t=0}^{T-1} \prod_{i=1}^{N} \frac{1}{1+\exp(-2\beta h_i^{(t)} S_i^{(t)})} \Big) e^{-i\sum_{t=0}^{T} \boldsymbol{S}^{(t)}\cdot\boldsymbol{\psi}^{(t)}} \\
& \times \exp\Big[i \sum_{t=0}^{T} \sum_{i=1}^{N} \hat{h}_i^{(t)} (h_i^{(t)} - \theta_i^{(t)}) \Big] \\
& \times \Big(\mathbb{E}_{\boldsymbol{\xi}^1} \exp\Big[-i \sum_{t=0}^{T-1} \sum_{i=1}^{N} \hat{h}_i^{(t)} \xi_i^1 \; p m_1^{p-1} \Big] \Big) \\
& \times \Big(\mathbb{E}_{\boldsymbol{\xi}^2,\cdots,\boldsymbol{\xi}^M} \exp\Big[-i \sum_{t=0}^{T-1} \sum_{i=1}^{N} \sum_{\mu=2}^{M} \hat{h}_i^{(t)} \xi_i^\mu \; p m_\mu^{p-1} \Big] \Big)
\end{aligned}$$

Here, we have separated the pattern being recalled, ξ_1, from the other $M-1$ patterns, where $M \propto N^{p-1}$. The analysis in the following sections are quoted directly from Mimura et al, with the exception of the update rule which includes thermal noise. For a full justification and derivation of the saddle point self consistency equations, see Appendices A and B of [11].

IV. MAIN THEORETICAL RESULT

First, we define the overlap, correlation function, and response function respectively as

$$m^{(t)} = \langle\!\langle \xi s^{(t)} \rangle\!\rangle, \tag{11}$$

$$C^{(t,t')} = \langle\!\langle s^{(t)} s^{(t')} \rangle\!\rangle, \tag{12}$$

$$G^{(t,t')} = \frac{\partial \langle\!\langle s^{(t)} \rangle\!\rangle}{\partial \theta^{(t')}}, \tag{13}$$

The $\langle\!\langle \ldots \rangle\!\rangle$ indicates an average calculated by the single site process of eqn 14. We note the response function is evaluated for $t > t'$ only, and zero otherwise to satisfy causality. Utilizing a saddle point approximation for the generating functional, we arrive at the following result.

$$\langle\!\langle f(\boldsymbol{s}) \rangle\!\rangle = \mathbb{E}_\xi \int \mathcal{D}\boldsymbol{v} \sum_{\boldsymbol{s}} f(\boldsymbol{s}) p[s^{(0)}] \prod_{t=0}^{T-1} \frac{\exp(\beta s^{t+1} H_t)}{2\cosh(\beta H_t)} \tag{14}$$

H_t is the effective field $\xi p (m^{(t)})^{p-1} + (\Gamma \boldsymbol{s})^{(t)} + v^{(t)}$, where the first term is a signal term, and the second and third term contribute cross-talk noise to the recall process. The random vector $\boldsymbol{v}$ follows a normal distribution with mean $\boldsymbol{0}$ and covariance $R = (R^{(t,t')})$ with elements

$$R^{(t,t')} = p^2 \alpha_p \sum_{k=0}^{p-1} A(p-1,k) (C^{(t,t')})^k \tag{15}$$

where $\alpha_p = \frac{M}{N^{p-1}}$. The function $A(l,k)$ is a combinatorial function that arises from correcting the Hamiltonian to ignore self coupling terms. The precise form is given by

$$A(\ell,k) = \begin{pmatrix} \ell \\ k \end{pmatrix}^2 k! \; ((\ell-k-1)!!)^2, \tag{16}$$

We note that within the Hopfield network literature, different conventions for the form of the Hamiltonian often result in such corrections, but do not tend to change the scaling of storage capacities and other key behaviors. Γ is a self-interacting term that relates the current state to its past states in a complicated manner. The matrix Γ is given by $\Gamma = D \circ G$, where the operator $\circ$ denotes the element-wise product of matrices. Each element of the matrix $D = (D^{(t,t')})$ is defined as

$$D^{(t,t')} = p^2 (p-1)^2 \alpha_p \sum_{k=0}^{p-2} A(p-2,k) (C^{(t,t')})^k \tag{17}$$

Mimura et al note that setting $\Gamma = 0$ recovers the result of the equilibrium analysis of the n-body Hopfield model when solving the fixed point equation for m.

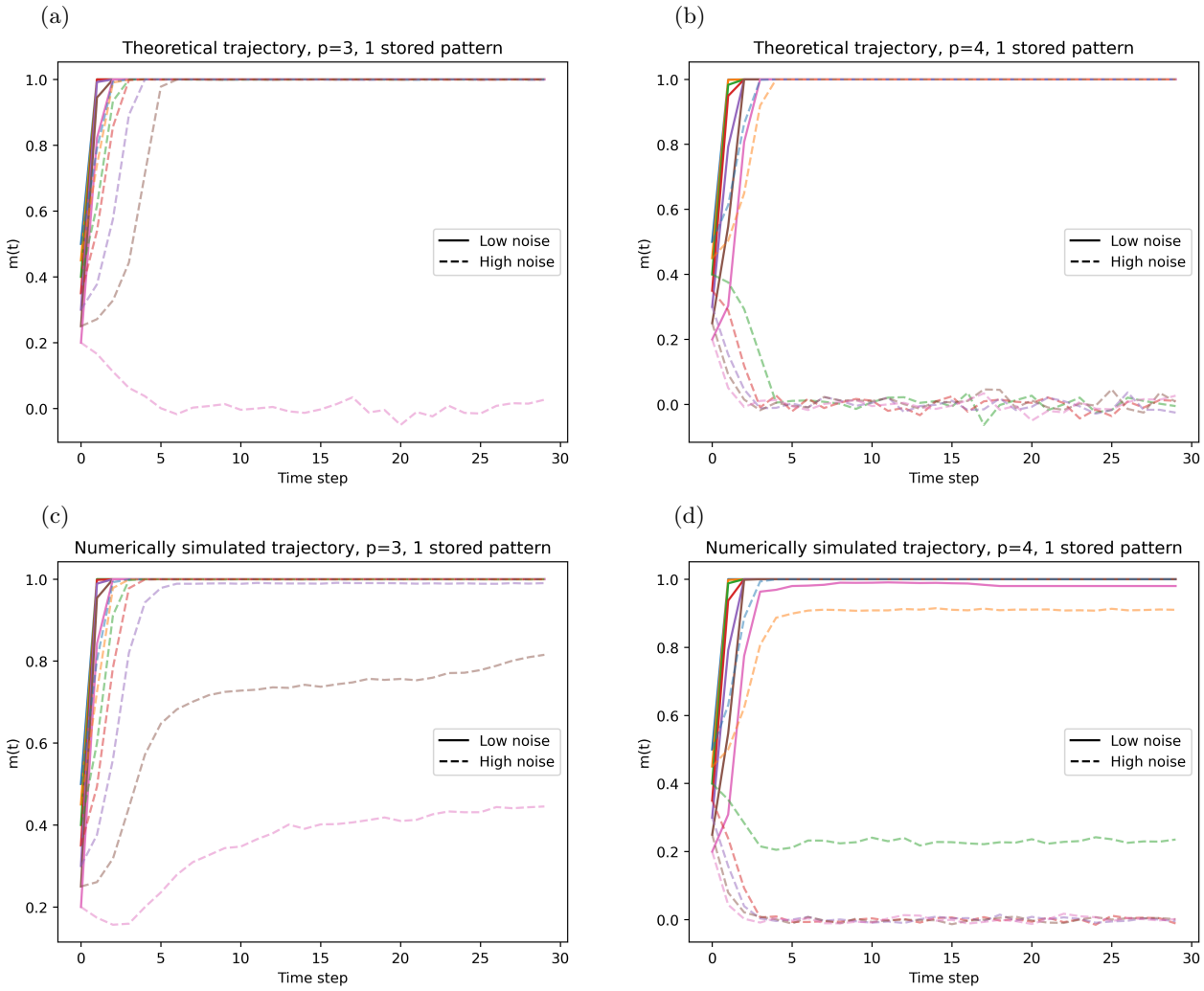

FIG. 1. Numerical and theoretical trajectories for a single stored pattern, with dotted lines showing high noise trajectories, and solid lines showing low noise trajectories. Initial overlaps ranging from 0.2 to 0.5 are shown to reduce visual clutter, and display where trajectories switch from divergence to convergence

V. SIMULATION RESULTS AND DISCUSSION

For each numerical simulation, 100 trials were run for $N = 200$, after which overlaps were averaged. Increasing N to higher values did not noticeably change any results. For the theoretical dynamics, the self consistency equations were iteratively solved using the Monte Carlo method. We defined low noise as $\beta = 10$, and high noise as $\beta = 1.5$. In addition, the bias $\theta(t)$ was set to 0. All simulations were implemented in JAX for random number generation reproducibility and computational speed [12]. We note that numerically simulating trajectories with patterns on the order of 10^5, which is close to the capacity for $p = 4$, was not tractable.

Simulation results agree quite well with the theoretical predictions. In both the regime of low pattern loading (Figs 1a-1d), and high pattern loading (Figs 2a-2d), noise is observed to decrease the radius of attraction and increase the convergence time as expected. The initial overlap which is required for convergence is also captured with accuracy, for both loading regimes and noise levels.

When one pattern is stored, there exists no cross talk noise, so the dominant factor affecting whether a state will converge is competition between the curvature of the energy landscape (i.e. cubic vs quartic) and the finite thermal energy scale. This is most readily seen in the high noise trajectories, where differences in the energy landscapes become exaggerated. For $p = 3$, the convergence time is faster for the initial conditions studied with respect to $p = 4$. In addition, the radius of attraction is larger for $p = 3$.

Interestingly, we observe a reversal behavior when going from a single stored pattern to a non-trivial amount of stored patterns. In this regime, suppressing cross talk relative to the signal term becomes important. For $p = 4$ and 1000 patterns stored, $\alpha_4 \approx 0$, and the so the convergence to the stored pattern is dominated by signal.

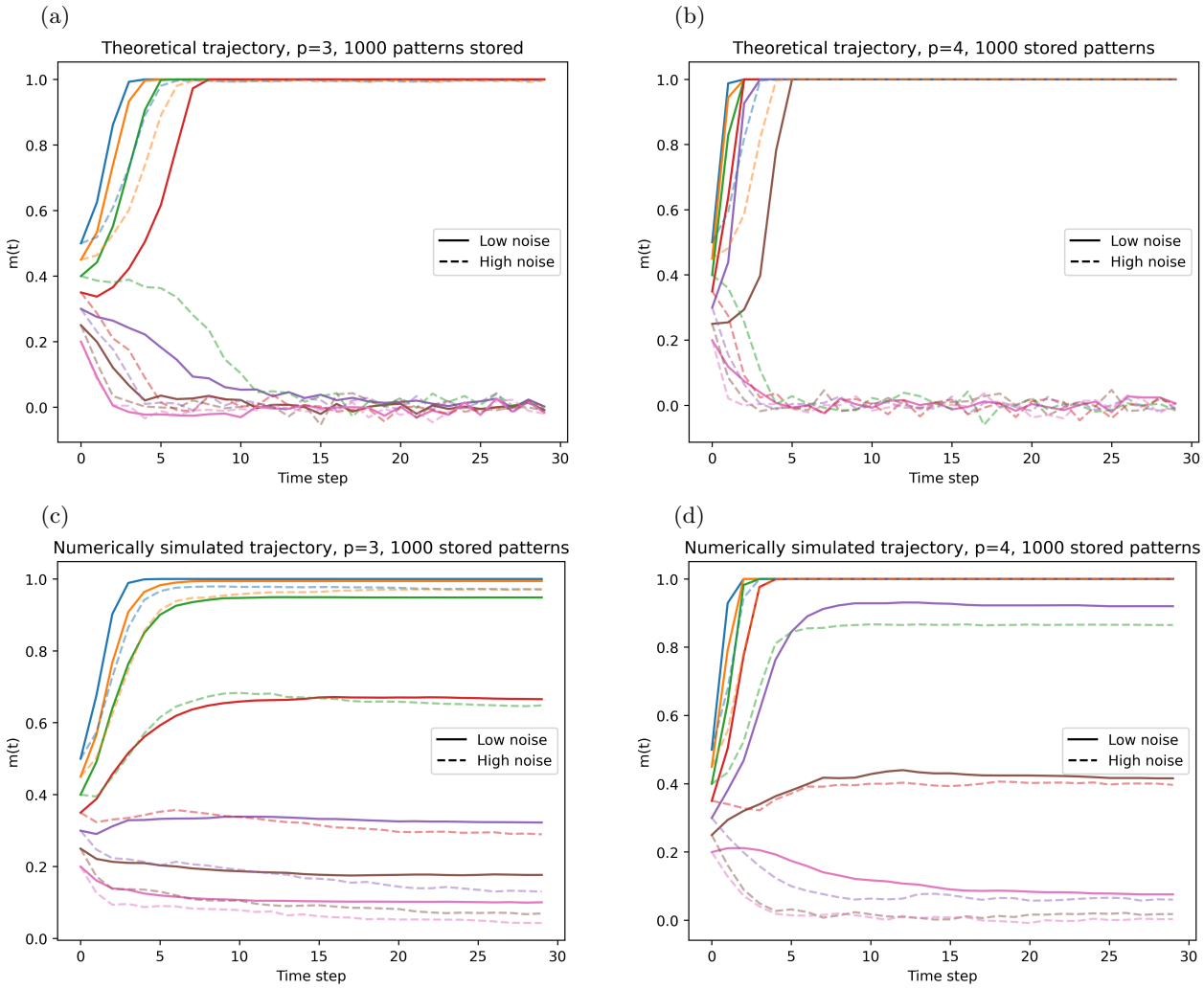

FIG. 2. Numerical and theoretical trajectories for 1000 stored patterns, with dotted lines showing high noise trajectories, and solid lines showing low noise trajectories. Initial overlaps ranging from 0.2 to 0.5 are shown to reduce visual clutter, and display where trajectories switch from divergence to convergence.

As such, we observe that the convergence time is faster for $p = 4$ and the radius of attraction is larger for $p = 4$, contrary to what was described before.

It is worthwhile to note that noise can not only slow convergence (or equivalently, accelerate divergence) towards the stored pattern, but it can also control whether or not an initial state can converge at all. In short, both the noise level and number of patterns stored will influence the desired architecture.

VI. FUTURE WORK

This methodology is broad, and can be applied to a variety of energy based network models that have been developed. Asynchronous updating, which is more biologically plausible [12], is also an interesting modification that may influence trajectory dynamics. In addition, mapping phase transitions in the (α_p, β) space may be useful for network design. An elegant feature of $p = 2$ networks is that they can add new stored patterns via Hebb's rule $J_{ij} \leftarrow J_{ij} + \xi_i^\mu \xi_j^\mu$, where the 'rewiring' process is simple[1]. An analysis of the cost to add more patterns to dense networks could be interesting, especially in its application to studying biological memory.

ACKNOWLEDGMENTS

I wish to thank Professor Kardar and the 8.334 TAs for an enjoyable and rigorous course, and Professor A. K. Chakraborty for an introduction to Hopfield models. Large language models (ChatGPT 4 and Claude) also contributed to the coding development for the theoretical calculation.

[1] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proceedings of the national academy of sciences **79**, 2554 (1982).

[2] C. Cortes, A. Krogh, and J. A. Hertz, Hierarchical associative networks, Journal of Physics A: Mathematical and General **20**, 4449 (1987).

[3] N. Parga and M. Virasoro, The ultrametric organization of memories in a neural network, Journal de Physique **47**, 1857 (1986).

[4] E. Gardner, Multiconnected neural network models, Journal of Physics A: Mathematical and General **20**, 3453 (1987).

[5] E. Gardner, B. Derrida, and P. Mottishaw, Zero temperature parallel dynamics for infinite range spin glasses and neural networks, Journal de Physique **48**, 741 (1987).

[6] D. Krotov and J. J. Hopfield, Dense associative memory for pattern recognition, CoRR **abs/1606.01164** (2016), 1606.01164.

[7] L. F. Abbott and Y. Arian, Storage capacity of generalized networks, Phys. Rev. A **36**, 5091(R) (1987).

[8] H. Ramsauer, B. Schäfl, J. Lehner, P. Seidl, M. Widrich, L. Gruber, M. Holzleitner, M. Pavlovic, G. K. Sandve, V. Greiff, D. P. Kreil, M. Kopp, G. Klambauer, J. Brandstetter, and S. Hochreiter, Hopfield networks is all you need, CoRR **abs/2008.02217** (2020), 2008.02217.

[9] D. Krotov, B. Hoover, P. Ram, and B. Pham, Modern methods in associative memory (2025), arXiv:2507.06211 [cs.LG].

[10] S. Rooke, D. Krotov, V. Balasubramanian, and D. Wolpert, Stochastic thermodynamics of associative memory (2026), arXiv:2601.01253 [cond-mat.stat-mech].

[11] K. Mimura, J. Takeuchi, Y. Sumikawa, Y. Kabashima, and A. C. C. Coolen, Dynamical properties of dense associative memory (2025), arXiv:2506.00851 [cond-mat.dis-nn].

[12] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, Y. Katariya, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).

[13] J. A. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the theory of neural computation* (CRC Press, Boca Raton, FL, 1991).