SYMMETRY-RESTORING RECTIFIED UPDATE RULES FOR HOPFIELD NEURAL NETWORKS: DYNAMICS, THEOREMS, AND HAMMING-DISTANCE AWARENESS

Anonymous authors
Paper under double-blind review

ABSTRACT

Hopfield neural networks (HNNs) provide a classical framework for associative memory, but the binary dynamics are not fully specified when a neuron's local field is zero. Common fixed-sign tie-breaking conventions can introduce an artificial sign preference and break the global sign symmetry inherent in the energy. We show that retaining the neuron's current state at zero field restores this symmetry and yields a principled **Rectified Update Rule**. This symmetry restoration exposes a Hamming-distance geometry underlying the resulting dynamics. For a single stored message, we obtain an exact characterization of convergence and symmetric two-cycles. For two stored messages, we formulate a **Hamming-Distance-Aware Rectified** representation and characterize the complete set of dynamical regimes, including **convergence, self-cycles, hetero-cycles, and symmetric-cycles**. These results further provide exact state counts and convergence-domain characterizations as functions of the network size and inter-memory Hamming distance. Exhaustive enumeration and Monte Carlo experiments agree with the theoretical predictions across a range of network sizes and memory separations.

1 INTRODUCTION

Hopfield neural networks (HNNs) provide a classical framework for associative memory, where stored messages are represented as stable configurations of a recurrent dynamical system. With symmetric interactions, HNNs admit an energy-based formulation that connects network structure, state transitions, and retrieval behavior. However, the energy function alone does not fully determine the evolution of a binary HNN, as the update mechanism can fundamentally affect its long-term dynamics. This is particularly evident under synchronous updates, where trajectories may converge to fixed points or enter cyclic states. Understanding binary HNN dynamics therefore requires not only characterizing the energy landscape, but also precisely specifying the state-transition rule.

Convergence analyses have long shown that HNN dynamics depend on the update schedule. With symmetric weights, Lyapunov-energy arguments establish convergence under asynchronous updates (Peretto, 1984; Komlós & Paturi, 1988); under synchronous updates, trajectories can instead reach fixed points or cycles, including two-step and longer cycles (McEliece et al., 1987). Related work has quantified stable-state counts, retrieval errors, attraction basins, and storage capacity, from classical HNNs to dense and modern variants (Abu-Mostafa & St-Jacques, 1985; Baldi & Venkatesh, 1987; Loukianova, 1997; Bovier, 1999; Krotov & Hopfield, 2016; Bao et al., 2022; Bao & Zhao, 2025). Structural modifications such as dilution and asymmetric couplings can further alter the attractor landscape, producing fixed-point, cyclic, or chaotic regimes (Folli et al., 2018; Hwang et al., 2019). These studies establish the central role of update mechanisms and network structure in HNN dynamics. Yet, a basic issue in synchronous binary HNNs remains insufficiently characterized: how the dynamics should be defined when a neuron's local field is exactly zero.

The standard binary update rule specifies the next state for positive and negative fields but leaves zero-field ties undefined; implementations may resolve them by forcing the neuron to $+1$ or -1. This choice can alter the qualitative dynamics. For example, with $N = 3$, a single stored memory $\xi = (-1, -1, +1)^\top$, and initial state $V = (+1, -1, +1)^\top$, forcing zero-field neurons to $+1$ pro-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

duces a two-cycle with $(-1,+1,+1)^\top$, whereas retaining the current state of the tied neuron sends V directly to the fixed point ξ. Sign reversal yields the analogous discrepancy under a rule that forces ties to -1. Thus, a fixed-sign tie convention can break the global sign symmetry $V \mapsto -V$ of the Hopfield model: sign-reversed states need not have corresponding successors or long-term behavior. The zero-field convention is therefore part of the dynamical definition of a binary HNN, rather than a mere implementation detail, and motivates a symmetry-consistent update rule.

To address this issue, we retain a neuron's current state whenever its local field is zero, while preserving the conventional sign-based update otherwise. We refer to this as the **Rectified Update Rule**. This rule preserves the global sign symmetry of the Hopfield dynamics and provides a structurally consistent state-transition mechanism. Under this formulation, the dynamics exhibit an explicit Hamming-distance structure. For a single stored message, the behavior of any state is characterized by its Hamming distances to the memory and its sign-reversed counterpart. For two stored messages, the corresponding geometry can be captured by a **Hamming-Distance-Aware Rectified (HDAR)** representation, which organizes states according to their distances from the two memories. This representation classifies the resulting dynamics into **convergence, self-cycle, hetero-cycle, and symmetric-cycle** regimes and supports exact characterization of their respective state regions.

Our contributions are summarized as follows:

- **Symmetry-consistent rectification of synchronous dynamics.** We identify the zero-field ambiguity in the classical binary HNN update rule as a source of artificial asymmetry and introduce the Rectified Update Rule, which retains the current state at zero local field while preserving the intrinsic sign symmetry of the model.
- **Hamming-distance-based theoretical characterization.** For HNNs storing one or two messages, we establish the underlying Hamming-distance structure of the rectified dynamics. In the two-message setting, the proposed HDAR formulation enables an exact characterization of convergence and cyclic regimes, including convergence domains and complete convergence domains.
- **Empirical verification of the theoretical results.** Through exhaustive enumeration on small networks and Monte Carlo sampling on larger networks, we verify the predicted state counts, cycle ratios, and dynamical regimes across network sizes and inter-message Hamming distances.

2 Rectified Update Rule

HNNs exemplify recurrent neural networks and are renowned globally as associative memory systems. The connection matrix W serves as the network's "core", encoding all stored memory messages, $\{\xi^1, \xi^2, ..., \xi^M\}$ and fully determining the network's dynamics. Following the original formulation of Hopfield (1982), these messages are embedded using the Hebbian learning rule (Hebb, 1949), with the diagonal elements of W explicitly set to zero to prevent self-connections.

$$W = \sum_{m=1}^{M} \xi^m (\xi^m)^\top - MI, \tag{1}$$

where I is the identity matrix of $N \times N$, N is the dimension of the memory message, M is the number of the memory message, and the element in the i-th row and j-th column of the weight matrix W is denoted as $w_{i,j}$. The original update rule of HNNs is as follows,

$$v_i(t+1) = \begin{cases} +1, & \text{if } \sum_{j=1}^{N} w_{i,j} v_j(t) > 0, \\ -1, & \text{if } \sum_{j=1}^{N} w_{i,j} v_j(t) < 0. \end{cases} \tag{2}$$

Each neuron computes its input summation $X_i(t) = \sum_{j=1}^{N} w_{i,j} v_j(t)$, and then updates its state by comparing $X_i(t)$ to a threshold (typically 0). However, the original rule leaves the case $X_i(t) = 0$ undefined. Some implementations introduce an ad hoc sign (sgn) function that defaults to a fixed state ($+1$ or -1) when encountering this scenario, but such heuristics will disrupt the network's symmetry and destabilize its dynamics, as noted in prior studies (Abu-Mostafa & St-Jacques, 1985;

Sompolinsky & Kanter, 1986; Kanter & Sompolinsky, 1987; McEliece et al., 1987; Komlós & Paturi, 1988; Mazza, 1997; Storkey, 1997; Storkey & Valabregue, 1999; Wang, 2005; Agliari et al., 2013; Krotov & Hopfield, 2016; Demircigil et al., 2017; Folli et al., 2018; Hwang et al., 2019).

$$v_i(t+1) = \begin{cases} +1, & X_i(t) \geq 0, \\ -1, & X_i(t) < 0 \end{cases} \quad \text{or} \quad v_i(t+1) = \begin{cases} +1, & X_i(t) > 0, \\ -1, & X_i(t) \leq 0. \end{cases} \tag{3}$$

To address this issue, we propose a **Rectified Update Rule** as Equation (4), which retains a neuron's previous state when $X_i(t) = 0$, thereby preserving both symmetry and convergence.

$$v_i(t+1) = \begin{cases} +1, & \text{if} X_i(t) > 0, \\ v_i(t), & \text{if} X_i(t) = 0, \\ -1, & \text{if} X_i(t) < 0. \end{cases} \tag{4}$$

The remainder of this paper is centered on this formulation. Rigorous theoretical analysis, supported by extensive simulations, confirms its effectiveness and structural symmetry.

3 Memorizing One Message

Specifically, we designate one randomly chosen pattern as the memory message ξ. We first formalize the notions of convergence and cycle after k steps, then present a theorem in case of memorizing single message. For any initial pattern $V(0)$, let $V(t)$ denote its state after t update steps.

Definition 1 (Fixed Point Convergence)**.** A network state sequence $V(0), V(1), V(2), \ldots$ converges to a fixed point V^* if there exists a minimum time step $t \geq 0$ such that: $V(t+1) = V(t) = V^*$. The pattern stabilizes at V^* after t steps.

Definition 2 (Limit Cycle)**.** A network state sequence $V(0), V(1), V(2), \ldots$ settles into a limit cycle $\mathcal{C}$ if there exists a minimum integer transient time $k \geq 0$ and a minimum integer period $P \geq 2$ such that: $V(k+P) = V(k)$. The cycle is the set $\mathcal{C} = \{V(k), V(k+1), \ldots, V(k+P-1)\}$. The pattern enters the cycle after k steps.

Theorem 1 (Dynamics with a Single Stored Message)**.** *Consider a HNN with N neurons storing a single random binary memory message ξ, and let the weight matrix be $W = \xi\xi^\top - I$, where I is the $N \times N$ identity matrix. Let V be a perturbed version of ξ at Hamming distance r. Under the Rectified Update Rule 4, the network exhibits the following dynamics:*

Case 1:** N **even.

(1) If $r < N/2$, the state V converges to the memory message ξ in a single update.

(2) If $r = N/2$, the state V immediately enters a symmetric 2-cycle between V and $-V$.

(3) If $r > N/2$, the state V converges to the negation $\bar{\xi}$ ($\bar{\xi} = -\xi$) in a single update.

Case 2:** N **odd.

(4) If $r < N/2$, the state V converges to ξ in a single update.

(5) If $r > N/2$, the state V converges to $\bar{\xi}$ in a single update.

Proof. The complete proof is given in Appendix A.1.

The above theorem is visualized in Figure 1 to facilitate readers' understanding. From this result, we immediately deduce the following corollary:

Corollary 2 (Quantity analysis)**.** *When memorizing one memory message ξ, the convergence behavior of HNNs under Rectified Update Rule is influenced by the parity of the network size N:*

- *When N is even, patterns at Hamming distance $N/2$ from the memory message become trapped in a symmetric-cycle, totaling $\binom{N}{N/2}$. The remaining $2^N - \binom{N}{N/2}$ patterns converge to either ξ or its symmetric anti-message $\bar{\xi}$, with each attracting half of these patterns.*
- *When N is odd, all 2^N patterns converge to either ξ or $\bar{\xi}$, each counting 2^{N-1}.*

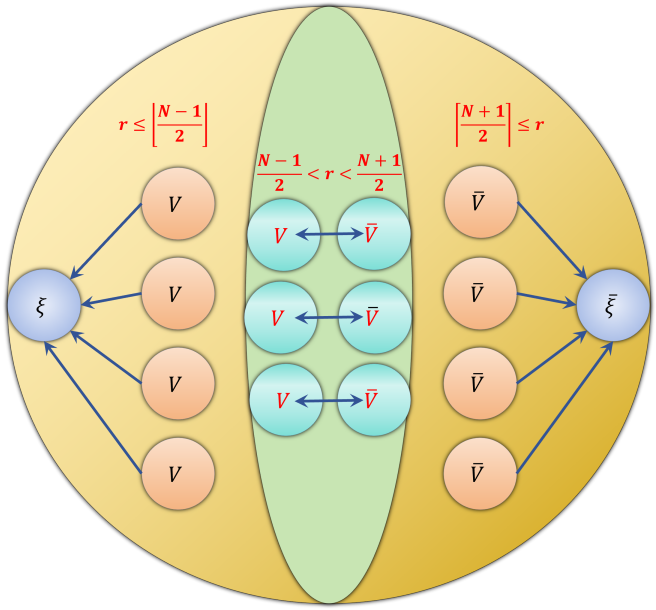

Figure 1: The evolution phenomena when memorizing one memory message. Notations: stable configurations (blue nodes), nodes converging to stable configurations (orange nodes), and nodes belonging to symmetric-cycles (cyan nodes).

Proof. A full derivation appears in Appendix A.2.

Theorem 1 determines the evolution of each pattern from its Hamming distance to the stored message. Corollary 2 extends this result to the complete state space, providing exact counts of convergent and cyclic patterns.

4 Memorizing Two Messages

4.1 Preparation

For any two messages ξ^1 and ξ^2 over N neurons, we decompose ξ^1 into two disjoint parts A_1, where ξ^1 and ξ^2 are identical; and A_2, where ξ^1 and ξ^2 are opposite. Formally denoted as:

$$\xi^1 = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad \xi^2 = \begin{bmatrix} A_1 \\ -A_2 \end{bmatrix}. \tag{5}$$

Given an arbitrary noise pattern V, we further refine A_1 and A_2 based on their alignment with ξ^1. Specifically, within A_1, A_{11} be the neurons where V matches ξ^1; and A_{12}, be the neurons where V is flipped relative to ξ^1. An identical decomposition is applied to A_2, yielding A_{21} and A_{22}. As a result, V can be decomposed into four distinct parts: V_{11}, V_{12}, V_{21} and V_{22}, expressed as:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}, \quad \xi^1 = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}, \quad \xi^2 = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{6}$$

Let r_1 denote the Hamming distance between the noise pattern V and the memory message ξ^1, r_2 the distance between V and ξ^2, and r_{12} the distance between ξ^1 and ξ^2. It is easily obtained that

$$\begin{cases} r_1 = |A_{12}| + |A_{22}|, \\ r_2 = |A_{12}| + |A_{21}|, \\ r_{12} = |A_{21}| + |A_{22}|. \end{cases} \tag{7}$$

$|A_{11}|, |A_{12}|, |A_{21}|, |A_{22}|$ respectively represent the cardinalities of the corresponding subsets introduced earlier.

4.2 Reformulation of the Rectified Update Rule Considering the Hamming Distance

When memorizing two messages ξ^1 and ξ^2, the weight matrix is $W = \xi^1(\xi^1)^\top + \xi^2(\xi^2)^\top - 2I$ in accordance with Equation (1). For any state V, the input to neuron i is $X_i = \sum_{j=1}^{N} w_{i,j} v_j$, thus,

$$X_i = \xi_i^1(\xi^1)^\top V + \xi_i^2(\xi^2)^\top V - 2v_i = \xi_i^1\big(N - 2r_1\big) + \xi_i^2\big(N - 2r_2\big) - 2v_i. \tag{8}$$

By partitioning the neuron indices into those on which ξ^1 and ξ^2 agree, and those where they differ, we decompose the summation into contributions from aligned and anti-aligned components.

$$X_i = \begin{cases} 2\xi_i^1\left(N - r_1 - r_2 - \xi_i^1 v_i\right), & \text{if } i \text{ in } A_1, \\ 2\xi_i^1\left(r_2 - r_1 - \xi_i^1 v_i\right), & \text{if } i \text{ in } A_2. \end{cases} \tag{9}$$

Focusing specifically on the interaction between ξ_i^1 and v_i, we derive an update rule that explicitly integrates Hamming distance into the neuron's transition dynamics. We refer to this refined mechanism as the **HDAR Update Rule**, which unifies our proposed rectified update strategy with a principled approach to managing pattern overlap considering the Hamming distance.

Definition 3 (HDAR update rule)**.** When memorizing two messages ξ^1 and ξ^2, let r_1 denotes the Hamming distance between V and ξ^1, r_2 the distance between V and ξ^2. The HDAR update rule is:

$$V_{11}(t+1) = \begin{cases} V_{11}(t), & \text{if } r_1 + r_2 \le N - 1, \\ -V_{11}(t), & \text{if } r_1 + r_2 > N - 1. \end{cases} \tag{10}$$

$$V_{12}(t+1) = \begin{cases} -V_{12}(t), & \text{if } r_1 + r_2 < N + 1, \\ V_{12}(t), & \text{if } r_1 + r_2 \ge N + 1. \end{cases} \tag{11}$$

$$V_{21}(t+1) = \begin{cases} V_{21}(t), & \text{if } r_2 - r_1 \ge 1, \\ -V_{21}(t), & \text{if } r_2 - r_1 < 1. \end{cases} \tag{12}$$

$$V_{22}(t+1) = \begin{cases} -V_{22}(t), & \text{if } r_2 - r_1 > -1, \\ V_{22}(t), & \text{if } r_2 - r_1 \le -1. \end{cases} \tag{13}$$

4.3 Four Phenomena of Evolution

Under the HDAR update rule, the network's dynamic behavior becomes both transparent and well-delineated: only **fixed-point convergence** and **cyclic trajectories** persist. When memorizing two messages, the network admits precisely four stable configurations—namely, the two original messages and their respective negations. Cyclic behaviors, in turn, can be classified into three distinct categories: **self-cycle, hetero-cycle, and symmetric-cycle**, whose definitions are as follows:

- Self-cycle: The noise pattern O enters a cycle between itself O and another pattern Q, i.e., $O \leftrightarrow Q$.
- Hetero-cycle: The noise pattern V updates in one step to a state O that would lead to a self-cycle, then continues cycling according to the self-cycle behavior, i.e., $V \to O \leftrightarrow Q$.
- Symmetric-cycle: The noise pattern P oscillates between itself P and its anti-pattern $-P$, where they are symmetric, i.e., $P \leftrightarrow -P$.

Based on the cycle classifications above, we can obtain **three conditions for cycle identification**:

Theorem 3 (Cycle in case of two messages)**.** *When memorizing two messages ξ_1 and ξ_2, r_1 denotes the distance between ξ_1 and the noise pattern V, r_2 distance between ξ_2 and V. There is a N neurons HNN, with the HDAR update rule. We can obtain the following three cycle conditions:*

1. *When $r_1 + r_2 \neq N, r_1 = r_2$ (**Condition I**):*
   - *When $r_1 + r_2 \le N - 1$, If the noise pattern V contains V_{12} components, it will enter into hetero-cycle. Otherwise, it will exhibit self-cycle.*
   - *When $r_1 + r_2 \ge N + 1$, If the noise pattern V contains V_{11} components, it will enter into hetero-cycle. Otherwise, it will exhibit self-cycle.*

2. *When $r_1 + r_2 = N, r_1 \neq r_2$ (**Condition II**):*
   - *When $r_2 - r_1 \leq -1$, if the noise pattern V contains V_{21} components, it will enter into hetero-cycle. Otherwise, it will exhibit self-cycle.*
   - *When $r_2 - r_1 \geq 1$, if the noise pattern V contains V_{22} components, it will enter into hetero-cycle. Otherwise, it will exhibit self-cycle.*
3. *When $r_1 + r_2 = N, r_1 = r_2$ (**Condition III**):*
   - *At this time, the noise pattern will enter into a symmetric-cycle.*

Proof. The argument is developed fully in Appendix A.3.

Moreover, we observe that noise pattern evolution exhibits distinct behaviors depending on the parity of N and the Hamming distance r_{12} between ξ^1 and ξ^2.

Corollary 4 (Evolution results with respect to N and r_{12})**.** *Cycle phenomena are governed by whether N and r_{12} are even or odd, leading to predictable cycles under the HDAR update rule, as follows:*

1. *N **even**, r_{12} **even**: All three cycle conditions can be realized by different noise patterns. Consequently, depending on V, the network can exhibit any of the three cycle types: self-cycle, hetero-cycle, or symmetric-cycle. The quantity C_{ee} (C_{ee} refers to the number of noise patterns in which the HNN eventually evolves into a cyclical state when N is even and r_{12} is even, C_{eo}, C_{oo}, C_{oo} is defined similarly) is:*

$$C_{\mathrm{ee}} = 2\left[\binom{N-r_{12}}{\frac{N-r_{12}}{2}} \sum_{i=\frac{N-r_{12}}{2}}^{\frac{N-2}{2}} \binom{r_{12}}{i-\frac{N-r_{12}}{2}} + \binom{r_{12}}{\frac{r_{12}}{2}} \sum_{i=\frac{r_{12}}{2}}^{\frac{N-2}{2}} \binom{N-r_{12}}{i-\frac{r_{12}}{2}}\right] + \binom{r_{12}}{\frac{r_{12}}{2}}\binom{N-r_{12}}{\frac{N-r_{12}}{2}}. \tag{14}$$

2. *N **even**, r_{12} **odd**: None of the cycle conditions can be satisfied. Thus the quantity C_{eo} is 0.*
3. *N **odd**, r_{12} **even**: Cycle Condition I can be satisfied, allowing self-cycles or hetero-cycles to occur. The quantity C_{oe} is:*

$$C_{\mathrm{oe}} = 2\binom{r_{12}}{\frac{r_{12}}{2}} \sum_{i=\frac{r_{12}}{2}}^{\frac{N-1}{2}} \binom{N-r_{12}}{i-\frac{r_{12}}{2}}. \tag{15}$$

4. *N **odd**, r_{12} **odd**: Cycle Condition II can be satisfied, allowing self-cycles or hetero-cycles to occur. The quantity C_{oo} is:*

$$C_{\mathrm{oo}} = 2\binom{N-r_{12}}{\frac{N-r_{12}}{2}} \sum_{i=\frac{N-r_{12}}{2}}^{\frac{N-1}{2}} \binom{r_{12}}{i-\frac{N-r_{12}}{2}}. \tag{16}$$

Proof. Complete details are deferred to Appendix A.4.

Theorem 3 identifies the conditions governing the three types of cycles, while Corollary 4 converts these conditions into exact cycle counts for different parities of N and r_{12}. These results provide the basis for the convergence-domain analysis that follows.

4.4 Convergence and Complete Convergence

We observe that the network's behavior transitions from fixed-point convergence to cyclic trajectories as the Hamming distance to the memory message increases. To capture this transition, we define two critical domains and present some corollaries that precisely quantify their boundaries.

Definition 4 (Convergence domain and complete convergence domain)**.** When memorizing two messages ξ^1 and ξ^2, each message and its anti-message defines a maximal Hamming distance $r_{\max}$, beyond which no pattern converges to that message under HNN dynamics. This delineates the boundary of the **convergence domain (CD)**. We further define the **complete convergence domain**

(CCD) as the set of all patterns within a critical Hamming distance r_{ccd}, where convergence to the corresponding memory message is guaranteed. For $r \leq r_{\text{ccd}}$, the network consistently settles into the correct attractor. In contrast, for $r_{\text{ccd}} < r \leq r_{\max}$, convergence behavior bifurcates—some patterns converge to a memory message, while others fall into cyclic trajectories.

Although prior works informally discussed this idea, it lacked a mathematical formulation. Notably, when $r > r_{\max}$, the evolution trajectories follow three scenarios: self-cycle, hetero-cycle, or symmetric-cycle. In this work, we derive exact expressions for CD and CDD of memory messages.

Corollary 5 (Bound of CCD). *If r_{ccd} satisfies*

$$2r_{\text{ccd}} + 1 \leq r_{12} \leq N - (2r_{\text{ccd}} + 1) \quad \text{and} \quad N \geq 4r_{\text{ccd}} + 2, \tag{17}$$

then any pattern V with $r \leq r_{\text{ccd}}$ will converge to its nearest memory message in finite update steps.

Proof. See Appendix A.5 for the full proof.

Corollary 6 (Relation between two bounds). *The minimum of r_{ccd} is 0, and $r_{\max} > 0$. When N is even and r_{12} is even: $r_{\max} = \frac{N}{2} - 2$. For all other cases: $r_{\max} = \lfloor \frac{N}{2} \rfloor - 1$. The relation between these two bounds satisfies*

$$r_{\text{ccd}} < r_{\max}. \tag{18}$$

Proof. The complete derivation is given in Appendix A.6.

To aid understanding, we visualize the full state space of 2^N patterns as a circle in Figure 2, with two messages and their negations acting as four anchor points. Each anchor defines a CD (shaded cyan) and a CCD (shaded pink). By definition, all patterns within a CCD deterministically converge to the nearest message. Patterns in the CD but outside the CCD display mixed behavior: some converge (pink V), while others enter self-cycles (gray V) or hetero-cycle (red V). Patterns outside any CD may still exhibit self-cycle or hetero-cycle, both within and beyond the CD. The central circular region corresponds to symmetric-cycle (green V), which emerge only when both N and the r_{12} are even. Notably, self-cycle caused by $r_1 = r_2$ occur strictly within the CD, whereas those induced by $r_1 + r_2 = N$ can occur inside or outside.

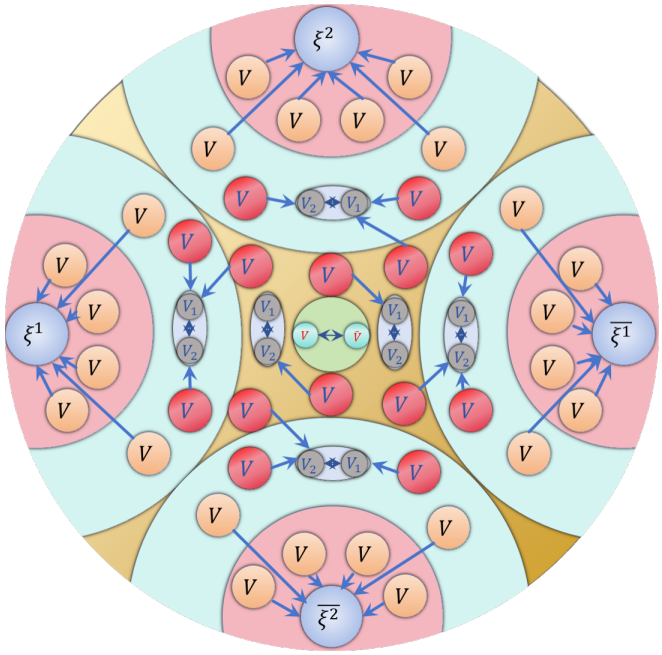

Figure 2: The evolution phenomena when memorizing two messages. Notations: stable configurations (blue V), nodes converging to stable configurations (orange V), nodes belonging to self-cycles (gray V), nodes belonging to hetero-cycles (red V), and nodes belonging to symmetric-cycles (green V); complete convergence domains (shaded pink), convergence domains (shaded cyan).

5 EXPERIMENTS

The experimental environment configuration is as follows: i7-10750H CPU at 2.60 GHz, 16 GB RAM, NVIDIA GTX 1650, and Python 3.7. In the experiment, for a noise pattern update denoted as $V(t)$ after t iterations, given the known cycle period of 2, we only track $V(t)$ and $V(t-1)$. Convergence is determined if $V(t) = V(t+1)$, indicating the noise pattern stabilizes at $V(t)$. A cycle is detected if $V(t-1) = V(t+1)$, signifying the noise is trapped in a cycle: $V(t-1) \to V(t) \to V(t-1) \to V(t)$.

5.1 EXPERIMENTS WHEN MEMORIZING ONE MESSAGE

When memorizing a single memory message, Theorem 1 establishes the existence of two stable configurations: the original message ξ and its symmetric anti-message $\bar{\xi}$. In the case where N is even, an intuitive observation arises: any pattern at a Hamming distance of $N/2$ from ξ is also exactly $N/2$ away from $\bar{\xi}$, rendering it equidistant from both attractors. Consequently, such a pattern lacks a preferential convergence direction and may fail to settle into either attractor. To verify this insight, we conduct the following experiments to confirm the theoretical conclusion.

A random N-bit binary vector is generated as the memory message to define the HNN and construct its weight matrix. All 2^N binary vectors—interpreted as noise patterns—are then individually updated using the **Rectified Update Rule**, and their final states (convergent or cyclic) are recorded. The purpose of the experiment is to verify whether the experimental cycling ratio equals the theoretical value calculated using Corollary 2. The cycling ratio is the number of noise patterns among all 2^N patterns in an HNN that enter a cyclic state, divided by 2^N.

For small networks ($N \leq 20$), exhaustive enumeration is feasible. By iterating over the full state space, we compute the exact number and proportion of patterns trapped in cycles. The results align precisely with theoretical calculations, validating our analytical framework. We also verify that a noise pattern enters a cyclic state precisely when it is at the same distance $N/2$ from ξ and $\bar{\xi}$; detailed examples are provided in **Case 1** of Appendix B.2.

For larger networks ($N > 20$), exhaustive analysis becomes infeasible, so we adopt Monte Carlo sampling. Here, each of the 10,000 sampled patterns is generated by independently setting each bit to ± 1 with equal probability. A randomly generated memory message defines the weight matrix, and our experiments adopt the **Rectified Update Rule**. For each fixed N, the trial is repeated 50 times, recording the cycling ratio each time (the number of sampled patterns out of 10,000 that enter a cyclic state, divided by 10,000). The reported results are the average values of these trials. To ensure statistical validity, both the memory message and the sample set are freshly randomized in each trial. Theoretical values are computed according to Corollary 2, retaining four decimal places. The observed cycling ratios consistently fall within the expected error margins in both regimes ($N \leq 20$ and $N > 20$), confirming the accuracy of our derived formula. For odd N, every tested pattern converges, yielding a convergence ratio of 1 and, equivalently, a cycle ratio of 0. For even N, theoretical and experimental cycle ratios are close; see Table 1(a).

5.2 EXPERIMENTS WHEN MEMORIZING TWO MESSAGES

The experimental settings follow Section 5.2. Each network instead stores two messages at a prescribed r_{12}; theoretical ratios are computed from Corollary 4, and exhaustive runs additionally compare each Rectified-rule outcome with its HDAR-predicted fixed point or cycle type. Every enumerated state reaches exactly the predicted fixed point or cycle type, confirming the theory at the level of individual trajectories rather than only aggregate ratios. The complete protocol is provided in Appendix B.1.

Table 1(b) and (c) demonstrate clear parity-dependent behavior. With odd $r_{12} = 3$, the cycle ratio for odd N decreases from 0.3125 to 0.0253, while it remains zero for even N. With even $r_{12} = 4$, the odd-N ratio remains 0.375, whereas the even-N ratio decreases from 0.5703 toward this value.

Tables 2 and 3 further show that distance scaling does not remove the parity effect. At $r_{12}/N \approx 1/3$, cycle ratios vanish when r_{12} is odd ($N = 100, 1000$) but remain nonzero when r_{12} is even ($N = 200, 500$). At fixed even $r_{12} = 30$, the odd-N ratio remains 0.1445, while the even-N ratio decreases

(a) One stored message

Odd N			Even N		
N	Th.	Exp.	N	Th.	Exp.
3	0	0	4	0.3750	0.3750
7	0	0	8	0.2734	0.2734
15	0	0	16	0.1964	0.1964
19	0	0	20	0.1762	0.1762
⋮	⋮	⋮	⋮	⋮	⋮
99	0	0	100	0.0796	0.0795
199	0	0	200	0.0563	0.0568
499	0	0	500	0.0357	0.0358
999	0	0	1000	0.0252	0.0250

(b) Two messages, $r_{12} = 3$

Odd N			Even N		
N	Th.	Exp.	N	Th.	Exp.
9	0.3125	0.3125	10	0	0
99	0.0812	0.0814	100	0	0
199	0.0569	0.0572	200	0	0
499	0.0358	0.0354	500	0	0
999	0.0253	0.0252	1000	0	0

(c) Two messages, $r_{12} = 4$

Odd N			Even N		
N	Th.	Exp.	N	Th.	Exp.
9	0.375	0.3750	10	0.5703	0.5703
99	0.375	0.3751	100	0.4258	0.4247
199	0.375	0.3753	200	0.4106	0.4095
499	0.375	0.3750	500	0.3974	0.3971
999	0.375	0.3752	1000	0.3908	0.3913

Table 1: Theoretical (Th.) and experimental (Exp.) cycle ratios for one and two stored messages at fixed Hamming distances.

Odd N			Even N		
N	Th.	Exp.	N	Th.	Exp.
99	0.0978	0.0977	100	0	0
199	0.0978	0.0976	200	0.1600	0.1597
499	0.0618	0.0620	500	0.1028	0.1023
999	0.0309	0.0310	1000	0	0

Table 2: Cycle ratios for two stored messages with $r_{12}/N \approx 1/3$.

Odd N			Even N		
N	Th.	Exp.	N	Th.	Exp.
99	0.1445	0.1445	100	0.2258	0.2258
199	0.1445	0.1440	200	0.1967	0.1970
499	0.1445	0.1452	500	0.1759	0.1762
999	0.1445	0.1441	1000	0.1664	0.1668

Table 3: Cycle ratios for two stored messages with $r_{12} = 30$.

from 0.2258 to 0.1664. The maximum theory–experiment discrepancy is only 1.1×10^{-3}; additional protocols, derivations, visualizations, and zero-field results are provided in Appendices B.1–B.5.

6 Conclusion

We address the zero-field ambiguity in synchronous binary HNNs with a Rectified Update Rule that preserves global sign symmetry. For one stored message, we characterize convergence and symmetric two-cycles through Hamming distance and derive the number of stable states. For two stored messages, the HDAR representation organizes the dynamics into convergence, self-cycle, heterocycle, and symmetric-cycle regimes, supporting exact descriptions of their state regions. Exhaustive enumeration and Monte Carlo experiments agree with the theoretical predictions. Together, these results provide a clearer account of how zero-field handling shapes the convergence landscape.

Limitation and future works. This study is limited to synchronous dynamics and networks storing at most two messages. Extending the analysis to more messages will require tracking a larger set of joint Hamming-distance profiles, whose number grows rapidly with the number of memories; asynchronous updates also remain to be examined. Such extensions could clarify how memory interactions and update schedules shape convergence domains and cycle conditions. More broadly, the symmetry-aware, distance-based framework developed here may help future studies compare update conventions on a common basis and formulate more precise analyses of convergence and cycling.

AI USE STATEMENT

Generative AI tools assisted with manuscript editing and organization, including language polishing. The authors reviewed the AI-assisted text and take responsibility for the final content.

REFERENCES

Yaser Abu-Mostafa and Jeannine St-Jacques. Information capacity of the Hopfield model. *IEEE Transactions on Information Theory*, 31(4):461–464, 1985.

Elena Agliari, Adriano Barra, Andrea De Antoni, and Andrea Galluzzi. Parallel retrieval of correlated patterns: From hopfield networks to boltzmann machines. *Neural Networks*, 38:52–63, 2013.

Elena Agliari, Francesco Alemanno, Miriam Aquaro, and Alberto Fachechi. Regularization, early-stopping and dreaming: a hopfield-like setup to address generalization and overfitting. *Neural Networks*, 177:106389, 2024.

Shun-Ichi Amari. Characteristics of random nets of analog neuron-like elements. *IEEE Transactions on systems, man, and cybernetics*, (5):643–657, 1972.

Pierre Baldi and Santosh S Venkatesh. Number of stable points for spin-glasses and neural networks of higher orders. *Physical Review Letters*, 58(9):913, 1987.

Han Bao and Zhongying Zhao. Binary associative memory networks: A review of mathematical framework and capacity analysis. *Information Sciences*, 694:121697, 2025.

Han Bao, Richong Zhang, and Yongyi Mao. The capacity of the dense associative memory networks. *Neurocomputing*, 469:198–208, 2022.

Adriano Barra, Alberto Bernacchia, Enrica Santucci, and Pierluigi Contucci. On the equivalence of Hopfield networks and Boltzmann machines. *Neural Networks*, 34:1–9, 2012.

Anton Bovier. Sharp upper bounds on perfect retrieval in the hopfield model. *Journal of applied probability*, 36(3):941–950, 1999.

Vivien Cabannes, Berfin Simsek, and Alberto Bietti. Learning associative memories with gradient descent. *arXiv preprint arXiv:2402.18724*, 2024.

Tzi-Dar Chiueh and Rodney M Goodman. Recurrent correlation associative memories. *IEEE Transactions on Neural Networks*, 2(2):275–284, 1991.

Mete Demircigil, Judith Heusel, Matthias Löwe, Sven Upgang, and Franck Vermet. On a model of associative memory with huge storage capacity. *Journal of Statistical Physics*, 168(2):288–299, 2017.

Saul José Rodrigues Dos Santos, Vlad Niculae, Daniel C Mcnamee, and Andre Martins. Sparse and structured hopfield networks. In *International Conference on Machine Learning*, pp. 43368–43388. PMLR, 2024.

Viola Folli, Giorgio Gosti, Marco Leonetti, and Giancarlo Ruocco. Effect of dilution in asymmetric recurrent neural networks. *Neural Networks*, 104:50–59, 2018.

Donald Olding Hebb. The organization of behavior: A neuropsychological theory. *New York: Wiley*, 1949.

John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554--2558, 1982.

John J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, 81(10):3088–3092, 1984.

Jerry Yao-Chieh Hu, Donglin Yang, Dennis Wu, Chenwei Xu, Bo-Yu Chen, and Han Liu. On sparse modern hopfield model. *Advances in Neural Information Processing Systems*, 36:27594–27608, 2023.

Jerry Yao-Chieh Hu, Pei-Hsuan Chang, Haozheng Luo, Hong-Yu Chen, Weijian Li, Wei-Po Wang, and Han Liu. Outlier-efficient hopfield layers for large transformer-based models. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 19123–19152, 2024a.

Jerry Yao-Chieh Hu, Bo-Yu Chen, Dennis Wu, Feng Ruan, and Han Liu. Nonparametric modern hopfield models. *arXiv preprint arXiv:2404.03900*, 2024b.

Jerry Yao-Chieh Hu, Thomas Lin, Zhao Song, and Han Liu. On computational limits of modern hopfield models: A fine-grained complexity analysis. *arXiv preprint arXiv:2402.04520*, 2024c.

Jerry Yao-Chieh Hu, Wei-Po Wang, Ammar Gilani, Chenyang Li, Zhao Song, and Han Liu. Fundamental limits of prompt tuning transformers: Universality, capacity and efficiency. *arXiv preprint arXiv:2411.16525*, 2024d.

Jerry Yao-Chieh Hu, Dennis Wu, and Han Liu. Provably optimal memory capacity for modern hopfield models: Transformer-compatible dense associative memories as spherical codes. *Advances in Neural Information Processing Systems*, 37:70693–70729, 2024e.

Jerry Yao-Chieh Hu, Weimin Wu, Yi-Chen Lee, Yu-Chao Huang, Minshuo Chen, and Han Liu. On statistical rates of conditional diffusion transformers: Approximation, estimation and minimax optimality. *arXiv preprint arXiv:2411.17522*, 2024f.

Sungmin Hwang, Viola Folli, Enrico Lanza, Giorgio Parisi, Giancarlo Ruocco, and Francesco Zamponi. On the number of limit cycles in asymmetric neural networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(5):053402, 2019.

Ido Kanter and Haim Sompolinsky. Associative recall of memory without errors. *Physical Review A*, 35(1):380, 1987.

János Komlós and Ramamohan Paturi. Convergence results in an associative memory model. *Neural Networks*, 1(3):239–250, 1988.

Dmitry Krotov and John J Hopfield. Dense associative memory for pattern recognition. In *Advances in Neural Information Processing Systems*, pp. 1172–1180, 2016.

Dmitry Krotov and John J Hopfield. Large associative memory problem in neurobiology and machine learning. In *International Conference on Learning Representations*, 2021.

Xiaoyu Li, Yuanpeng Li, Yingyu Liang, Zhenmei Shi, and Zhao Song. On the expressive power of modern hopfield networks. *arXiv preprint arXiv:2412.05562*, 2024.

Yuchen Liang, Chaitanya K Ryali, Benjamin Hoover, Leopold Grinberg, Saket Navlakha, Dmitry Krotov, and Mohammed J Zaki. Can a fruit fly learn word embeddings? In *International Conference on Learning Representations*, 2021.

William A Little. The existence of persistent states in the brain. *Mathematical biosciences*, 19(1-2): 101–120, 1974.

Daria Loukianova. Lower bounds on the restitution error in the hopfield model. *Probability theory and related fields*, 107(2):161–176, 1997.

Carlo Lucibello and Marc Mézard. Exponential capacity of dense associative memories. *Physical Review Letters*, 132(7):077301, 2024.

Christian Mazza. On the storage capacity of nonlinear neural networks. *Neural Networks*, 10(4): 593–597, 1997.

Robertj McEliece, Edwardc Posner, Eugener Rodemich, and Santoshs Venkatesh. The capacity of the Hopfield associative memory. *IEEE Transactions on Information Theory*, 33(4):461–482, 1987.

Beren Millidge, Tommaso Salvatori, Yuhang Song, Thomas Lukasiewicz, and Rafal Bogacz. Universal Hopfield networks: A general framework for single-shot associative memory models. In *International Conference on Machine Learning*, pp. 15561–15583. PMLR, 2022.

Chang Niu, Huanyu Zhang, Chuanlong Xu, Wenjie Hu, Yunzhuo Wu, Yu Wu, Yadi Wang, Tong Wu, Yi Zhu, Yinyan Zhu, et al. A self-learning magnetic hopfield neural network with intrinsic gradient descent adaption. *Proceedings of the National Academy of Sciences*, 121(51):e2416294121, 2024.

Toshihiro Ota and Ryo Karakida. Attention in a family of boltzmann machines emerging from modern hopfield networks. *Neural Computation*, 35(8):1463–1480, 2023.

Pierre Peretto. Collective properties of neural networks: a statistical physics approach. *Biological cybernetics*, 50(1):51–62, 1984.

Edmund T Rolls. The memory systems of the human brain and generative artificial intelligence. *Heliyon*, 10(11), 2024.

Bernhard Schäfl, Lukas Gruber, Angela Bitto-Nemling, and Sepp Hochreiter. Hopular: Modern Hopfield networks for tabular data. In *NeurIPS 2023 Second Table Representation Learning Workshop*, 2023.

Matthew Smart and Anton Zilman. On the mapping between Hopfield networks and restricted Boltzmann machines. In *International Conference on Learning Representations*, 2021.

Haim Sompolinsky and Ido Kanter. Temporal association in asymmetric neural networks. *Physical review letters*, 57(22):2861, 1986.

Amos Storkey. Increasing the capacity of a hopfield network without sacrificing functionality. In *Artificial Neural Networks—ICANN'97: 7th International Conference Lausanne, Switzerland, October 8–10, 1997 Proceeedings 7*, pp. 451–456. Springer, 1997.

Amos J Storkey and Romain Valabregue. The basins of attraction of a new hopfield learning rule. *Neural Networks*, 12(6):869--876, 1999.

Danil Tyulmankov, Ching Fang, Annapurna Vadaparty, and Guangyu Robert Yang. Biological learning in key-value memory networks. In *Advances in Neural Information Processing Systems*, volume 34, pp. 22247–22258, 2021.

Shouhong Wang. Classification with incomplete survey data: a hopfield neural network approach. *Computers & operations research*, 32(10):2583--2594, 2005.

Dennis Wu, Jerry Yao-Chieh Hu, Teng-Yun Hsiao, and Han Liu. Uniform memory retrieval with larger capacity for modern hopfield models. *arXiv preprint arXiv:2404.03827*, 2024.

Chenwei Xu, Yu Chao Huang, Jerry Yao Chieh Hu, Weijian Li, Ammar Gilani, Hsi Sheng Goan, and Han Liu. Bishop: Bi-directional cellular learning for tabular data with generalized sparse modern hopfield model. *Proceedings of Machine Learning Research*, 235:55048–55075, 2024.

Supplementary Material

This supplementary materials comprises Sections A and Section B. Section A presents detailed proofs of the theorems introduced in the main text. Section B provides additional experiments related to the proposed Rectified and HDAR update rules, demonstrating strong agreement between theoretical predictions and empirical results, while also revealing several interesting phenomena explainable by our theory.

Sections B.3 and B.4 reproduce the complete fixed-ratio and fixed-distance experiments moved from the main text, including their protocols, numerical tables, observations, and derivations.

Section B.5 investigates zero-field probabilities as functions of network size and the number of independently generated stored memories under synchronous binary updates.

Section C collects the full background discussion on HNN memory capacity and convergence dynamics, followed by the broader frontier research discussion. References for these sections are included in the bibliography following the AI Use Statement.

Appendix Contents

A **Proofs of the Main Text** . . . 14
A.1 Proof of Theorem 1: Dynamics with a Single Stored Message . . . 14
A.2 Proof of Corollary 2: Quantity Analysis . . . 15
A.3 Proof of Theorem 3: Cycle in Case of Two Messages . . . 15
A.4 Proof of Corollary 4: Evolution Results with Respect to N and r_{12} . . . 30
A.5 Proof of Corollary 5: Bound of CCD . . . 30
A.6 Proof of Corollary 6: Relation Between Two Bounds . . . 31
B **More Experiments** . . . 33
B.1 Statistical Analyses . . . 33
B.2 Visual Presentation . . . 38
B.2.1 Full-State Outcome Visualizations . . . 42
B.3 Experiments with a Fixed r_{12}/N Ratio . . . 48
B.4 Experiments with Fixed $r_{12} = 30$ . . . 48
B.4.1 Experimental Setup and Results . . . 48
B.4.2 Derivation of the Invariant Cycle Ratios . . . 49
B.5 Zero-Field Probability versus Network Size and Number of Stored Memories . . . 50
C **Extended Related Work** . . . 51
C.1 HNNs and Memory Capacity . . . 51
C.2 Convergence and Cyclic Dynamics . . . 52
C.3 Other Frontier Research . . . 52

A PROOFS OF THE MAIN TEXT

A.1 PROOF OF THEOREM 1: DYNAMICS WITH A SINGLE STORED MESSAGE

Proof. First, let's see the ordinary situation $r = 0$ and $r = N$.

- **Case 1:** $r = 0$. At this point, V is the message ξ. Regardless of whether N is even or odd, ξ will converge to itself after 0 steps.
- **Case 2:** $r = N$. At this point, V is the message $\bar{\xi}$. Regardless of whether N is even or odd, $\bar{\xi}$ will converge to itself after 0 steps.

From these two ordinary situations, it can be observed that ξ and $\bar{\xi}$ are stable configurations. Then we study the other patterns' evolution behaviors. Because the Hamming distance between V and ξ is r, without loss of generality, assume that the initial state V satisfies $V_1 = -\xi^1, \ldots, V_r = -\xi^r$, and $V_i = \xi_i$, for $i = r+1, \ldots, N$. When $r \leq \lfloor N/2 \rfloor$, consider the i-th component of WV under the initialization. For clarity, consider the case $i = 1$ (i.e., $i \leq r$), and observe how the rectified update rule applies.

$$\begin{aligned}(WV)_1 =& 0 + \xi_1\xi_2(-\xi_2) + \cdots + \xi_1\xi_r(-\xi_r) + \\ & \xi_1\xi_{r+1}\xi_{r+1} + \cdots + \xi_1\xi_n\xi_n \\ =& \xi_1\left(0 - \xi_2^2 - \cdots - \xi_r^2 + \xi_{r+1}^2 + \cdots + \xi_n^2\right) \\ =& [N - 1 - 2(r-1)]\xi_1 = (N - 2r + 1)\xi_1. \end{aligned} \tag{A.1}$$

When $i > r$, consider $i = r + 1$ for clarity.

$$\begin{aligned}(WV)_{r+1} =& \xi_{r+1}\xi_1(-\xi_1) + \xi_{r+1}\xi_2(-\xi_2) + \cdots + \\ & \xi_{r+1}\xi_r(-\xi_r) + 0 + \cdots + \xi_{r+1}\xi_n\xi_n \\ =& \xi_{r+1}\left(-\xi_1^2 - \xi_2^2 - \cdots - \xi_r^2 + 0 + \cdots + \xi_n^2\right) \\ =& (N - 1 - 2r)\xi_{r+1}. \end{aligned} \tag{A.2}$$

Since $\lfloor N/2 \rfloor$ depends on whether V is even or odd, we treat these cases separately:

- N **even**: From Equation (A.1), when $1 \leq i \leq \frac{N}{2}$, $[W\,V]_i = \xi_i$. From Equation (A.2), when $\frac{N}{2} + 1 \leq i \leq N$, $[W\,V]_i = -\xi_i$. Hence, the noise pattern V at Hamming distance $N/2$ from the memory message evolves as

$$\begin{aligned} V(0) &= (-\xi_1, \cdots, -\xi_{\frac{N}{2}}, \xi_{\frac{N}{2}+1}, \cdots, \xi_N)^\top \\ \to V(1) &= (\xi_1, \cdots, \xi_{\frac{N}{2}}, -\xi_{\frac{N}{2}+1}, \cdots, -\xi_N)^\top \\ \to V(2) &= (-\xi_1, \cdots, -\xi_{\frac{N}{2}}, \xi_{\frac{N}{2}+1}, \cdots, \xi_N)^\top. \end{aligned} \tag{A.3}$$

- N **odd**: For $1 \leq i \leq \lfloor N/2 \rfloor$, $[W\,V]_i = \xi_i$, and for $\lfloor N/2 \rfloor + 2 \leq i \leq N$, $[W\,V]_i = -\xi_i$, The central neuron $i = \lfloor N/2 \rfloor + 1$ satisfies $[W\,V]_i = 0$, so under the Rectified Update Rule it retains its previous value, preventing cyclic behavior.

The Equation (A.3) shows that when a noisy pattern lies at Hamming distance $N/2$ from the memory message, it becomes trapped in a two-cycle between itself and its negation. Intuitively, this cyclic behavior can be written as $V \to -V \to V \to -V \to \cdots$.

When N is odd, regardless of the initial configuration, one update suffices to resolve the network's state. Any pattern with a Hamming distance less than $\lfloor \frac{N}{2} \rfloor$ from the original memory message ξ converges to the original memory message ξ itself, while patterns with a distance more than $\lfloor \frac{N}{2} \rfloor$ converge to its negation $\bar{\xi}$. □

A.2 Proof of Corollary 2: Quantity Analysis

Proof. According to the theorem *Dynamics with a Single Stored Message* in the main paper, when N is odd, the entire state space of 2^N patterns splits evenly: exactly 2^{N-1} patterns converge to the memory message ξ, and the remaining 2^{N-1} patterns converge to its symmetric anti-message $\bar{\xi}$.

When N is even, all patterns at Hamming distance $N/2$ from ξ become trapped in symmetric-cycle, totaling $\binom{N}{N/2}$ such patterns. The remaining $2^N - \binom{N}{N/2}$ patterns exhibit perfect symmetry: half converge to ξ and half to $\bar{\xi}$. □

A.3 Proof of Theorem 3: Cycle in Case of Two Messages

To prove the theorem *Cycle in Case of Two Messages*, we will enumerate all possible evolutionary outcomes when memorizing two messages, and then organize the three cyclic conditions in the theorem by combining the same conditions. For the convenience of reading, we will rephrase the theorem below.

Theorem 3 (Cycle in case of two messages)**.** *When memorizing two messages ξ^1 and ξ^2, r_1 denotes the distance between ξ^1 and the noise pattern V, r_2 distance between ξ^2 and V. There is a N neurons HNN, with the HDAR update rule. We can obtain the following three cycle conditions:*

1. *When $r_1 + r_2 \neq N, r_1 = r_2$ (**Condition I**):*
   - *When $r_1 + r_2 \leq N - 1$, If the noise pattern V contains V_{12} components, it will enter hetero-cycle. Otherwise, it will exhibit self-cycle.*
   - *When $r_1 + r_2 \geq N + 1$, If the noise pattern V contains V_{11} components, it will enter hetero-cycle. Otherwise, it will exhibit self-cycle.*
2. *When $r_1 + r_2 = N, r_1 \neq r_2$ (**Condition II**):*
   - *When $r_2 - r_1 \leq -1$, if the noise pattern V contains V_{21} components, it will enter hetero-cycle. Otherwise, it will exhibit self-cycle.*
   - *When $r_2 - r_1 \geq 1$, if the noise pattern V contains V_{22} components, it will enter hetero-cycle. Otherwise, it will exhibit self-cycle.*
3. *When $r_1 + r_2 = N, r_1 = r_2$ (**Condition III**):*
   - *At this time, the noise pattern will enter a symmetric-cycle.*

Proof. First, let us clarify the notations:

- $V(t)$ represents the noise pattern after t updates, where $V(0)$ is the original noise pattern. Similarly, $V_{11}(t)$ denotes the V_{11} component after t updates.
- $r(t)$ represents the Hamming distance after t updates, with r being the initial Hamming distance.
- $|A_{11}|, |A_{12}|, |A_{21}|, |A_{22}|$ represent the number of neurons in each part $A_{11}, A_{12}, A_{21}, A_{22}$ respectively.

Please note the following conventions for the ensuing proof:

- **Update Rules:** Every update employs our HDAR update rules (Equations 10–13 in the main text).
- **Repartition After Each Update:** After each application of the update rule, the noise pattern must be re-partitioned according to Equation 6 in the main text. As a result, some subsets $V_{11}, V_{12}, V_{21}, V_{22}$ may "disappear"—i.e., certain indices no longer satisfy that subset's original definition and are absorbed into other subsets.
- **Three-Step Iteration at Each Time Step t:**

1. **Partition:** Divide the current state $V(t)$ into subsets via Equation 6 in the main text.
2. **Compute Distances:** Calculate the current Hamming distances $r_1(t)$ and $r_2(t)$.
3. **Update:** Apply the HDAR update rules to each subset.

- **Terminology:** It should be noted that "unchanged" indicates that the subset retains its previous value during that update; "flipped" indicates that every bit in the subset is inverted.

CASE 1: WHEN THE NOISE PATTERN V CONTAINS ONLY TWO OF THE FOUR COMPONENTS $V_{11}, V_{12}, V_{21}, V_{22}$

Since:

- V_{11} and V_{12} represent the identical parts between ξ^1 and ξ^2.
- V_{21} and V_{22} represent the opposite parts between ξ^1 and ξ^2.
- With $1 \leq r_{12} \leq N-1$.

We have the constraints:

$$\begin{aligned} |V_{11}| + |V_{12}| &\geq 1, \\ |V_{21}| + |V_{22}| &\geq 1. \end{aligned} \tag{A.4}$$

Therefore, the two components must consist of:

- One part from $\{V_{11}, V_{12}\}$ (identical parts).
- And one part from $\{V_{21}, V_{22}\}$ (opposite parts).

1.1 When the pattern V only contains V_{11} and V_{21}: $V = \begin{bmatrix} V_{11} \\ V_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \to \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$.

In this case, $V = \xi^1$, $r_1 = 0$, and $r_2 = |A_{21}|$. Then:

- $r_2 + r_1 = |A_{21}| < N$, i.e., $r_2 + r_1 \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(1) = A_{11}$ (remains unchanged).
- $r_2 - r_1 = |A_{21}| > 0$, i.e., $r_2 - r_1 \geq 1$. By Equations 10–13 in the main text, $V_{21}(1) = A_{21}$ (remains unchanged).

Therefore, $V(1) = V(0)$, which means the noise pattern V converges to ξ^1 after zero update step.

1.2 When the pattern V only contains V_{11} and V_{22}: $V = \begin{bmatrix} V_{11} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{22} \end{bmatrix} \to \begin{bmatrix} A_{11} \\ -A_{22} \end{bmatrix}$.

In this case, $V = \xi^2$, $r_1 = |A_{22}|$, and $r_2 = 0$. Then:

- $r_2 + r_1 = |A_{22}| < N$, i.e., $r_2 + r_1 \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(1) = A_{11}$ (remains unchanged).
- $r_2 - r_1 = -|A_{22}| < 0$, i.e., $r_2 - r_1 \leq -1$. By Equations 10–13 in the main text, $V_{22}(1) = -A_{22}$ (remains unchanged).

Therefore, $V(1) = V(0)$, which means the noise pattern V converges to ξ^2 after zero update step.

1.3 When the pattern V only contains V_{12} and V_{21}: $V = \begin{bmatrix} V_{12} \\ V_{21} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \end{bmatrix} \to \begin{bmatrix} -A_{12} \\ A_{21} \end{bmatrix}$.

In this case, $V = \bar{\xi^2}$, $r_1 = |A_{12}|$, and $r_2 = |A_{12}| + |A_{21}|$. Then:

- $r_2 + r_1 = |A_{12}| + |A_{21}| + |A_{12}| = N + |A_{12}| \geq N + 1$. By Equations 10–13 in the main text, $V_{12}(1) = -A_{12}$ (remains unchanged).

- $r_2 - r_1 = |A_{21}| > 0$, i.e., $r_2 - r_1 \geq 1$. By Equations 10–13 in the main text, $V_{21}(1) = A_{21}$ (remains unchanged).

Therefore, $V(1) = V(0)$, which means the noise pattern V converges to $\bar{\xi^2}$ after zero update step.

1.4 When the pattern V only contains V_{12} and V_{22}: $V = \begin{bmatrix} V_{12} \\ V_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{12} \\ -A_{22} \end{bmatrix}$.

In this case, $V = \bar{\xi^1}$, $r_1 = |A_{12}| + |A_{22}|$, and $r_2 = |A_{12}|$. Then:

- $r_2 + r_1 = |A_{12}| + |A_{12}| + |A_{22}| = |A_{12}| + N \geq N + 1$. By Equations 10–13 in the main text, $V_{12}(1) = -A_{12}$ (remains unchanged).
- $r_2 - r_1 = -|A_{22}| < 0$, i.e., $r_2 - r_1 \leq -1$. By Equations 10–13 in the main text, $V_{22}(1) = -A_{22}$ (remains unchanged).

Therefore, $V(1) = V(0)$, which means the noise pattern V converges to $\bar{\xi^1}$ after zero update step.

CASE 2: WHEN THE NOISE PATTERN V CONTAINS ONLY THREE OF THE FOUR COMPONENTS $V_{11}, V_{12}, V_{21}, V_{22}$

Since any three components can be selected, there are a total of $\binom{4}{3} = 4$ possible combinations.

2.1 When $V = \begin{bmatrix} V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}, \xi^1 = \begin{bmatrix} A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}, \xi^2 = \begin{bmatrix} A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}$.

In this case, $r_1 = |A_{12}| + |A_{22}|$, and $r_2 = |A_{12}| + |A_{21}|$. Then:

- $r_2 + r_1 = |A_{12}| + |A_{21}| + |A_{12}| + |A_{22}| = |A_{12}| + N \geq N + 1$. By Equations 10–13 in the main text, $V_{12}(1) = -A_{12}$ (remains unchanged)

The components of V_{21} and V_{22} need to be categorized and discussed separately:

2.1.1 If $r_2 + r_1 \geq N + 1, -1 < r_2 - r_1 < 1$, i.e., $r_1 = r_2,\ |A_{21}| = |A_{22}|$:

$$V = \begin{bmatrix} V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.5}$$

- $r_2 - r_1 = |A_{21}| - |A_{22}| = 0$, i.e., $-1 < r_2 - r_1 < 1$. By Equations 10–13 in the main text, $V_{21}(1) = -A_{21}$ (flipped); $V_{22}(1) = A_{22}$ (flipped).

Therefore, $V(1) = \begin{bmatrix} -A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix}$,where $r_1(1) = |A_{12}| + |A_{21}|$ and $r_2(1) = |A_{12}| + |A_{22}|$. Then:

- $r_2(1) + r_1(1) = |A_{12}| + |A_{22}| + |A_{12}| + |A_{21}| = |A_{12}| + N \geq N + 1$.. By Equations 10–13 in the main text, $V_{12}(2) = -A_{12}$ (remains unchanged).
- $r_2(1) - r_1(1) = |A_{21}| - |A_{22}| = 0,$ i.e., $-1 < r_2(1) - r_1(1) < 1$.. By Equations 10–13 in the main text, $V_{21}(2) = A_{21}$ (flipped); $V_{22}(2) = -A_{22}$ (flipped).

Therefore, $V(2) = \begin{bmatrix} -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} = V(0)$. The noise pattern V enters a self-cycle after zero update step.

2.1.2 If $r_2 + r_1 \geq N + 1,\ r_2 - r_1 \geq 1$, i.e. $|A_{21}| > |A_{22}|$:

$$V = \begin{bmatrix} V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}. \tag{A.6}$$

- $r_2 - r_1 = |A_{21}| - |A_{22}| > 0$, i.e., $r_2 - r_1 \geq 1$. By Equations 10–13 in the main text, $V_{21}(1) = A_{21}$ (remains unchanged); $V_{22}(1) = A_{22}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}$. At this point, due to the flipping of the V_{22} component, the situation becomes identical to Case 1.3. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to $\bar{\xi^2}$ after zero update step.

2.1.3 If $r_2 + r_1 \geq N + 1, r_2 - r_1 \leq -1$, i.e. $|A_{21}| < |A_{22}|$:

$$V = \begin{bmatrix} V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.7}$$

- $r_2 - r_1 = |A_{21}| - |A_{22}| < 0$, i.e., $r_2 - r_1 \leq -1$. By Equations 10–13 in the main text, $V_{21}(1) = -A_{21}$ (flipped); $V_{22}(1) = -A_{22}$ (remains unchanged).

In conclusion, $V(1) = \begin{bmatrix} -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}$. At this point, due to the flipping of the V_{21} component, the situation becomes identical to Case 1.4. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to $\bar{\xi^1}$ after zero update step.

2.2 When $V = \begin{bmatrix} V_{11} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ -A_{22} \end{bmatrix}, \xi^1 = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{22} \end{bmatrix}, \xi^2 = \begin{bmatrix} A_{11} \\ -A_{21} \\ -A_{22} \end{bmatrix}$.

In this case, $r_1 = |A_{22}|$, and $r_2 = |A_{21}|$. Then:

- $r_2 + r_1 = |A_{21}| + |A_{22}| = N - |A_{11}| \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(1) = A_{11}$ (remains unchanged).

The components of V_{21} and V_{22} need to be categorized and discussed separately:

2.2.1 If $r_2 + r_1 \leq N - 1, -1 < r_2 - r_1 < 1$, i.e., $r_1 = r_2,\ |A_{21}| = |A_{22}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{21} \\ A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.8}$$

- $r_2 - r_1 = |A_{21}| - |A_{22}| = 0$, i.e., $-1 < r_2 - r_1 < 1$. By Equations 10–13 in the main text, $V_{21}(1) = -A_{21}$ (flipped); $V_{22}(1) = A_{22}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ -A_{21} \\ A_{22} \end{bmatrix}$, where $r_1(1) = |A_{21}|$ and $r_2(1) = |A_{22}|$. Then:

- $r_2 + r_1 = |A_{22}| + |A_{21}| = N - |A_{11}| \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(2) = A_{11}$ (remains unchanged).
- $r_2(1) - r_1(1) = |A_{21}| - |A_{22}| = 0$, i.e., $-1 < r_2(1) - r_1(1) < 1$.. By Equations 10–13 in the main text, $V_{21}(2) = A_{21}$ (flipped); $V_{22}(2) = -A_{22}$ (flipped).

Therefore, $V(2) = \begin{bmatrix} A_{11} \\ A_{21} \\ -A_{22} \end{bmatrix} = V(0)$. The noise pattern V enters a self-cycle after zero update step.

2.2.2 If $r_2 + r_1 \leq N - 1,\ r_2 - r_1 \geq 1$, i.e. $|A_{21}| > |A_{22}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{22} \end{bmatrix}. \tag{A.9}$$

- $r_2 - r_1 = |A_{21}| - |A_{22}| > 0$, i.e., $r_2 - r_1 \geq 1$. By Equations 10–13 in the main text, $V_{21}(1) = A_{21}$ (remains unchanged); $V_{22}(1) = A_{22}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ A_{21} \\ A_{22} \end{bmatrix}$. At this point, due to the flipping of the V_{22} component, the situation becomes identical to Case 1.1. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to ξ^1 after zero update step.

2.2.3 If $r_2 + r_1 \leq N - 1,\ r_2 - r_1 \leq -1$, i.e. $|A_{21}| < |A_{22}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{21} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.10}$$

- $r_2 - r_1 = |A_{21}| - |A_{22}| < 0$, i.e., $r_2 - r_1 \leq -1$. By Equations 10–13 in the main text, $V_{21}(1) = -A_{21}$ (flipped); $V_{22}(1) = -A_{22}$ (remains unchanged).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ -A_{21} \\ -A_{22} \end{bmatrix}$. At this point, due to the flipping of the V_{21} component, the situation becomes identical to Case 1.2. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to ξ^2 after zero update step.

2.3 When $V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix}, \xi^1 = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{22} \end{bmatrix}, \xi^2 = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{22} \end{bmatrix}$.

In this case, $r_1 = |A_{12}| + |A_{22}|$, and $r_2 = |A_{12}|$. Then:

- $r_2 - r_1 = -|A_{22}| \leq -1$. By Equations 10–13 in the main text, $V_{22}(1) = -A_{22}$ (remains unchanged).

The components of V_{11} and V_{12} need to be categorized and discussed separately:

2.3.1 If $N - 1 < r_2 + r_1 < N + 1, r_2 - r_1 \leq -1$, i.e., $r_1 + r_2 = N,\ |A_{11}| = |A_{12}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix}. \tag{A.11}$$

- $r_2 + r_1 = |A_{12}| + |A_{22}| + |A_{12}| = |A_{11}| + |A_{22}| + |A_{12}| = N$, i.e., $N - 1 < r_2 + r_1 < N + 1$. By Equations 10–13 in the main text, $V_{11}(1) = -A_{11}$ (flipped); $V_{12}(1) = A_{12}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{22} \end{bmatrix}$, where $r_1(1) = |A_{11}| + |A_{22}|$ and $r_2(1) = |A_{11}|$. Then:

- $r_2(1) - r_1(1) = -|A_{22}| \leq -1$. By Equations 10–13 in the main text, $V_{22}(2) = -A_{22}$ (remains unchanged).

- $r_2(1) + r_1(1) = |A_{11}| + |A_{11}| + |A_{22}| = |A_{11}| + |A_{12}| + |A_{22}| = N$, i.e., $N - 1 < r_2(1) + r_1(1) < N + 1$.. By Equations 10–13 in the main text, $V_{11}(2) = A_{11}$ (flipped); $V_{12}(2) = -A_{12}$ (flipped).

Therefore, $V(2) = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix} = V(0)$. The noise pattern V enters a self-cycle after zero update step.

2.3.2 If $r_2 + r_1 \leq N - 1,\ r_2 - r_1 \leq -1$**, i.e.** $|A_{11}| > |A_{12}|$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{22} \end{bmatrix}. \tag{A.12}$$

- $r_2 + r_1 = |A_{12}| + |A_{22}| + |A_{12}| < |A_{11}| + |A_{22}| + |A_{12}| = N$, i.e., $r_2 + r_1 \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(1) = A_{11}$ (remains unchanged); $V_{12}(1) = A_{12}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{22} \end{bmatrix}$. At this point, due to the flipping of the V_{12} component, the situation becomes identical to Case 1.2. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to ξ^2 after zero update step.

2.3.3 If $r_2 + r_1 \geq N + 1, r_2 - r_1 \leq -1$**, i.e.** $|A_{11}| < |A_{12}|$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix} = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix}. \tag{A.13}$$

.

- $r_2 + r_1 = |A_{12}| + |A_{22}| + |A_{12}| > |A_{11}| + |A_{22}| + |A_{12}| = N$, i.e., $r_2 + r_1 \geq N + 1$. By Equations 10–13 in the main text, $V_{11}(1) = -A_{11}$ (flipped); $V_{12}(1) = -A_{12}$ (remains unchanged).

In conclusion, $V(1) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{22} \end{bmatrix}$. At this point, due to the flipping of the V_{11} component, the situation becomes identical to Case 1.2. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to $\bar{\xi^1}$ after zero update step.

2.4 When $V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix}, \xi^1 = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \end{bmatrix}, \xi^2 = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \end{bmatrix}$.

In this case, $r_1 = |A_{12}|$, and $r_2 = |A_{12}| + |A_{21}|$. Then:

- $r_2 - r_1 = |A_{21}| \geq 1$. By Equations 10–13 in the main text, $V_{21}(1) = A_{21}$ (remains unchanged).

The components of V_{11} and V_{12} need to be categorized and discussed separately:

2.4.1 If $N - 1 \leq r_2 + r_1 \leq N + 1, r_2 - r_1 \geq 1$**, i.e.,** $r_1 + r_2 = N, |A_{11}| = |A_{12}|$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix} = \begin{bmatrix} -A_{11} \\ A_{12} \\ A_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix}. \tag{A.14}$$

- $r_2+r_1 = |A_{12}|+|A_{21}|+|A_{12}| = |A_{11}|+|A_{21}|+|A_{12}| = N$, i.e., $N-1 < r_2+r_1 < N+1$. By Equations 10–13 in the main text, $V_{11}(1) = -A_{11}$ (flipped); $V_{12}(1) = A_{12}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} -A_{11} \\ A_{12} \\ A_{21} \end{bmatrix}$, where $r_1(1) = |A_{11}|$ and $r_2(1) = |A_{11}| + |A_{21}|$. Then:

- $r_2(1) - r_1(1) = |A_{21}| \geq 1$. By Equations 10–13 in the main text, $V_{21}(2) = A_{21}$ (remains unchanged).
- $r_2(1) + r_1(1) = |A_{11}| + |A_{11}| + |A_{21}| = |A_{11}| + |A_{12}| + |A_{21}| = N$, i.e., $N - 1 < r_2(1) + r_1(1) < N + 1$.. By Equations 10–13 in the main text, $V_{11}(2) = A_{11}$ (flipped); $V_{12}(2) = -A_{12}$ (flipped).

Therefore, $V(2) = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix} = V(0)$. The noise pattern V enters a self-cycle after zero update step.

2.4.2 If $r_2 + r_1 \leq N - 1, r_2 - r_1 \geq 1$, i.e. $|A_{11}| > |A_{12}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \end{bmatrix}. \tag{A.15}$$

- $r_2 + r_1 = |A_{12}| + |A_{21}| + |A_{12}| < |A_{11}| + |A_{21}| + |A_{12}| = N$, i.e., $r_2 + r_1 \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(1) = A_{11}$ (remains unchanged); $V_{12}(1) = A_{12}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \end{bmatrix}$. At this point, due to the flipping of the V_{12} component, the situation becomes identical to Case 1.1. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to ξ^1 after zero update step.

2.4.3 If $r_2 + r_1 \geq N + 1, r_2 - r_1 \geq 1$, i.e. $|A_{11}| < |A_{12}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix} = \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix} = \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix}. \tag{A.16}$$

- $r_2 + r_1 = |A_{12}| + |A_{21}| + |A_{12}| > |A_{11}| + |A_{21}| + |A_{12}| = N$, i.e., $r_2 + r_1 \geq N + 1$. By Equations 10–13 in the main text, $V_{11}(1) = -A_{11}$ (flipped); $V_{12}(1) = -A_{12}$ (remains unchanged).

In conclusion, $V(1) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \end{bmatrix}$. At this point, due to the flipping of the V_{11} component, the situation becomes identical to Case 1.3. After one more update, the pattern remains unchanged. Therefore, $V(2) = V(1)$, and the noise pattern V will converge to ξ^2 after zero update step.

CASE 3: ALL FOUR SEGMENTS $V_{11}, V_{12}, V_{21}, V_{22}$ OF THE NOISE PATTERN V APPEAR

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}, \quad \xi^1 = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}, \quad \xi^2 = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.17}$$

We notice that $r_1 = |A_{12}| + |A_{22}|$, $r_2 = |A_{12}| + |A_{21}|$ and $r_{12} = |A_{21}| + |A_{22}|$. As mentioned earlier:

- V_{11} represents the case where noise V is the same as ξ^1, but ξ^1 is the same as ξ^2.
- V_{12} represents the case where noise V is different from ξ^1, but ξ^1 is the same as ξ^2.
- V_{21} represents the case where noise V is the same as ξ^1, but ξ^1 is different from ξ^2.
- V_{22} represents the case where noise V is different from ξ^1, but ξ^1 is different from ξ^2.

When all four components of the noise V exist, each component has two possibilities: remaining unchanged or flipping. According to the HDAR update rule, there should be $2^4 = 16$ possible combinations. However, seven of these cases cannot exist due to contradictory inequality conditions. For example, when V_{11}, V_{12}, V_{21}, V_{22} all remain unchanged, we have

$$1 \leq r_2 - r_1 \leq -1,\ N + 1 \leq r_2 + r_1 \leq N - 1. \tag{A.18}$$

Clearly, this system of inequalities is inconsistent, therefore such cases cannot exist. Here are seven impossible cases:

- $V_{11}(1) = A_{11}, V_{12}(1) = -A_{12}, V_{21}(1) = A_{21}, V_{22}(1) = -A_{22}$ (all unchanged).
- $V_{11}(1) = A_{11}, V_{12}(1) = -A_{12}, V_{21}(1) = A_{21}$ (unchanged). $V_{22}(1) = A_{22}$ (flipped).
- $V_{11}(1) = A_{11}, V_{12}(1) = -A_{12}, V_{22}(1) = -A_{22}$ (unchanged). $V_{21}(1) = -A_{21}$ (flipped).
- $V_{11}(1) = A_{11}, V_{12}(1) = -A_{12}$ (unchanged). $V_{21}(1) = -A_{21}, V_{22}(1) = A_{22}$ (flipped).
- $V_{11}(1) = A_{11}, V_{21}(1) = A_{21}, V_{22}(1) = -A_{22}$ (unchanged). $V_{12}(1) = A_{12}$ (flipped).
- $V_{12}(1) = -A_{12}, V_{21}(1) = A_{21}, V_{22}(1) = -A_{22}$ (unchanged). $V_{11}(1) = -A_{11}$ (flipped).
- $V_{21}(1) = A_{21}, V_{22}(1) = -A_{22}$ (unchanged). $V_{11}(1) = -A_{11}, V_{12}(1) = A_{12}$ (flipped).

This leaves us with the remaining nine cases.

3.1 If $r_2 + r_1 \leq N - 1, r_2 - r_1 \geq 1$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}. \tag{A.19}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = A_{11}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{12}(1) = A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(1) = A_{21}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{22}(1) = A_{22}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}$, where $r_1(1) = 0$ and $r_2(1) = |A_{21}| + |A_{22}|$.

- $r_2(1) + r_1(1) = |A_{21}| + |A_{22}| < N$, i.e., $r_2(1) + r_1(1) \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(2) = A_{11},\ V_{12}(2) = A_{12}$ (remain unchanged).
- $r_2(1) - r_1(1) = |A_{21}| + |A_{22}| > 0$, i.e., $r_2(1) - r_1(1) \geq 1$. By Equations 10–13 in the main text, $V_{21}(2) = A_{21},\ V_{22}(2) = -A_{22}$ (remain unchanged).

Therefore, $V(2) = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} = V(1) = \xi^1$. The noise pattern V will converge to ξ^1 after zero update step.

In fact, for $V(1)$: The flipping of V_{12} and V_{22} leads to the disappearance of these two parts. A_{12} is incorporated into the V_{11} part, and A_{22} is incorporated into the V_{21} part, this situation has effectively become the case described in Case 1.1. After one more update, the pattern remains unchanged.

3.2 If $r_2 + r_1 \leq N - 1, r_2 - r_1 \leq -1$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.20}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = A_{11}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{12}(1) = A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(1) = -A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(1) = -A_{22}$ (remains unchanged).

In conclusion, $V(1) = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}$, where $r_1(1) = |A_{21}| + |A_{22}|$ and $r_2(1) = 0$.

- $r_2(1) + r_1(1) = |A_{21}| + |A_{22}| < N$, i.e., $r_2(1) + r_1(1) \leq N - 1$. By Equations 10–13 in the main text, $V_{11}(2) = A_{11},\ V_{12}(2) = A_{12}$ (remain unchanged).
- $r_2(1) - r_1(1) = -(|A_{21}| + |A_{22}|) < 0$, i.e., $r_2(1) - r_1(1) \leq -1$.By Equations 10–13 in the main text, $V_{21}(2) = -A_{21},\ V_{22}(2) = -A_{22}$ (remain unchanged).

Therefore, $V(2) = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} = V(1) = \xi^2$. The noise pattern V will converge to ξ^2 after zero update step.

In fact, for $V(1)$: The flipping of V_{12} and V_{21} leads to the disappearance of these two parts. A_{12} is incorporated into the V_{11} part, and A_{21} is incorporated into the V_{22} part, this situation has effectively become the case described in Case 1.2. After one more update, the pattern remains unchanged.

3.3 If $r_2 + r_1 \geq N + 1, r_2 - r_1 \leq -1$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.21}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = -A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(1) = -A_{12}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{21}(1) = -A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(1) = -A_{22}$ (remains unchanged).

In conclusion, $V(1) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}$, where $r_1(1) = |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}|$ and $r_2(1) = |A_{11}| + |A_{12}|$.

- $r_2(1) + r_1(1) = |A_{11}| + |A_{12}| + |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}| = |A_{11}| + |A_{12}| + N \geq N + 1$. By Equations 10–13 in the main text, $V_{11}(2) = -A_{11},\ V_{12}(2) = -A_{12}$ (remain unchanged).
- $r_2(1) - r_1(1) = |A_{11}| + |A_{12}| - (|A_{11}| + |A_{12}| + |A_{21}| + |A_{22}|) = -(|A_{21}| + |A_{22}|) \leq -2 < -1$. By Equations 10–13 in the main text, $V_{21}(2) = -A_{21},\ V_{22}(2) = -A_{22}$ (remain unchanged).

Therefore, $V(2) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} = V(1) = \bar{\xi^1}$. The noise pattern V will converge to $\bar{\xi^1}$ after zero update step.

In fact, for $V(1)$: The flipping of V_{11} and V_{21} leads to the disappearance of these two parts. A_{11} is incorporated into the V_{12} part, and A_{21} is incorporated into the V_{22} part, this situation has effectively become the case described in Case 1.4. After one more update, the pattern remains unchanged.

3.4 If $r_2 + r_1 \geq N + 1, r_2 - r_1 \geq 1$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}. \tag{A.22}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = -A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(1) = -A_{12}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{21}(1) = A_{21}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{22}(1) = A_{22}$ (flipped).

In conclusion, $V(1) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}$, where $r_1(1) = |A_{11}| + |A_{12}|$ and $r_2(1) = |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}|$.

- $r_2(1) + r_1(1) = |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}| + |A_{11}| + |A_{12}| = N + |A_{11}| + |A_{12}| \geq N + 1$. By Equations 10–13 in the main text, $V_{11}(2) = -A_{11},\ V_{12}(2) = -A_{12}$ (remain unchanged).

- $r_2(1) - r_1(1) = |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}| - (|A_{11}| + |A_{12}|) = |A_{21}| + |A_{22}| \geq 2 > 1$. By Equations 10–13 in the main text, $V_{21}(2) = A_{21},\ V_{22}(2) = A_{22}$ (remain unchanged).

Therefore, $V(2) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} = V(1) = \bar{\xi^2}$. The noise pattern V will converge to $\bar{\xi^2}$ after zero update step.

In fact, for $V(1)$: The flipping of V_{11} and V_{22} leads to the disappearance of these two parts. A_{11} is incorporated into the V_{12} part, and A_{22} is incorporated into the V_{21} part, this situation has effectively become the case described in Case 1.3. After one more update, the pattern remains unchanged.

3.5 If $N - 1 \leq r_2 + r_1 \leq N + 1, -1 < r_2 - r_1 < 1$**, i.e.** $r_1 = r_2 = N/2$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.23}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = -A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(1) = A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(1) = -A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(1) = A_{22}$ (flipped).

Thus, $V(1) = -V(0)$, where $r_1(1) = |A_{11}| + |A_{21}|$, $r_2(1) = |A_{11}| + |A_{22}|$. Then:

- $r_2(1) + r_1(1) = N$.
- $r_2(1) - r_1(1) = 0$.

At this point:

- By Equations 10–13 in the main text,$V_{11}(2) = A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(2) = -A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(2) = A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(2) = -A_{22}$ (flipped).

Therefore, $V(2) = -V(1) = V(0)$, and the noise pattern enters a symmetric-cycle between its original state and its negated state after zero update step.

3.6 If $r_2 + r_1 \leq N - 1, -1 < r_2 - r_1 < 1$**, i.e.** $r_1 = r_2, |A_{21}| = |A_{22}|$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix}. \tag{A.24}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = A_{11}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{12}(1) = A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(1) = -A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(1) = A_{22}$ (flipped).

Noted that $V(1) = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix}$, where $r_1(1) = |A_{21}|$ and $r_2(1) = |A_{22}|$.

- $r_2(1) + r_1(1) = |A_{22}| + |A_{21}| = N - (|A_{11}| + |A_{12}|) \leq N - 2 < N - 1$. By Equations 10–13 in the main text, $V_{11}(2) = A_{11},\ V_{12}(2) = A_{12}$ (remain unchanged).
- $r_2(1) - r_1(1) = |A_{22}| - |A_{21}| = 0$, i.e., $-1 < r_2(1) - r_1(1) < 1$. By Equations 10–13 in the main text, $V_{21}(2) = A_{21},\ V_{22}(2) = -A_{22}$ (flipped).

In conclusion, $V(2) = \begin{bmatrix} A_{11} \\ A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}$, where $r_1(2) = |A_{22}|$ and $r_2(2) = |A_{21}|$.

- $r_2(2) + r_1(2) = |A_{21}| + |A_{22}| = N - (|A_{11}| + |A_{12}|) \leq N - 2 < N - 1$. By Equations 10–13 in the main text, $V_{11}(3) = A_{11},\ V_{12}(3) = A_{12}$ (remain unchanged).
- $r_2(2) - r_1(2) = |A_{21}| - |A_{22}| = 0$, i.e., $-1 < r_2(2) - r_1(2) < 1$. By Equations 10–13 in the main text, $V_{21}(3) = -A_{21},\ V_{22}(3) = A_{22}$ (flipped).

Therefore, $V(3) = \begin{bmatrix} A_{11} \\ A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix} = V(1)$. The noise pattern V enters a hetero-cycle after one update step.

In fact, for $V(1)$: The flipping of V_{12} leads to the disappearance of this part itself. A_{12} is incorporated into the V_{11} part, given that $|A_{21}| = |A_{22}|$, this situation has effectively become the case described in Case 2.2.1.

3.7 If $r_2 + r_1 \geq N + 1, -1 < r_2 - r_1 < 1$, i.e. $r_1 = r_2, |A_{21}| = |A_{22}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \to \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix} \to \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \to \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix}. \tag{A.25}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = -A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(1) = -A_{12}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{21}(1) = -A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(1) = A_{22}$ (flipped).

Noted that, $V(1) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix}$. where $r_1(1) = |A_{11}| + |A_{12}| + |A_{21}|$ and $r_2(1) = |A_{11}| + |A_{12}| + |A_{22}|$.

- $r_2(1) + r_1(1) = |A_{11}| + |A_{12}| + |A_{22}| + |A_{11}| + |A_{12}| + |A_{21}| = N + |A_{11}| + |A_{12}| \geq N + 2 > N + 1$. By Equations 10–13 in the main text, $V_{11}(2) = -A_{11},\ V_{12}(2) = -A_{12}$ (remain unchanged).
- $r_2(1) - r_1(1) = |A_{22}| - |A_{21}| = 0$, i.e., $-1 < r_2(1) - r_1(1) < 1$. By Equations 10–13 in the main text, $V_{21}(2) = A_{21},\ V_{22}(2) = -A_{22}$ (flipped).

In conclusion, $V(2) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix}$, where $r_1(2) = |A_{11}| + |A_{12}| + |A_{22}|$ and $r_2(2) = |A_{11}| + |A_{12}| + |A_{21}|$.

- $r_2(2) + r_1(2) = |A_{11}| + |A_{12}| + |A_{21}| + |A_{11}| + |A_{12}| + |A_{22}| = N + |A_{11}| + |A_{12}| \geq N + 2 > N + 1$. By Equations 10–13 in the main text, $V_{11}(3) = -A_{11},\ V_{12}(3) = -A_{12}$ (remain unchanged).
- $r_2(2) - r_1(2) = |A_{21}| - |A_{22}| = 0$, i.e., $-1 < r_2(2) - r_1(2) < 1$. By Equations 10–13 in the main text, $V_{21}(3) = -A_{21},\ V_{22}(3) = A_{22}$ (flipped).

Therefore, $V(3) = \begin{bmatrix} -A_{11} \\ -A_{12} \\ -A_{21} \\ A_{22} \end{bmatrix} = V(1)$. The noise pattern V enters a hetero-cycle after one update step.

3.8 If $N - 1 < r_2 + r_1 < N + 1, r_2 - r_1 \geq 1$, i.e. $r_1 + r_2 = N, |A_{11}| = |A_{12}|$:

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}. \tag{A.26}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = -A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(1) = A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(1) = A_{21}$ (remains unchanged).
- By Equations 10–13 in the main text,$V_{22}(1) = A_{22}$ (flipped).

Noted that $V(1) = \begin{bmatrix} -A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}$, where $r_1(1) = |A_{11}|$ and $r_2(1) = |A_{11}| + |A_{21}| + |A_{22}|$.

- $r_2(1) + r_1(1) = |A_{11}| + |A_{21}| + |A_{22}| + |A_{11}| = |A_{11}| + |A_{21}| + |A_{22}| + |A_{12}| = N$, i.e., $N - 1 < r_2(1) + r_1(1) < N + 1$. By Equations 10–13 in the main text, $V_{11}(2) = A_{11}$, $V_{12}(2) = -A_{12}$ (flipped).

- $r_2(1) - r_1(1) = |A_{21}| + |A_{22}| \geq 2 \geq 1$. By Equations 10–13 in the main text, $V_{21}(2) = A_{21}, V_{22}(2) = A_{22}$ (remain unchanged).

In conclusion, $V(2) = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix}$, where $r_1(2) = |A_{12}|$ and $r_2(2) = |A_{12}| + |A_{21}| + |A_{22}|$.

- $r_2(2) + r_1(2) = |A_{12}| + |A_{21}| + |A_{22}| + |A_{12}| = |A_{12}| + |A_{21}| + |A_{22}| + |A_{11}| = N$, i.e., $N - 1 < r_2(2) + r_1(2) < N + 1$. By Equations 10–13 in the main text, $V_{11}(3) = -A_{11}$, $V_{12}(3) = A_{12}$ (flipped).
- $r_2(2) - r_1(2) = |A_{21}| + |A_{22}| \geq 2 \geq 1$. By Equations 10–13 in the main text, $V_{21}(3) = A_{21}, V_{22}(3) = A_{22}$ (remain unchanged).

Therefore, $V(3) = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ A_{22} \end{bmatrix} = V(1)$. The noise pattern V enters a hetero-cycle after one update step.

In fact, for $V(1)$: The flipping of V_{22} leads to the disappearance of this part itself. A_{22} is incorporated into the V_{21} part, given that $|A_{11}| = |A_{12}|$, this situation has effectively become the case described in Case 2.4.1.

3.9 If $N - 1 < r_2 + r_1 < N + 1, r_2 - r_1 \leq -1$**, i.e.** $r_1 + r_2 = N, |A_{11}| = |A_{12}|$**:**

$$V = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{21} \\ V_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ -A_{12} \\ A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}. \tag{A.27}$$

At this point:

- By Equations 10–13 in the main text,$V_{11}(1) = -A_{11}$ (flipped).
- By Equations 10–13 in the main text,$V_{12}(1) = A_{12}$ (flipped).
- By Equations 10–13 in the main text,$V_{21}(1) = -A_{21}$ (flipped).
- By Equations 10–13 in the main text,$V_{22}(1) = -A_{22}$ (remains unchanged).

Noted that $V(1) = \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}$, where $r_1(1) = |A_{11}| + |A_{21}| + |A_{22}|$ and $r_2(1) = |A_{11}|$.

- $r_2(1) + r_1(1) = |A_{11}| + |A_{11}| + |A_{21}| + |A_{22}| = |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}| = N$, i.e., $N - 1 < r_2(1) + r_1(1) < N + 1$. By Equations 10–13 in the main text, $V_{11}(2) = A_{11}$, $V_{12}(2) = -A_{12}$ (flipped).
- $r_2(1) - r_1(1) = -(|A_{21}| + |A_{22}|) \leq -2 \leq -1$. By Equations 10–13 in the main text, $V_{21}(2) = -A_{21}, V_{22}(2) = -A_{22}$ (remain unchanged).

In conclusion, $V(2) = \begin{bmatrix} A_{11} \\ -A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix}$, where $r_1(2) = |A_{12}| + |A_{21}| + |A_{22}|$ and $r_2(2) = |A_{12}|$.

- $r_2(2) + r_1(2) = |A_{12}| + |A_{12}| + |A_{21}| + |A_{22}| = |A_{11}| + |A_{12}| + |A_{21}| + |A_{22}| = N$, i.e., $N - 1 < r_2(2) + r_1(2) < N + 1$. By Equations 10–13 in the main text, $V_{11}(3) = -A_{11}$, $V_{12}(3) = A_{12}$ (flipped).
- $r_2(2) - r_1(2) = -(|A_{21}| + |A_{22}|) \leq -2 \leq -1$. By Equations 10–13 in the main text, $V_{21}(3) = -A_{21}$, $V_{22}(3) = -A_{22}$ (remain unchanged).

Therefore, $V(3) = \begin{bmatrix} -A_{11} \\ A_{12} \\ -A_{21} \\ -A_{22} \end{bmatrix} = V(1)$. The noise pattern enters a hetero-cycle after one update step.

In fact, for $V(1)$: The flipping of V_{21} leads to the disappearance of this part itself. A_{21} is incorporated into the V_{22} part, given that $|A_{11}| = |A_{12}|$, this situation has effectively become the case described in Case 2.3.1.

From these results, we have identified three fundamental conditions required for the noise pattern to enter a cyclical regime. The complete analytical results are documented in Theorem 3. □

A.4 Proof of Corollary 4: Evolution Results with Respect to N and r_{12}

Proof. **1. N even, r_{12} even.**

Explanation: The derivation approach for the first and second types of cycling is the same as described above, with the only difference being that the upper limit for i can only go up to $N-2/2$. This is because when N is even and r_{12} is even, symmetric cycling exists, thus requiring the upper limit to be reduced by one. For the third type of cycling (symmetric-cycle), the conditions $r_1 = r_2 = N/2$ must be satisfied. To achieve $r_1 = r_2$, we need $|V_{21}| = |V_{22}| = r_{12}/2$. This means that among the r_{12} terms where ξ^1 and ξ^2 differ, exactly $r_{12}/2$ terms must be selected. Similarly, among the $N - r_{12}$ terms where ξ^1 and ξ^2 are identical, exactly $(N - r_{12})/2$ terms should be selected.

2. N even, r_{12} odd.

Explanation: It is obviously that $C_{\text{eo}} = 0$.

3. N odd, r_{12} even.

Explanation: Since $r_1 = |V_{12}| + |V_{22}|$ and $r_2 = |V_{12}| + |V_{21}|$, while $|V_{21}| + |V_{22}| = r_{12}$. To enter the first type of circling, the condition $r_1 = r_2$ must be satisfied, which implies $|V_{21}| = |V_{22}| = r_{12}/2$. Therefore, if we use i to represent the Hamming distance from the noise to ξ^1, namely r_1, it is clear that i should start from $r_{12}/2$ and end at $\frac{N-1}{2}$. Regardless of the value of i, among the r_{12} terms where ξ^1 and ξ^2 differ, $r_{12}/2$ terms should be selected. Among the $N - r_{12}$ terms where ξ^1 and ξ^2 are identical, $i - r_{12}/2$ terms should be selected. Due to the symmetry of the network, the final result multiplied by 2 gives the number of noise configurations that enter the circling state.

4. N odd, r_{12} odd.

Explanation: Given $r_1 = |V_{12}| + |V_{22}|$ and $r_2 = |V_{12}| + |V_{21}|$, while $|V_{21}| + |V_{22}| = r_{12}$. To enter the second type of circling, the condition $r_1 + r_2 = N$ must be satisfied, which implies $|V_{11}| = |V_{12}| = (N - r_{12})/2$. Therefore, if we use i to represent the Hamming distance from the noise to ξ^1, namely r_1, it is clear that i should start from $(N - r_{12})/2$ and end at $(N - 1)/2$. Regardless of the value of i, among the $N - r_{12}$ terms where ξ^1 and ξ^2 are identical, $(N - r_{12})/2$ terms should be selected. Among the r_{12} terms where ξ^1 and ξ^2 differ, $i-(N - r_{12})/2$ terms should be selected. Due to the symmetry of the network, the final result multiplied by 2 gives the number of noise configurations that enter the circling state. □

A.5 Proof of Corollary 5: Bound of CCD

Proof. From the subsection *Preparation* and the definition *HDAR update rule* in the main paper, we know that for a noise pattern to converge to a specific memory message, the components opposite to the memory message should be flipped while the identical components remain unchanged. Taking ξ_1 as an example, this requires keeping v_{11} and v_{21} unchanged while flipping v_{12} and v_{22}. Based on this condition, we obtain four inequality groups:

$$r_2 - r_1 \geq 1, r_2 + r_1 \leq N - 1, r_2 - r_1 > -1, r_2 + r_1 < N + 1. \tag{A.28}$$

Taking the intersection yields:

$$r_2 - r_1 \geq 1, \quad r_2 + r_1 \leq N - 1. \tag{A.29}$$

Since

$$r_1 = |V_{12}| + |V_{22}|, r_2 = |V_{12}| + |V_{21}|, r_{12} = |V_{21}| + |V_{22}|. \tag{A.30}$$

The inequalities can then be transformed as follows:

$$\begin{aligned} r_2 - r_1 \geq 1 &\rightarrow |V_{21}| - |V_{22}| \geq 1 \\ &\rightarrow |V_{21}| + |V_{22}| \geq 1 + 2|V_{22}| \\ &\rightarrow r_{12} \geq 1 + 2|V_{22}|. \end{aligned} \tag{A.31}$$

$$\begin{aligned} r_2 + r_1 \leq N - 1 &\rightarrow |V_{21}| + |V_{22}| + 2|V_{12}| \leq N - 1 \\ &\rightarrow r_{12} \leq N - 1 - 2|V_{12}|. \end{aligned} \tag{A.32}$$

Taking the intersection gives:

$$1 + 2\left|V_{22}\right| \leq r_{12} \leq N - 1 - 2\left|V_{12}\right|. \tag{A.33}$$

To ensure all noise patterns with Hamming distance r_1 from ξ_1 converge to ξ_1, we consider the minimal range case:

- When all opposite components are in $\left|V_{22}\right|$, r_{12} attains its maximal lower bound $1 + 2r_1$.
- When all opposite components are in $\left|V_{12}\right|$, r_{12} attains its minimal upper bound $N-1-2r_1$.

Thus we obtain the tightest convergence bound r_1, of which is r_{ccd}:

$$1 + 2r_{\mathrm{ccd}} \leq r_{12} \leq N - 1 - 2r_{\mathrm{ccd}}. \tag{A.34}$$

Furthermore, this requires the system size constraint:

$$N \geq 4r_{\mathrm{ccd}} + 2. \tag{A.35}$$

From this, we can conclude that if r_{ccd} satisfies $1 + 2r_{\mathrm{ccd}} \leq r_{12} \leq N - 1 - 2r_{\mathrm{ccd}}$, $N \geq 4r_{\mathrm{ccd}} + 2$, then all noise patterns with a Hamming distance $r \leq r_{\mathrm{ccd}}$ from the target message will converge to it. Here, r corresponds to the critical threshold r_{ccd} mentioned in the definition *Convergence domain and complete convergence domain* in the main paper. □

A.6 Proof of Corollary 6: Relation Between Two Bounds

Proof. First, we obtain the bounds of r. According to Equation (A.33), for the noise pattern V to converge to ξ^1, it holds:

$$\begin{aligned} 1 + 2|V_{22}| \leq N - 1 - 2|V_{12}| &\rightarrow 2|V_{12}| + 2|V_{22}| + 2 \leq N \\ &\rightarrow 2r_1 + 2 \leq N \rightarrow r_1 \leq \frac{N}{2} - 1. \end{aligned} \tag{A.36}$$

Only when $r_1 \leq \frac{N}{2} - 1$ is satisfied can the noise pattern V potentially converge to ξ^1. Considering the case where N is odd, we have: $r_1 \leq \lfloor \frac{N}{2} \rfloor - 1$. However, there is an exception: when N is even and r_{12} is even, the noise pattern V can only converge to ξ^1 if $r_1 \leq \frac{N}{2} - 2$. The proof is as follows:

From the above Equations (A.31) and A.32,

$$|V_{21}| - |V_{22}| \geq 1, \quad |V_{21}| + |V_{22}| + 2|V_{12}| \leq N - 1. \tag{A.37}$$

Assuming $|V_{21}| - |V_{22}| = 1$, then $|V_{21}| = |V_{22}| + 1$, and $r_{12} = |V_{21}| + |V_{22}| = 2|V_{22}| + 1$. In this case, r_{12} must be odd, which contradicts the assumption. Therefore,

$$\begin{aligned} & |V_{21}| + |V_{22}| + 2|V_{12}| < N - 1 \\ \rightarrow \quad & |V_{21}| + |V_{22}| + 2|V_{12}| \leq N - 2 \\ \rightarrow \quad & r_{12} \leq N - 2 - 2|V_{12}|. \end{aligned} \tag{A.38}$$

Similarly, assuming $|V_{21}| + |V_{22}| + 2|V_{12}| = N - 1$, then $r_{12} = |V_{21}| + |V_{22}| = N - 1 - 2|V_{12}|$. Here, r_{12} is also odd, which contradicts the assumption. Thus,

$$\begin{aligned} & |V_{21}| + |V_{22}| + 2|V_{12}| < N - 1 \\ \rightarrow \quad & |V_{21}| + |V_{22}| + 2|V_{12}| \leq N - 2 \\ \rightarrow \quad & r_{12} \leq N - 2 - 2|V_{12}|. \end{aligned} \tag{A.39}$$

In conclusion,

$$2 + 2|V_{22}| \leq r_{12} \leq N - 2 - 2|V_{12}|. \tag{A.40}$$

Using a similar proof method, it can be concluded that in this case, the condition $r_1 \leq \frac{N}{2} - 2$ must be satisfied. Therefore,

- When N is even and r_{12} is even, the noise pattern can only converge to the message if the Hamming distance r satisfies $r \leq \frac{N}{2} - 2$.

- In all other cases, the condition $r \leq \left\lfloor \frac{N}{2} \right\rfloor - 1$ must be satisfied.

Here, r corresponds to the $r_{\max}$ mentioned in the definition *Convergence domain and complete convergence domain* in the main paper. That is:

When N is even and r_{12} is even:

$$r_{\max} = \frac{N}{2} - 2. \tag{A.41}$$

For all other cases:

$$r_{\max} = \left\lfloor \frac{N}{2} \right\rfloor - 1. \tag{A.42}$$

According to the definition of CD and CCD, the memory message ξ converges to itself. Thus the minimum of r_{ccd} is 0, and $r_{\max} > 0$. To prove $r_{\text{ccd}} < r_{\max}$, we enumerate all the situations of r_{ccd}.

(1) When $r_{\text{ccd}} = 0$:

Since the convergence domain $r_{\max}$ is clearly greater than 0, we have $r_{\text{ccd}} < r_{\max}$.

(2) When $r_{\text{ccd}} = 1$:

According to the corollary *Bound of CCD* in the main paper, N satisfies

$$N \geq 6, \quad 3 \leq r_{12} \leq N - 3. \tag{A.43}$$

For the case $N = 6$, we should have $r_{12} = 3$. Then according to Equation (A.42):

$$r_{\max} = \left\lfloor \frac{N}{2} \right\rfloor - 1 = 2 > r_{\text{ccd}} = 1. \tag{A.44}$$

For the case when $N > 6$:

- If N is even and r_{12} is even, according to Equation (A.41):

$$r_{\max} = \frac{N}{2} - 2 > 1 > r_{\text{ccd}} = 1. \tag{A.45}$$

- For other combinations of parity for N and r_{12}, according to Equation (A.42):

$$r_{\max} = \left\lfloor \frac{N}{2} \right\rfloor - 1 > 2 > r_{\text{ccd}} = 1. \tag{A.46}$$

Therefore, when the complete convergence domain r_{ccd} is 1, we have $r_{\text{ccd}} < r_{\max}$.

(3) When $r_{\text{ccd}} \geq 2$:

According to Equation (A.40), N satisfies:

$$N \geq 4r_{\text{ccd}} + 2, \quad 1 + 2r_{\text{ccd}} \leq r_{12} \leq N - 1 - 2r_{\text{ccd}}. \tag{A.47}$$

For cases when N is even and r_{12} is even, according to Equation (A.41):

$$r_{\max} = \frac{N}{2} - 2 \geq 2r_{\text{ccd}} - 1 \geq r_{\text{ccd}} + 1 > r_{\text{ccd}}. \tag{A.48}$$

For other cases:

$$r_{\max} = \left\lfloor \frac{N}{2} \right\rfloor - 1 \geq 2r_{\text{ccd}} > r_{\text{ccd}}. \tag{A.49}$$

Therefore, when the complete convergence domain r_{ccd} is any number greater than or equal to 2, we have $r_{\text{ccd}} < r_{\max}$.

In conclusion, for any complete convergence domain r_{ccd}, the inequality $r_{\text{ccd}} < r_{\max}$ always holds true. □

B MORE EXPERIMENTS

B.1 STATISTICAL ANALYSES

The experiments with two stored messages verify whether theoretical cycle ratios from Corollary 4 agree with Rectified-rule simulations across network sizes N and message distances r_{12}, and whether the HDAR rule correctly predicts the complete outcome landscape. For small networks ($N \leq 20$), we first prescribe an odd or even r_{12}, randomly generate two N-bit messages at that distance, and construct the weight matrix. We exhaustively evolve all 2^N noise patterns under the Rectified Update Rule and compare each observed fixed point or cycle type (self-cycle, hetero-cycle, or symmetric-cycle) with its HDAR prediction.

For large networks ($N > 20$), we use the same Monte Carlo setting as in Section 6.1 of the main text. Each of 50 independent trials regenerates a message pair with the prescribed r_{12} and 10,000 noise vectors whose bits are independently sampled from $\{-1, +1\}$ with equal probability. We evolve the noise vectors under the Rectified Update Rule, record the fraction entering cycles, and report the average over the 50 trials. Thus, the two-message experiments differ from the one-message experiments only in the prescribed relation between the two stored messages and the additional HDAR-based outcome verification.

We verify our experimental results using statistical error testing methods. Below are some statistics of essential calculation formulas:

Formula 1:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}. \tag{B.1}$$

This formula calculates the sample standard deviation of experimental values, where:

- x_i represents the i-th experimental value.
- n represents the sample size.
- $\bar{x}$ represents the mean of n experimental values.

Formula 2:

$$\mathrm{SE} = \frac{s}{\sqrt{n}}. \tag{B.2}$$

This formula calculates the standard error of experimental values, where:

- s represents the sample standard deviation of experimental values.
- n represents the sample size.

Formula 3:

$$\mathrm{Z} = \frac{\bar{x} - x_{\mathrm{t}}}{\mathrm{SE}}. \tag{B.3}$$

This formula calculates the standardized error (Z-score) of experimental values, where:

- $\bar{x}$ represents the mean of n experimental values.
- x_{t} represents the theoretical value for the experiment.
- SE represents the standard error of experimental values.

Formula 4:

$$\mathrm{d} = \frac{|\bar{x} - x_{\mathrm{t}}|}{\sigma}. \tag{B.4}$$

This formula calculates Cohen's d effect size for experimental values, where:

- $\bar{x}$ represents the mean of n experimental values.
- x_{t} represents the theoretical value for the experiment.
- σ represents the population standard deviation of experimental values.
- Effect size interpretation:
  - $\mathrm{d} < 0.2$: negligible difference.
  - $0.2 \leq \mathrm{d} < 0.5$: small difference.
  - $0.5 \leq \mathrm{d} < 0.8$: medium difference.
  - $\mathrm{d} \geq 0.8$: large difference.

The original definition of Cohen's d uses population standard deviation (σ). However, since sample standard deviation s is an unbiased estimator of σ (especially when n is large), and in practice σ is usually unknown, we typically substitute it with sample standard deviation s. In our experiments, since we cannot exhaust all possible noise patterns, we also use sample standard deviation s instead of population standard deviation (σ). Therefore, the practical version of Formula 4 becomes:

$$\mathrm{d} = \frac{|\bar{x} - x_{\mathrm{t}}|}{s}. \tag{B.5}$$

Taking $N = 9$, $r_{12} = 3$ in Table I(b) in the main text as an example for analysis:

- Null hypothesis H_0: No significant difference between experimental observations and theoretical values, i.e., observed deviations are caused by random fluctuations.
- Alternative hypothesis H_1: Significant difference exists between experimental and theoretical values (possibly systematic bias).

Calculated results:

- Sample standard deviation: 0.0038.
- Standard error: 0.0005.
- Standardized error Z: 1.0802.

Since $|\mathrm{Z}| \approx 1.0802 < 1.96 = \mathrm{Z}_{\alpha=0.05}$, the difference is not statistically significant. We therefore fail to reject the null hypothesis (H_0): the observed discrepancy between experimental and theoretical values can be attributed to random fluctuations, with no evidence of a meaningful deviation.

To assess the practical relevance of this difference, we further compute Cohen's d effect size. The obtained Cohen's d is approximately $0.1817 < 0.2$, indicates a negligible effect, confirming that the deviation is not only statistically insignificant but also practically minor.

From Table 4 to Table 11, the standardized errors and Cohen's d effect sizes are calculated using the aforementioned methodology. Table 4 corresponds to the even-N data in Table I(a) of the main text; Table 5 corresponds to the odd-N data in Table I(b); Tables 6 and 7 correspond to the odd- and even-N data in Table I(c); Tables 8 and 9 correspond to the data in Table 13; and Tables 10 and 11 correspond to the data in Table 14.

Table 4: Comparison between theoretical and experimental values under even N when memorizing single memory message.

N**(EVEN)**	**THEORY**	**EXPERIMENT**	Z	Cohen's d
100	0.0796	0.0795	-0.1703	0.0239
200	0.0563	0.0568	1.1872	0.1662
500	0.0357	0.0358	0.6359	0.0890
1000	0.0252	0.0250	-1.1071	0.1550

Table 5: Comparison between theoretical and experimental values when memorizing two messages with an odd Hamming distance $r_{12} = 3$.

N(ODD)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
99	3	0.0812	0.0820	0.3405	0.0477
199	3	0.0569	0.0557	0.8112	0.1137
499	3	0.0358	0.0359	-1.3206	0.1849
999	3	0.0253	0.0252	-0.3945	0.0552

Table 6: Comparison between theoretical and experimental values when memorizing two messages with an even Hamming distance $r_{12} = 4$.

N(ODD)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
99	4	0.375	0.3751	0.1619	0.0223
199	4	0.375	0.3753	0.3576	0.0501
499	4	0.375	0.3750	-0.0332	0.0046
999	4	0.375	0.3752	0.4294	0.0601

Table 7: Comparison between theoretical and experimental values when memorizing two messages with an even Hamming distance $r_{12} = 4$.

N(EVEN)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
100	4	0.4258	0.4247	-1.5549	0.2177
200	4	0.4106	0.4095	-1.8358	0.2571
500	4	0.3974	0.3971	-0.3403	0.0476
1000	4	0.3908	0.3919	0.6429	0.0900

Table 8: Comparison between theoretical and experimental values with different odd N under the same r_{12} ratio.

N(ODD)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
99	33	0.0978	0.0977	-0.4832	0.0677
199	66	0.0978	0.0976	-0.6299	0.0882
499	166	0.0618	0.0620	0.3571	0.0500
999	333	0.0309	0.0310	0.5085	0.0712

From the data presented in Table 4 to Table 11, all standardized errors Z) fall below the critical value of 1.96, indicating no statistically significant differences between experimental and theoretical results. Moreover, the majority of Cohen's d values are below 0.2, suggesting that the observed differences are practically negligible. A few instances with $d > 0.2$ are expected and not unusual.

These slight deviations can be attributed to sampling variability and measurement noise. Under the null hypothesis (H_0), the presence of standard error due to finite sample sizes leads to natural dispersion in the sampling distribution of the mean difference. Even in the absence of a true effect, Cohen's d may occasionally exceed conventional thresholds (e.g., $d \approx 0.2$) due to random fluctuation—especially in low-power settings with limited samples or high variance. This phenomenon is consistent with asymptotic statistical theory: across repeated trials, such deviations will occur with probability α (Type I error rate), even without underlying effects.

Table 9: Comparison between theoretical and experimental values with different even N under the same r_{12} ratio.

N(EVEN)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
100	33	0	0	None	None
200	66	0.1600	0.1597	-0.4243	0.0594
500	166	0.1028	0.1023	-1.1737	0.1643
1000	333	0	0	None	None

Table 10: Comparison between theoretical and experimental values with different odd N under even $r_{12} = 30$.

N(ODD)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
99	30	0.1445	0.1445	0.0179	0.0025
199	30	0.1445	0.1440	-0.8602	0.1205
499	30	0.1445	0.1452	1.3036	0.1825
999	30	0.1445	0.1441	-0.8564	0.1199

Table 11: Comparison between theoretical and experimental values with different even N under even $r_{12} = 30$.

N(EVEN)	r_{12}	THEORY	EXPERIMENT	Z	Cohen's d
100	30	0.2258	0.2258	0.1191	0.0167
200	30	0.1967	0.1970	0.5556	0.0779
500	30	0.1759	0.1762	0.4179	0.0585
1000	30	0.1664	0.1668	0.7757	0.1086

In addition, we have converted the tabular data in the main text into corresponding line charts for clearer visualization. The left panel shows results for small-scale networks ($N \leq 20$), where the theoretical and experimental values coincide perfectly since exhaustive enumeration was applied. The right panel displays results for larger networks ($N > 20$), where Monte Carlo sampling was employed. As illustrated, the theoretical and experimental values remain closely aligned, thereby validating the correctness of our theoretical analysis.

Figures 3–6 show, respectively, the even-N data in Table I(a), the odd-N data in Table I(b), and the odd- and even-N data in Table I(c).

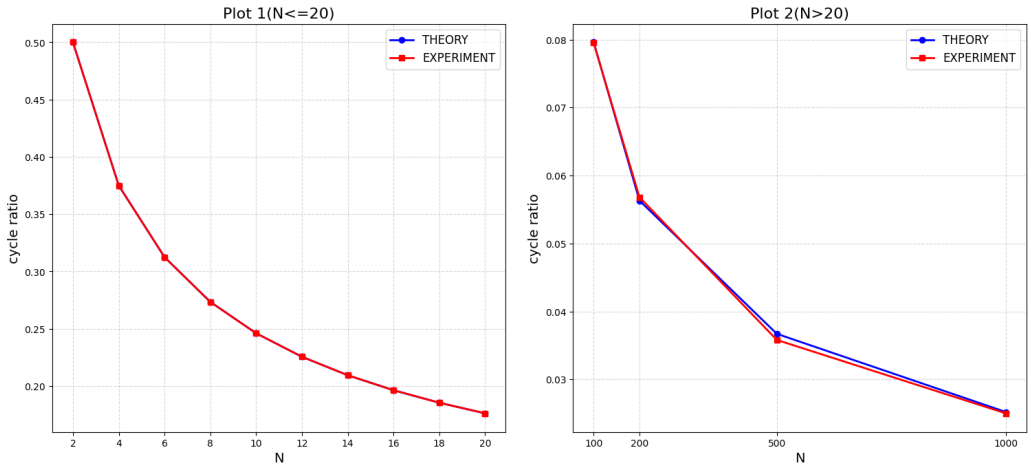

Figure 3: Even-N results from main-text Table I(a) (left: $N \leq 20$; right: $N > 20$).

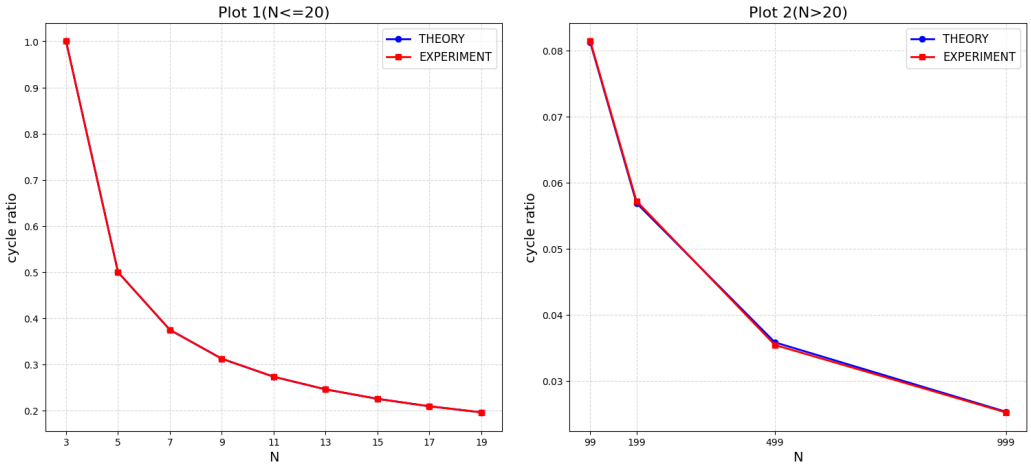

Figure 4: Odd-N, $r_{12} = 3$, results from main-text Table I(b) (left: $N \leq 20$; right: $N > 20$).

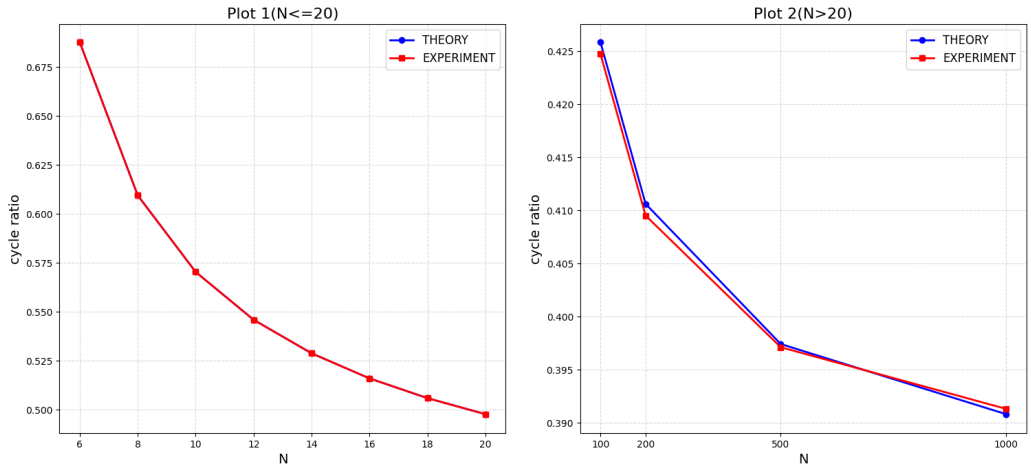

Figure 6: Even-N, $r_{12} = 4$, results from main-text Table I(c) (left: $N \leq 20$; right: $N > 20$).

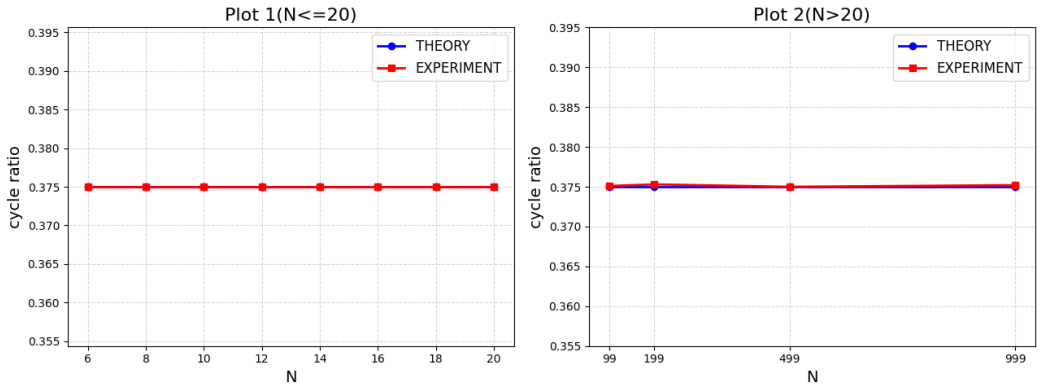

Figure 5: Odd-N, $r_{12} = 4$, results from main-text Table I(c) (left: $N \leq 20$; right: $N > 20$).

B.2 Visual Presentation

To facilitate understanding of the content and conclusions in this paper, the dynamic behavior of the Hopfield Neural Network will be demonstrated through specific examples and visual illustrations.

Memorizing One Message

When memorizing one message, the convergence of the Hopfield network is only related to the parity of N. If N is odd, there is no symmetric-cycle, and there are no cyan nodes in the middle part of Figure I in the text. If N is even, the symmetric-cycle occurs. The following two examples demonstrate the relationship between the dynamic behavior of the network and the Hamming distance when N is even and odd.

Case 1: N even.

For example, when the number of neurons, $N = 4$, and one message $[-1 \quad 1 \quad 1 \quad -1]^{\mathrm{T}}$ is memorized (converted to decimal as the number 6), its symmetric message is 9. According to Case1 in Theorem 1, we can conclude that the noise patterns with a Hamming distance less than 2 (i.e., a Hamming distance of 1 from 6) will converge to 6, while the noise patterns with a Hamming distance less than 2 from 9 will converge to 9. Noise patterns with a Hamming distance equal to 2 from both 6 and 9 will fall into a symmetric-cycle.

The visual simulation of the experiment is shown in Figure 7. As seen, the central circle in the left side of the diagram represents the memorized message, and the central circle on the right represents the symmetric message. Starting from the center, each expanding ring indicates an increase in the Hamming distance of the noise pattern from the message in central circle. The red arrows start from noise patterns and point to the stable fixed points (the memorized message), while the green arrows start from noise patterns and point to the stable fixed points (the symmetric message). The blue arrows represent noise patterns that fall into symmetric-cycle.

It is clear that all noise patterns in the first ring (Hamming distance of 1) around the memorized message 6 converge to 6, and similarly, the noise patterns in the first ring around the symmetric message 9 converge to 9. Noise patterns with a Hamming distance of 2 from both 6 and 9 fall into symmetric-cycle. The experimental simulation is fully consistent with the theoretical derivation.

Additionally, the new symmetries in the noise pattern landscape are reflected in the following aspects:

- Two noise patterns trapped in symmetric-cycle are mutually symmetric.
- Noise patterns converging to message 6 are $(2, 4, 7, 14)$, while those converging to message 9 are $(13, 11, 8, 1)$. The attraction basins of the memory message and its symmetric counterpart are identical in size (both of size 4), and the corresponding patterns are mutually symmetric (for instance, 2 and 13, 4 and 11).

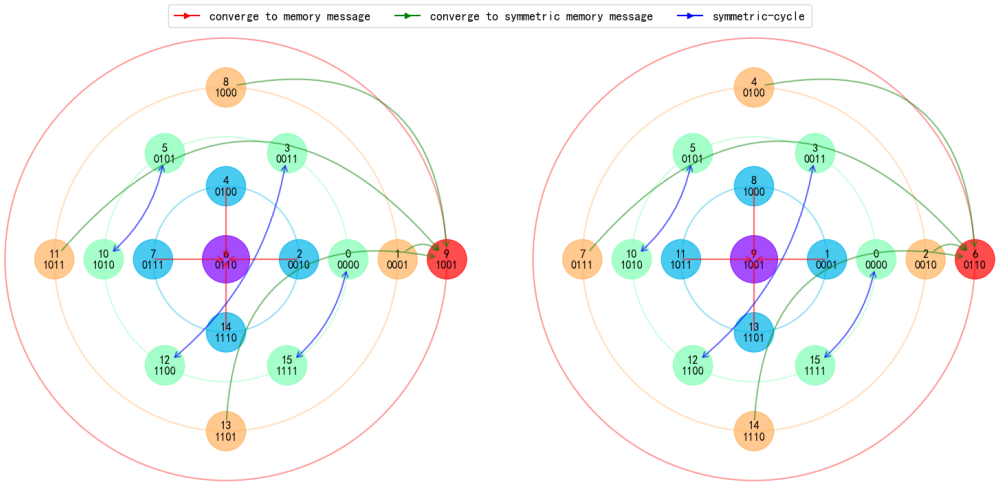

Figure 7: Evolution results of a HNN with $N = 4$, memorizing one message $[-1 \quad 1 \quad 1 \quad -1]^{\mathrm{T}}$ (The decimal representation is 6). Clear symmetry emerges along the evolution.

(Note: For ease of binary conversion in the diagram, the messages is composed of $\{0, 1\}$; however, in the actual experiments and theoretical derivations, all memories and noise patterns are binary vectors composed of $\{1,\text{-}1\}$. This will not be reiterated.)

Case 2: N odd.

For example, when the number of neurons, $N = 5$, and one memory $[-1 \quad -1 \quad 1 \quad 1 \quad -1]^{T}$ is stored (converted to decimal as number 6), its symmetric memory is 25. According to Case 2 in Theorem I, noise patterns with a Hamming distance less than 2.5 (i.e., Hamming distances 1 and 2) from 6 will converge to 6, and those with Hamming distances 1 and 2 from 25 will converge to 25.

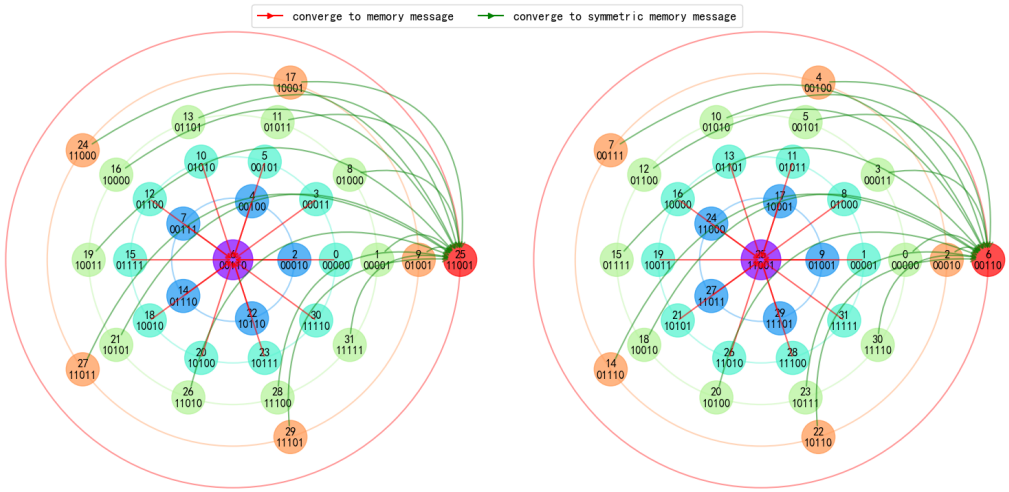

Figure 8: Evolution results of a HNN with $N = 5$, memorizing one message $[-1 \quad -1 \quad 1 \quad 1 \quad -1]^{T}$ (The decimal representation is 6). Clear symmetry emerges along the evolution.

The Figure 8 above shows the visualization of the simulation. The central circle in the left circular area represents the stored memory message, and the central circle in the right represents its symmetric memory message. Moving outward from the center, each ring indicates an increase of 1 in

the Hamming distance. Red arrows start from a noise pattern and end at the final stable fixed point. Clearly, all noise nodes on the first and second rings (Hamming distances 1 and 2) around memory 6 converge to 6, and similarly, all noise nodes on the first and second rings around symmetric memory 25 converge to 25. The simulation results are entirely consistent with the theoretical derivation.

Additionally, the new symmetries in the noise pattern landscape are reflected in the following aspects:

- The sizes of the attraction basins for memory message 6 and its symmetric counterpart message 25 are identical (both are 16, including the messages themselves). Moreover, all noise patterns with a Hamming distance of 1 or 2 from message 6 converge to message 6, while all noise patterns with a Hamming distance of 1 or 2 from message 25 converge to message 25.
- Among the noise patterns that converge to message 6, those with a Hamming distance of 1 include $(2, 4, 7, 14, 22)$, while those with a Hamming distance of 1 that converge to symmetric message 25 include $(29, 27, 24, 17, 9)$. The corresponding patterns are symmetric to each other (for instance, 2 and 29, 4 and 27, etc.). Similarly, among the noise patterns that converge to message 6, those with a Hamming distance of 2 include $(0, 3, 5, 10, 12, 15, 18, 20, 23, 30)$, while those with a Hamming distance of 2 that converge to message 25 include $(31, 28, 26, 21, 19, 16, 13, 11, 8, 1)$. The corresponding patterns are also symmetric to each other (for instance, 0 and 31, 3 and 28, etc.).

Memorizing Two Messages

Case 1: N even, r_{12} even.

For example, when the number of neurons N is 6, two messages $[-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad -1]^T$ (decimal 6) and $[-1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1]^T$ (decimal 10) are stored. The Hamming distance between them is 2. Their symmetric messages are 57 and 53, respectively. According to Corollary 5 and Corollary 6, there is no complete convergence domain, and $r_{max} = 1$. This means only noise patterns with a Hamming distance less than or equal to 1 from a memory message or its symmetric memory message can converge. Furthermore, according to the first case in Corollary 4, self-cycle, hetero-cycle, and symmetric-cycle occur.

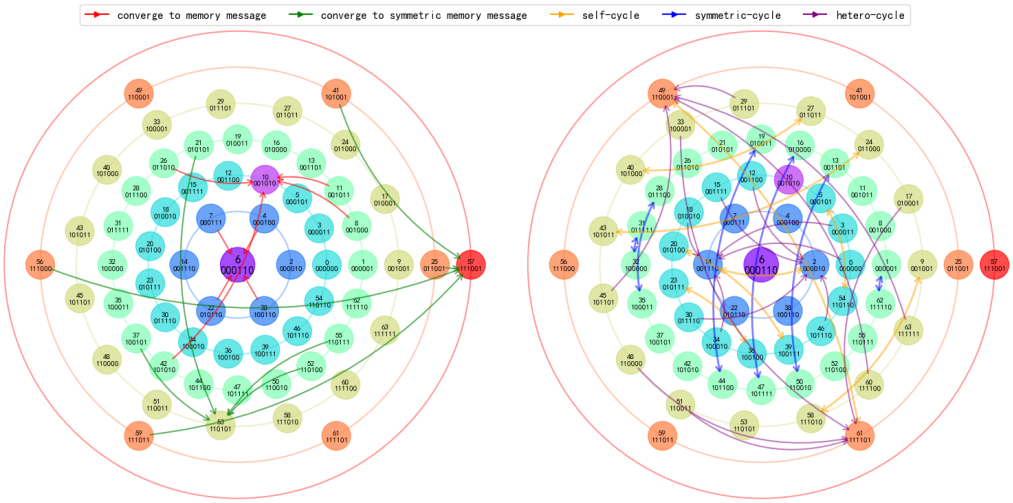

Figure 9: Evolution results of a HNN with $N = 6$, memorizing two messages $\xi_1 = [-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad -1]^T$ (decimal 6) and $\xi_2 = [-1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1]^T$ (decimal 10). Clear symmetry emerges along the evolution.

The Figure 9 above shows the visualization of the simulation. To avoid clutter, the left panel only shows convergence trajectories, and the right panel only shows cycle trajectories. In the left panel,

noise patterns converging to message 6, 10 and their symmetric message 57, 53 all have a Hamming distance of 1 from the respective memory. Not all six noise patterns on the first ring around message 6 converge to 6, indicating the absence of complete convergence domain, consistent with the theoretical derivation. In the right panel, orange lines represent self-cycles (e.g., between 9 and 58, 40 and 27, 49 and 61). Purple lines represent hetero-cycles (e.g., 29, 45, 60, 63 all update to 49, thus falling into the cycle between 49 and 61). Blue lines represent symmetric-cycles (e.g., between 31 and 32, 28 and 35, which are mutual symmetric noise patterns).

The specific cycle type a noise pattern falls into can be determined using Theorem 3. For example, consider noise pattern 45, represented by $V = [1 \quad -1 \quad 1 \quad 1 \quad -1 \quad 1]^T$, and the memories $\xi_1 = [-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad -1]^T$, $\xi_2 = [-1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1]^T$. The Hamming distance between V and ξ_1 is $r_1 = 4$, and with ξ_2 is $r_2 = 4$. Since $r_1 + r_2 = 8 \neq 6 = N$, the **Condition I** in Theorem 3 is met. Furthermore, since $r_1 + r_2 = 8 \geq 7 = N + 1$, we need to check if V contains a V_{11} part. As explained in Section V-A of the main text, the V_{11} part consists of components where $V = \xi_1 = \xi_2$. Observing the second component, all three values are -1, so the V_{11} part exists, and thus the pattern falls into a hetero-cycle. Another example: noise pattern 35 has Hamming distances $r_1 = 3$ from ξ_1 and $r_2 = 3$ from ξ_2. Here, $r_1 + r_2 = 6 = N$ and $r_1 = r_2$, satisfying the **Condition III** in Theorem 3, resulting in a symmetric-cycle.

Additionally, the new symmetries in the noise pattern landscape are reflected in the following aspects:

- Memory messages 6, 10 and their symmetric counterparts 57, 53 possess attraction basins of identical size (each of size 5, including the pattern itself). Moreover, the noise patterns within these basins all have a Hamming distance of 1 from the memory pattern to which they converge.
- Two noise patterns trapped in symmetric-cycle are mutually symmetric(for instance, 1 and 62, 13 and 50, etc.).
- Noise patterns that enter self-cycle maintain equal Hamming distances to the original memory message, and consequently also to its symmetric counterpart. For example, patterns 2 and 14 both have a Hamming distance of 1 from memory message 6, while patterns 49 and 61 both have a Hamming distance of 5 from memory message 6. Similarly, patterns 5 and 54 both have a Hamming distance of 4 from memory message 10, as do patterns 20 and 39.
- Patterns entering hetero-cycle also exhibit equal Hamming distances between corresponding members. For instance, patterns 29, 45, 60, and 63 all first evolve to pattern 49 before entering the self-cycle between 49 and 61, while patterns 17, 33, 48, and 51 all first evolve to pattern 61 before entering the same cycle. Between these two groups of noise patterns, corresponding pairs have identical Hamming distances (for instance, the Hamming distance between 29 and 17 is 2, as is that between 45 and 33, 60 and 48, and 63 and 51).

Case 2: N even, r_{12} odd.

For example, when N=6, two messages $[-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad -1]^T$ (decimal 6) and $[-1 \quad -1 \quad 1 \quad -1 \quad 1 \quad 1]^T$ (decimal 11) are stored. The Hamming distance between them is 3. Their symmetric memories are 57 and 52, respectively. According to Corollary 5 and Corollary 6, $r_{ccd} = 1$ and $r_{max} = 2$. This means all noise patterns with a Hamming distance of 1 from the message or its symmetric message will converge to that message; only noise patterns with Hamming distance less than or equal to 2 might converge. Furthermore, according to the second case in Corollary 4, no cycle exist; all noise patterns will converge.

The Figure 10 below shows the visualization of the simulation. To avoid clutter, the left panel shows convergence for message 6 and its symmetric message 57, and the right for message 11 and its symmetric message 52. All noise patterns converge; no cycle occur. All noise patterns on the first ring (Hamming distance 1) around message 6 converge to 6, and similarly for message 11 in the right panel. (To save space, visualizations for symmetric message 57 and 52 are not shown, but the observation is consistent: all noise patterns on the first ring around a symmetric message converge to that symmetric message). Furthermore, in the visualization centered on message 6, only noise patterns on the first two rings (Hamming distances 1 or 2) can converge to 6, consistent with the $r_{max} = 2$.

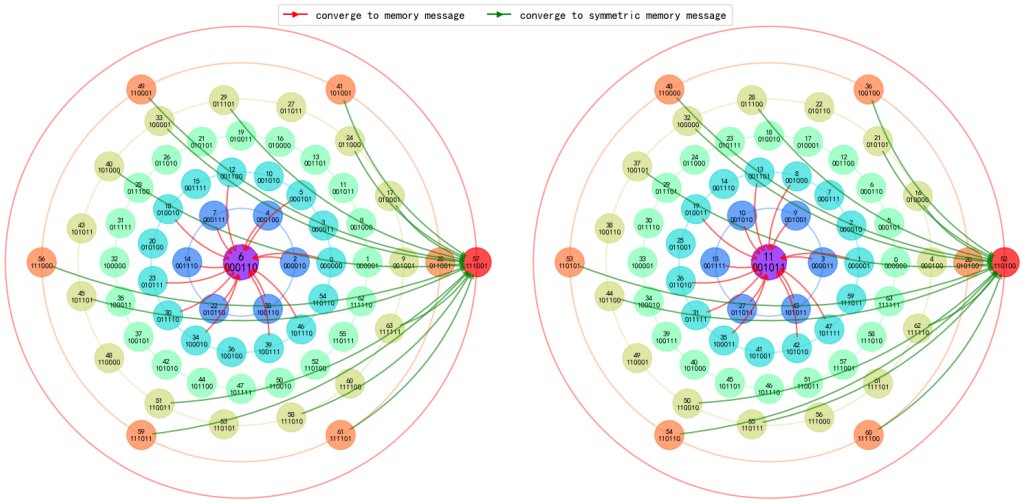

Figure 10: Evolution results of a HNN with $N = 6$, memorizing two messages $[-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad -1]^T$ (decimal 6) and $[-1 \quad -1 \quad 1 \quad -1 \quad 1 \quad 1]^T$ (decimal 11). Clear symmetry emerges along the evolution.

Why do some noise patterns within the convergence domain of a message not converge to it? For example, pattern 15 has a Hamming distance of 2 from message 6 but ultimately converges to message 11. There are two explanations:

- Noise pattern 15 has a Hamming distance of 1 from message 11, placing it within the **complete convergence domain** of message 11, hence it converges to 11.
- Consider noise pattern $V = [-1 \quad -1 \quad 1 \quad 1 \quad 1 \quad 1]^T$ and the message $\xi_1 = [-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad -1]^T$, $\xi_2 = [-1 \quad -1 \quad 1 \quad -1 \quad 1 \quad 1]^T$. Noise pattern V contains a V_{11} part (components 1, 2, 5), a V_{21} part (component 4), and a V_{22} part (components 3, 6). The Hamming distances are $r_1 = 2$, $r_2 = 1$. Since $r_2 + r_1 = 3 \leq 5 = N - 1$, $r_2 - r_1 = -1 \leq -1$, and it only contains V_{11}, V_{21}, and V_{22} parts, this matches case 2.2.3 in the Proof of Theorem 3 in Appendix A.3. According to the conclusion, it will eventually converge to ξ_2, i.e., 11.

Additionally, the new symmetries in the noise pattern landscape are reflected in the following aspects:

- The attraction basins of memory messages 6, 11 and their symmetric counterparts 57, 52 are identical in size (each containing 16 patterns, including the memory message itself). Moreover, all noise patterns in these basins have a Hamming distance of either 1 or 2 from the memory message to which they converge.
- Among the noise patterns converging to memory message 6, those with a Hamming distance of 1 are $(2, 4, 7, 14, 22, 38)$. The corresponding noise patterns with a Hamming distance of 1 that converge to its symmetric counterpart 57 are $(61, 59, 56, 49, 41, 25)$. Each corresponding pair of patterns is symmetric to the other (for instance, 2 and 61, 4 and 59, etc.). The same rule applies to memory message 11 and its symmetric counterpart 52.

Cases where N is odd, and r_{12} is either odd or even, can be analyzed similarly, corresponding to the third and fourth cases in Corollary 4, respectively. In these cases, there are no symmetric-cycle, only self-cycle and hetero-cycle.

B.2.1 Full-State Outcome Visualizations

The four full-state atlases below complement the original Hamming-ring visualizations by displaying every possible noise pattern in each of the four concrete examples. The atlases with one message

report the Hamming distance d from the stored message, whereas the atlases with two messages report the distance pair (r_1, r_2). In every atlas, a cell contains the decimal state, its binary representation, the relevant distance information, the number of synchronous updates before the terminal component is reached, and the final fixed point or cycle. The four figures use the same visual grammar and color encoding and are placed on separate full-width pages for direct inspection.

Table 12: Overview of the four full-state outcome atlases. Each state count includes the stored memories and their symmetric counterparts.

Case	**State space**	**Terminal behavior**
One memory, $N = 4$ (Figure 11)	$2^4 = 16$	Two fixed-point basins of size 5 and 6 symmetric-cycle states.
One memory, $N = 5$ (Figure 12)	$2^5 = 32$	Two fixed-point basins of size 16, with no cycle.
Two messages, $N = 6$, $r_{12} = 2$ (Figure 13)	$2^6 = 64$	20 fixed-point, 16 self-cycle, 16 hetero-cycle, and 12 symmetric-cycle states.
Two messages, $N = 6$, $r_{12} = 3$ (Figure 14)	$2^6 = 64$	Four fixed-point basins of size 16, with no cycle.

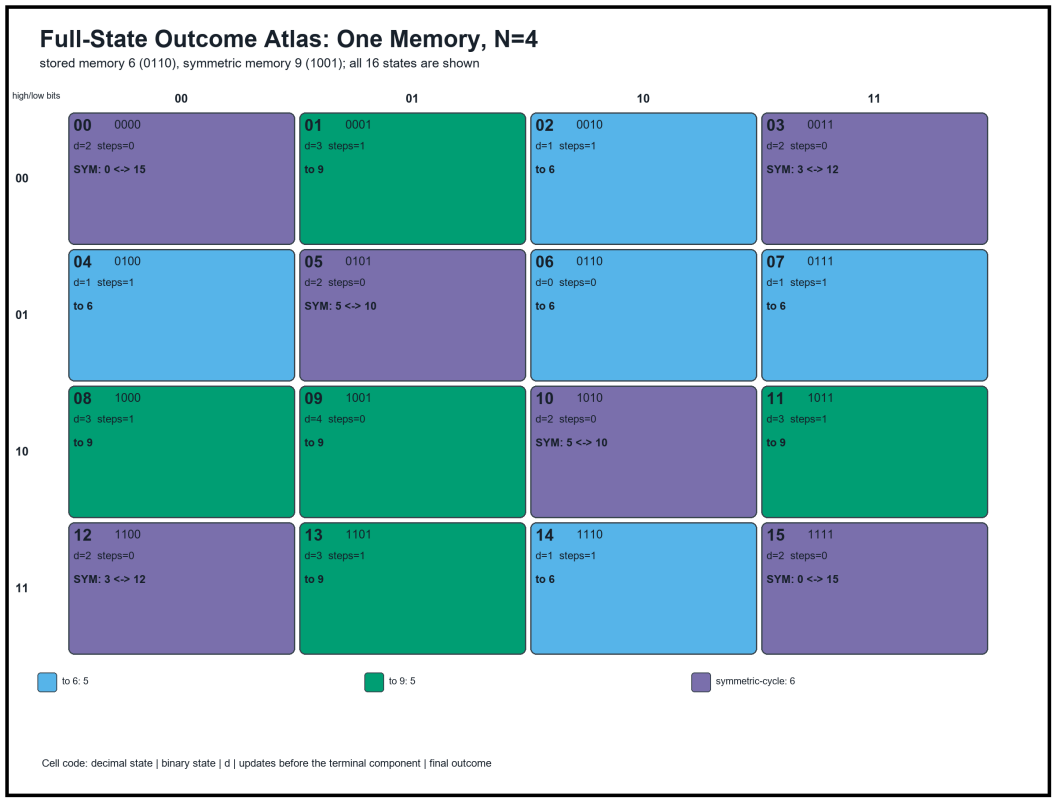

Figure 11: Full-state outcome atlas for the one-memory case $N = 4$, with stored memory $\xi = 6$ and symmetric memory $-\xi = 9$. All $2^4 = 16$ states appear once. Each cell reports the decimal and binary labels, the Hamming distance d from ξ, the number of synchronous updates before reaching the terminal component, and the final outcome. Blue cells converge to 6, green cells converge to 9, and purple cells belong to a symmetric-cycle. The states with $d = 0$ or 1 form the basin of 6, those with $d = 3$ or 4 form the basin of 9, and the six states with $d = 2$ form three complementary symmetric-cycle pairs. Thus, the two fixed-point basins each contain 5 states and the remaining 6 states are cyclic.

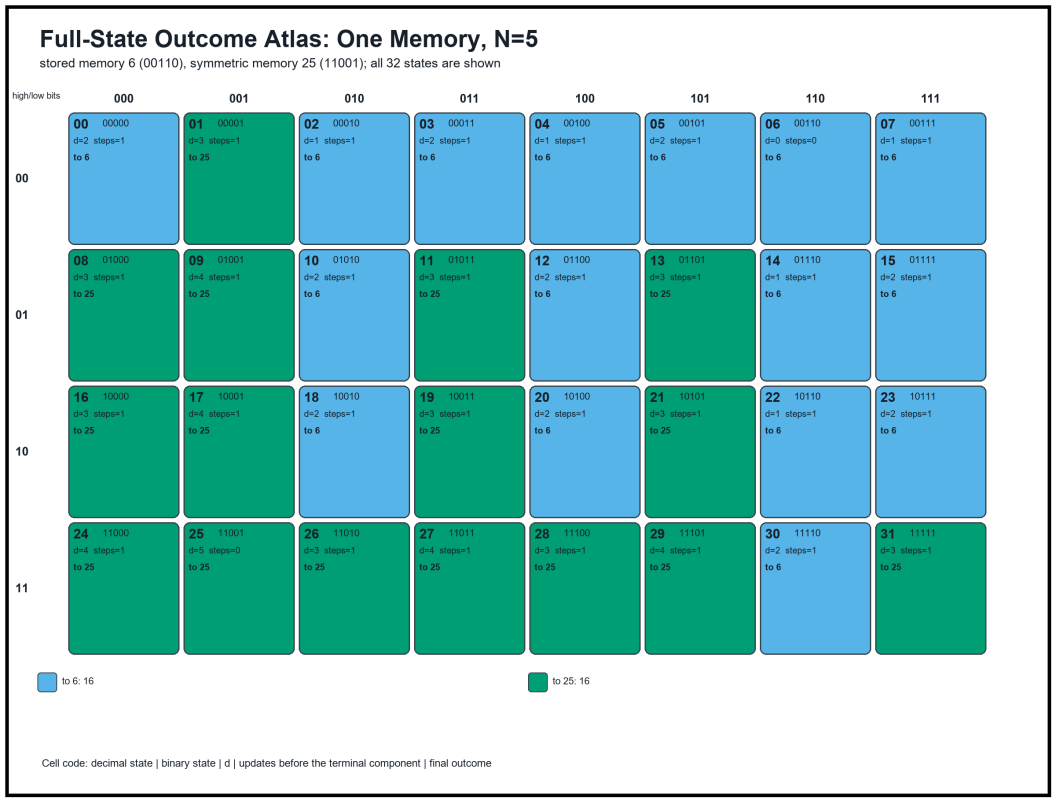

Figure 12: Full-state outcome atlas for the one-memory case $N = 5$, with stored memory $\xi = 6$ and symmetric memory $-\xi = 25$. All $2^5 = 32$ states appear once. Each cell reports the decimal and binary labels, the Hamming distance d from ξ, the number of synchronous updates before the terminal fixed point, and the final outcome. Blue cells converge to 6, whereas green cells converge to 25. All states with $d \leq 2$ converge to 6, and all states with $d \geq 3$ converge to 25. Each attraction basin therefore contains 16 states, and no cycle occurs, consistently with the odd value of N.

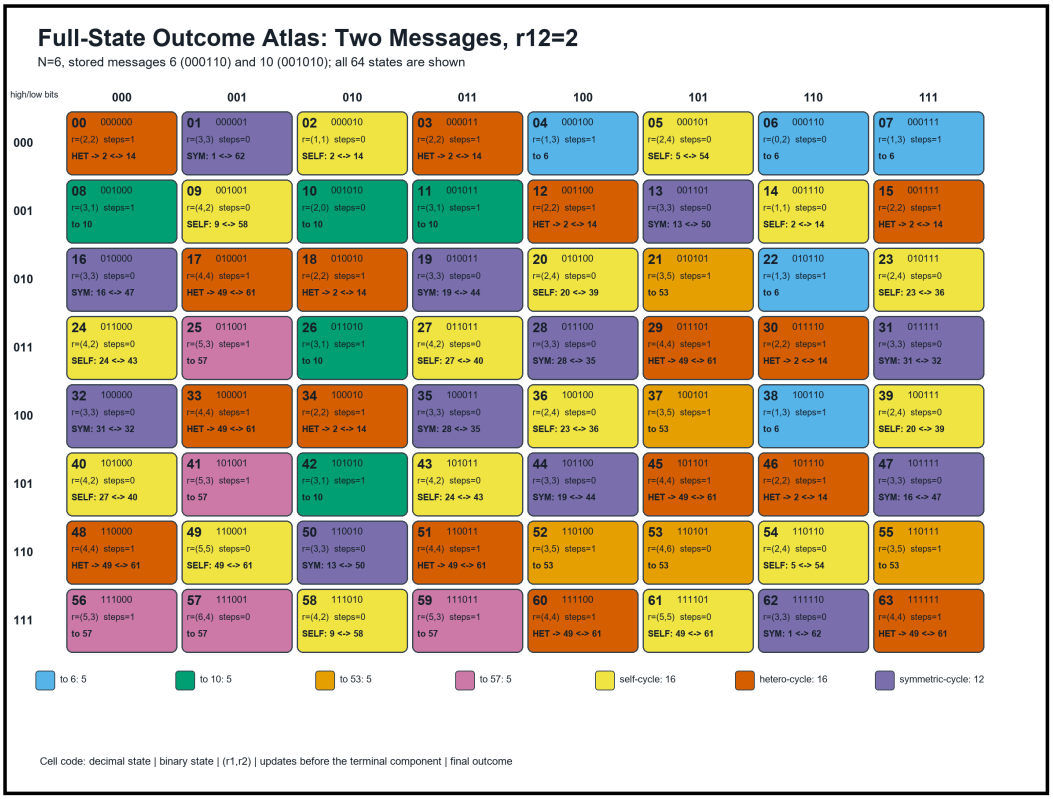

Figure 13: Full-state outcome atlas for the case with two messages, $N = 6$, with $\xi_1 = 6$, $\xi_2 = 10$, and $r_{12} = 2$. All $2^6 = 64$ states appear once. Each cell reports the decimal and binary labels, the distance pair (r_1, r_2), the number of synchronous updates before reaching the terminal component, and the final outcome. Blue, green, orange, and pink cells converge to fixed points 6, 10, 53, and 57, respectively; each fixed-point basin contains 5 states. Yellow cells are the 16 states lying on non-symmetric self-cycle cores, dark-orange cells are the 16 hetero-cycle states entering those cores, and purple cells are the 12 states forming six symmetric-cycle pairs. The resulting $20 + 16 + 16 + 12 = 64$ partition reproduces the complete convergence and cycle behavior reported for the even-r_{12} example.

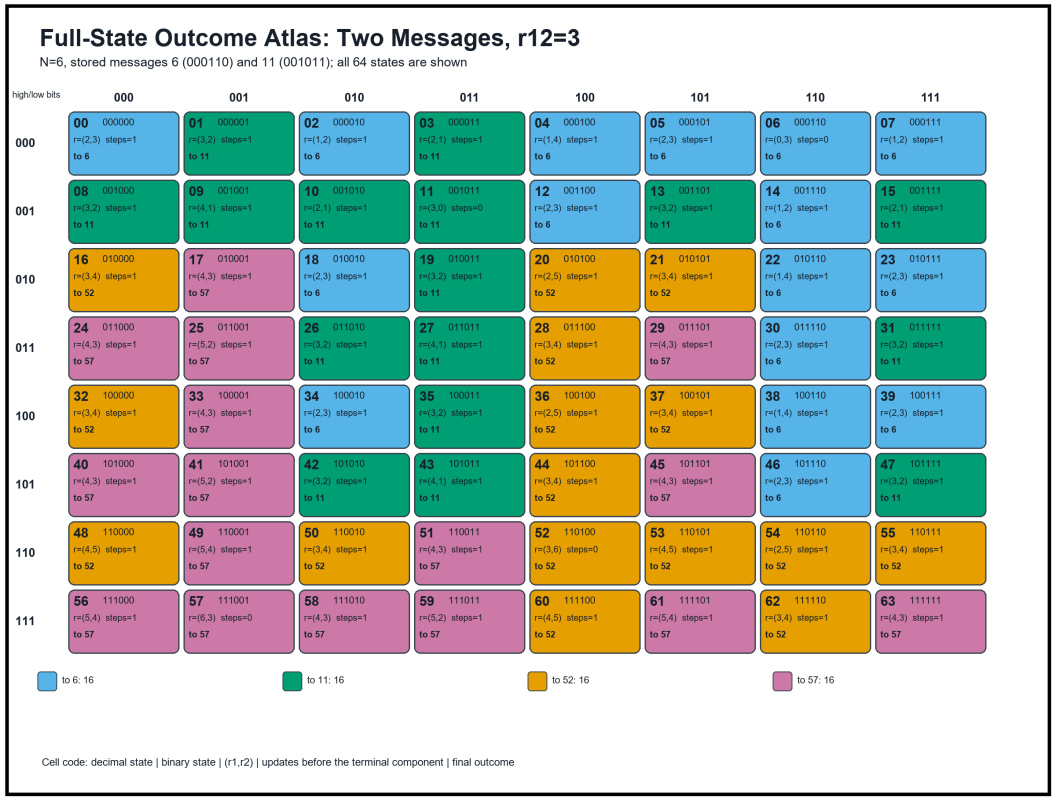

Figure 14: Full-state outcome atlas for the case with two messages, $N = 6$, with $\xi_1 = 6$, $\xi_2 = 11$, and $r_{12} = 3$. All $2^6 = 64$ states appear once. Each cell reports the decimal and binary labels, the distance pair (r_1, r_2), the number of synchronous updates before reaching the terminal fixed point, and the final outcome. Blue, green, orange, and pink cells converge to fixed points 6, 11, 52, and 57, respectively. Each fixed-point basin contains 16 states, so the four basins partition the entire state space. No self-cycle, hetero-cycle, or symmetric-cycle occurs, in agreement with $r_{\mathrm{ccd}} = 1$, $r_{\max} = 2$, and the theoretical conclusion for odd r_{12}.

B.3 Experiments with a Fixed r_{12}/N Ratio

We set r_{12} to the nearest admissible integer satisfying $r_{12}/N \approx m$, where m is fixed.

For small-scale networks ($N \leq 10$), we perform exhaustive enumeration. Two N-bit patterns with Hamming distance r_{12} are randomly generated, and the corresponding weight matrix is constructed. All 2^N possible noise patterns are evolved under the HDAR update rule, and the proportion of patterns entering cyclic trajectories is recorded. In all trials, the empirical cycle ratios match the theoretical values exactly.

For large-scale networks ($N \geq 99$), we use 50 Monte Carlo trials. Each trial generates a new message pair with fixed r_{12} and 10,000 random noise vectors, then records their asymptotic behavior under the HDAR update rule.

The averaged cycle ratios closely match theory. Results for $r_{12}/N \approx 1/3$ are presented in Table 13.

Table 13: Comparison between theoretical and experimental values under the same r_{12} ratio.

N **(parity)**	r_{12}	**THEORY**	**EXPERIMENT**
99 (odd)	33	0.0978	0.0977
199 (odd)	66	0.0978	0.0976
499 (odd)	166	0.0618	0.0620
999 (odd)	333	0.0309	0.0310
100 (even)	33	0	0
200 (even)	66	0.1600	0.1597
500 (even)	166	0.1028	0.1023
1000 (even)	333	0	0

B.4 Experiments with Fixed $r_{12} = 30$

Experimental Setup and Results

We randomly select two N-bit binary vectors as memory messages, ensuring a fixed Hamming distance r_{12} between them (with $N > r_{12}$). Due to the infeasibility of exhaustive noise enumeration, we adopt Monte Carlo sampling, using the same noise generation procedure as in the previous experiment.

For each fixed N, we construct a weight matrix from the selected memory messages and evolve randomly generated noise samples under the HDAR update rule, recording whether each pattern converges or enters a cycle. This process is repeated 50 times, with both the memory messages and noise samples regenerated in each trial. The average cycle ratio is computed over all repetitions and compared against theoretical predictions.

To investigate how network size influences dynamics, we fix $r_{12} = 30$ and vary N, conducting experiments across multiple odd values of N. As shown in Table 14, the empirical results reveal a striking phenomenon: the cycle ratio remains constant across all odd values of N, aligning precisely with theoretical expectations.

Based on experimental observations, we conclude that for any fixed even value of r_{12}, the cycle ratio remains invariant across all odd values of N, and is precisely given by $\frac{\binom{r_{12}}{r_{12}/2}}{2^{r_{12}}}$. This implies that the cycle ratio is solely determined by the even parameter r_{12}, independent of the network size.

Table 14: Comparison between theoretical and experimental values under $r_{12} = 30$.

N **(parity)**	r_{12}	**Theory**	**Experiment**
99 (odd)	30	0.1445	0.1445
199 (odd)	30	0.1445	0.1440
499 (odd)	30	0.1445	0.1452
999 (odd)	30	0.1445	0.1441
100 (even)	30	0.2258	0.2258
200 (even)	30	0.1967	0.1970
500 (even)	30	0.1759	0.1762
1000 (even)	30	0.1664	0.1668

Derivation of the Invariant Cycle Ratios

The derivation is presented below, based on Corollary 4 in the main text, under the condition that N is odd and r_{12} is even.

$$C_{\text{oe}} = 2\binom{r_{12}}{r_{12}/2} \sum_{i=r_{12}/2}^{\frac{N-1}{2}} \binom{N-r_{12}}{i-r_{12}/2}$$
$$= 2\binom{r_{12}}{r_{12}/2} \sum_{i=0}^{\frac{N-1-r_{12}}{2}} \binom{N-r_{12}}{i}. \tag{B.6}$$

From the binomial theorem $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$, when $a = b = 1$:

$$2^n = \sum_{k=0}^{n} \binom{n}{k}. \tag{B.7}$$

Using the symmetry property of binomial coefficients $\binom{n}{k} = \binom{n}{n-k}$ and for odd n:

$$2^n = 2\sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} \Rightarrow \sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} = 2^{n-1}. \tag{B.8}$$

Since N is odd and r_{12} is even ($N - r_{12}$ is odd):

$$C_{\text{oe}} = 2\binom{r_{12}}{r_{12}/2} \sum_{i=0}^{\frac{N-r_{12}-1}{2}} \binom{N-r_{12}}{i}$$
$$= 2\binom{r_{12}}{r_{12}/2} \sum_{k=0}^{\frac{N-r_{12}-1}{2}} \binom{N-r_{12}}{k}$$
$$= 2\binom{r_{12}}{r_{12}/2} 2^{N-r_{12}-1}. \tag{B.9}$$

Thus the cycling ratio becomes:

$$\frac{2\binom{r_{12}}{r_{12}/2} 2^{N-r_{12}-1}}{2^N} = \frac{\binom{r_{12}}{r_{12}/2}}{2^{r_{12}}}. \tag{B.10}$$

Similarly, we can derive another corollary: when r_{12} is any even number and we add any odd number to it to form N (set N odd), the resulting cycling ratio remains identical to the original case with

even r_{12}. The derivation proceeds as follows:

$$C_{\text{oo}} = 2\binom{N-r_{12}}{(N-r_{12})/2} \sum_{i=\frac{N-r_{12}}{2}}^{\frac{N-1}{2}} \binom{r_{12}}{i-\frac{N-r_{12}}{2}}$$
$$= 2\binom{N-r_{12}}{(N-r_{12})/2} \sum_{i=0}^{\frac{r_{12}-1}{2}} \binom{r_{12}}{i}$$
$$= \binom{N-r_{12}}{(N-r_{12})/2} 2^{r_{12}}. \tag{B.11}$$

Thus, the cycling ratio becomes:

$$\frac{\binom{N-r_{12}}{(N-r_{12})/2}2^{r_{12}}}{2^N} = \frac{\binom{N-r_{12}}{(N-r_{12})/2}}{2^{N-r_{12}}}. \tag{B.12}$$

By comparing Equations (B.10) and (B.12), it can be observed that: for an odd $r_{12}^{(\text{odd})}$, an even $r_{12}^{(\text{even})}$, and odd N, if they satisfy

$$N - r_{12}^{(\text{odd})} = r_{12}^{(\text{even})}, \tag{B.13}$$

then the cycle ratios under both conditions will be identical.

B.5 Zero-Field Probability versus Network Size and Number of Stored Memories

Experimental objective and setup. We estimate how the probability of a zero local field depends on the two experimental variables: the number of neurons N and the number of stored memories M. For every independent trial, we generate a fresh set of M memory messages $\{\xi^1, \ldots, \xi^M\}$, where each $\xi^\mu \in \{-1, +1\}^N$ and

$$\Pr(\xi_i^\mu = +1) = \Pr(\xi_i^\mu = -1) = \frac{1}{2}. \tag{B.14}$$

All entries are sampled independently across neurons, memories, and trials. The initial state $X(0) \in \{-1, +1\}^N$ is generated independently by the same Rademacher distribution. We impose no balancing constraint, Hamming-distance constraint, or duplicate-rejection rule. The weight matrix has zero diagonal and follows the standard Hebbian construction,

$$W_{ij} = \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu \quad (i \neq j), \qquad W_{ii} = 0. \tag{B.15}$$

At every iteration, all local fields $h_i(t) = \sum_{j=1}^N W_{ij} X_j(t)$ are computed from the same current state, after which all neurons are updated synchronously according to the Rectified rule:

$$X_i(t+1) = \begin{cases} +1, & h_i(t) > 0, \\ -1, & h_i(t) < 0, \\ X_i(t), & h_i(t) = 0. \end{cases} \tag{B.16}$$

Theoretical reference. For a fixed neuron at the random initial state,

$$h_i(0) = \sum_{\mu=1}^{M} \sum_{j \neq i} \xi_i^\mu \xi_j^\mu X_j(0) \tag{B.17}$$

is a sum of $K = M(N-1)$ independent Rademacher variables. Consequently, the exact initial per-neuron zero-field probability is

$$p_0(N, M) = \begin{cases} \binom{K}{K/2} 2^{-K}, & K \text{ is even}, \\ 0, & K \text{ is odd}. \end{cases} \tag{B.18}$$

This parity condition also shows that when N is even and M is odd, every local field is a sum of an odd number of ± 1 terms and therefore cannot be zero at any synchronous update.

Monte Carlo protocol and measures. We evaluate

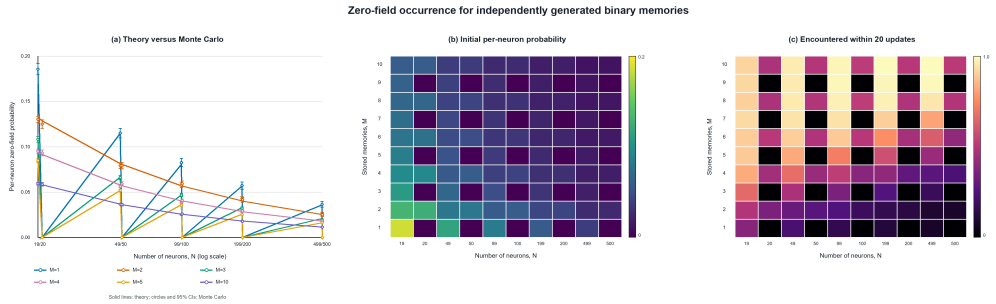

Figure 15: Zero-field probabilities under independently generated $\{-1, +1\}$ memories and synchronous Rectified updates. (a) The initial per-neuron probability: solid curves give the exact value in Equation (B.18), circles give Monte Carlo estimates, and vertical bars give 95% confidence intervals. For readability, representative values $M \in \{1, 2, 3, 4, 5, 10\}$ are shown; the paired horizontal labels compare adjacent odd and even network sizes. (b) Initial per-neuron zero-field probability over the complete N-by-M grid. (c) Probability that a trajectory encounters at least one zero field during 20 synchronous field evaluations. White cells denote probability zero. The checkerboard pattern is the exact parity effect: zero fields are impossible whenever N is even and M is odd.

$$N \in \{19, 20, 49, 50, 99, 100, 199, 200, 499, 500\}, \qquad M \in \{1, \ldots, 10\}.$$

For each (N, M) pair, the initial-state statistics use $6{,}144$ independent trials, each with a newly generated complete memory set and initial state. The dynamic statistic uses another $1{,}152$ independently generated trajectories and evaluates the local fields for 20 consecutive synchronous updates. We record (i) the fraction of neurons with $h_i(0) = 0$, (ii) whether at least one neuron has zero field at initialization, and (iii) whether a trajectory encounters at least one zero field during the 20 field evaluations. Reported error bars are 95% Monte Carlo confidence intervals across independent trials. All experiments use pseudorandom seed 20260723.

Results. Figure 15(a) shows close agreement between the exact initial per-neuron probability and simulation: across all 100 parameter pairs, the mean absolute error is 4.94×10^{-4}, and the maximum absolute error is 3.79×10^{-3}. No zero-field event was observed in any of the even-N, odd-M conditions, consistent with the exact parity restriction. For parity-compatible settings, the per-neuron probability generally decreases as N or M increases, as also predicted by Equation (B.18). However, the network-level dynamic probability can remain high because each update tests N local fields and the trajectory supplies repeated opportunities for a zero-field event. For example, at $M = 10$, the probability of encountering at least one zero field within 20 updates is 0.9931 for $N = 499$ and 0.5104 for $N = 500$, although their initial per-neuron probabilities are only 0.01139 and 0.01126, respectively. Thus, a small per-neuron zero-field probability need not imply that zero-field handling is negligible at the network or trajectory level.

C Extended Related Work

C.1 HNNs and Memory Capacity

A key conceptual transition occurred when Amari (1972) incorporated Hebbian learning (Hebb, 1949) into the Ising framework, reinterpreting it as an associative memory model. This idea was further developed by Little (1974) and culminated in Hopfield's seminal work (Hopfield, 1982), which formally established HNNs as spin-glass-inspired neural systems with symmetric interactions. The framework was later extended to continuous-time dynamics (Hopfield, 1984), enriching its theoretical scope.

Hopfield empirically observed that reliable retrieval is maintained when the number of stored patterns is below approximately $0.14N$ (Hopfield, 1982). Abu-Mostafa & St-Jacques (1985) subsequently formalized the notion of memory capacity for associative memories, showing that an HNN

with N neurons can store up to N^3 bits of information, with the number of stable attractors bounded by N. In 1987, McEliece et al. (1987) established that the storage capacity of the classical HNN scales as $N/(2 \ln N)$, while Kanter & Sompolinsky (1987) demonstrated that linear capacity in N is achievable under optimal conditions.

Chiueh & Goodman (1991) showed that, under specific model assumptions, HNNs can exhibit exponential memory capacity, and Storkey (1997) proposed learning rules that improve capacity to $N/\sqrt{2 \ln N}$. In parallel, Krotov & Hopfield (2016) introduced dense associative memory (DAM) networks with generalized energy functions and substantially higher capacity. Bao et al. (2022) subsequently proved polynomial memory capacity and derived sharp capacity bounds and convergence radii for binary DAMs, MHNs, and spherical HNNs (Bao & Zhao, 2025).

The statistical and thermodynamic equivalence between hybrid Boltzmann machines (HBMs) and HNNs was established by Barra et al. (2012), and subsequent work linked HBMs to glassy phase transitions in HNNs (Agliari et al., 2013). General formulations based on local neural computations and generalized energy functions unify classical and MHNs (Krotov & Hopfield, 2021; Millidge et al., 2022). Ota & Karakida (2023) introduced the Attentional Boltzmann Machine, connecting it to restricted Boltzmann machines and denoising autoencoders.

C.2 Convergence and Cyclic Dynamics

Early studies of HNNs established fundamental links between symmetry, energy functions, and convergence. Peretto (1984) unified HNNs and Little networks within a statistical-mechanics framework, showing that symmetric synaptic weights guarantee the existence of a Lyapunov energy function and ensure convergence. Subsequent experimental and theoretical investigations characterized richer dynamical behaviors under different update schemes. McEliece et al. (1987) identified distinct convergence patterns under synchronous updates, including one-step, two-step, and multi-step convergence, while Baldi & Venkatesh (1987) analyzed the number of stable states in related spin-glass systems. Rigorous mathematical treatments soon followed. Komlós & Paturi (1988) proved strong convergence and error-correction guarantees for Hopfield's associative memory, establishing linear error-correcting capability and large basins of attraction, as well as the existence of exponentially many spurious stable states. Later work refined these bounds: Loukianova (1997) derived lower bounds on capacity in the presence of errors, and Bovier (1999) provided sharp upper bounds on perfect retrieval and convergence.

More recent studies have focused on how structural modifications alter the attractor landscape. Folli et al. (2018) systematically analyzed the impact of dilution and asymmetric couplings, revealing phase-transition-like changes between fixed-point, cyclic, and chaotic regimes. Hwang et al. (2019) further quantified the exponential growth rates of limit cycles in asymmetric networks and identified thresholds at which synchronous dynamics transition to long-period or chaotic behavior. Collectively, these results demonstrate that convergence and cyclicity in HNNs are highly sensitive to symmetry and structural perturbations. In contrast, our work isolates the effect of the update rule itself—specifically the long-standing ambiguity in zero-field updates—and shows that resolving this ambiguity restores dynamical symmetry, yields fully characterizable regimes, and enables exact enumeration of fixed-point and cyclic behaviors.

Broader developments oriented HNNs are reviewed in Appendix C.3, *Other Frontier Research*.

C.3 Other Frontier Research

Beyond classical capacity and convergence analyses, Hopfield Neural Networks (HNNs) have continued to inspire diverse lines of frontier research spanning theory, biology, and modern machine learning. Early work by Barra et al. Barra et al. (2012) established a statistical-mechanics equivalence between HNNs and Boltzmann machines, deepening the conceptual link between associative memory and probabilistic models. Building on this connection, Smart Smart & Zilman (2021) showed that HNNs with correlated patterns can be exactly mapped to restricted Boltzmann machines via QR decomposition, overcoming the limitations of earlier analyses restricted to uncorrelated regimes.

In parallel, biologically inspired studies explored associative memory from a neuroscience perspective. Tyulmankov et al. Tyulmankov et al. (2021) investigated biologically plausible learning rules

in key–value memory networks, while Liang et al. Liang et al. (2021) examined whether insect neural systems could implement mechanisms analogous to word embeddings, highlighting intriguing intersections between biological cognition and machine learning.

Recent years have also seen rapid progress in applying modern Hopfield networks to practical learning problems. Schäfl et al. Schäfl et al. (2023) demonstrated strong performance of modern Hopfield networks on tabular data, and Xu et al. Xu et al. (2024) further addressed sparsity and non-rotational invariance via generalized sparse Hopfield layers. Lucibello et al. Lucibello & Mézard (2024) established conditions for exponential capacity scaling in dense associative memories, while Agliari et al. Agliari et al. (2024) leveraged Hopfield-like dynamics for regularization and improved generalization. Complementing these advances, Rolls et al. Rolls (2024) provided a comparative analysis between hippocampal memory systems and generative AI models, and Cabannes et al. Cabannes et al. (2024) analyzed training dynamics of associative memory networks under gradient descent.

On the theoretical front, modern Hopfield networks have been enriched by attention-based and kernelized formulations. Hu et al. Hu et al. (2023) introduced sparsemax-based attention to improve capacity and retrieval guarantees, later extending this framework to nonparametric, kernelized, and computationally efficient variants with provable sub-quadratic or sub-linear complexity Hu et al. (2024b;e;c). Related work clarified expressive limits via circuit complexity Li et al. (2024), proposed in-material learning rules for energy-efficient memory Niu et al. (2024), and unified Hopfield energies with Fenchel–Young losses and SparseMAP transformations Dos Santos et al. (2024). Further refinements addressed capacity enhancement, robustness, and outlier efficiency in large-scale Hopfield and Transformer-based models Wu et al. (2024); Hu et al. (2024a). Most recently, Hu et al. Hu et al. (2024f;d) provided a statistical characterization of conditional diffusion transformers, elucidating efficiency limits of prompt-based adaptation in foundation models.

The awarding of the 2024 Nobel Prize in Physics to John J. Hopfield and Geoffrey E. Hinton for their foundational contributions to artificial neural networks has further revitalized interest in HNNs. This recognition underscores the enduring influence of Hopfield's ideas and highlights the continued relevance of HNNs across theoretical neuroscience, statistical physics, and modern machine learning.