Associative Memory in Superconducting Optoelectronic Networks

A Deep Dive into Stage 8: Capacity, Landscapes, Hardware, and the Road to Attention

Micheal Ohagwu
obi@nani-inc.com
New York

June 16, 2026

Abstract

These notes expand Stage 8 of the SOEN tour — the network run *backward* as a content-addressable memory. We treat four facets in turn. First, **capacity and breakdown**: why a Hopfield network of N neurons stores only $\alpha_c N \approx 0.138\, N$ random patterns before crosstalk noise triggers a catastrophic collapse. Second, the **energy landscape**: stored patterns are local minima with basins of attraction, and the asynchronous dynamics can only descend, so a corrupted input rolls into the nearest memory. Third, the **superconducting realization**: a weight is a mutual inductance, its reciprocity supplies the symmetry Hopfield needs for free, and frustration in the resulting spin glass is exactly the origin of spurious memories. Fourth, the **modern dense Hopfield network**: sharpening the energy wells lifts capacity to exponential in N, and the one-step retrieval rule is precisely the attention operation at the heart of transformers. Each facet corresponds to one panel of the companion interactive explorer.

1 Recap: the classical Hopfield network

A Hopfield network [1] is N binary units $s_i \in \{-1, +1\}$ with symmetric weights $W_{ij} = W_{ji}$, $W_{ii} = 0$. To store a set of P patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^P$ we use the Hebbian rule

$$W_{ij} = \frac{1}{N} \sum_{\mu=1}^{P} \xi_i^\mu\, \xi_j^\mu, \qquad E(\mathbf{s}) = -\frac{1}{2} \sum_{i,j} W_{ij} s_i s_j. \tag{1}$$

Units update asynchronously by aligning with their local field $h_i = \sum_j W_{ij} s_j$:

$$s_i \leftarrow \operatorname{sgn}(h_i). \tag{2}$$

Energy never increases. Flip one unit $s_i \to s_i' = \operatorname{sgn}(h_i)$. Because W is symmetric with zero diagonal, only terms containing i change, and

$$\Delta E = -\,(s_i' - s_i)\, h_i \le 0, \tag{3}$$

since s_i' has the same sign as h_i. The energy is bounded below, so the dynamics must halt at a local minimum — an attractor. *Recall is just descent.* This is the same relaxation the SOEN dendrites and soma already perform; made recurrent and symmetric, it becomes memory.

2 Capacity and breakdown

Why can't we store unlimited patterns? Probe the field at unit i when the state equals a stored pattern $\boldsymbol{\xi}^\nu$:

$$h_i = \sum_j W_{ij} \xi_j^\nu = \underbrace{\xi_i^\nu}_{\text{signal}} + \underbrace{\frac{1}{N} \sum_{\mu \ne \nu} \sum_{j \ne i} \xi_i^\mu \xi_j^\mu \xi_j^\nu}_{\text{crosstalk } C_i}. \tag{4}$$

The signal term wants $s_i = \xi_i^\nu$ — the pattern is a fixed point if the crosstalk never overwhelms it. For random ± 1 patterns, C_i is a sum of $\approx (P-1)(N-1)/N$ independent $\pm 1/N$ terms, so it is approximately Gaussian with zero mean and variance

$$\sigma^2 \approx \frac{P}{N} \equiv \alpha. \tag{5}$$

A bit is unstable when $\xi_i^\nu C_i < -1$, with probability $P_{\text{err}} = \frac{1}{2}\operatorname{erfc}\big(1/\sqrt{2\alpha}\big)$. As the *load* $\alpha = P/N$ rises, P_{err} climbs; a full replica analysis [2] sharpens this into a phase transition at

$$\boxed{\alpha_c \approx 0.138} \tag{6}$$

Below α_c the stored patterns are stable attractors with only a few flipped bits; above it they abruptly cease to be minima and the network collapses into a *spin-glass* phase of exponentially many spurious states uncorrelated with anything stored (Fig. 1). Recall does not gently fade — it falls off a cliff. Besides this hard limit, even below it the network harbors **spurious memories**: sign-reverses $-\boldsymbol{\xi}^\mu$ (always present), and odd *mixture* states such as $\operatorname{sgn}(\boldsymbol{\xi}^1+\boldsymbol{\xi}^2+\boldsymbol{\xi}^3)$.

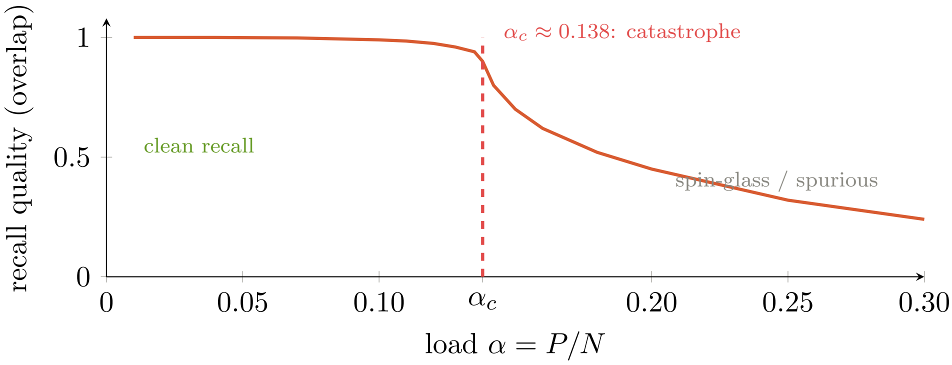

Figure 1: Recall quality versus memory load for random patterns. Recall is near-perfect until α approaches $\alpha_c \approx 0.138$, where the stored patterns abruptly stop being energy minima and the network dissolves into glassy spurious states.

3 The energy landscape and basins of attraction

Geometry makes recall intuitive. Each stored pattern is a local minimum of E; the set of initial states that flow to it under (2) is its *basin of attraction*. A corrupted input is a point displaced from a minimum; because energy can only fall, it slides into whichever basin it started in.

A clean way to see this for two stored patterns A, B is to track the *overlaps* (Mattis magnetizations)

$$m_A = \frac{1}{N}\,\boldsymbol{\xi}^A \cdot \mathbf{s}, \qquad m_B = \frac{1}{N}\,\boldsymbol{\xi}^B \cdot \mathbf{s}, \tag{7}$$

and project the dynamics onto the (m_A, m_B) plane. The four stored states $\pm A, \pm B$ sit at $(\pm 1, 0)$ and $(0, \pm 1)$; the energy, $E \approx -\frac{N}{2}(m_A^2 + m_B^2)$, has a well at each. The diagonals $m_A = \pm m_B$ are the *basin boundaries*: which memory you recover is decided by which side of the diagonal the corrupted start lands on (Fig. 2). Deepen the corruption past a boundary and the network confidently returns the *wrong* memory — a feature of all associative memories.

4 The superconducting realization

Hopfield's model demands two things that ordinary hardware gives only with effort: *symmetric* couplings and a physical *energy* that the dynamics minimize. Superconducting optoelectronic hardware supplies both for free.

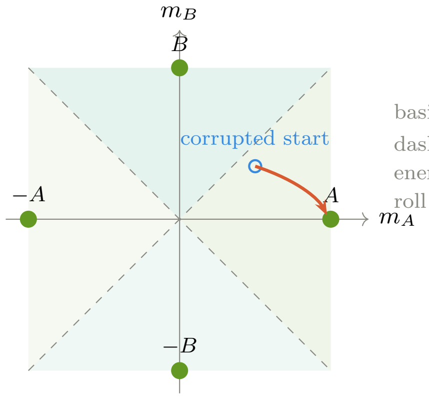

Figure 2: The two-pattern overlap plane. Stored memories $\pm A, \pm B$ are wells at the axis tips; the diagonals are basin boundaries. A corrupted state descends into the basin it starts in; crossing a boundary flips which memory is recalled.

A SOEN weight is implemented as a **mutual inductance** M_{ij} between two superconducting flux loops: the circulating current s_j in loop j couples a flux $\phi_i = \sum_j M_{ij} s_j$ into loop i. Mutual inductance is *reciprocal* — a theorem of magnetostatics, $M_{ij} = M_{ji}$ — so the weights are automatically symmetric, exactly the condition that guarantees the energy descent of Section 1. The loop currents are Ising spins, and the magnetic coupling energy is the Ising / spin-glass Hamiltonian $E = -\frac{1}{2} \sum M_{ij} s_i s_j$.

This identification also *explains* spurious memories. When couplings have mixed signs, a loop of bonds may be *frustrated*: no spin configuration can satisfy every bond simultaneously (Fig. 3b). The classic example is a triangle of antiferromagnetic bonds — any assignment leaves one bond unsatisfied, creating multiple degenerate metastable minima. These frustration-induced minima are the spin-glass states that, in memory language, are the spurious attractors not corresponding to any stored pattern.

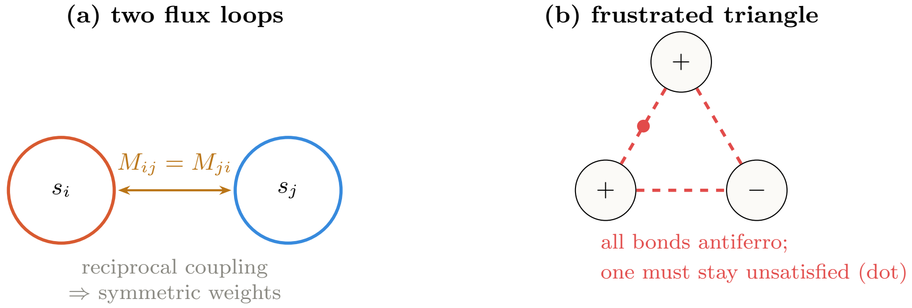

Figure 3: (a) A Hopfield weight is a mutual inductance between flux loops; reciprocity $M_{ij} = M_{ji}$ makes SOEN weights symmetric by construction. (b) Frustration: a triangle of antiferromagnetic bonds cannot be fully satisfied, producing the extra spin-glass minima that appear as spurious memories.

Hopfield model	**Superconducting realization**
weight W_{ij}	mutual inductance M_{ij} between flux loops
symmetry $W_{ij} = W_{ji}$	reciprocity $M_{ij} = M_{ji}$ (automatic)
spin $s_i = \pm 1$	loop-current direction
$E = -\frac{1}{2}\mathbf{s}^\top W \mathbf{s}$	magnetic coupling energy (Ising Hamiltonian)
recall	relaxation to an energy minimum
spurious memory	spin-glass state from frustration

5 Dense (modern) Hopfield networks and attention

The $0.138N$ ceiling is a consequence of the *quadratic* energy: with only pairwise terms, the wells are broad and overlap, so they crowd each other out. Modern ("dense") Hopfield networks [3, 4] replace the quadratic interaction with a sharply peaked one,

$$E(\mathbf{s}) = -\sum_{\mu=1}^{M} F\big(\boldsymbol{\xi}^{\mu} \cdot \mathbf{s}\big), \tag{8}$$

where F grows fast (a high polynomial, or for continuous states the log-sum-exp $F = \frac{1}{\beta} \log \sum_{\mu} e^{\beta\, \boldsymbol{\xi}^{\mu} \cdot \mathbf{s}}$). Sharper F digs each stored pattern its own deep, narrow, well-separated well, so the number of storable patterns grows *exponentially* in N rather than linearly. For the log-sum-exp energy, minimizing by a concave–convex step gives a *one-shot* retrieval update for a query $\mathbf{q}$:

$$\boxed{\mathbf{s}_{\text{new}} = X\,\text{softmax}\big(\beta\, X^{\top}\mathbf{q}\big)} \qquad X = [\boldsymbol{\xi}^{1}, \ldots, \boldsymbol{\xi}^{M}]. \tag{9}$$

The parameter β is an inverse temperature / sharpness: small β averages many patterns into a blurry mixture; large β collapses the softmax onto the single closest pattern and returns it exactly (Fig. 4).

This is attention. Equation (9) is term-for-term the transformer attention operation. Identify the query $\mathbf{q}$ with Q, the stored patterns X with both keys K and values V, and β with the inverse $1/\sqrt{d}$ scaling:

$$\text{Attention}(Q, K, V) = V\,\text{softmax}\Big(\tfrac{1}{\sqrt{d}} K^{\top} Q\Big). \tag{10}$$

A single attention layer is therefore one step of retrieval in a dense associative memory: the network "looks up" the stored item most similar to its query. For a cryogenic SOEN substrate — where the synapse already computes weighted inner products in flux and a soma already performs a soft threshold — this is an inviting target: the same loops that give classical Hopfield recall can, with sharper source functions, implement attention-like dense memory.

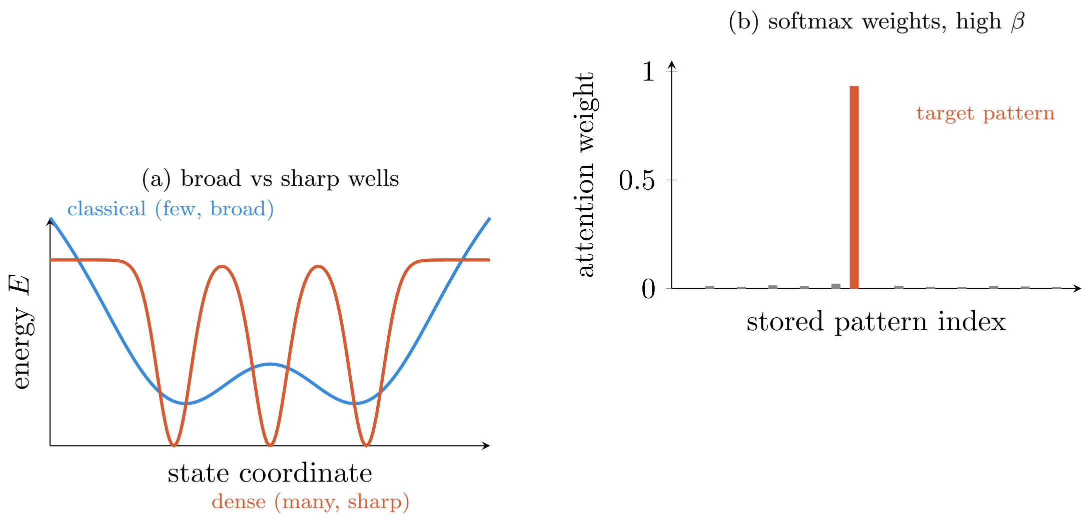

Figure 4: (a) Sharpening the energy (broad blue → sharp coral) separates the wells, lifting capacity from $\sim 0.14N$ to exponential in N. (b) At large β the softmax retrieval weights collapse onto the single closest stored pattern — exactly an attention lookup.

Synthesis

Associative memory is the network's relaxation dynamics viewed as descent on an energy landscape. Classically, Hebbian wells hold only $\alpha_c N \approx 0.138N$ random patterns before crosstalk noise tips the system into a spin glass; geometrically, recall is rolling into the nearest basin, and basin boundaries decide which memory wins. Superconducting hardware is an unusually natural host: mutual-inductance reciprocity delivers the required weight symmetry automatically, the coupling energy *is* the Ising Hamiltonian the dynamics minimize, and frustration in that spin glass is precisely the source of spurious memories. Finally, sharpening the energy turns the same machine into a dense memory of exponential capacity whose retrieval rule is the attention mechanism — suggesting that an energy-based, cryogenic, optically-communicating substrate could implement both classical recall and modern attention with one set of loops.

Four take-aways. (1) Capacity $\alpha_c \approx 0.138$, then catastrophic collapse. (2) Recall = energy descent into the nearest basin. (3) SOEN gives symmetric weights for free (reciprocal M_{ij}); frustration = spurious states. (4) Sharper wells $\Rightarrow$ exponential capacity, and the retrieval rule $X\,\mathrm{softmax}(\beta X^\top q)$ is attention.

References

[1] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, PNAS **79**, 2554 (1982).

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

[3] D. Krotov, J. J. Hopfield, *Dense Associative Memory for Pattern Recognition*, NeurIPS (2016).

[4] H. Ramsauer *et al.*, *Hopfield Networks is All You Need*, ICLR (2021); arXiv:2008.02217.

[5] J. M. Shainline, B. A. Primavera, R. O'Loughlin, *Relating Superconducting Optoelectronic Networks to Classical Neurodynamics*, arXiv:2409.18016 (2024), NIST.