Symmetric Hopfield storage: round 2

Claude Fable 5.1 campaign

2026-11-09

Problem and notation

Let n be the number of neurons, $[n] = \{1, \dots, n\}$, and $D = \{s^1, \dots, s^m\}$ a uniformly chosen set of m distinct vectors in $\{-1, 1\}^n$. The coordinate s_i^μ is neuron i in pattern μ. A single network, with symmetric weights $J \in \mathbb{R}^{n \times n}$, $J_{ii} = 0$, and biases $h \in \mathbb{R}^n$, stores D if every margin $s_i^\mu((Js^\mu)_i + h_i)$ is positive. Write $p_n(m)$ for its storage probability and u for a fixed load, with $m = \lfloor un \rfloor$. Then

$$\boxed{\alpha_- = \sup\{u > 0 : p_n(\lfloor un \rfloor) \longrightarrow 1\}.}$$

Current bounds: $1 \le \alpha_- \le 2$**.** To raise the lower bound, prove storage with probability tending to one at some $u > 1$. To lower the upper bound, prove a nonvanishing failure probability at some $u < 2$.

Theorems E and F: building a certificate of failure

A **certificate** is a nonzero array of nonnegative weights $\lambda = (\lambda_{i\mu}) \in \mathbb{R}^{n \times m}$ satisfying

$$\sum_{\mu=1}^{m} \lambda_{i\mu} s_i^\mu = 0, \qquad \sum_{\mu=1}^{m} (\lambda_{i\mu} + \lambda_{j\mu}) s_i^\mu s_j^\mu = 0 \quad (i < j).$$

The first equation holds for every i. Such an array exists exactly when D cannot be stored (round 1, Theorem B). Its **active set** consists of neurons whose row of λ is nonzero.

Theorem E (Exterior balance; S3-14)**.** Fix $1 < u < 2$ and $1 - u/2 < \beta < 1$. Choose a set $I_n \subseteq [n]$ of $\lfloor \beta n \rfloor$ neurons independently of D. With probability tending to one, every $i \in I_n$ admits weights $\pi_\mu \ge 0$ such that

$$\sum_\mu \pi_\mu = 1, \qquad \sum_\mu \pi_\mu s_i^\mu = 0, \qquad \sum_\mu \pi_\mu s_i^\mu s_k^\mu = 0 \quad (k \notin I_n).$$

Meaning. β is the active fraction; π is a probability vector on the patterns. These equations balance the bias and all connections to neurons outside I_n. They leave the connections between active neurons to be matched. The set I_n must be chosen before seeing D.

Contribution to α_-. E supplies the exterior part of a failure certificate; an upper bound still needs compatible interior weights.

Theorem F (Restricted certificates; S3-15)**.** For $I \subseteq [n]$, a certificate with $\lambda_{i\mu} = 0$ outside I exists if and only if no symmetric zero-diagonal network can make all margins positive on the neurons in I. This event is decided by one linear program and implies failure to store D.

Meaning. The network still has all n neurons, but only the storage constraints indexed by I are imposed. When $I = [n]$, this is the full storage problem. A proper subset can already certify failure.

Contribution to α_-. If this event has probability bounded away from zero at a fixed $u < 2$, then $\alpha_- \le u$.

Theorem G: failure requires many active neurons

Theorem G (Extensive certificates; S4-11)**.** For every fixed $0 < u < 2$, there is a constant $\beta(u) > 0$ such that, for every fixed $0 < \beta < \beta(u)$, with probability tending to one no certificate has at most βn active neurons. In particular, certificates supported on two neurons, or on $o(n)$ neurons, disappear in this limit.

Meaning. $\beta(u)$ is a positive threshold independent of n, and β is any smaller active fraction. The notation $o(n)$ means a number whose ratio to n tends to zero. Thus any surviving obstruction must involve a linear number of neurons. This is a necessary condition for failure; it does not assert that such certificates exist.

Contribution to α_-. G directs an upper-bound proof toward collective obstructions; excluding only small certificates cannot establish storage above load one.

Theorem H: a limitation of projection corrections

Theorem H (Complement corrections; S2-10, S2-11; verification pending)**.** Let D_0 be stored by its projection network. Let $V_0 \subseteq \mathbb{R}^{n+1}$ be the span of its augmented patterns $\hat{s} = (s, 1)$, and set $c = n + 1 - \dim V_0$. Add a symmetric $Q \in \mathbb{R}^{(n+1)\times(n+1)}$ to its augmented projector, with $Qv = 0$ for every $v \in V_0$. Extract J from the upper-left block with its diagonal removed, and h from the first n entries of the last column. This correction family has dimension $c(c+1)/2$. For $k \geq 1$ further independent uniform patterns,

$$\Pr(\text{some network in this family stores all patterns}) \leq \Pr\Big(\mathrm{Bin}(k-1, \tfrac{1}{2}) \leq c-1\Big)\,.$$

Starting from $m_0 = \lfloor (1-\varepsilon)n \rfloor$ patterns, with fixed $0 < \varepsilon < 1$, this probability tends to zero when the total load is any fixed $u > 1 + \varepsilon$.

Meaning. Augmentation includes the bias as one extra coordinate. The number c counts unused directions orthogonal to V_0. $\mathrm{Bin}(k-1, \frac{1}{2})$ counts successes in $k-1$ independent fair trials. The bound concerns this correction family, rather than all symmetric networks.

Contribution to α_-. If verified, H rules out this family as a route to a fixed improvement above one by taking $\varepsilon \downarrow 0$; it gives no upper bound on α_- itself.

Evidence and the next estimate

Exact finite experiments found symmetry-only failures in 11/15 samples at $n = 48$ and 24/28 at $n = 64$, at load $u = 1.7$, conditional on every neuron being separately feasible. These counts motivate an upper-bound estimate; they do not prove its asymptotic premise.

Next target. For some fixed $u < 2$ and $\beta < 1$, prove a nonvanishing probability of a certificate supported on $\lfloor \beta n \rfloor$ neurons. E supplies exterior balance, F links the certificate to failure, and G requires an extensive support. Matching the interior weights remains open. The recorded bounds remain $1 \leq \alpha_- \leq 2$.

Round 2 status. Three solver–verifier lanes and a shared experimenter ran for about 3.5 hours: 100.6 million tokens, approximately $219 at list prices. E–G were checked; H's verification was still in progress at this snapshot. A fresh final audit remains part of the campaign.