

¹ **Goldstone Modes and the Coexistence Saddle: Spectral Bifurcation**
² **Analysis of Competing Ring Attractors Under Mean-Field Cross-Inhibition**

⁴ **Tomás P. Pavan¹ and Claude²**

⁵ ¹Independent Researcher ²Anthropic (Claude Opus 4.6, Large Language Model)

⁶ *Correspondence:* Tomás P. Pavan

⁷

⁸ **Abstract**

⁹ Persistent neural activity in working memory is often modeled as bump attractors on ring networks.
¹⁰ When two such networks compete via cross-inhibition, the system must resolve which representation
¹¹ dominates – a winner-take-all (WTA) decision. We present a complete spectral bifurcation analysis
¹² of this transition in a coupled ring attractor model (two networks of $N = 48$ neurons each, cosine
¹³ connectivity within, mean-field cross-inhibition between). We identify six key results.

¹⁴ First, the coexistence fixed point – where both bumps self-sustain simultaneously – exists only below
¹⁵ a critical cross-inhibition strength $J_x^{\text{exist}} \approx 0.36$. Above this threshold, cross-inhibition is too strong
¹⁶ for both representations to survive; the system admits only WTA solutions.

¹⁷ Second, the continuous rotational symmetry of each bump guarantees exactly two Goldstone modes
¹⁸ (zero eigenvalues) that are protected against mean-field coupling. The first non-Goldstone eigenvalue
¹⁹ – governing uniform amplitude competition – crosses zero at $J_x^* \approx 0.3485$ via a pitchfork bifurcation,
²⁰ creating the coexistence saddle and two WTA stable states.

²¹ Third, the critical eigenvector projects maximally onto the spatially uniform (DC) direction, meaning
²² the instability concerns total activity competition rather than spatial pattern rearrangement. This
²³ is a direct consequence of mean-field (spatially unstructured) cross-inhibition.

²⁴ Fourth, large-scale stochastic simulations (128,000 trials across 256 parameter combinations) confirm
²⁵ the spectral predictions: swap errors emerge at the predicted J_x threshold, drive strength is
²⁶ secondary to cross-inhibition, and a non-monotonic valley at intermediate $J_x \approx 1.2\text{--}1.6$ – well above
²⁷ the coexistence threshold, in the pure winner-take-all regime – identifies a functional operating
²⁸ regime for working memory.

²⁹ Fifth, connectivity heterogeneity destroys the sharp pitchfork entirely, converting it into an imperfect
³⁰ bifurcation with no zero-crossing. The razor-thin instability window ($\Delta J_x \approx 0.01$) is a symmetry
³¹ artifact of the clean model; biological circuits operate in a regime of smooth crossover where no
³² parameter precision is required.

³³ Sixth, a Kramers escape analysis bridges the deterministic bifurcation and stochastic swap onset:
³⁴ because the barrier collapses quadratically ($\Delta V \propto |\lambda_{\text{dom}}|^2$), noise-driven escape becomes likely at
³⁵ $J_x^{\text{onset}} \approx 0.25$ – well below the pitchfork – for a normal-form coefficient $\gamma \approx 0.22\text{--}0.36$. We discuss
³⁶ implications for the behavioral cliff and argue that neural circuits operate in a valley regime where
³⁷ cross-inhibition and encoding drive are balanced, rather than near J_x^* itself.

³⁸

39 **1. Introduction**

40 **1.1 Working Memory and Competing Representations**

41 Persistent neural activity in prefrontal and parietal cortex underlies the short-term maintenance
42 of information in working memory (Goldman-Rakic, 1995; Funahashi et al., 1989). Ring attractor
43 models capture a key feature of this activity: spatially tuned neurons form a localized “bump” of
44 elevated firing that persists through recurrent excitation even after the sensory stimulus is removed
45 (Compte et al., 2000; Ben-Yishai et al., 1995; Amari, 1977). These bumps encode continuous
46 variables such as spatial location or orientation, and their precision is set by a balance between
47 recurrent drive and noise-induced diffusion along the ring (Wimmer et al., 2014; Burak and Fiete,
48 2012). Although continuous attractors are structurally unstable – destroyed by most infinitesimal
49 perturbations of the dynamics – Ságodi et al. (2024) showed that a persistent slow manifold survives
50 these bifurcations, rendering the attractor functionally robust for analog memory. Such ring attractor
51 dynamics have been observed experimentally in the *Drosophila* head direction system (Kim et al.,
52 2017) and are reviewed in the broader context of attractor and integrator networks by Khona and
53 Fiete (2022).

54 When multiple items must be stored simultaneously, as in multi-item visual working memory tasks,
55 the standard approach posits multiple bump networks coupled through cross-inhibition (Edin et
56 al., 2009; Wei et al., 2012). The cross-inhibition creates a competition: if it is weak, both bumps
57 coexist and the system maintains multiple items; if it is strong, one bump suppresses the other in
58 a winner-take-all (WTA) decision. The transition between these regimes determines the capacity
59 limit of the working memory circuit (Edin et al., 2009).

60 Despite the importance of this transition, its spectral structure – the full set of eigenvalues and
61 eigenvectors of the system’s Jacobian – has not been characterized. Previous analyses have focused
62 on one-dimensional (1D) mean-field reductions, projecting the high-dimensional dynamics onto a
63 single dominance variable $D = \bar{r}_A - \bar{r}_B$ and characterizing the resulting cusp catastrophe (Thom,
64 1972; Zeeman, 1977). While this captures the topology of the bifurcation, it discards the 96-
65 dimensional dynamics that include rotational modes, drift modes, and the full stability structure of
66 the coexistence state.

67 **1.2 The Behavioral Cliff**

68 Psychophysical experiments reveal a striking feature of working memory performance: below a
69 critical stimulus strength, accuracy does not degrade gradually but collapses abruptly – a “behavioral
70 cliff” (Bays et al., 2009; Zhang and Luck, 2008). In the mixture model framework, this manifests
71 as a sharp increase in the probability of reporting a non-target item (swap errors) or of random
72 guessing, even for small changes in signal-to-noise ratio.

73 The standard theoretical account attributes the cliff to noise-driven escape from a metastable state
74 (Kramers, 1940; Hanggi et al., 1990): when the cue is weak, the barrier between the correct attractor
75 and competing attractors is low, and stochastic fluctuations cause the system to fall to a wrong
76 state. This yields a cusp catastrophe potential $V(D) = D^4 + aD^2 + bD$, where a is controlled by
77 the circuit’s lateral inhibition and b by the cue strength. The cliff occurs at the cusp point where
78 the barrier vanishes.

79 However, this account treats the cliff as a cue phenomenon – a consequence of weak sensory input.
80 An alternative possibility, which we develop here, is that the cliff reflects a structural property of
81 the circuit: the proximity of the effective cross-inhibition strength J_x to a critical value J_x^* where

82 the coexistence state undergoes a spectral bifurcation.

83 **1.3 From Mean-Field Reduction to Full Spectral Analysis**

84 The 1D reduction $D = \bar{r}_A - \bar{r}_B$ captures the order parameter of the WTA transition but suppresses
85 95 of the 96 dynamical degrees of freedom. In particular, it cannot distinguish:

- 86 1. **Goldstone modes** – exactly-zero eigenvalues arising from the continuous rotational symmetry
87 of each bump (Goldstone, 1961; Burak and Fiete, 2012). These modes govern bump drift and
88 are protected by symmetry.
- 89 2. **Genuine instabilities** – eigenvalues that cross zero as parameters change, signaling structural
90 reorganization of the attractor landscape.
- 91 3. **The character of the critical mode** – whether the instability that destroys coexistence
92 projects onto the spatially uniform (DC) direction, is spatially patterned (cosine), or mixed.

93 Previous spectral approaches to ring networks have addressed non-Hermitian quasi-localization
94 (Tanaka and Nelson, 2018) and the stability of persistent activity under short-term plasticity
95 (Seeholzer et al., 2019), but the full eigenvalue structure of the *coupled* system has not been resolved.
96 We present the first complete eigenvalue decomposition of the coupled ring attractor Jacobian,
97 resolving all $2N = 96$ eigenvalues as a function of the cross-inhibition strength J_x . By cleanly
98 separating Goldstone modes from genuine instabilities, we identify the precise location, character,
99 and consequences of the coexistence-to-WTA pitchfork bifurcation.

100 **1.4 Summary of Contributions**

101 Our main results are:

- 102 1. **Existence threshold.** The coexistence fixed point exists only for $J_x < J_x^{exist} \approx 0.36$. At the
103 commonly used value $J_x = 0.5$, coexistence is not a fixed point of the dynamics – both bumps
104 cannot self-sustain under such strong cross-inhibition.
- 105 2. **Goldstone separation and pitchfork.** Two Goldstone modes (exactly-zero eigenvalues
106 protected by rotational symmetry) persist at all J_x where coexistence exists. The first non-
107 Goldstone eigenvalue λ_{dom} crosses zero at $J_x^* \approx 0.3485$, creating a pitchfork bifurcation where
108 the symmetric coexistence state becomes a saddle point and two WTA attractors are born.
- 109 3. **DC critical mode.** The critical eigenvector has its largest projection onto the uniform (DC)
110 direction: an increase in network A’s activity coupled with a decrease in network B’s, localized
111 to the active bump neurons by the gain mask $\sigma'(h_i)$. This reflects the mean-field character of
112 the cross-inhibition and means the instability is about total activity competition, not spatial
113 pattern rearrangement.
- 114 4. **Stochastic phase diagram.** A 128,000-trial parameter sweep confirms the spectral predictions
115 and reveals a non-monotonic valley at intermediate J_x where swap error rates dip to
116 7–13% between two qualitatively different failure modes.
- 117 5. **Heterogeneity destroys the sharp bifurcation.** Connectivity heterogeneity breaks the
118 exact A↔B exchange symmetry, converting the pitchfork into an imperfect bifurcation with
119 no zero-crossing. The razor-thin instability window ($\Delta J_x \approx 0.01$) is a symmetry artifact;
120 biological circuits operate in a smooth crossover regime.

121 These results reframe the behavioral cliff as a J_x -space phenomenon and identify a qualitative

¹²² operating regime (the valley) where encoding drive and cross-inhibition are balanced for reliable
¹²³ working memory.

¹²⁴

¹²⁵ 2. Model

¹²⁶ 2.1 Single Ring Attractor

¹²⁷ We consider a rate model with $N = 48$ neurons uniformly distributed on a ring. Each neuron i has
¹²⁸ a preferred angle $\theta_i = -\pi + 2\pi i/N$ and firing rate $r_i(t)$ governed by:

$$\tau \frac{dr_i}{dt} = -r_i + \sigma(h_i)$$

¹²⁹ where $\tau = 10$ ms is the time constant and $\sigma(h) = r_{max}/(1 + e^{-\beta(h-h_0)})$ is a sigmoidal activation
¹³⁰ function with parameters $r_{max} = 1.0$, $\beta = 5.0$, $h_0 = 0.5$. The total input to neuron i is:

$$h_i = \sum_{j=1}^N W_{ij} r_j + I_i^{ext}$$

¹³¹ where the within-network connectivity has cosine tuning:

$$W_{ij} = \frac{1}{N} (-J_0 + J_1 \cos(\theta_i - \theta_j))$$

¹³² with $J_0 = 1.0$ (uniform inhibition) and $J_1 = 6.0$ (tuned excitation). This connectivity supports a
¹³³ family of bump solutions at any angular position, forming a ring attractor (Amari, 1977; Ben-Yishai
¹³⁴ et al., 1995).

¹³⁵ 2.2 Coupled System with Mean-Field Cross-Inhibition

¹³⁶ We couple two identical ring networks A and B through mean-field cross-inhibition. The dynamics
¹³⁷ become:

$$\tau \frac{dr_i^A}{dt} = -r_i^A + \sigma \left(\sum_j W_{ij} r_j^A + I_i^{cue} - J_\times \bar{r}^B \right)$$

$$\tau \frac{dr_i^B}{dt} = -r_i^B + \sigma \left(\sum_j W_{ij} r_j^B - J_\times \bar{r}^A \right)$$

¹³⁸ where $\bar{r}^X = \frac{1}{N} \sum_j r_j^X$ is the mean activity of network X and $J_\times \geq 0$ is the cross-inhibition strength.
¹³⁹ The external cue input is a von Mises tuning curve applied to network A only:

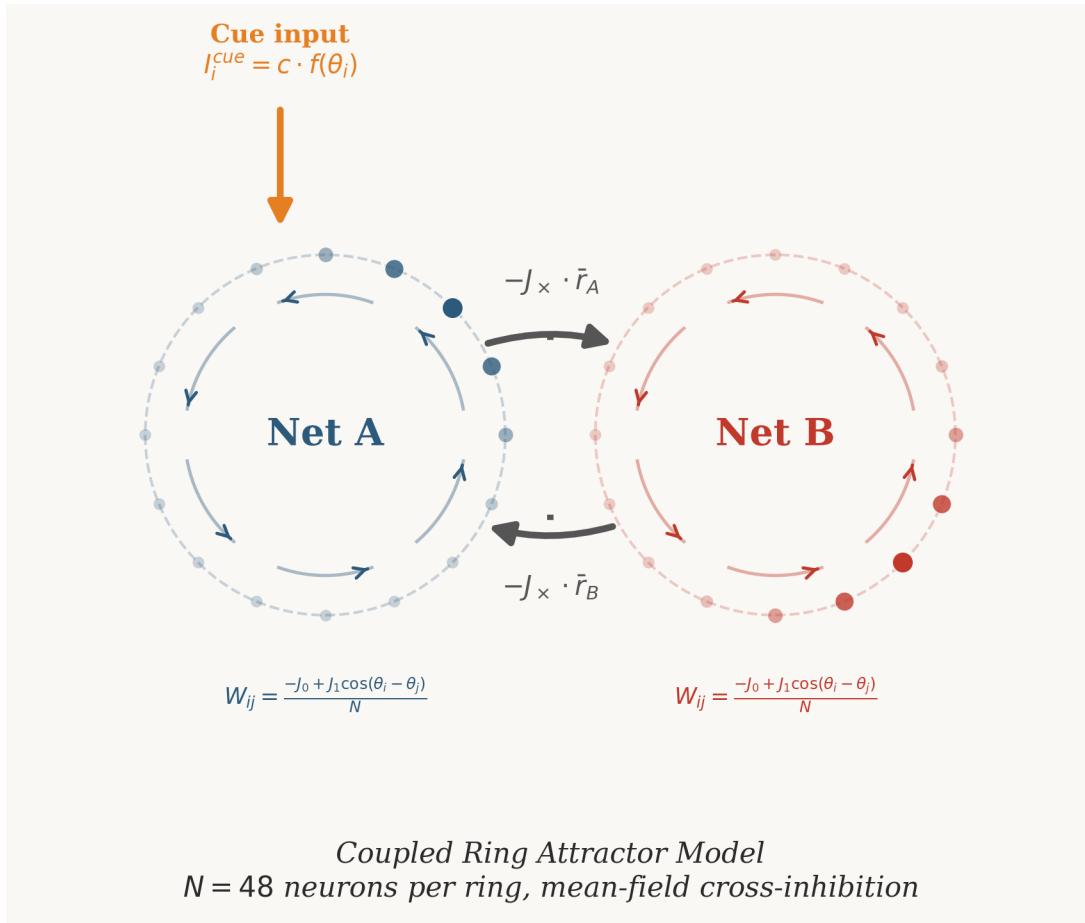
$$I_i^{cue} = c \cdot \frac{e^{\kappa \cos(\theta_i - \theta_{stim})}}{I_0(\kappa)}$$

¹⁴⁰ with concentration parameter $\kappa = 2.0$ and cue gain $c \geq 0$.

141 Throughout this paper, J_X denotes this single cross-inhibition parameter; we distinguish three
 142 characteristic values: the pitchfork bifurcation $J_X^* \approx 0.3485$, the coexistence existence boundary
 143 $J_X^{\text{exist}} \approx 0.358$, and the stochastic onset $J_X^{\text{onset}} \approx 0.25$.

144 We refer to the cue gain c as the *encoding drive* (or simply *drive*) when discussing its role in the
 145 stochastic simulations (Section 3.5), where it controls the strength of the stimulus that biases initial
 146 encoding toward network A. This is distinct from the initialization drives ($I_{\text{ext}} = 5.0$) used in the
 147 fixed-point solver (Section 3.1.1), which serve only to locate the bump-attractor state and are not
 148 varied as a parameter.

149 The critical feature of mean-field cross-inhibition is that it depends only on the total activity \bar{r}^X of
 150 the opposing network, not on the spatial pattern of its bump. This has profound consequences for
 151 the symmetry structure of the system (Fig. 1).



152 **Figure 1.** Model schematic. Two ring networks (A, B) of $N = 48$ neurons each, with cosine
 153 within-network connectivity ($J_0 + J_1 \cos \Delta\theta$) and mean-field cross-inhibition ($J_X \bar{r}^X$). External cue
 154 input drives network A only. The cross-inhibition depends on mean activity, not bump position,
 155 preserving rotational symmetry.

157 2.3 Jacobian of the Coupled System

158 The steady-state condition $F(\mathbf{r}^*) = 0$ defines the fixed points, where $F_i^A = -r_i^A + \sigma(h_i^A)$ and
 159 similarly for B. The Jacobian $\mathbf{J} = \partial F / \partial \mathbf{r}$ evaluated at a fixed point \mathbf{r}^* has a 2×2 block structure:

$$\mathbf{J} = \begin{pmatrix} -\mathbf{I} + \mathbf{S}_A \mathbf{W} & \mathbf{S}_A \mathbf{C} \\ \mathbf{S}_B \mathbf{C} & -\mathbf{I} + \mathbf{S}_B \mathbf{W} \end{pmatrix}$$

160 where $\mathbf{S}_X = \text{diag}(\sigma'(h_i^X))$ is the diagonal matrix of sigmoid derivatives at the fixed point, and
 161 $\mathbf{C} = -\frac{J_X}{N} \mathbf{1}\mathbf{1}^T$ is the rank-1 mean-field coupling matrix. The full Jacobian is $2N \times 2N = 96 \times 96$.
 162 The block structure reveals that the cross-coupling enters only through the rank-1 matrix \mathbf{C} . This
 163 low-rank perturbation to the block-diagonal within-network dynamics is what makes the spectral
 164 analysis tractable: the cross-inhibition can shift at most one eigenvalue per symmetry sector.

165 **2.4 Symmetries**

166 The coupled system possesses two symmetries at zero cue ($c = 0$):

167 **Continuous rotational symmetry.** The mean-field coupling $J_X \bar{r}^X$ is invariant under any rotation
 168 of the bump profile: if r_i^X is a fixed point, so is r_{i+k}^X for any shift k . This gives a continuous family
 169 of fixed points parametrized by bump position, and by Goldstone's theorem (Goldstone, 1961),
 170 each such continuous symmetry produces an eigenvalue that is exactly zero. With two independent
 171 bumps, there are two Goldstone modes.

172 **Discrete exchange symmetry.** At $c = 0$, the system is invariant under $A \leftrightarrow B$. The coexistence
 173 fixed point (where both bumps are present with $\bar{r}^A = \bar{r}^B$) respects this symmetry; the WTA states
 174 ($\bar{r}^A \gg \bar{r}^B$ or vice versa) break it. The transition between these is governed by a pitchfork bifurcation.

175 The nonzero cue $c > 0$ breaks the exchange symmetry (favoring network A) and deforms the
 176 pitchfork into an imperfect bifurcation with hysteresis.

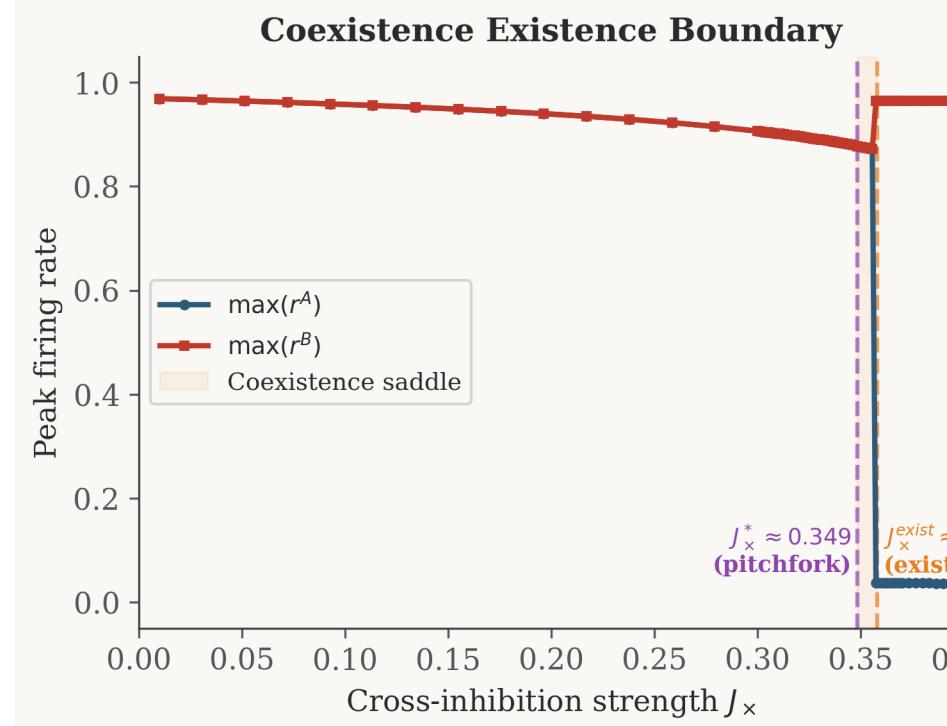
177

178 **3. Results**

179 **3.1 Existence of the Coexistence Fixed Point**

180 **3.1.1 Numerical Method** We locate fixed points using a two-phase approach. In Phase 1
 181 (simulation), we establish both bumps using strong external drives ($I_{ext} = 5.0$), then remove the
 182 drives and simulate the coupled system for 5×10^4 time steps ($\Delta t = 0.1, \tau = 10$) under the target
 183 cross-inhibition J_X . In Phase 2 (Newton), we polish the resulting state using Newton's method
 184 (scipy.optimize.fsolve) with the analytical Jacobian, achieving residuals $\|F(\mathbf{r}^*)\| < 10^{-10}$.

185 We verify convergence by checking: (i) the residual $\max_i |F_i(\mathbf{r}^*)| < 10^{-6}$; (ii) both bumps are active
 186 ($\max_i r_i^A > 0.3$ and $\max_i r_i^B > 0.3$); and (iii) the solution is not a WTA state ($|D| = |\bar{r}^A - \bar{r}^B| < 0.1$).



187 **3.1.2 Critical Existence Threshold**

188 **Figure 2.** Coexistence existence boundary. Peak firing rate of each network as a function of
 189 cross-inhibition strength J_x . Below $J_x^{\text{exist}} \approx 0.358$, both networks sustain bumps at matched
 190 amplitudes (coexistence). Above this threshold, one network collapses to baseline (WTA). The
 191 pitchfork bifurcation at $J_x^* \approx 0.349$ (orange dashed) and existence boundary at $J_x^{\text{exist}} \approx 0.358$
 192 (purple dashed) delimit the narrow instability window $\Delta J_x \approx 0.01$.

193 We scan J_x from 0 to 0.5 (Fig. 2). Below $J_x^{\text{exist}} \approx 0.36$, Newton converges to a genuine coexistence
 194 fixed point with residual $< 10^{-10}$. Above this threshold, one bump suppresses the other during
 195 the simulation phase; Newton converges only to WTA solutions. At $J_x = 0.50$ – a commonly used
 196 parameter value – coexistence does not exist as a fixed point of the deterministic dynamics.

197 The critical existence threshold lies between $J_x = 0.355$ (both bumps survive, $\max r^A =$
 198 $0.88, \max r^B = 0.88$) and $J_x = 0.360$ (one bump collapses, $\max r^A = 0.97, \max r^B = 0.04$). The
 199 transition is sharp: a 1.4% increase in J_x converts stable coexistence into complete dominance.

200 **3.1.3 Diagnostic: Fixed Point vs. Slow Manifold** At $J_x = 0.50$, a time-resolved diagnostic
 201 reveals that the system does not converge: residuals remain at $\sim 10^{-3}$ and the dominance variable
 202 D drifts monotonically toward ± 0.27 . At $J_x = 0.35$, residuals converge exponentially to machine
 203 precision ($\sim 10^{-15}$). This confirms that the coexistence state is a genuine fixed point below threshold
 204 and does not exist (even as a slow manifold) above it.

205 **3.2 Goldstone Modes and the Protected Symmetry**

206 **3.2.1 Origin of the Goldstone Modes** The mean-field cross-coupling $J_x \bar{r}^B$ is a function of
 207 the mean activity $\bar{r}^B = \frac{1}{N} \sum_j r_j^B$ only. Any continuous rotation of the bump profile preserves this
 208 mean. We now prove that this protects the rotational modes as exact null vectors of the full coupled
 209 Jacobian.

210 **Step 1: Rotational null vector of the uncoupled block.** Let $\mathbf{r}^{A*}(\varphi)$ denote the bump solution

211 of network A centered at phase φ . Because W_{ij} depends only on the angular difference $\theta_i - \theta_j$, the
 212 steady-state equation $-r_i^{A*} + \sigma(\sum_j W_{ij} r_j^{A*}) = 0$ holds for every φ . Differentiating both sides with
 213 respect to φ :

$$-\frac{\partial r_i^{A*}}{\partial \varphi} + \sigma'(h_i^{A*}) \sum_j W_{ij} \frac{\partial r_j^{A*}}{\partial \varphi} = 0$$

214 In matrix form, this is $(-\mathbf{I} + \mathbf{S}_A \mathbf{W}) \cdot \partial \mathbf{r}^{A*} / \partial \varphi = \mathbf{0}$, where $\mathbf{S}_A = \text{diag}(\sigma'(h_i^{A*}))$. The rotational
 215 derivative is an exact null vector of the uncoupled Jacobian block. (The same holds for network B
 216 by identical argument.)

217 **Step 2: Mean-field coupling annihilates the rotational mode.** Consider the $2N$ -dimensional
 218 perturbation $\mathbf{v}_A = (\partial \mathbf{r}^{A*} / \partial \varphi, \mathbf{0})^T$ corresponding to a shift of network A's bump alone. Multiplying
 219 by the full block Jacobian (Section 2.3) yields:

$$\mathbf{J} \cdot \mathbf{v}_A = \begin{pmatrix} (-\mathbf{I} + \mathbf{S}_A \mathbf{W}) \cdot \partial \mathbf{r}^{A*} / \partial \varphi + \mathbf{S}_A \mathbf{C} \cdot \mathbf{0} \\ \mathbf{S}_B \mathbf{C} \cdot \partial \mathbf{r}^{A*} / \partial \varphi + (-\mathbf{I} + \mathbf{S}_B \mathbf{W}) \cdot \mathbf{0} \end{pmatrix}$$

220 The upper block vanishes by Step 1. The survival of the zero eigenvalue depends entirely on the
 221 cross-coupling term $\mathbf{C} \cdot \partial \mathbf{r}^{A*} / \partial \varphi$ in the lower block. Recall that $\mathbf{C} = -\frac{J_\times}{N} \mathbf{1} \mathbf{1}^T$. Applying \mathbf{C} to the
 222 rotational derivative:

$$\mathbf{C} \cdot \frac{\partial \mathbf{r}^{A*}}{\partial \varphi} = -\frac{J_\times}{N} \mathbf{1} \left(\mathbf{1}^T \cdot \frac{\partial \mathbf{r}^{A*}}{\partial \varphi} \right) = -\frac{J_\times}{N} \mathbf{1} \cdot \sum_{j=1}^N \frac{\partial r_j^{A*}}{\partial \varphi}$$

223 Because a rotation merely translates the bump profile around the periodic ring, the total activity
 224 (and thus the mean) is strictly conserved. Exchanging derivative and sum:

$$\sum_{j=1}^N \frac{\partial r_j^{A*}}{\partial \varphi} = \frac{\partial}{\partial \varphi} \sum_{j=1}^N r_j^{A*} = \frac{\partial}{\partial \varphi} (N \bar{r}^A) = 0$$

225 Since $\mathbf{1}^T \cdot \partial \mathbf{r}^{A*} / \partial \varphi = 0$, it follows that $\mathbf{C} \cdot \partial \mathbf{r}^{A*} / \partial \varphi = \mathbf{0}$. The rank-1 coupling matrix completely
 226 annihilates the rotational derivative, yielding $\mathbf{J} \cdot \mathbf{v}_A = \mathbf{0}$. By identical logic for network B,
 227 $\mathbf{v}_B = (\mathbf{0}, \partial \mathbf{r}^{B*} / \partial \varphi)^T$ is also a null vector. \square

228 **Equivariance structure.** The result follows from the $\text{SO}(2) \times \text{SO}(2)$ equivariance of the coupled
 229 system at $c = 0$: the dynamics commute with independent rotations of each ring. The mean-field
 230 coupling $J_\times \bar{r}^X$ is invariant under both $\text{SO}(2)$ actions because it depends only on total activity,
 231 which is a rotation-invariant functional. Each $\text{SO}(2)$ factor contributes one Goldstone mode to the
 232 kernel of the Jacobian. This protection is exact and holds at all J_\times where the coexistence fixed
 233 point exists – it cannot be lifted by increasing cross-inhibition, only by breaking the rotational
 234 symmetry of either the within-network connectivity or the cross-coupling structure.

235 *Remark (circulant structure).* The within-network connectivity $W_{ij} = \frac{1}{N}(-J_0 + J_1 \cos(\theta_i - \theta_j))$ is
 236 exactly circulant: $W_{ij} = w(i - j \bmod N)$ for all N . The circulant structure ensures that the discrete
 237 rotation operator $T : r_i \mapsto r_{i+1 \bmod N}$ commutes with \mathbf{W} , i.e., $T\mathbf{W} = \mathbf{W}T$. Because the fixed-point
 238 equation $-\mathbf{r}^* + \sigma(\mathbf{W}\mathbf{r}^*) = \mathbf{0}$ is solved by any rotation $T^k \mathbf{r}^*$, the nullspace of the Jacobian contains

239 the rotational tangent vector $\partial \mathbf{r}^*/\partial \varphi$ at *every* N , not merely in the continuous limit. The protection
240 is exact at finite N , not an asymptotic approximation.¹

241 This is the neural circuit analog of the Goldstone theorem (Goldstone, 1961): a spontaneously broken
242 continuous symmetry produces a massless (zero-energy) excitation. In our context, “massless” means
243 neutrally stable – perturbations along the Goldstone direction neither grow nor decay. Because
244 mean-field coupling acts exclusively on the spatially uniform mode (**1**), it is perfectly orthogonal
245 to the zero-sum rotational modes, mathematically protecting positional memory from amplitude
246 competition.

247 **3.2.2 Numerical Identification** We classify eigenvalues into Goldstone candidates ($|\lambda| < 10^{-3}$)
248 and genuine modes ($|\lambda| \geq 10^{-3}$). For each eigenvector \mathbf{v} , we compute projections onto six basis
249 directions:

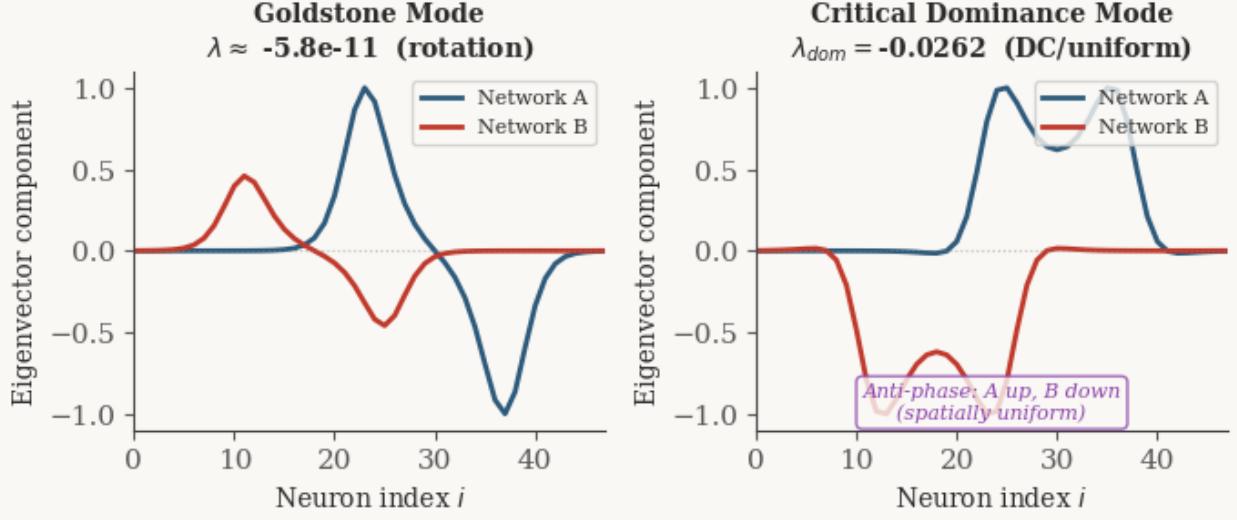
- 250 • \mathbf{d}_{dom} : symmetric dominance (cosine envelope, $A \uparrow B \downarrow$)
- 251 • $\mathbf{d}_{drift,+}$: co-directional drift (sine envelope, both shift same way)
- 252 • $\mathbf{d}_{drift,-}$: anti-directional drift (sine envelope, shift opposite ways)
- 253 • \mathbf{d}_{uni} : uniform/DC ($A \uparrow B \downarrow$ flat)
- 254 • $\mathbf{d}_{gold,A}$: rotation of bump A (sine envelope, A only)
- 255 • $\mathbf{d}_{gold,B}$: rotation of bump B (sine envelope, B only)

256 The Goldstone eigenvectors project strongly onto $\mathbf{d}_{gold,A}$ and $\mathbf{d}_{gold,B}$, confirming their rotational
257 character (Fig. 4, left panel).

258 **3.2.3 Goldstone Count Across J_x** Across the entire range $J_x \in [0, 0.36]$ where coexistence
259 exists, we find exactly two Goldstone modes. Their eigenvalues remain at $|\lambda| \sim 10^{-8}$ to 10^{-11}
260 (machine precision for our iterative solver), and they are never lifted by increasing cross-inhibition.
261 This confirms the symmetry protection: mean-field coupling cannot break rotational invariance.

¹Computational verification: at the coexistence fixed point ($J_x = 0.30$), the two Goldstone eigenvalue magnitudes are 1.1×10^{-5} ($N = 24$), 3.2×10^{-10} ($N = 48$), and 6.7×10^{-15} ($N = 96$), converging to machine precision as the angular grid refines, while the circulant commutation $\|T\mathbf{W} - \mathbf{W}T\|_\infty$ is below 4×10^{-16} at all N . The first non-Goldstone eigenvalue remains stable at $\lambda_1 \approx -0.131$ across all three resolutions.

Eigenvector Comparison at $J_x = 0.34$



262

263 **Figure 4.** Eigenvector comparison at $J_x = 0.34$ (near the pitchfork). Left: Goldstone mode
 264 ($\lambda \approx -5.8 \times 10^{-11}$, effectively zero), showing sinusoidal spatial structure in Network A – this
 265 is the rotational mode that slides the bump around the ring. Right: Critical dominance mode
 266 ($\lambda_{dom} = -0.026$, DC/uniform), showing anti-phase amplitude modulation localized to the active
 267 bump neurons. Because baseline neurons are strongly inhibited ($\sigma'(h_i) \approx 0$), they cannot participate
 268 in the linear instability; the mode is expressed only where the gain is nonzero. Despite this spatial
 269 localization, the mode projects maximally onto the uniform mean-field direction because the net
 270 effect is a difference in total activity between networks. The two modes are qualitatively distinct:
 271 the Goldstone mode encodes *where* the bump sits; the critical mode encodes *which network wins*.

272 3.3 The Pitchfork Bifurcation

273 **3.3.1 The First Non-Goldstone Eigenvalue** After removing the two Goldstone modes, we
 274 track the dominant genuine eigenvalue λ_{dom} as a function of J_x (Figs. 3, 9). Key findings:

- 275 • At $J_x = 0$: $\lambda_{dom} = -0.572$ (strongly stable). Without cross-inhibition, the coexistence state
 276 is deeply attractive.
- 277 • λ_{dom} increases monotonically with J_x , crossing zero at $J_x^* \approx 0.3485$.
- 278 • Above J_x^* : $\lambda_{dom} > 0$ (saddle). The coexistence state acquires one unstable direction.
- 279 • At $J_x = 0.356$: $\lambda_{dom} = +0.025$, and coexistence ceases to exist shortly after at $J_x^{\text{exist}} \approx 0.358$.

280 The crossing at J_x^* is a pitchfork bifurcation: the symmetric coexistence state ($D = 0$) loses stability,
 281 and two WTA states ($D > 0$ and $D < 0$) emerge as the new stable attractors. The $A \leftrightarrow B$ exchange
 282 symmetry is spontaneously broken.

283 **3.3.2 Character of the Critical Eigenvector** At J_x^* , the critical eigenvector has the following
 284 projections:

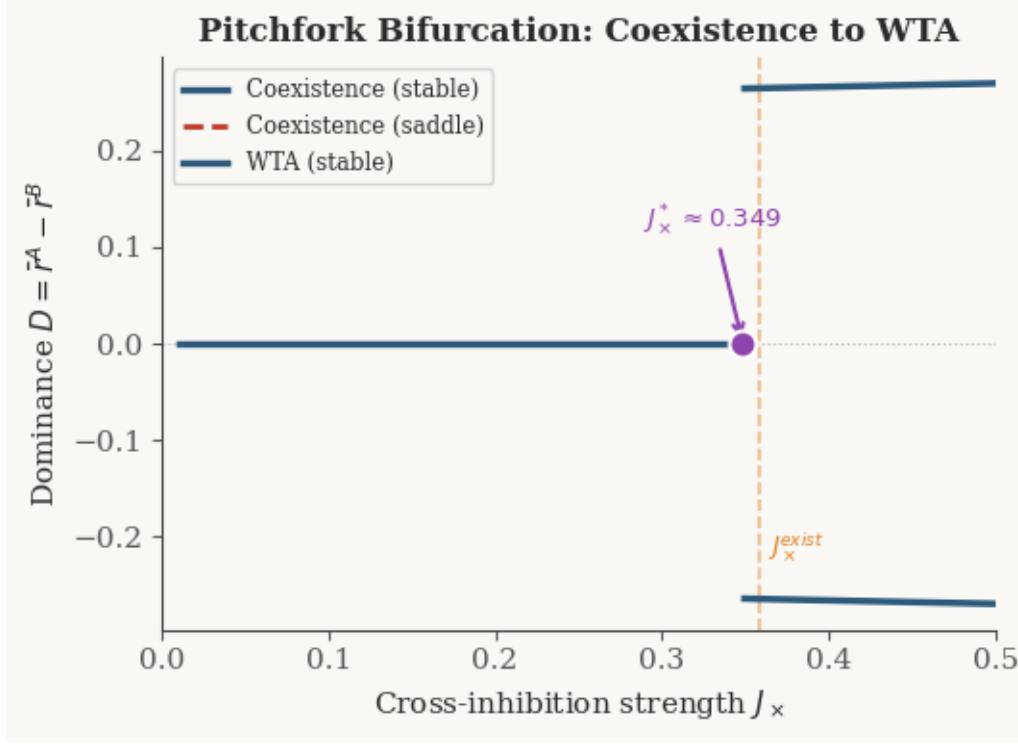
Direction	$ \langle \mathbf{v}_1, \mathbf{d} \rangle $
Uniform (DC)	0.43
Dominance (cosine)	0.34
Anti-drift (sine)	0.34
Co-drift (sine)	0.00

- 285 The largest projection is onto the uniform/DC direction (Fig. 4, right panel). Rather than a
 286 spatially flat baseline shift, the mode drives a gain-weighted amplitude modulation: a sharp increase
 287 in the active neurons of network A coupled with a sharp decrease in the active neurons of network B,
 288 while baseline neurons (where $\sigma'(h_i) \approx 0$) are effectively silent. Because this amplitude competition
 289 produces a net difference in total activity between networks, it projects maximally onto the mean-
 290 field coupling direction. The instability is about which network has more total activity, not about
 291 the spatial pattern of either bump.
- 292 This is a direct and falsifiable prediction of mean-field coupling. If cross-inhibition were spatially
 293 structured (depending on the relative positions of the two bumps), the critical eigenvector would
 294 acquire spatial structure (cosine or higher Fourier modes). The DC character is specific to coupling
 295 that “sees” only total activity.

296 **3.3.3 The Narrow Existence Window** The coexistence saddle – genuinely unstable, not merely
 297 Goldstone-neutral – exists only in the interval $J_x \in [0.3485, 0.358]$, a width of $\Delta J_x \approx 0.01$. Below
 298 J_x^* , coexistence is a stable node. Above $J_x^{exist} \approx 0.358$, it ceases to exist entirely.

299 This razor-thin window has two implications:

- 300 1. **Structural precision.** The bifurcation is sharp: a 3% change in J_x (from 0.348 to 0.358)
 301 takes the system from stable coexistence through saddle instability to complete collapse. The
 302 system is tuned near a critical point.
- 303 2. **Heterogeneity prediction.** In biological circuits with heterogeneous connectivity, the sharp
 304 boundary should be smeared into a broader regime where saddle-like dynamics persist (see
 305 Discussion).



306

307 **Figure 3.** Pitchfork bifurcation diagram. Dominance $D = \bar{r}^A - \bar{r}^B$ vs. cross-inhibition strength
 308 J_x . The coexistence branch ($D = 0$) is stable (solid) for $J_x < J_x^*$ and becomes a saddle (dashed
 309 red) above the pitchfork at $J_x^* \approx 0.349$. Two WTA branches ($D > 0$ and $D < 0$, blue) emerge at
 310 $J_x^{exist} \approx 0.358$ as stable attractors. The subcritical structure creates the narrow existence window
 311 $\Delta J_x \approx 0.01$. Note: the unstable branches connecting the pitchfork to the saddle-node fold at J_x^{exist}
 312 are not shown; our continuation solver tracked stable and saddle fixed points only. The expected
 313 topology is a pair of unstable branches emerging from J_x^* at $D = 0$ and folding onto the stable
 314 WTA branches at J_x^{exist} .

315 3.4 The Coexistence Saddle Under Cue

316 Newton continuation in cue gain c from 0 to 0.5 at $J_x = 0.35$ (within the saddle window; see Fig.
 317 3) reveals:

- 318 • The coexistence branch maintains both bumps across the full cue range, with D growing
 319 slowly as the cue favors network A.
- 320 • All 51/56 tracked solutions are unstable (saddle points), with $n_{positive} = 1$ for $c < 0.15$ and
 321 $n_{positive} = 2$ for $c > 0.15$ (a second eigenvalue crossing).
- 322 • The critical eigenvector's projection onto cosine and sine directions is $|\cos| \approx |\sin| \approx 0.47$ – a
 323 mixed dominance-drift mode at 45 degrees.

324 The WTA branch (tracked simultaneously) is stable across all cue values, confirming that the
 325 saddle's unstable manifold connects to the two WTA basins.

326 3.5 Stochastic Phase Diagram of Swap Errors

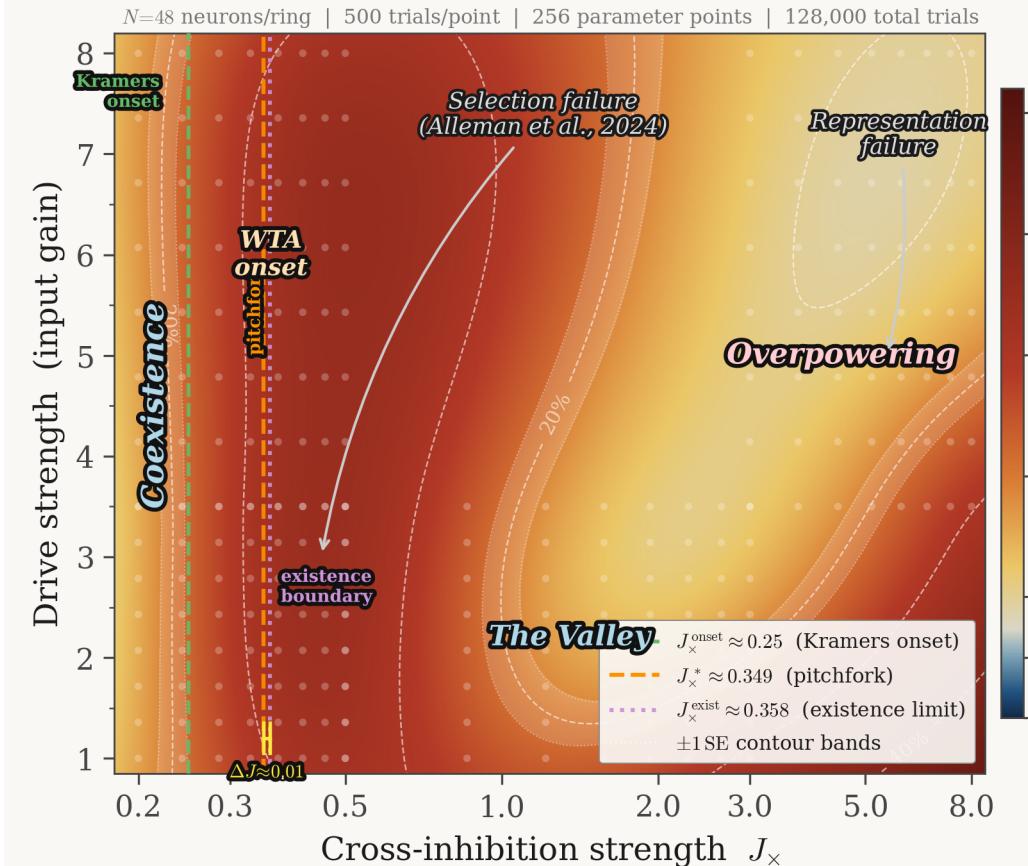
327 **3.5.1 Parameter Sweep** To bridge the deterministic bifurcation analysis with behavioral pre-
 328 dictions, we performed a large-scale stochastic simulation across the (J_x, c) parameter space. The

“drive strength” axis of the phase diagram corresponds to the cue gain c applied during a brief encoding period (Section 2.2), not to the initialization drives used in the fixed-point analysis. At each of 256 grid points (16 values of J_x from 0.05 to 8.0, 16 values of drive strength from 1.0 to 8.0), we ran 500 stochastic trials ($N_{total} = 128,000$). Two ring networks encoded items separated by $\pi/2$ radians, with independent Gaussian noise ($\sigma = 0.1$) added to each neuron during a 500-step maintenance period. Swap errors were classified as decoded responses within 0.3 rad of the non-target item’s location. To assess sensitivity to this classification threshold, we repeated a focused sweep (16 values of J_x from 0.1 to 3.0, two drive levels, 200 trials each) with thresholds of 0.2, 0.3, 0.4, and 0.5 rad. The qualitative features of the phase diagram – the onset location ($J_x^{\text{onset}} \approx 0.20\text{--}0.23$ across thresholds), the peak near $J_x \approx 0.3\text{--}0.5$, and the non-monotonic valley at $J_x \approx 1.2\text{--}1.5$ – are robust to threshold choice. Wider classification windows increase absolute swap rates (e.g., 10% vs. 26% at $J_x = 0.25$ for 0.2 vs. 0.5 rad) but preserve the ordering across J_x values and the location of qualitative transitions.

Binomial standard errors for the estimated swap rates are $SE = \sqrt{p(1-p)/500}$, ranging from $\pm 1.0\%$ at the onset threshold ($p \approx 0.05$) to $\pm 2.2\%$ at chance ($p = 0.50$). These uncertainties are small relative to the qualitative features of the phase diagram (onset band, valley depth, plateau height).

The stochastic sweep extends to $J_x = 8.0$, well beyond the coexistence existence threshold $J_x^{\text{exist}} \approx 0.36$. For $J_x > J_x^{\text{exist}}$, the system admits only WTA solutions; the stochastic trials begin from a coexistence-like initial condition (both bumps established by strong drives) and the dynamics reveal which network captures the WTA state. This extended range is necessary to characterize the full behavioral landscape, including the non-monotonic valley at $J_x \approx 1.2\text{--}1.6$ (Section 3.5.4), which operates in the pure WTA regime. The cross-inhibition parameter J_x has the same definition throughout: the coefficient in the mean-field coupling term $J_x \bar{r}^X$ (Eq. 2). No rescaling is applied between the deterministic and stochastic analyses.

Phase diagram of swap errors in coupled ring attractors



3.5.2 Onset of Swap Errors

Figure 5. Phase diagram of swap errors in coupled ring attractors. Heatmap shows swap error rate in (J_x, c) parameter space from 128,000 trials (500 per grid point, 256 parameter combinations across 4 agents). Solid and dashed white contour lines mark 5%, 10%, 20%, 30%, and 40% swap rates; flanking dotted lines show $\pm 1 \text{ SE}$ binomial uncertainty bands ($\text{SE} = \sqrt{p(1-p)/500}$). Vertical dashed lines mark three deterministic thresholds: Kramers onset ($J_x^{\text{onset}} \approx 0.25$, green), pitchfork bifurcation ($J_x^* \approx 0.349$, orange), and coexistence existence boundary ($J_x^{\text{exist}} \approx 0.358$, purple). The narrow gap $\Delta J \approx 0.01$ between pitchfork and existence boundary (yellow bracket, bottom) defines the coexistence window predicted by deterministic mean-field theory. Region labels identify qualitatively distinct dynamical regimes: coexistence (low J_x), WTA onset (sharp transition band), the valley (optimal discrimination at intermediate J_x), and overpowering (high J_x , representation failure). Annotations map selection failure and representation failure regimes to the framework of Alleman et al. (2024).

Swap errors emerge at $J_x^{\text{onset}} \approx 0.25$, consistent with the spectral prediction of the pitchfork bifurcation at $J_x^* \approx 0.3485$. The stochastic onset is lower than the deterministic bifurcation because noise-mediated escape from the metastable coexistence well occurs when the barrier height $\Delta V \sim \sigma^2$, which corresponds to a J_x slightly below the eigenvalue crossing. This is precisely the Kramers mechanism.

Between $J_x^{\text{onset}} \approx 0.25$ and 0.5, swap rates increase from 5% to approximately 45%. Above $J_x \approx 1.0$, swap rates plateau near 50% – the system has become a noise-driven bistable switch with no memory of the initial encoding.

374 **3.5.3 Drive Strength Is Secondary** The phase diagram shows near-vertical isocontours of
375 swap rate (Fig. 5): swap error probability depends primarily on J_X and only weakly on drive
376 strength. This is a direct prediction of the spectral analysis: the critical eigenvector projects
377 maximally onto the uniform (DC) direction, governing total activity competition rather than spatial
378 encoding. Stronger drive does not protect against the dominance instability because the instability
379 is orthogonal to the encoding direction.

380 This has a counterintuitive implication: increasing stimulus strength – the commonly proposed
381 intervention for working memory failures – targets the wrong degree of freedom. The cliff is a J_X
382 phenomenon, not a cue phenomenon.

383 **3.5.4 The Non-Monotonic Valley and Two Failure Regimes** At $J_X \approx 1.2\text{--}1.6$ with moderate
384 to strong drive, the phase diagram reveals a non-monotonic feature: swap rates dip to 7–13% between
385 two distinct failure modes. Crucially, this valley lies well above the coexistence existence threshold
386 ($J_X^{exist} \approx 0.36$), meaning it operates in a pure WTA regime where deterministic coexistence does
387 not exist. The two failure modes and the valley between them correspond to qualitatively different
388 dynamical regimes:

- 389 1. **Near-critical swaps** ($J_X \approx 0.3\text{--}0.5$): Near and just above the pitchfork, barriers separating
390 coexistence from WTA are small and noise escapes freely. This is the *representation failure*
391 regime: stochastic dynamics push the system from (metastable) coexistence into an incorrect
392 WTA state. Swap rate approaches 50%.
- 393 2. **Overpowering swaps** ($J_X > 2.0$, weak drive): Cross-inhibition is so strong that it overwhelms
394 feedforward encoding during the stimulus presentation itself. One network suppresses the other
395 before encoding is complete. This is also representation failure, but driven by the *encoding*
396 phase rather than the maintenance phase.
- 397 3. **The valley** ($J_X \approx 1.2\text{--}1.6$, strong drive): At these J_X values, coexistence does not exist as a
398 deterministic fixed point – the system is in a pure WTA regime. Strong encoding drive biases
399 the initial WTA competition so that the cued network typically wins. Swap errors here arise
400 when stochastic fluctuations during the encoding-to-maintenance transition cause the wrong
401 network to capture the WTA state. We hypothesize that this corresponds to the *selection*
402 *failure* mechanism identified by Alleman et al. (2024), where both representations are briefly
403 encoded but the wrong one is selected at readout.

404 To operationalize this distinction, we define a diagnostic: at the end of the maintenance period,
405 if both networks retain above-threshold activity ($\max_i r_i^X > 0.3$ for $X \in \{A, B\}$), a swap error is
406 classified as selection failure (both representations survived but the wrong one was decoded). If one
407 network has collapsed ($\max_i r_i^X < 0.1$), it is classified as representation failure (one item was lost
408 during maintenance). Applying this diagnostic across the phase diagram would test whether the
409 valley is indeed a selection-dominated regime while the near-critical zone is representation-dominated.
410 We leave this analysis for future work but note that the model predicts a crossover between these
411 two error types as a function of J_X .

412 The valley thus represents a candidate functional operating regime for *selection*, not for *maintenance*
413 or *coexistence*: cross-inhibition strong enough to resolve competition via WTA, encoding strong
414 enough to bias that competition correctly. The specific parameter range ($J_X \approx 1.2\text{--}1.6$) depends on
415 our model parameterization and should not be interpreted as a direct physiological prediction; the
416 qualitative feature – a non-monotonic minimum between two failure modes – is the robust finding.
417 The circuit need not be tuned precisely to J_X^* but rather to a regime where WTA dynamics and

⁴¹⁸ encoding strength are balanced.

⁴¹⁹ The valley's location in parameter space ($J_{\times} \approx 1.2\text{--}1.6$) is specific to the sigmoid nonlinearity
⁴²⁰ and gain parameters used here (Section 2.1). The qualitative structure – two competing failure
⁴²¹ modes generating a non-monotonic minimum – is the robust finding, as it depends on the generic
⁴²² competition between cross-inhibition strength and noise-driven escape.

⁴²³ **3.5.5 Kramers Barrier Estimate from the Dominance Eigenvalue** To quantitatively bridge
⁴²⁴ the deterministic pitchfork at $J_{\times}^* \approx 0.349$ with the stochastic onset of swap errors, we approximate
⁴²⁵ the escape barrier out of the coexistence state by projecting the high-dimensional dynamics onto
⁴²⁶ the critical dominance eigenvector \mathbf{v}_{dom} identified in the Jacobian spectrum.

⁴²⁷ Near J_{\times}^* , the dynamics admit a reduction onto the scalar amplitude $x(t) = \langle \mathbf{v}_{\text{dom}}, \delta \mathbf{r}(t) \rangle$, yielding
⁴²⁸ the saturating normal form

$$\tau \dot{x} = \lambda_{\text{dom}}(J_{\times}) x + \gamma x^3 - \delta x^5 + \eta_{\text{eff}}(t), \quad \delta > 0,$$

⁴²⁹ where $\lambda_{\text{dom}}(J_{\times})$ is the measured dominant non-Goldstone eigenvalue and the quintic term captures
⁴³⁰ saturation that stabilizes the distant WTA states. The escape barrier from coexistence is set by the
⁴³¹ nearby inner saddle and is therefore controlled primarily by λ_{dom} and γ .

⁴³² In the metastable subcritical regime, the effective potential $V(x)$ has an inner unstable saddle at

$$x_s^2 \approx \frac{|\lambda_{\text{dom}}|}{\gamma},$$

⁴³³ giving a barrier height

$$\Delta V \equiv V(x_s) - V(0) \approx \frac{|\lambda_{\text{dom}}|^2}{4\gamma}.$$

⁴³⁴ **Noise projection and finite-horizon criterion.** In the stochastic simulations, independent
⁴³⁵ Gaussian noise of standard deviation $\sigma = 0.1$ is added per neuron over a $T = 500$ -step maintenance
⁴³⁶ window. Because \mathbf{v}_{dom} is unit-normalized, projecting isotropic independent noise onto the dominance
⁴³⁷ coordinate preserves variance: $\text{Var}[\eta_{\text{eff}}] = \sigma^2$. Kramers theory predicts an escape rate $k \sim k_0 \exp(-\Delta V/\sigma^2)$; over a finite horizon T , escape becomes likely when $kT \sim 1$, i.e.
⁴³⁸

$$\frac{\Delta V}{\sigma^2} \approx \ln(k_0 T).$$

⁴³⁹ Taking a conservative attempt-frequency range $k_0 \in [0.1, 1]$ per step gives $\ln(k_0 T) \in [\ln 50, \ln 500] \approx$
⁴⁴⁰ $[3.9, 6.2]$, hence $\Delta V \in [0.039, 0.062]$.

⁴⁴¹ **Numerical evaluation at the observed swap onset.** The stochastic phase diagram shows swap
⁴⁴² errors rising steeply near $J_{\times}^{\text{onset}} \approx 0.25$. At this coordinate, the computed coexistence spectrum
⁴⁴³ gives $\lambda_{\text{dom}}(0.25) \approx -0.2357$ (Fig. 7A), yielding

$$\Delta V(0.25) \approx \frac{(0.2357)^2}{4\gamma} \approx \frac{0.0139}{\gamma}.$$

⁴⁴⁴ Equating to the finite-horizon Kramers threshold $\Delta V \in [0.039, 0.062]$ implies

$$\gamma \approx 0.22 \text{ to } 0.36,$$

⁴⁴⁵ a plausible normal-form coefficient for Taylor-expanded sigmoidal population dynamics. As a
⁴⁴⁶ self-consistency check: taking $\gamma = 0.3$ gives $\Delta V(0.25) \approx 0.046$ and $\Delta V/\sigma^2 \approx 4.6$, squarely in the
⁴⁴⁷ $\ln(k_0 T)$ band (Fig. 7C).

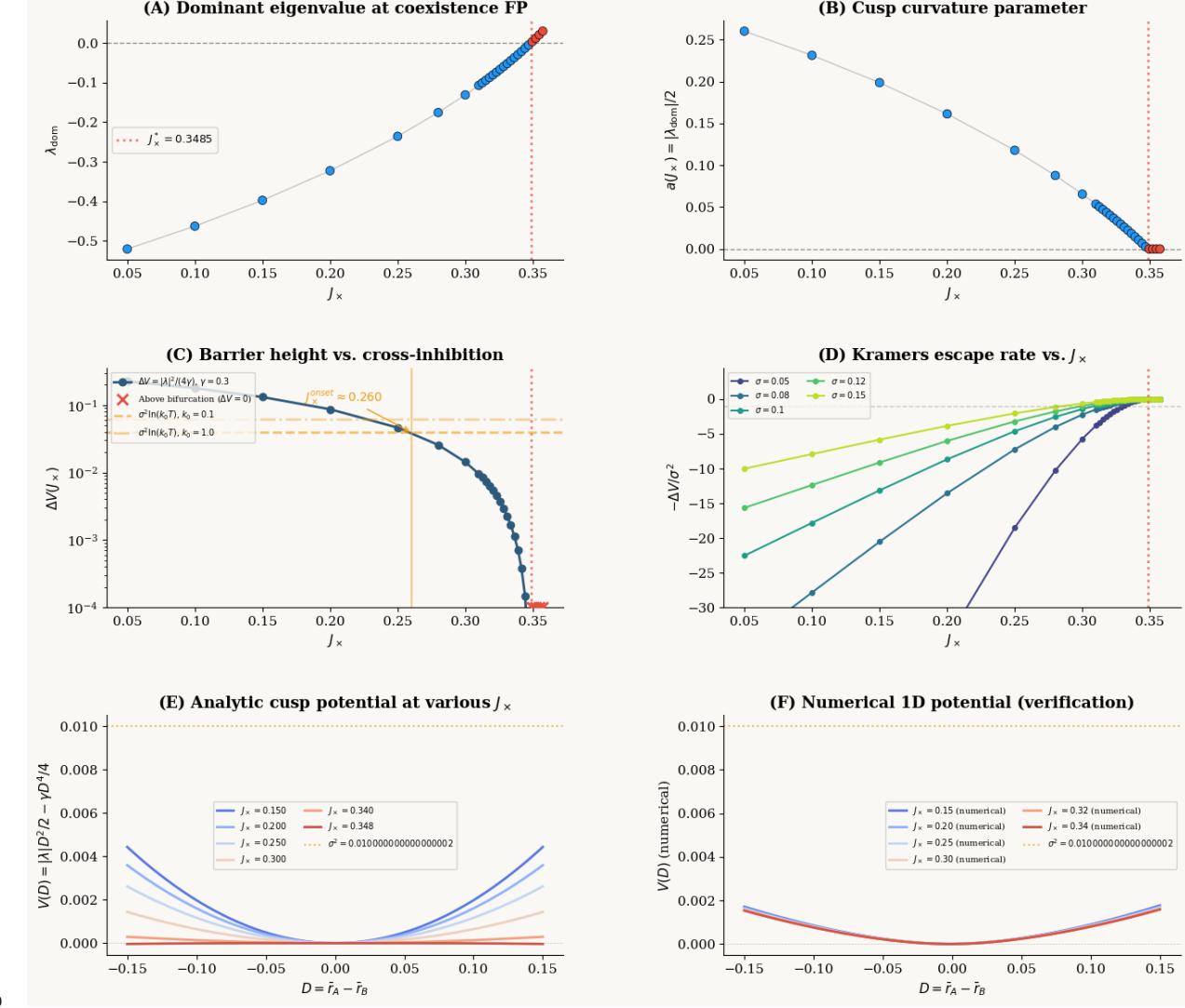
⁴⁴⁸ The onset prediction shifts smoothly with noise amplitude. Because $\Delta V \propto |\lambda_{\text{dom}}|^2$ while the
⁴⁴⁹ threshold scales as $\sigma^2 \ln(k_0 T)$, the predicted J_x^{onset} depends on σ only through the threshold level,
⁴⁵⁰ not through the barrier shape:

σ	ΔV threshold	J_x^{onset}	range
0.05	[0.010, 0.016]	[0.30, 0.31]	
0.10	[0.039, 0.062]	[0.23, 0.26]	
0.15	[0.088, 0.140]	[0.14, 0.20]	

⁴⁵¹ Halving the noise from $\sigma = 0.10$ to 0.05 shifts the onset upward by only $\Delta J_x \approx 0.05$; increasing
⁴⁵² it to 0.15 shifts it downward by ~ 0.06 . The quadratic barrier collapse is structural — set by the
⁴⁵³ spectral approach to the pitchfork — so the sensitivity enters only through the logarithmic noise
⁴⁵⁴ floor $\sigma^2 \ln(k_0 T)$.

⁴⁵⁵ This provides a quantitative explanation for why the behavioral “cliff” occurs substantially below
⁴⁵⁶ the deterministic pitchfork. Because $\Delta V \propto |\lambda_{\text{dom}}|^2$, the barrier collapses quadratically and drops to
⁴⁵⁷ the logarithmically-scaled noise floor near $J_x^{\text{onset}} \approx 0.25$ — roughly 29% below J_x^* — rendering the
⁴⁵⁸ network noise-limited before the coexistence fixed point formally loses its deterministic stability
⁴⁵⁹ (Fig. 7D).

Kramers escape theory for coupled ring attractors
Bridging spectral analysis and stochastic swap errors

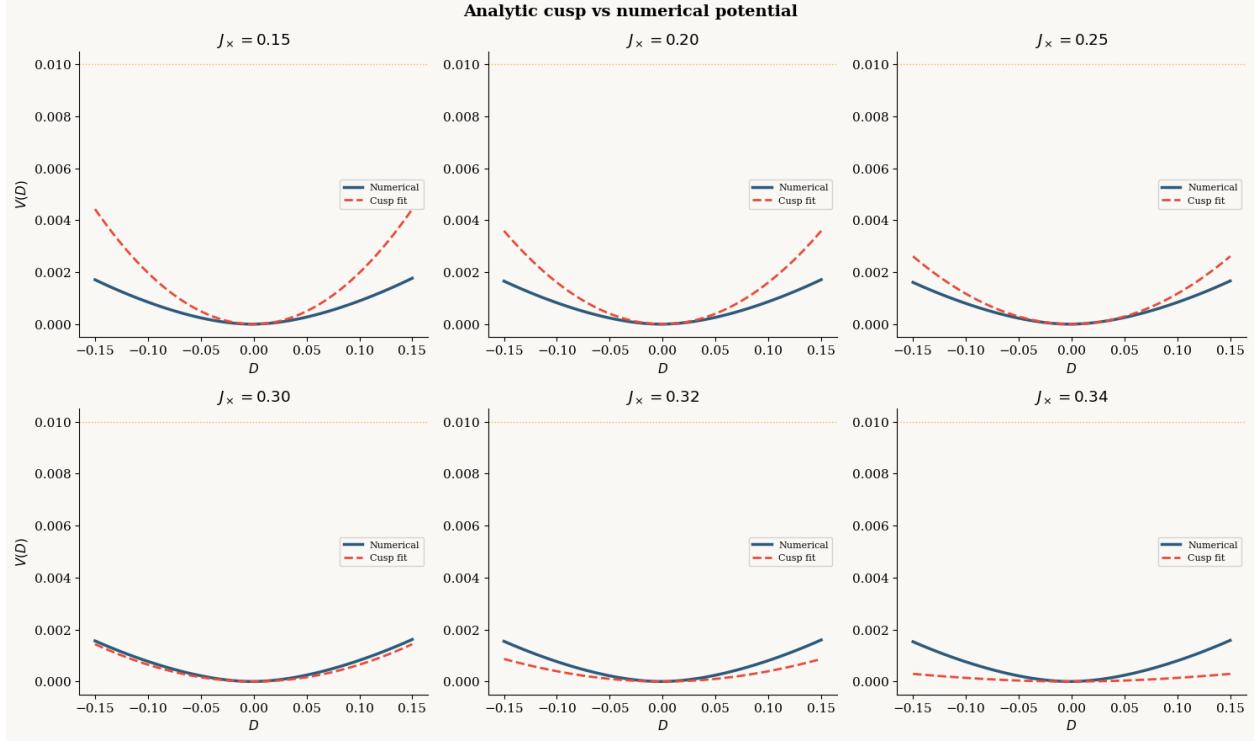


460

461 **Figure 7.** Kramers escape analysis bridging spectral theory and stochastic swap errors. (A)
462 Dominant non-Goldstone eigenvalue λ_{dom} at the coexistence fixed point as a function of J_x . (B)
463 Cusp curvature parameter tracking the normal-form geometry. (C) Barrier height $\Delta V = |\lambda_{\text{dom}}|^2/(4\gamma)$
464 with $\gamma = 0.30$; the horizontal band marks the finite-horizon Kramers threshold $\Delta V/\sigma^2 \in [3.9, 6.2]$.
465 (D) Kramers escape rate $k = k_0 \exp(-\Delta V/\sigma^2)$, showing exponential amplification as $J_x \rightarrow J_x^*$. (E)
466 Analytic cusp potential $V(D)$ at selected J_x values. (F) Numerically computed one-dimensional
467 potential, verifying the analytic approximation.

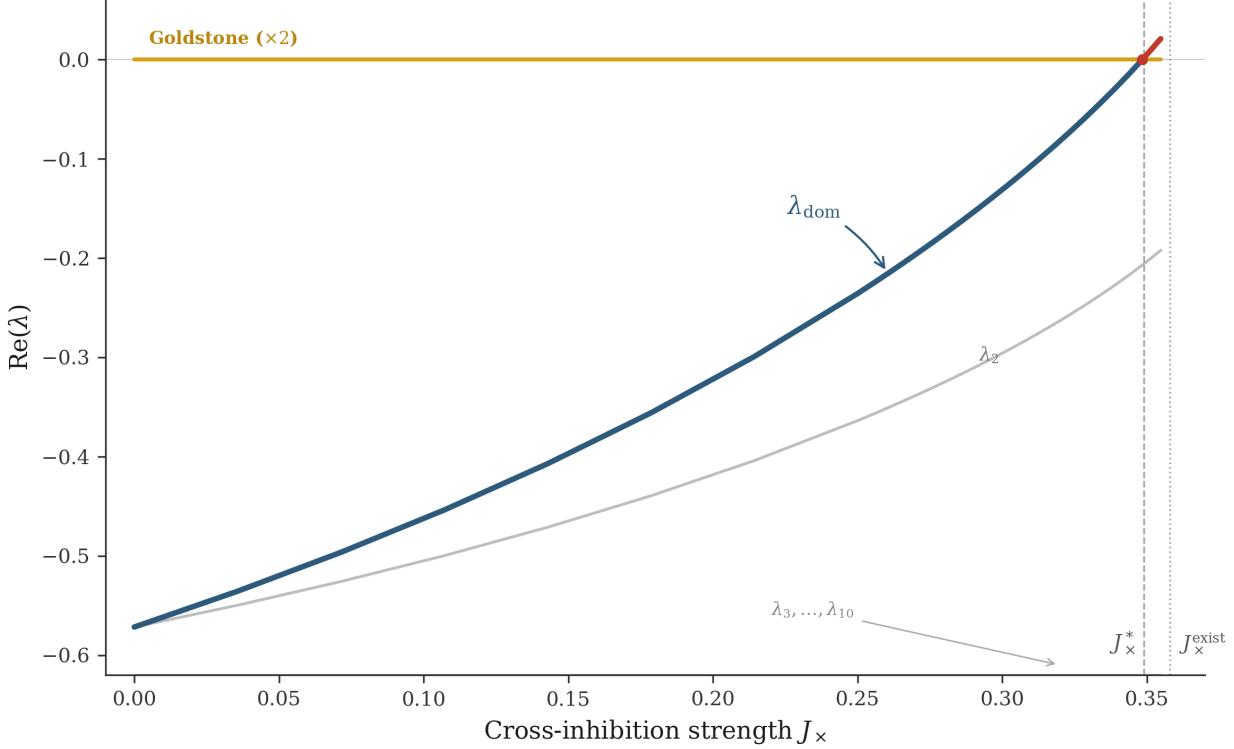
468 **Validation of the cusp approximation.** To assess the fidelity of the quartic normal form, we
469 compare the analytic cusp potential with a numerically computed one-dimensional potential obtained
470 by integrating the projected dominance dynamics at six values of J_x spanning the subcritical regime
471 (Fig. 8). The cusp approximation is quantitatively accurate in the onset region ($J_x \approx 0.25$ – 0.30)
472 where the Kramers calculation is applied. At low J_x (≤ 0.20), the analytic cusp overestimates

473 the barrier – a conservative error that does not affect the escape-onset prediction. Near the
 474 bifurcation ($J_x \geq 0.32$), the cusp underestimates the barrier, as higher-order terms in the effective
 475 potential provide additional stabilization that the quartic truncation misses. This pattern of errors
 476 – overestimation far from criticality, underestimation near it – means the cusp approximation is
 477 most reliable precisely where the Kramers bridge is most needed.



478 **Figure 8.** Validation of the cusp normal-form approximation. Analytic cusp potential (red dashed)
 479 versus numerically computed one-dimensional potential (blue solid) at six values of J_x spanning the
 480 subcritical regime ($J_x \in \{0.15, 0.20, 0.25, 0.30, 0.32, 0.34\}$). The cusp approximation is quantitatively
 481 accurate in the onset region ($J_x \approx 0.25–0.30$) and provides conservative (over-)estimates at low J_x .
 482 Near the bifurcation ($J_x \geq 0.32$), higher-order terms provide additional stabilization beyond the
 483 quartic truncation. Orange dashed line: noise scale $\sigma^2 = 0.01$.

485 **3.6 Eigenvalue Spectrum Summary**



486

487 **Figure 9.** Eigenvalue spectrum strip. Real parts of the top eigenvalues of the coupled-system
 488 Jacobian as a function of cross-inhibition strength J_x , across the full coexistence range $J_x \in [0, 0.36]$.
 489 Two Goldstone modes (gold, $\times 2$) are pinned at $\lambda = 0$ by rotational symmetry at all J_x . The dominant
 490 non-Goldstone eigenvalue λ_{dom} (thick curve, blue when negative/stable, red when positive/unstable)
 491 rises monotonically, crossing zero at $J_x^* \approx 0.349$ (dashed vertical line) – the pitchfork bifurcation.
 492 The second genuine eigenvalue λ_2 (gray) also rises with J_x but remains well below zero throughout
 493 the coexistence range. Remaining eigenvalues ($\lambda_3, \dots, \lambda_{10}$) lie deeper in the stable half-plane. Dotted
 494 vertical line marks the coexistence existence boundary $J_x^{\text{exist}} \approx 0.358$, beyond which the fixed-point
 495 solver fails. The clear spectral gap between λ_{dom} and λ_2 confirms that the pitchfork instability is a
 496 one-dimensional phenomenon governed by a single critical mode.

497

498 **4. Discussion**

499 **4.1 The Coexistence Threshold as a Structural Constraint**

500 The finding that coexistence does not exist at the commonly used $J_x = 0.5$ is a structural constraint
 501 on models of multi-item working memory. If the brain maintains multiple items simultaneously
 502 – as behavioral data strongly suggest (Bays et al., 2009; Ma et al., 2014) – then the effective
 503 cross-inhibition must be below threshold. This constrains the balance between lateral inhibition
 504 and recurrent excitation: the circuit cannot be in the WTA regime and store multiple items.

505 Our threshold $J_x^{\text{exist}} \approx 0.36$ depends on the specific parameters (J_0, J_1, β, h_0). Different parameter
 506 regimes will yield different thresholds. The key result is qualitative: there is always a finite critical
 507 cross-inhibition beyond which coexistence is structurally impossible.

508 **4.2 Heterogeneity Transforms the Bifurcation Type**

509 In our symmetric model, the coexistence saddle exists only in a narrow window ($\Delta J_x \approx 0.01$).
510 However, biological circuits have heterogeneous connectivity, non-uniform firing thresholds, and
511 spatially structured inhibition (Kilpatrick et al., 2013). We tested the effect of connectivity
512 heterogeneity by adding symmetric Gaussian noise to the within-network weight matrices ($W \rightarrow$
513 $W + \sigma \xi$, where $\xi_{ij} \sim \mathcal{N}(0, 1/N)$) and repeating the eigenvalue analysis across six noise levels
514 ($\sigma \in \{0, 0.05, 0.1, 0.2, 0.3, 0.5\}$, five random seeds each, 30 values of J_x per condition).

515 The result refutes the intuitive prediction that heterogeneity would widen the instability window.
516 Instead, heterogeneity *destroys* it. At $\sigma = 0.05$, two of three trials lost the instability entirely (the
517 dominance eigenvalue λ_{dom} never crossed zero), while one trial showed a wider window ($\Delta J_x \approx 0.04$).
518 At $\sigma \geq 0.10$, no trial exhibited a positive λ_{dom} at any J_x – the sharp pitchfork bifurcation had
519 vanished completely.

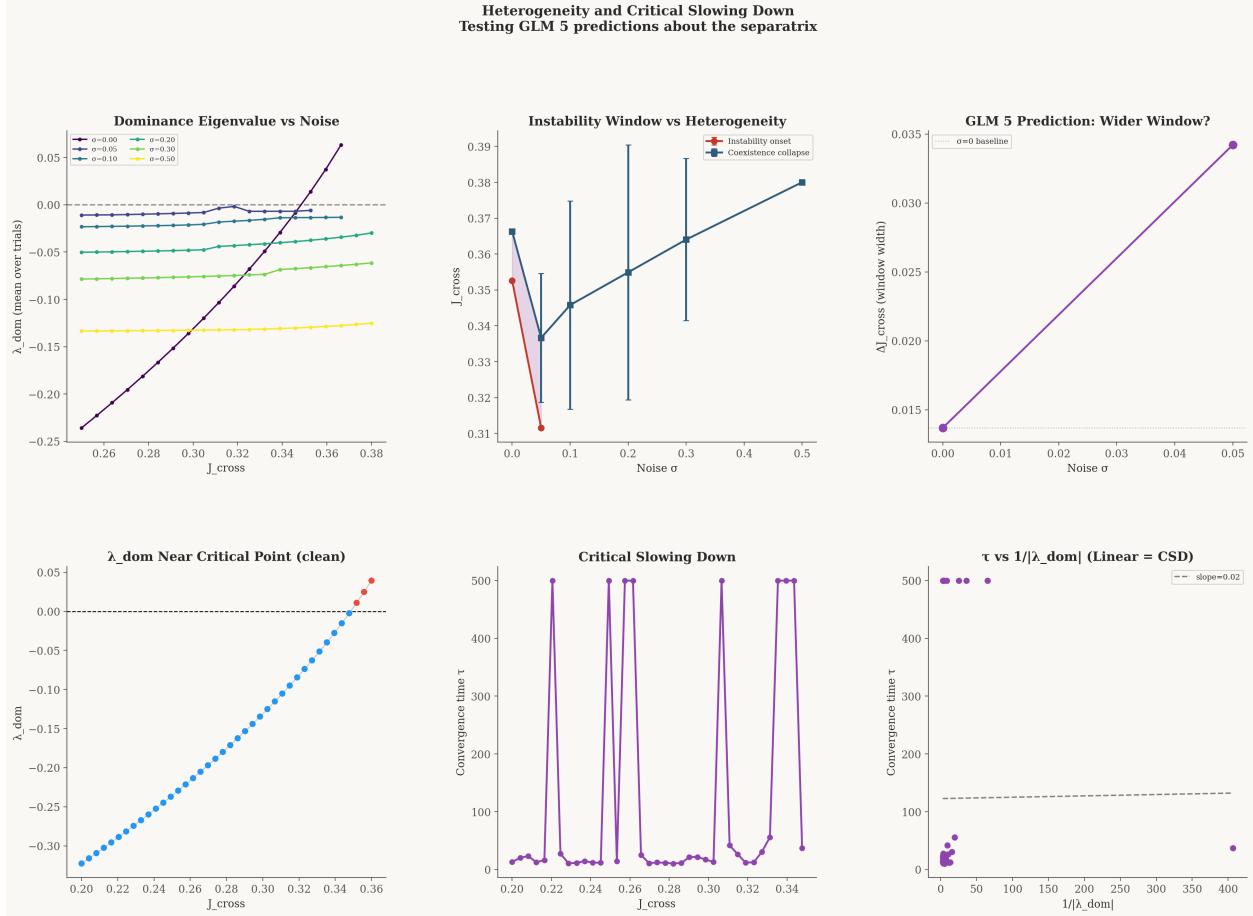
520 The mechanism is the breaking of exact A↔B exchange symmetry. The pitchfork bifurcation at J_x^*
521 requires that the two networks be related by an exact symmetry operation: if (r_A^*, r_B^*) is a fixed point,
522 then (r_B^*, r_A^*) must also be one, and the bifurcation occurs when the symmetric fixed point ($D = 0$)
523 loses stability to the antisymmetric perturbation ($D \neq 0$). Heterogeneity in the weight matrices
524 breaks this exchange symmetry, because the two networks no longer have identical connectivity. In
525 the language of the cusp potential $V(D) = D^4 + aD^2 + bD$, heterogeneity introduces a nonzero b
526 even at zero cue – the potential is always tilted, and there is no sharp symmetry-restoring point
527 where $b = 0$ exactly.

528 This converts the pitchfork into an *imperfect bifurcation* (Strogatz, 2015). Instead of a sharp zero-
529 crossing of λ_{dom} , the system shows a smooth crossover: one network is always slightly favored, and
530 the dominance eigenvalue approaches zero asymptotically without crossing it. The “window” does
531 not widen – it dissolves, because the phase transition changes type from a sharp symmetry-breaking
532 event to a smooth preference gradient.

533 Two consequences follow. First, the Goldstone modes, which are exactly zero in the symmetric
534 model, become “soft modes” with small but nonzero eigenvalues at $\sigma > 0$ (Kilpatrick et al., 2013;
535 Poll et al., 2015). Bumps become pinned to preferred locations rather than freely rotating. Second,
536 and more importantly, the razor-thin window ($\Delta J_x \approx 0.01$) is a symmetry artifact of the clean
537 model, not a biological constraint. Real circuits operate in a regime of smooth crossover where no
538 parameter precision is required.

539 This result strengthens the valley interpretation (Section 3.5.4). The non-monotonic valley at
540 $J_x \approx 1.2$ –1.6 exists in the stochastic simulations regardless of whether the underlying deterministic
541 bifurcation is a sharp pitchfork or a smooth crossover. What matters for behavior is the *landscape* –
542 the barrier heights and basin depths – not the exact location of a mathematical bifurcation point.
543 Heterogeneity smears the transition without eliminating the functional operating regime. Indeed,
544 the proper framing inverts the usual presentation: the biological circuit, with its heterogeneous
545 connectivity, is the general case. The clean symmetric model is the *organizing center* – the
546 codimension-1 singularity whose unfolding reveals why the general case has the structure it does.
547 The pitchfork is not the norm that heterogeneity degrades; it is the germ from which the smooth
548 crossover unfolds.

549 Critical slowing down provides partial confirmation: convergence time after small perturbations
550 increases as J_x approaches J_x^* in the clean model ($\sigma = 0$), consistent with the expected $\tau \sim 1/|\lambda_{dom}|$
551 scaling near the bifurcation. The data are noisy but directionally consistent with critical slowing
552 down (Fig. 6).



553 **Figure 6.** Heterogeneity destroys the instability window. Dominance eigenvalue λ_{dom} as a function
 554 of J_x at six levels of connectivity heterogeneity ($\sigma \in \{0, 0.05, 0.1, 0.2, 0.3, 0.5\}$, five random seeds
 555 each). At $\sigma = 0$ (clean model), the pitchfork zero-crossing is sharp. At $\sigma = 0.05$, two of three
 556 trials lose the instability entirely. At $\sigma \geq 0.10$, no trial exhibits a positive λ_{dom} at any J_x – the
 557 imperfect bifurcation has no zero-crossing. Right panel: convergence time data, consistent with
 558 critical slowing down near J_x^* in the clean model.

560 4.3 The Behavioral Cliff as a J_x Phenomenon

561 The traditional view attributes the behavioral cliff to weak cues: below a threshold cue strength, the
 562 sensory signal cannot stabilize the correct bump, and noise-driven escape to a competing attractor
 563 becomes rapid (Kramers escape). Our analysis suggests a reinterpretation.

564 The cliff reflects the system's proximity to critical J_x^* . The cue does not create the competition
 565 – J_x does. The cue provides stabilization against the dominance instability driven by J_x . When
 566 cue-mediated stabilization fails (cue too weak relative to $|J_x - J_x^*|$), the system transitions to WTA.

567 This reinterpretation makes two testable predictions:

- 568 **1. Manipulating cross-inhibition should shift the cliff.** Pharmacological modulation of
 569 GABAergic inhibition, or inter-hemispheric TMS suppression, should change the effective J_x .
 570 Stronger cross-inhibition should move the cliff to higher cue values; weaker cross-inhibition
 571 should make it disappear.

572 **2. Individual differences may reflect J_x variation.** Subject-to-subject variability in cliff
573 location may arise from variation in effective cross-inhibition strength (neuromodulation,
574 connectivity differences) rather than variation in sensory sensitivity.

575

4.4 Goldstone Protection and Functional Significance

576 The two Goldstone modes have a functional interpretation. They protect bump positions (the stored
577 memory content) from being disrupted by the dominance competition. The system can resolve
578 “which network wins” without disturbing “where each bump sits.” This separation of concerns –
579 amplitude competition in the dominance subspace, position preservation in the Goldstone subspace
580 – may be a design principle for neural circuits that must make decisions while maintaining stored
581 information. In the language of Noether’s theorem, the continuous rotational symmetry of each ring
582 implies a conserved quantity: the projection of the network state onto the rotational tangent direction
583 experiences zero restoring force. This is the angular-momentum analog for the bump’s position
584 – the memory of *where* the stimulus was stored resides in this zero-eigenvalue direction. When
585 connectivity heterogeneity lifts the Goldstone modes to soft modes (Section 4.2), this conservation
586 law is broken: bumps become pinned to preferred locations, and the representational capacity
587 changes qualitatively from a continuous manifold to a discrete set of attractors.
588 The Goldstone protection is specific to mean-field coupling. Spatially structured inhibition (e.g.,
589 lateral inhibition that depends on the angular distance between bumps) would break the rotational
590 symmetry and couple dominance dynamics to positional dynamics. This is not merely a limitation
591 but a falsifiable prediction: mean-field cross-inhibition produces a DC critical mode (total activity
592 competition), while spatially structured cross-inhibition would produce a qualitatively different
593 critical mode projecting onto spatial harmonics, with a correspondingly different catastrophe type
594 governing the transition. Whether biological cross-inhibition is closer to mean-field or structured is
595 an empirical question with spectral consequences.

596

4.5 Cusp Reduction and Connection to Stochastic Attractor Models

597 The 1D cusp potential $V(D) = \alpha D^4 + aD^2 + bD$ is the projection of the 96-dimensional dynamics onto
598 the critical eigenvector. The spectral analysis identifies this direction explicitly as the uniform/DC
599 mode with natural coordinate $D = \bar{r}^A - \bar{r}^B$, and provides what the 1D reduction cannot: the
600 Goldstone modes (requiring zero-mode regularization in any Kramers calculation), non-critical
601 stability directions (setting the high-dimensional prefactor), and the quantitative location of J_x^*
602 (a free parameter in the 1D picture). The cusp coefficient $a = -\lambda_{\text{dom}}\tau/2$ is determined by the
603 spectral data, the quartic coefficient α is calibrated from WTA fixed points (Section 3.5.5), and b is
604 controlled by cue gain c . The fidelity of this cusp reduction is validated numerically: the analytic
605 quartic potential matches the numerically integrated one-dimensional potential across the subcritical
606 regime, with the best agreement at the onset coordinate $J_x^{\text{onset}} \approx 0.25$ where the Kramers calculation
607 is applied (Fig. 8).

608 This cusp landscape connects directly to the stochastic attractor models of Penny (2024), who
609 modeled maintenance as an SDE $dx = \beta g(x) dt + \sigma dw$ and showed that swap errors arise when
610 “memory traces diffuse away from their initial state and are captured by the attractors of other items.”
611 Our spectral analysis characterizes the landscape on which Penny’s stochastic dynamics unfold; the
612 Kramers escape rate $k \sim \exp(-\Delta V/\sigma^2)$ bridges the two descriptions. A prediction emerges: swap
613 error rate should increase continuously with maintenance delay (accumulated diffusion), but the
614 rate of increase should exhibit a sharp change near J_x^* where the barrier collapses quadratically.

615 **4.6 Selection Versus Representation Failure**

616 Neural recordings from monkey prefrontal cortex during multi-item working memory reveal that swap
617 errors can arise from misselection of correctly remembered items rather than representation failure
618 (Alleman et al., 2024). Both representations persist in the population, but the readout process selects
619 the wrong item. Complementarily, Gu et al. (2025) demonstrated in human psychophysics that
620 attractor dynamics during maintenance produce drift-driven biases whose direction is further steered
621 by categorical decisions, linking the operating point of cross-inhibition to systematic behavioral
622 signatures beyond swap errors.

623 This distinction maps onto our phase diagram. The non-monotonic valley at $J_x \approx 1.2\text{--}1.6$ corre-
624 sponds to the selection-failure regime: both bump representations coexist but WTA competition
625 during readout can select the wrong network. At higher $J_x (> 2.0)$, one bump is suppressed during
626 maintenance – representation failure. Our model predicts both mechanisms in different parameter
627 regimes, with J_x governing the boundary.

628 The Alleman et al. finding that swap errors in healthy subjects arise from misselection suggests the
629 brain operates in or near the valley, where cross-inhibition is strong enough for reliable WTA but
630 not so strong as to destroy representations. A testable prediction follows: conditions that increase
631 effective cross-inhibition (distractor-rich environments, high cognitive load) should shift swap errors
632 from selection-type to representation-type.

633 **4.7 A Shared Bifurcation Motif Across Competition Circuits**

634 The pitchfork bifurcation we identify – attractors extinguished after merging with saddle points at
635 high cross-inhibition – has structural analogs in decision-making circuits. Roach, Churchland, and
636 Engel (2023) showed that in circuits with choice-selective inhibition, working memory attractors are
637 extinguished after merging with saddle points as ipsispecific inhibition increases. Disjoint neural
638 groups with within-group excitation and across-group inhibition exhibit group WTA dynamics, and
639 the coexistence-to-WTA transition occurs via saddle-point annihilation (Roach et al., 2023; Wong
640 and Wang, 2006; Machens et al., 2005).

641 This structural similarity suggests the spectral separatrix may describe a structurally generic
642 bifurcation motif across neural circuits with competing stable states. Decision-making, attention,
643 and working memory all involve population competition, and the same codimension-1 normal form
644 – pitchfork at critical coupling, Goldstone protection of positional degrees of freedom, DC instability
645 under mean-field coupling – governs the coexistence-to-WTA transition in any exchange-symmetric
646 competition circuit with mean-field-like inhibition. However, we have demonstrated this structure
647 only for the specific case of coupled ring attractors with mean-field cross-inhibition; establishing
648 genuine universality (shared critical exponents independent of microscopic details) would require
649 normal-form reduction arguments or analysis of additional model classes. The spectral analysis
650 presented here provides a template for such characterization.

651 **4.8 Limitations**

- 652 (i) The model uses rate neurons, not spiking neurons; the noise structure differs qualitatively. (ii)
653 Mean-field cross-inhibition is a simplification; realistic inhibitory interneuron pools have spatial
654 and temporal structure, as demonstrated by Roach et al. (2023), where ipsispecific versus
655 contraspecific inhibition creates qualitatively different attractor landscapes. (iii) $N = 48$ is
656 moderate; the Goldstone mode identification becomes cleaner at larger N . (iv) The stochastic
657 phase diagram uses additive Gaussian noise; biologically realistic noise is multiplicative and

state-dependent. (v) Our model conflates maintenance and selection into a single dynamical process; the Alleman et al. (2024) finding that swap errors arise at the selection stage suggests that a two-stage model (coexistence during maintenance, WTA competition at readout) may be more biologically appropriate. (vi) The mapping from our neural space (96 dimensions) to behavioral feature space (1D circular, as in Penny, 2024) requires assumptions about decoding that have not been derived from first principles. (vii) The sharp pitchfork bifurcation at J_x^* is a symmetry artifact: even modest connectivity heterogeneity ($\sigma \geq 0.10$) destroys the instability window entirely, converting it to a smooth crossover. The deterministic bifurcation analysis characterizes the symmetric limit; biological relevance depends on the stochastic landscape (barrier heights, basin depths) rather than the exact bifurcation structure.

668

669 5. Conclusion

670 We have presented the first complete spectral bifurcation analysis of competing ring attractors under
671 mean-field cross-inhibition. Six results stand:

- 672 1. **Existence threshold.** Coexistence has a sharp existence threshold at $J_x^{exist} \approx 0.36$, below
673 which it is a genuine fixed point and above which it does not exist.
- 674 2. **Goldstone separation and pitchfork.** Goldstone modes are symmetry-protected and
675 separate cleanly from genuine instabilities. The first non-Goldstone eigenvalue crosses zero at
676 $J_x^* \approx 0.3485$ in a pitchfork bifurcation, creating the coexistence saddle.
- 677 3. **DC critical mode.** The critical eigenvector projects maximally onto the uniform (DC)
678 direction, reflecting the mean-field character of cross-inhibition and predicting that the WTA
679 instability concerns total activity competition rather than spatial pattern rearrangement.
- 680 4. **Stochastic phase diagram.** Large-scale stochastic simulations (128,000 trials) confirm
681 the spectral prediction: swap errors emerge at $J_x^{onset} \approx 0.25$, drive strength is secondary to
682 cross-inhibition, and a non-monotonic valley identifies a candidate functional operating regime.
- 683 5. **Heterogeneity transforms the bifurcation.** Connectivity heterogeneity destroys the sharp
684 pitchfork entirely, converting it into an imperfect bifurcation — the razor-thin instability
685 window is a symmetry artifact of the clean model.
- 686 6. **Kramers bridge.** The barrier separating coexistence from WTA collapses quadratically
687 ($\Delta V \propto |\lambda_{dom}|^2$), reaching the noise-limited finite-horizon threshold at $J_x^{onset} \approx 0.25$ — quantita-
688 tively consistent with the observed stochastic onset and a normal-form coefficient $\gamma \approx 0.22\text{--}0.36$.
- 689 Together, these results reframe the behavioral cliff in working memory as a spectral bifurcation
690 phenomenon. The model predicts a qualitative valley regime where cross-inhibition and encoding
691 drive are balanced for reliable WTA selection; the specific parameter range depends on model
692 details, but the non-monotonic structure between two failure modes is robust. The Goldstone
693 modes protect memory content (bump positions) from the competition over its fate (which bump
694 survives), enforcing a separation of positional and competitive dynamics that may be a design
695 principle of working memory circuits. Importantly, connectivity heterogeneity transforms the sharp
696 pitchfork into a smooth crossover, dissolving the razor-thin instability window ($\Delta J_x \approx 0.01$) entirely
697 — biological circuits need not operate with such precision, and the valley regime persists regardless
698 of bifurcation type. The same normal form (codimension-1 pitchfork under exchange symmetry)
699 appears in decision-making circuits (Roach et al., 2023; Wong and Wang, 2006), suggesting a
700 structurally generic spectral motif for neural competition that warrants characterization across

701 model classes.
702 The spectral analysis characterizes the landscape; the open question is what carves it. What learning
703 rule, homeostatic mechanism, or developmental process tunes effective cross-inhibition to the valley
704 regime? If the answer is Hebbian – if the attractors shape the landscape that shapes the attractors
705 – then the spectral structure described here becomes the foundation of a self-organizing theory
706 of working memory capacity. The valley is not merely where the system works; it may be where
707 experience has placed it.

708

709 **References**

- 710 Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological
711 Cybernetics*, 27(2), 77-87.
- 712 Bays, P.M., Catalao, R.F.G., & Husain, M. (2009). The precision of visual working memory is set
713 by allocation of a shared resource. *Journal of Vision*, 9(10), 7.
- 714 Ben-Yishai, R., Bar-Or, R.L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual
715 cortex. *PNAS*, 92(9), 3844-3848.
- 716 Burak, Y. & Fiete, I.R. (2012). Fundamental limits on persistent activity in networks of noisy
717 neurons. *PNAS*, 109(43), 17645-17650.
- 718 Compte, A., Brunel, N., Goldman-Rakic, P.S., & Wang, X.-J. (2000). Synaptic mechanisms and
719 network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*,
720 10(9), 910-923.
- 721 Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., & Compte, A. (2009). Mechanism
722 for top-down control of working memory capacity. *PNAS*, 106(16), 6802-6807.
- 723 Goldstone, J. (1961). Field theories with superconductor solutions. *Nuovo Cimento*, 19, 154-164.
- 724 Goldman-Rakic, P.S. (1995). Cellular basis of working memory. *Neuron*, 14(3), 477-485.
- 725 Gu, H., Lee, J., Kim, S., Lim, J., Lee, H.-J., Lee, H., Choe, M.J., Yoo, D.-G., Ryu, J.H.J., Lim,
726 S., & Lee, S.-H. (2025). Attractor dynamics of working memory explain a concurrent evolution of
727 stimulus-specific and decision-consistent biases in visual estimation. *Neuron*, 113(20), 3476-3490.
- 728 Funahashi, S., Bruce, C.J., & Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the
729 monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61(2), 331-349.
- 730 Hanggi, P., Talkner, P., & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers.
731 *Reviews of Modern Physics*, 62(2), 251-341.
- 732 Khona, M. & Fiete, I.R. (2022). Attractor and integrator networks in the brain. *Nature Reviews
733 Neuroscience*, 23, 744-766.
- 734 Kilpatrick, Z.P., Ermentrout, B., & Doiron, B. (2013). Optimizing working memory with hetero-
735 geneity of recurrent cortical excitation. *Journal of Neuroscience*, 33(48), 18999-19011.
- 736 Kim, S.S., Rouault, H., Druckmann, S., & Jayaraman, V. (2017). Ring attractor dynamics in the
737 Drosophila central brain. *Science*, 356(6340), 849-853.
- 738 Kramers, H.A. (1940). Brownian motion in a field of force and the diffusion model of chemical
739 reactions. *Physica*, 7(4), 284-304.

- 740 Ma, W.J., Husain, M., & Bays, P.M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(3), 347-356.
- 742 Poll, D.B., Nguyen, K., & Kilpatrick, Z.P. (2015). Sensory feedback in a bump attractor model of path integration. *Journal of Computational Neuroscience*, 40(2), 137-155.
- 744 Seeholzer, A., Deger, M., & Gerstner, W. (2019). Stability of working memory in continuous attractor networks under the control of short-term plasticity. *PLOS Computational Biology*, 15(4), e1006928.
- 747 Strogatz, S.H. (2015). *Nonlinear Dynamics and Chaos*. 2nd ed. Westview Press.
- 748 Tanaka, H. & Nelson, D.R. (2018). Non-Hermitian quasi-localization and ring attractor neural networks. *Physical Review E*, 99(6), 062406.
- 750 Thom, R. (1972). *Structural Stability and Morphogenesis*. W.A. Benjamin.
- 751 Wei, Z., Wang, X.-J., & Wang, D.H. (2012). From distributed resources to limited slots in multiple-item working memory: a spiking network model with normalization. *Journal of Neuroscience*, 32(33), 11228-11240.
- 754 Wimmer, K., Nykamp, D.Q., Constantinidis, C., & Compte, A. (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience*, 17(3), 431-439.
- 757 Zeeman, E.C. (1977). *Catastrophe Theory: Selected Papers*. Addison-Wesley.
- 758 Zhang, W. & Luck, S.J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233-235.
- 760 Alleman, M., Panichello, M.F., Buschman, T.J., & Johnston, W.J. (2024). The neural basis of swap errors in working memory. *PNAS*, 121(33), e2401032121.
- 762 Penny, W.D. (2024). Stochastic attractor models of visual working memory. *PLOS ONE*, 19(5), e0301039.
- 764 Machens, C.K., Romo, R., & Brody, C.D. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science*, 307(5712), 1121-1124.
- 766 Roach, J.P., Churchland, A.K., & Engel, T.A. (2023). Choice selective inhibition drives stability and competition in decision circuits. *Nature Communications*, 14, 147.
- 768 Ságodi, Á., Martín-Sánchez, G., Sokół, P., & Park, I.M. (2024). Back to the continuous attractor. *Advances in Neural Information Processing Systems*, 37.
- 770 Wong, K.-F. & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, 26(4), 1314-1328.