Interpretable Coupling Structure Beyond Deep Learning: Probabilistic and Energy-Based Modelling of Multivariate (Neural) Dynamics

Poznańskie Centrum Superkomputerowo-Sieciowe
Poznan Supercomputing and Networking Center

Julian Kędys [1*] Cezary Mazurek [1]

[1] Department of Digital Medicine, Poznań Supercomputing and Networking Center, Polish Academy of Sciences
* jkedys@man.poznan.pl

ML in PL CONFERENCE 2025
15 - 17 OCTOBER / WARSAW, POLAND

Introduction

- **Why interpretable EBMs for brain dynamics:** Pairwise maximum-entropy (Ising) models recover large-scale brain activity structure from neuroimaging data, balancing parsimony with fidelity and offering explicit couplings and energies for mechanistic interpretation.
- **Energy landscapes as a unifying lens:** Disconnectivity graphs and landscape views expose stable states and barriers—useful across physics, chemistry, and now neural systems.
- **Phase-diagram perspective:** Locating subjects in (μ,σ) coupling space reveals ordered-disordered/spin-glass-like regimes and distance to criticality; this has been demonstrated on human resting-state data and linked to behaviour.
- **Generality beyond neuroscience:** PMEM/Ising-based ELA applies to any discretisable multivariate system (genomics, ecology, social systems, and more) with meaningful interactions of the nodes.
- **Positioning vs. deep nets:** Deep models excel with big data but can be opaque; the Ising/PMEM family gives interpretable parameters, works in small-N / small-T regimes, and is generative (MCMC).

Pairwise Maximum-Entropy (Ising) model (PMEM)

Modelling the energy landscape for binary brain states: $\sigma = (\sigma_1, \dots, \sigma_n) \in \{-1, +1\}^n$, with local fields $h = (h_i) \in \mathbb{R}^n$ (bias activity at node *i*), and symmetric couplings $J = (J_{ij}) \in \mathbb{R}^{n \times n}$ (pairwise interactions; zero diagonal and symmetry: $J_{ii}=0$, $J=J^\top$). The model defines a probability over whole-brain configurations via the Boltzmann law:

$$p(\sigma) = \frac{1}{Z}\exp\left(h^\top \sigma + \frac{1}{2}\sigma^\top J \sigma\right)$$

(1) Distribution (Boltzmann form; β=1) - probability of a brain-wide configuration

$$E(\sigma) = -\,h^\top \sigma - \frac{1}{2}\sigma^\top J \sigma$$

(2) Energy (Hamiltonian; quadratic form over fields/couplings) - scalar score that shapes the landscape by defining minima, transition barriers

$$Z = \sum_{\sigma \in \{-1,+1\}^n} \exp\left(h^\top \sigma + \frac{1}{2}\sigma^\top J \sigma\right)$$

(3) Partition function: normalises p over all 2^n; exact form is exponential in n

Pseudo-likelihood (PL) Ising — scalable surrogate to exact likelihood

- **Idea:** Replace the intractable joint likelihood by the product of node-wise conditionals - each node is a logistic regression on the remaining spins:

$$p(\sigma_i \mid \sigma_{-i}) = \frac{\exp(\sigma_i f_i)}{2\cosh f_i}, \quad f_i = h_i + \sum_{j \neq i} J_{ij}\sigma_j$$

(4) Conditional probability (logistic)

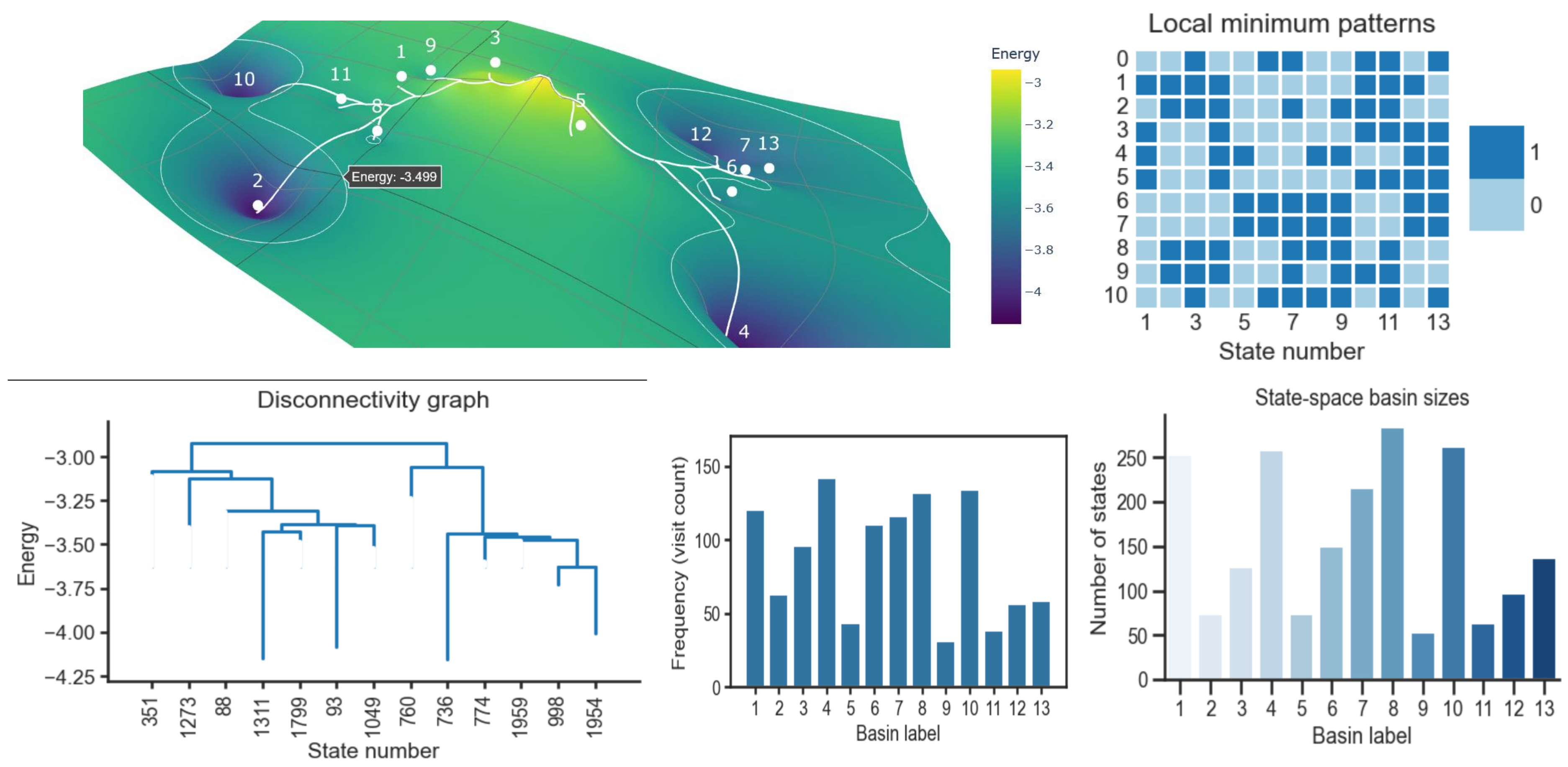
- **Objective (with L2):** Maximise the sum of conditional log-likelihoods, over all time points and nodes, with L2 penalties (ridge) on h and the off-diagonal J; Keep **J symmetric** and **zero diagonal** throughout optimisation.
- **When to use:** fast, stable point estimate (h, J) for moderate/large n; no need to evaluate Z; suitable for downstream energy-landscape analysis and phase-diagram positioning, even when exact likelihood is infeasible.

Variational Bayes (VB) Ising — posterior over parameters with uncertainty

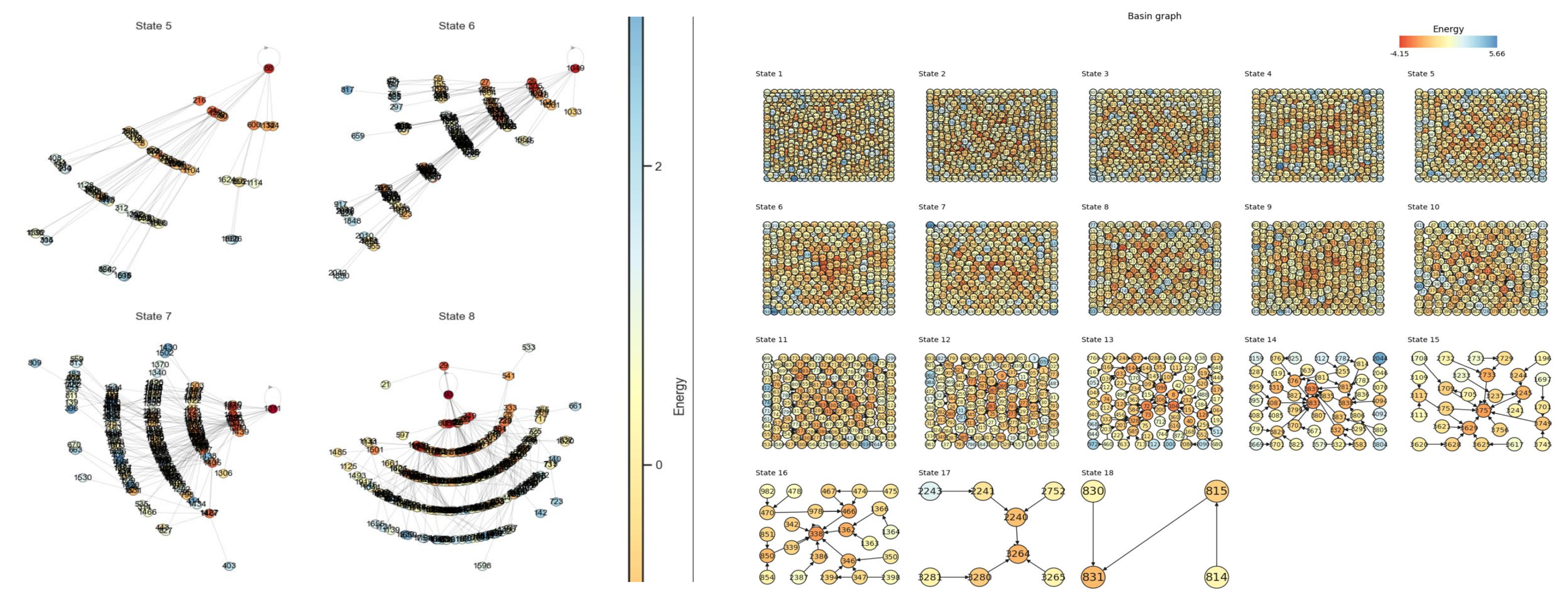
- **Idea:** Place Gaussian priors on h and the off-diagonal couplings J (separate precisions for h and J). Approximate the parameter posterior with a Gaussian q(θ) = N(μ, Σ). (5)
- **Moments (E-step):** Estimate model feature means and covariances via Monte Carlo, applying a small diagonal floor for numerical stability. (6)
- **Parameter update (M-step):** Update posterior precision λ and mean μ using the moment covariance and the data–model moment mismatch; prior precision stays fixed. We monitor the ELBO and stop when both the relative parameter change and the ELBO change are small. (7)
- **When to use:** Small-T or small-N regimes when uncertainty matters - provides credible-interval-ready parameters
- **Constraints & output:** Enforce J symmetry and zero diagonal; return h and J from the posterior mean (plus uncertainties from the posterior covariance/precisions).

- (5) $$\theta = [h; J_{i<j}], \quad \theta \sim \mathcal{N}(0, \Lambda_0^{-1}), \quad q(\theta) = \mathcal{N}(\mu, \Sigma)$$
- (6) $$m_\eta = \mathbb{E}_{p_\eta}[\Phi], \quad C_\eta = \mathrm{Cov}_{p_\eta}(\Phi)$$
- (7) $$\Sigma^{-1} = \Lambda_0 + T C_\eta, \quad \mu \leftarrow \eta + \Sigma T(\bar{\Phi} - m_\eta)$$

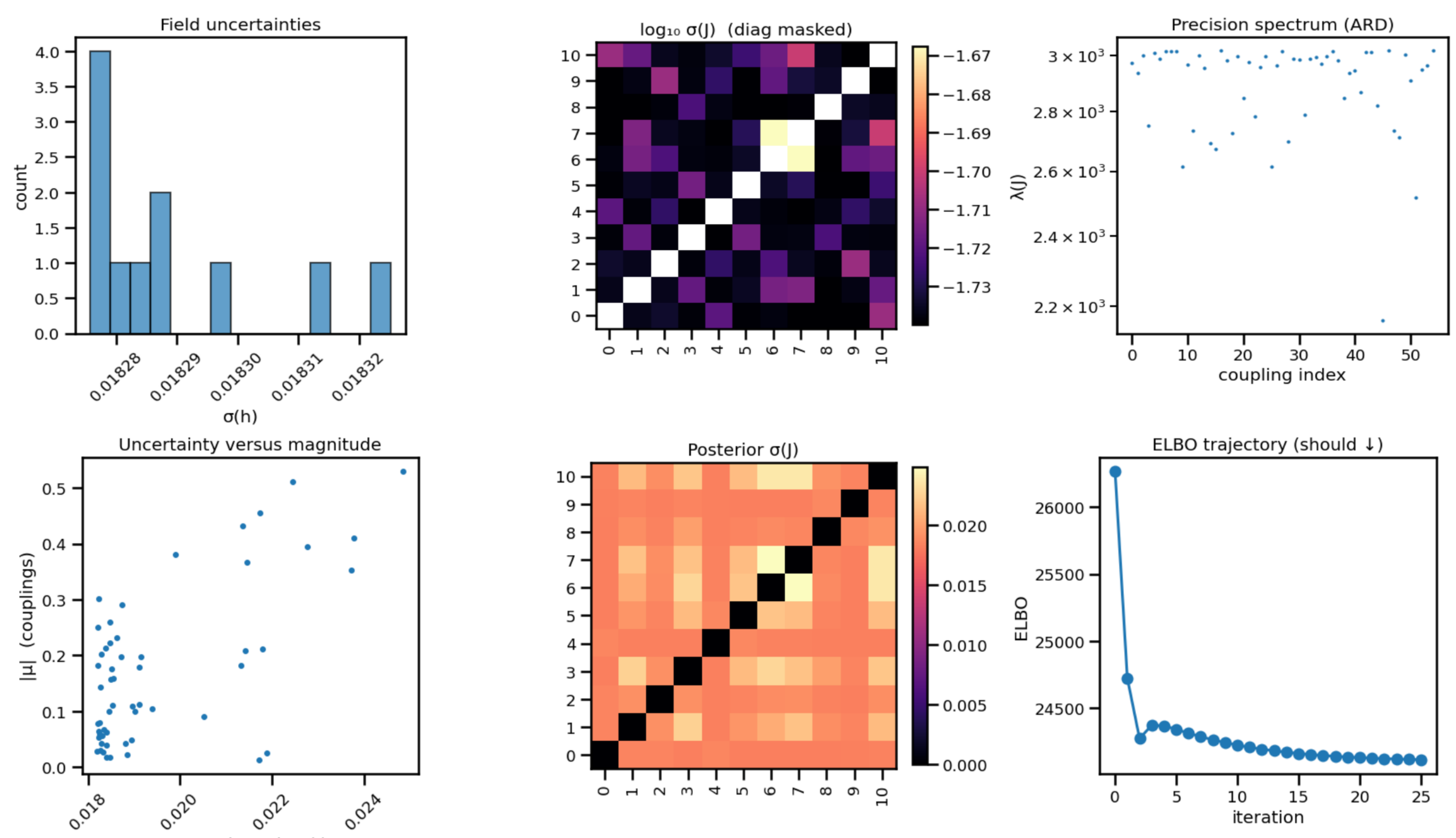
Select descriptors and metrics of the Ising-model pipeline



[Fig. 1] Top-left to bottom-right: 3D representation of individual energy landscape; Node-to-attractor map; Isolated graph branches as attractors; Visit counts per basin (all attractor-bound states; Distribution of basin sizes

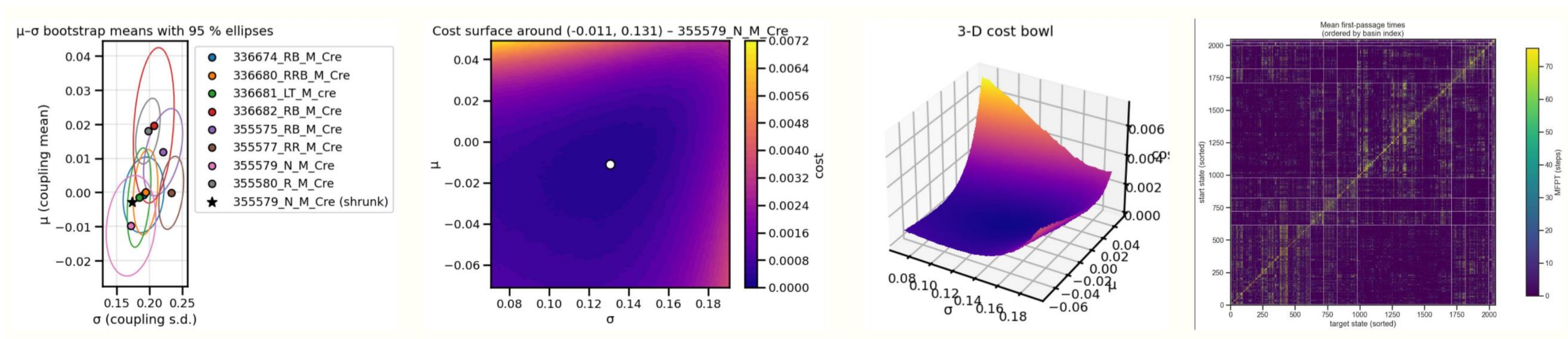


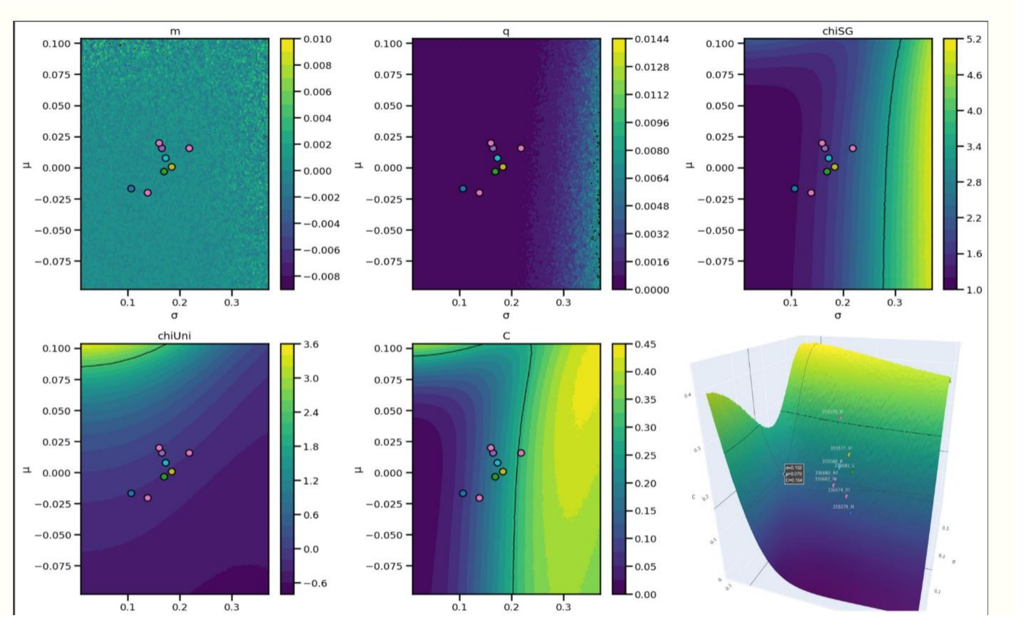
[Fig. 2] Per-state basin graphs derived from the gradient flow on the hypercube (single-spin flips). Nodes are patterns; colour = energy; edges follow descent towards the attractor (labelled node). **Left:** Fast Kamada-Kawai graph + radial layering: attractors centred; concentric rings mark shortest-path distance (number of single-flip steps) to the attractor **Right**: Overlaps-free per-state basin graphs with spring layout



[Fig. 3]Top row (left to right): (1) Field uncertainties - histogram of the posterior standard deviations of the node biases (h); biases are tightly estimated (most bars near zero) (2) Coupling uncertainty (log scale) - heat map of $\sigma(J_{ij})$; brighter squares mean more uncertainty; the diagonal is hidden (self-couplings are fixed to zero (3) Precision per coupling - scatter of precisions λ(J) (≈ inverse variance). Higher dots mean more confidence; a wide spread shows the model shrinks weak edges more than strong ones Bottom row (left to right): (4) Size vs uncertainty - each dot is a coupling; left & high (large magnitude, low σ) = most reliable; right & low = small and uncertain (5) Coupling uncertainty (linear scale) - heat map of σ(J); brighter = less certain; symmetric with zero diagonal (6) ELBO over iterations - falls quickly and then levels off, indicating the VB fit has converged

Select results of the phase diagram analysis





[Fig. 4] Common phase-diagram placement and landscape kinetics
Subjects are positioned on a single (μ , σ) reference via multi-observable fitting: top-left shows bootstrap means with 95% CIs; centre panels show the cost surface (2-D and 3-D "bowl") around an example optimum. Right: MFPT matrix summarises transition times between landscape states. Bottom row: reference surfaces for m, q, χ_SG, χ_UNI, and heat capacity C with subject points overlaid; far-right: 3-D view of the shared phase frame