UNIVERSITÄT ZU KÖLN

Metastability in Spiking Attractor Networks: Scalable Simulation and Cortical Decision Dynamics

Inaugural-Dissertation zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Felix Johannes Schmitt

2026

Lightly, child, lightly. You've got to learn to do everything lightly.
Think lightly, act lightly, feel lightly…

– Aldous Huxley , Island

Abstract

Metastability in spiking attractor networks denotes a regime of neural computation defined by noise-driven transitions among discrete coexisting population states. This thesis studies that regime in spiking attractor networks with locally balanced excitatory-inhibitory clustered connectivity and asks whether it can be simulated at scale, whether metastable computation depends on how clustering is structurally implemented, and whether it can account for cortical decision dynamics under uncertainty.

GPU- and CPU-based simulation and calibration methods achieve real-time performance for networks of up to 100,000 neurons. A unified framework introduces a single mixing parameter that interpolates continuously between clustering by connection probability and clustering by synaptic weight while preserving the mean input of a balanced random network. Mean-field analysis and network simulations show that metastable dynamics persist across this continuum, although correlation structure, co-activation patterns, and bifurcation structure depend on the specific implementation. When tested against motor-cortical recordings from monkeys performing a delayed-reach task under graded target uncertainty, the model reproduces uncertainty-dependent modulation of trial-to-trial variability, constant spiking irregularity, and condition-dependent reaction times. In this way, it links metastable attractor dynamics to neural and behavioral variability during competitive action selection.

These results establish spiking attractor networks with locally balanced excitatory-inhibitory clustered connectivity as a scalable, analytically tractable, and experimentally grounded model class for metastable cortical decision dynamics. Their explanatory reach can therefore be assessed without committing to a unique microscopic wiring motif.

Zusammenfassung

Metastabilität in spikenden Attraktornetzwerken bezeichnet ein Regime neuronaler Berechnung, das durch rauschgetriebene Übergänge zwischen diskreten, koexistierenden Populationszuständen gekennzeichnet ist. Die Arbeit untersucht dieses Regime in spikenden Attraktornetzwerken mit lokal balancierter exzitatorisch-inhibitorischer (E/I) Cluster-Konnektivität und fragt, ob es sich skalierbar simulieren lässt, ob metastabile Berechnung davon abhängt, wie Clustering strukturell realisiert wird, und ob sich damit kortikale Entscheidungsdynamik unter Unsicherheit erklären lässt.

GPU- und CPU-basierte Simulations- und Kalibrierungsmethoden erreichen Echtzeitfähigkeit für Netzwerke mit bis zu 100.000 Neuronen. In einem einheitlichen Rahmen interpoliert ein einzelner Mischungsparameter kontinuierlich zwischen Clustering über Verbindungswahrscheinlichkeit und Clustering über synaptische Gewichte, wobei der mittlere Input eines balancierten Zufallsnetzwerks erhalten bleibt. Mean-Field-Analyse und Netzwerksimulationen zeigen, dass Metastabilität über dieses Kontinuum hinweg bestehen bleibt, während Korrelationen, Koaktivierungsmuster und Bifurkationsstruktur von der konkreten Implementierung abhängen. Im Abgleich mit motorkortikalen Ableitungen von Makaken während einer Delayed-Reach-Aufgabe unter abgestufter Zielunsicherheit reproduziert das Modell die unsicherheitsabhängige Modulation der Trial-to-Trial-Variabilität, konstante Spiking-Irregularität und bedingungsabhängige Reaktionszeiten. Damit verknüpft es metastabile Attraktordynamik mit der neuronalen und der Verhaltensvariabilität bei kompetitiver Handlungsselektion.

Diese Ergebnisse etablieren spikende Attraktornetzwerke mit lokal balancierter E/I-Cluster-Konnektivität als skalierbare, analytisch zugängliche und experimentell fundierte Modellklasse für metastabile kortikale Entscheidungsdynamik. Ihre Erklärungsreichweite lässt sich damit beurteilen, ohne sich auf ein eindeutiges mikroskopisches Konnektivitätsmotiv festzulegen.

Contents

List of Figures

Note: This list includes only those figures that are not part of the results chapter.

Use of AI-based tools

ChatGPT and Claude were used in a limited supporting role during the preparation of this thesis. The preparation of this thesis involved limited use of ChatGPT and Claude for text editing and identifying relevant literature. The selection and interpretation of the literature, the development of the scientific arguments, and the final wording of the manuscript remain the sole responsibility of the author.

List of Abbreviations

ACh Acetylcholine

ALM anterolateral motor cortex

BRN balanced random network

CANN continuous attractor neural network

CV_2 local coefficient of variation of inter-spike intervals

E/I excitatory–inhibitory

FF Fano factor

HMM hidden Markov model

ISI interspike interval

ISN inhibition-stabilized network

LFP local field potential

LIF leaky integrate-and-fire

M1 primary motor cortex

RNN recurrent neural network

SBI simulation-based inference

SFA spike frequency adaptation

SSN stabilized supralinear network

STDP spike-timing-dependent plasticity

WTA winner-take-all

Chapter 1

Introduction

1.1 A dynamical systems view of neural computation

Even in well-controlled laboratory tasks, the neural activity underlying a given behavior can appear strikingly complex. It varies across repetitions, is distributed across large populations, and is shaped by context and history in ways that are not obvious from the task itself (Churchland, Cunningham, et al., 2012; Saxena and Cunningham, 2019; Renart and Machens, 2014; Nawrot et al., 2008). Understanding how variable, distributed neural activity nevertheless implements reliable computation by transforming inputs into distinguishable network states and outputs remains a central challenge in neuroscience. One influential approach to understanding neural function has been to relate the activity of individual neurons to experimentally controlled variables (Barlow, 1972; Yuste et al., 2024), for example, head direction (Taube, 1998), movement direction during reaching (Georgopoulos et al., 1986; Rickert et al., 2009), or the orientation of a visual stimulus (Hubel and Wiesel, 1962). This approach has revealed systematic relationships between single-neuron responses and behaviorally relevant variables.

Single neurons, however, do not operate in isolation. Their response properties are shaped by recurrent interactions with other neurons and thus by the state of the surrounding network. The dependence of single-neuron responses on network state has motivated a shift toward understanding neural computation through the collective dynamics of neural populations (Saxena and Cunningham, 2019; Vyas et al., 2020). In this view, the activity of N neurons evolves in an N-dimensional state space, tracing out a trajectory over time (Shenoy et al., 2013). Depending on the underlying dynamics, trajectories may converge toward stable states, transition between transiently stable states, or remain constrained to low-dimensional manifolds (Perich et al., 2025; Khona and Fiete, 2022). This perspective provides a useful framework for describing context dependence, history effects, and flexible computation under changing inputs (Mante et al., 2013). All of these phenomena are difficult to capture from single-neuron response properties alone. For example, preparatory activity in the motor cortex is clearly structured at the population level yet is often not readily interpretable from single-neuron analyses (Kaufman et al., 2014). Such observations suggest that the relevant unit of description is often the population state trajectory rather than the individual neuron.

These two perspectives are not contradictory. Rather, population dynamics provide a framework in which single-neuron response properties can be understood as consequences of recurrent interactions and the evolving network state (Saxena and Cunningham, 2019). Explaining such processes requires models that connect biological observations to explicit dynamical mechanisms. Computational neuroscience serves this role by combining experimental data, theory, and simulation to test how neural circuits can generate the computations of interest. Models developed for this purpose can also be studied as dynamical systems in their own right. For example, the Hopfield network (Hopfield, 1982) was originally introduced as a model of biological associative memory, but it has also become important in

statistical physics (Amit, Gutfreund, et al., 1985), combinatorial optimization (Hopfield and Tank, 1985), and modern deep learning architectures in which Hopfield-like layers serve as associative memory modules (Ramsauer et al., 2021). The question then shifts from what neural circuits compute to what a class of dynamical systems can compute in principle.

This thesis adopts that mechanistic dynamical-systems perspective and asks how structured recurrent spiking networks can realize metastable computation, that is, dynamics in which population activity transiently dwells near discrete population states and switches between them, and whether such dynamics can support selection among competing internal representations or actions. The specific hypothesis is that locally balanced excitatory–inhibitory (E/I)-clustered networks provide a useful and interpretable substrate for this form of computation[1]. The goal is not to claim that the cortex directly implements this exact connectivity architecture, but to test it as a mechanistic model whose computational properties, robustness, and explanatory value can be assessed. The primary application domain is cortical decision dynamics during movement preparation, where decision is understood as the competitive selection among partially specified actions by the evolving network state. The following sections introduce the modeling choices and dynamical concepts needed to formulate and evaluate this hypothesis.

1.2 Levels of modeling in neuroscience

Understanding neural computation requires choosing a level of description at which the relevant mechanisms remain both interpretable and appropriate to the question being asked (Marr, 1982). This creates a fundamental trade-off in neuroscientific modeling. The more abstract and general a model is, the easier it becomes to analyze and relate to broad computational principles, but the harder it is to relate it to biological observations. Conversely, the more biologically detailed a model is, the easier it becomes to connect it to known anatomy and physiology, but the harder it becomes to determine which aspects are actually necessary for the computation of interest. Increasing detail enlarges the parameter space and can obscure the mechanism of interest within a multitude of interacting components. Even a highly detailed whole-brain simulation, while potentially valuable as an *in silico* experimental platform, would not by itself explain which circuit motifs and dynamical interactions are responsible for its behavior.

One useful framing distinguishes between abstract, mechanistic, and statistical models (Levenstein et al., 2023). Abstract models aim to isolate general computational motifs with minimal biological commitment and connect naturally to dynamical systems theory and statistical physics. Mechanistic models specify circuit architecture and neural dynamics explicitly enough to relate theoretical hypotheses to identifiable biological processes and experimentally testable predictions. Statistical approaches remain important as tools for constraining model parameters and interpreting model behavior. In the context of this trade-off, spiking network models provide a useful intermediate level of description. They remain simple enough to analyze at the circuit level, yet concrete enough to retain an interpretable relation to biological circuit organization and dynamics.

This is the level at which the computations of interest can be formulated in terms of recurrent circuit dynamics while retaining a direct link to biological interpretation. It is therefore the main model class used to study how structured recurrent connectivity can generate metastable dynamics and support selection among competing representations or actions. Reduced models nevertheless remain essential. Binary networks isolate the recurrent interaction rules required for metastability while suppressing detail that is not needed for

[1] We abbreviate excitatory–inhibitory as E/I throughout this thesis. The included publications variously use the notations E/I and EI; both refer to the same architecture.

that question. Mean-field theory provides a population-level description that exposes fixed points, bifurcations, and effective control parameters that are difficult to infer directly from full spiking simulations. These descriptions are therefore used complementarily. The spiking network is the model in which the computational mechanism is formulated, whereas binary and mean-field models are used to simplify, explain, and generalize the same dynamical mechanism.

1.3 Excitatory–inhibitory balance and its computational limits

At this level of description, the relevant biological question is not which full set of cortical details can be reproduced, but which features of cortical organization must be retained to capture the neural computation of interest. The mammalian neocortex exhibits substantial diversity across areas, layers, and cell types (Harris and Shepherd, 2015). This diversity is functionally meaningful. Inhibitory interneuron classes, for example, differ in their morphology, physiology, and patterns of synaptic connectivity, and these differences support distinct circuit functions (Kepecs and Fishell, 2014). This diversity is organized around recurring circuit motifs. Across cortical areas, a common feature is the local recurrence between excitatory and inhibitory populations (Douglas et al., 1989). In this thesis, this recurrent excitatory–inhibitory interaction defines the relevant computational substrate, while finer-grained differences across cell types, layers, and regions provide the biological grounding for that framework.

A central property of cortical recurrent excitatory–inhibitory circuits is the approximate balance between excitation and inhibition (Okun and Lampl, 2008). At the level of synaptic input, excitatory and inhibitory currents are often co-tuned so that their means largely cancel, leaving the net input to a neuron subthreshold or only weakly suprathreshold. Neurons in this regime therefore fire mainly because fluctuations transiently push the membrane potential across threshold. This fluctuation-driven regime produces irregular spiking and low firing rates despite strong synaptic bombardment (Amit and Brunel, 1997; Softky and Koch, 1993). It also supports divisive gain modulation, because co-varying excitatory and inhibitory background input can rescale the slope of the neuronal input–output function without strongly changing the mean drive (Chance et al., 2002). For these reasons, excitatory–inhibitory balance is a plausible generic operating regime for recurrent cortical circuits.

The computational consequences of balance depend not only on whether excitation and inhibition cancel on average, but also on the granularity at which this cancellation is maintained. In a network with random, unstructured connectivity, each neuron receives a statistically equivalent mixture of recurrent inputs, so balance holds only in a statistical sense, as an average across neurons—a form of global balance (Vreeswijk and Sompolinsky, 1996). If excitatory connectivity is structured such that subsets of neurons are more strongly coupled to one another, excitatory drive becomes locally inhomogeneous. Whether this local excess is compensated depends on the organization of inhibition. If inhibitory neurons connect independently of excitatory structure, they provide inhibition that matches population-averaged excitation but does not track these local deviations. If inhibition is instead co-organized with excitatory structure, the cancellation of excitation and inhibition is preserved within each subgroup, establishing local excitatory–inhibitory balance (Schaub et al., 2015; Rost et al., 2018). This distinction is crucial because it determines whether structured excitation generates fragile local imbalances or a controlled dynamical organization, and thereby whether structured networks can sustain robust non-stationary dynamics without fine tuning.

Balanced random networks (BRNs)[2] formalize the principle of global balance in recurrent network models under minimal structural assumptions, including sparse recurrent connectivity, statistically homogeneous neurons, and no imposed functional substructure (Vreeswijk and Sompolinsky, 1996; Brunel, 2000). They have become the canonical baseline for cortical circuit modeling precisely because they reproduce these statistical signatures of spontaneous activity. They are also computationally nontrivial (Jaeger and Haas, 2004; Maass et al., 2002). When viewed as reservoirs, such networks transform inputs into high-dimensional, temporally expanded representations, allowing trained readout weights to perform classification, temporal pattern generation, or nonlinear function approximation. In that sense, the computational power of BRNs derives from the rich transient responses generated by recurrent dynamics.

What BRNs do not provide is intrinsic recurrent computation of the kind considered here, namely computation realized by the network's own recurrent dynamics rather than delegated to a trained readout operating on transient responses. Although homogeneous BRNs can undergo bifurcations between different collective regimes, their connectivity does not by itself generate multiple robust, stimulus-selectable attractors corresponding to distinct internal representations (Brunel, 2000). The BRN itself therefore cannot maintain distinct activity configurations, autonomously select among competing representations, or sustain a state change after stimulus offset. Computation remains delegated to the readout, not realized in the network dynamics[3]. As a consequence, while BRNs can generate the rich transient dynamics that make reservoir computation possible, a BRN cannot produce categorical output. Persistent maintenance, selection among competing representations, and context-dependent modulation all require a trained readout to interpret the transient responses (Churchland, B. M. Yu, et al., 2010; Litwin-Kumar and Doiron, 2012). Accounting for computation that is realized within recurrent dynamics requires architectures with additional structure that generate richer dynamics.

1.4 Attractor networks and metastable computation

The richer state-space structure required for intrinsic recurrent computation is naturally described in the language of dynamical systems and attractors (Khona and Fiete, 2022). Consider a dynamical system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t))$$

with state[4] $\mathbf{x}(t) \in \mathbb{R}^N$ and input $\mathbf{u}(t) \in \mathbb{R}^U$. The vector field $\mathbf{f}$ assigns to each point in the state space a direction and a rate of change. For a fixed input $\mathbf{u}^*$, it induces a flow $\boldsymbol{\phi}_t^{\mathbf{u}^*}(\mathbf{x}_0)$, that is, a trajectory generated from an initial condition $\mathbf{x}_0$ over time t. An attractor $\mathcal{A}$ is an invariant set toward which trajectories from a neighborhood converge under that fixed input. Its basin of attraction,

$$\mathcal{B}_{\mathbf{u}^*}(\mathcal{A}) = \left\{ \mathbf{x}_0 : \boldsymbol{\phi}_t^{\mathbf{u}^*}(\mathbf{x}_0) \to \mathcal{A} \text{ as } t \to \infty \right\},$$

[2]Throughout this thesis we use balanced random network (BRN) as the canonical term. The first included study (Schmitt et al., 2023) follows the alternative convention random balanced network (RBN); both terms refer to the same model class (Vreeswijk and Sompolinsky, 1996; Brunel, 2000).

[3]In efficient-coding frameworks (Denève and Machens, 2016), tight excitatory–inhibitory balance is itself interpreted as internal computation. Even there, a readout is required to access the representation, and the computational capacity arises because connectivity is structured by the coding objective rather than drawn at random.

[4]In discrete models such as the Hopfield network, attractors are static configurations of the microscopic state. In spiking networks, the microscopic state is never stationary; the attractors discussed here are statistically stationary mesoscopic states, defined by averaging across neurons or over time.

is the set of initial conditions from which the system reaches $\mathcal{A}$. Attractors are the configurations toward which the system tends to settle. Their basins of attraction describe from which initial states they are reached and how robustly they are recovered after perturbations.

Attractors differ in their geometry. The simplest case is a fixed-point attractor: an equilibrium state $\mathbf{x}^*$ satisfying $\mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) = \mathbf{0}$. A fixed point is locally stable when small perturbations decay back toward it, which for smooth systems is determined by the eigenvalues of the Jacobian of $\mathbf{f}$ at $\mathbf{x}^*$ (Khalil, 2002). Other geometries such as continuous attractor manifolds, limit cycles, or strange attractors are possible, but the focus here is on multistable regimes composed of coexisting stable macroscopic states and the noise-driven transitions between them.

In theoretical neuroscience, attractor networks are recurrent neural networks (RNNs) whose structured connectivity supports a particular set of attractors and basins of attraction, so that computation is implemented by the network dynamics themselves (Khona and Fiete, 2022). Different attractor geometries support different computational functions: discrete fixed-point attractors underlie associative memory and categorical choice (X.-J. Wang, 2002; Wong and X.-J. Wang, 2006; X.-J. Wang, 2008), continuous attractor neural networks (CANNs) encode analog variables such as head direction or spatial position (Zhang, 1996), and limit cycles can generate rhythmic motor patterns (Marder and Calabrese, 1996). Attractor networks are therefore not merely computational abstractions. They make quantitative predictions about population dynamics, stimulus selectivity, and variability that can be tested against neural recordings. In favorable cases, they also describe identified circuit mechanisms, as demonstrated by ring-attractor models of the insect head-direction system (Turner-Evans et al., 2020).

The low-dimensional population structure recovered by dimensionality-reduction methods can reflect these same underlying state-space objects (Cunningham and B. M. Yu, 2014; Langdon et al., 2023). However, low-dimensional structure alone does not uniquely determine the mechanism that generates it. Similar trajectories can arise from continuous attractors, metastable switching among coexisting states, deterministic heteroclinic flow, or low-rank and non-normal recurrent dynamics (Khona and Fiete, 2022; Rabinovich, Huerta, et al., 2008; Mastrogiuseppe and Ostojic, 2018; Bondanelli and Ostojic, 2020; Christodoulou et al., 2022). Low-rank models are especially analytically tractable because they provide comparatively direct links between connectivity structure and the resulting low-dimensional dynamics (Mastrogiuseppe and Ostojic, 2018; Bondanelli and Ostojic, 2020). These frameworks differ less in the dimensionality of the resulting activity than in the circuit mechanisms that generate it, and therefore in their predictions for stability, variability, and responses to perturbation.

An important alternative relevant for this thesis is the stochastic stabilized supralinear network (SSN), in which variability is quenched around a single stimulus-dependent stable state rather than by noise-driven switching among multiple coexisting states (Hennequin, Ahmadian, et al., 2018; Ahmadian and Miller, 2021). The SSN operates in the inhibitory-stabilized regime, in which recurrent excitation is strong enough to be unstable on its own but is held in check by inhibitory feedback (Tsodyks et al., 1997; Ozeki et al., 2009; Ahmadian and Miller, 2021). This inhibition-stabilized network (ISN) regime is not specific to the SSN; it is a general property of balanced cortical circuits and, as discussed in the third study, is also relevant to the E/I-clustered network. Accordingly, this thesis focuses on one specific mechanistic hypothesis within this broader landscape, namely noise-driven metastable switching in locally balanced E/I-clustered spiking networks. Rather than treating this as the only plausible realization of cortical population dynamics, the thesis examines it as a mechanistically explicit hypothesis whose robustness and explanatory reach can be tested. More generally, the block structure of E/I-clustered networks can be viewed as one structured motif within the low-rank framework (Mastrogiuseppe and Ostojic, 2018), and

context-dependent reconfiguration of the effective couplings between clusters could embed different computational functions within the same circuit (Dubreuil et al., 2022; Beiran et al., 2023).

The regime of interest here is that of multistable networks supporting several discrete stable states under a single input condition. Recurrent connectivity determines the set of coexisting attractors, including their number, identity, and locations in state space. External input does not create this structure from scratch, but biases transitions among the available states by deforming the vector field and thereby shifting attractors and reshaping their basins of attraction.

Whether the network dwells near one stable state or transitions to another depends on the interplay between basin geometry and stochastic fluctuations, including finite-size effects, synaptic variability, and stochastic state updates (Faisal et al., 2008). When basins are deep, perturbations are absorbed and the current state is maintained; when they are shallow, noise-driven escape across basin boundaries becomes possible. The characteristic escape time scales as $\tau \sim \exp(\Delta/\sigma^2)$, where Δ is an effective barrier height and σ^2 the noise intensity, a relation derived from Kramers' theory for potential systems and often a useful approximation beyond that setting (Kramers, 1940). The metastable[5] regime arises when Δ is small enough relative to σ^2 that escapes occur on behaviorally relevant timescales: the network dwells for a stochastic dwell time near one state before transitioning abruptly to another. Multiple locally stable states thus coexist at the mesoscopic level, but noise drives itinerant transitions between them on timescales that are long relative to the intrinsic network dynamics yet short enough to be computationally meaningful.

Metastable dynamics support a mode of computation in which the state-space trajectory becomes a sequence of discrete population states and transitions between them implement changes in representation. This makes metastability a natural candidate mechanism for decision dynamics, working-memory updating, and other computations requiring selection among internal alternatives (La Camera et al., 2019). The stochastic character of the transitions is not merely a by-product of imperfect neural hardware, but reflects the geometry of the underlying basins of attraction. Dwell-time distributions, transition preferences, and trial-to-trial variability are all shaped by effective barrier heights and basin structure, thereby linking attractor dynamics to cortical population activity and behavior (Brinkman et al., 2022).

Given these considerations, the question is which forms of structured recurrent connectivity in spiking circuits can generate a metastable regime that is both robust and biologically interpretable.

1.5 Clustered architectures for metastable computation

One way to generate a multistable regime in RNNs is through structured connectivity that groups neurons into subpopulations with stronger internal coupling. When neurons that tend to co-activate also share potentiated connections, the recurrent feedback within each subgroup can sustain an elevated activity state that persists after the triggering input is removed. Such dynamical patterns of co-activation are commonly referred to as cell assemblies (Hebb, D. O., 1949; Harris, 2005). The Hopfield network provides the classical precedent. Connection strengths between neurons are set proportional to the correlation of their activity across stored patterns, so that each pattern becomes a fixed-point attractor of the

[5]The term metastability is used differently in coordination dynamics (Kelso, 2012), where it refers to dwelling near remnants of attractors after a bifurcation, and in heteroclinic frameworks (Rabinovich, Huerta, et al., 2008), where deterministic orbits connecting saddle equilibria produce sequential dynamics without noise. The models studied here operate in the noise-driven regime.

network dynamics, and noise can induce transitions between them (Hopfield, 1982)[6]. The essential insight is that non-random, correlation-structured connectivity can create multiple coexisting stable states. The question here is how this principle carries over to cortical-like recurrent spiking circuits subject to biological constraints such as sparse connectivity, excitatory–inhibitory balance, and directed synapses.

A standard approach embeds this idea in a BRN by partitioning its excitatory population into clusters with potentiated recurrent synaptic efficacies, while weakening cross-cluster connections to preserve average balance (Litwin-Kumar and Doiron, 2012; Deco and Hugues, 2012; Mazzucato et al., 2015). The resulting network supports multiple coexisting activity configurations that can be occupied spontaneously or stabilized by external input. However, inhibition remains homogeneous, so balance holds only globally. When an excitatory cluster activates, local inhibition cannot track the increased recurrent drive, and the activated population enters a mean-driven regime with high rates and regular spiking (Rost et al., 2018).

A way to resolve both difficulties is to pair each excitatory cluster with a dedicated inhibitory counterpart, forming E/I-cluster pairs (Rost et al., 2018; Schaub et al., 2015). Potentiating the couplings within each excitatory–inhibitory cluster pair allows local inhibition to track local excitation and preserve cancellation of mean input at the population level. As a result, neurons remain in a fluctuation-driven regime in both active and quiescent states. Cluster activation then produces only moderate rate increases, irregular spiking is preserved, and the effective barriers remain shallow enough for noise-driven transitions to occur across a broad parameter range. This principle of local excitatory–inhibitory balance is consistent with experimental evidence for co-tuned excitation and inhibition (Wehr and Zador, 2003; Okun and Lampl, 2008). It is also consistent with growing evidence that inhibitory connectivity is more structured than would be expected if inhibition were distributed uniformly and non-specifically across the local circuit in a blanket of inhibition (Znamenskiy et al., 2024; Najafi et al., 2020).

The E/I-clustered network thus serves as the central circuit hypothesis, a structured RNN class that supports robust metastable computation while preserving cortical-like spiking statistics.

1.6 Open questions

Evaluating this hypothesis requires distinguishing between the different kinds of claims it can support (Marr, 1982).

An existence claim concerns whether the kind of local clustered connectivity assumed in the present models is in fact present in cortex. This is the hardest claim to establish, since it requires direct anatomical or physiological evidence at the appropriate scale. Structured local connectivity, fine-scale subnetworks, and selective excitatory and inhibitory connectivity have been reported in cortical microcircuits (Song et al., 2005; Perin et al., 2011; Morishima et al., 2017; Najafi et al., 2020), and recent millimeter-scale connectomics data confirm clearly non-random, functionally organized connectivity (Bae et al., 2025; Znamenskiy et al., 2024). Yet the observed structure often follows continuous, graded similarity-based connectivity rather than strictly discrete clusters, and similar signatures may also arise from distance-dependent connectivity or broad degree heterogeneity (Vegué et al., 2017). It therefore remains open whether the discrete clustering assumed in the present E/I-clustered models is a useful idealization of a more graded reality.

[6]The Hopfield network uses symmetric weights ($J_{ij} = J_{ji}$), guaranteeing an energy function whose minima correspond to stored patterns. Cortical connectivity is directed and asymmetric, so the existence of a comparable effective landscape must be established by other means.

A utility claim concerns whether the model class can implement useful computations in principle, regardless of whether the specific motif exists in cortex. In the present context, this means supporting multistable and metastable dynamics, winner-take-all (WTA) selection, and structured sequences of internal states while preserving biologically plausible spiking statistics. Such utility matters not only for cortical modeling but also, more generally, for identifying computational principles that may be realized in other physical substrates, including neuromorphic systems.

An explanatory claim concerns whether the model accounts for specific neural and behavioral observations under defined task conditions. It generates quantitative predictions about firing rates, variability dynamics, and decision-related behavior that can be tested against experimental data. Explanatory claims are more tightly constrained by data than utility claims but less so than existence claims, because a model may account for observations through mechanisms that are functionally correct while remaining structurally approximate.

These three claims are not stages of a single argument, but distinct axes along which a model can be evaluated. A model may score highly on utility and explanatory power while the existence claim remains unresolved. Conflating these axes, for example, by treating explanatory success as evidence for the literal presence of the hypothesized motif, is a common interpretive error that this thesis aims to avoid. On this basis, the central open questions concern scalability, structural robustness, and explanatory reach. First, systematic simulation and calibration of clustered spiking networks at scale remain computationally demanding, limiting both exploration of parameter space and extension to larger, more biologically detailed systems. Second, the theoretical relationship between different implementations of clustering, through connection probability (structural clustering), synaptic efficacy (weight clustering), or mixtures of both, and the resulting dynamical signatures is not yet understood. As a result, it remains unclear how tightly function depends on a particular structural realization. Third, the model must be confronted with neural and behavioral data to determine its explanatory reach, especially with respect to variability dynamics and decision-related behavior.

Aims and contributions

This thesis addresses the above mentioned questions through three studies that examine, in turn, the scalability, structural robustness, and explanatory reach of the proposed model class.

The first study benchmarks GPU-based and CPU-based simulation of large-scale spiking attractor networks, demonstrating real-time capability for networks of up to 100,000 neurons and repurposing a batch-simulation mode for systematic parameter-space exploration. Because the metastable decision-dynamics regime is sensitive to network parameters, systematic calibration across high-dimensional parameter spaces is a prerequisite for the modeling work that follows.

Whether the metastable regime depends on how clustering is implemented structurally is the subject of the second study. It introduces a unified framework in which clustering contrast can be realized through connection probability, synaptic efficacy, or any mixture of both, controlled by a single parameter[7]. Mean-field analysis together with binary and spiking simulations show that metastable dynamics persist across this continuum, while observable signatures, including correlations, co-activation patterns, and bifurcation structure, depend on the implementation. This dissociation between robust function and

[7]The second study parameterizes clustering contrast as R_{E+} and the structural-weight mixing ratio as $\kappa \in [0, 1]$. At $\kappa = 1$ (pure weight clustering), R_{E+} coincides with the synaptic potentiation factor J_{E+} used in the first and third studies.

implementation-dependent phenomenology matters for biological interpretation because it shows that statistical signatures alone do not uniquely identify the underlying architecture, thereby reinforcing the distinction between explanatory and existence claims.

The third study turns from the model's internal structure to its explanatory reach, applying the E/I-clustered network to motor-cortical decision dynamics during movement preparation under uncertainty. The model reproduces task-dependent changes in neural variability, preserves irregular spiking, and accounts for condition-dependent reaction times observed in behaving monkeys. By linking metastable attractor dynamics to context-dependent neural variability and decision behavior, this study tests whether the utility demonstrated in the second study translates into quantitative explanatory power.

Chapter 2

Results

2.1 Efficient parameter calibration and real-time simulation of large-scale spiking neural networks with GeNN and NEST

Authors: Felix J. Schmitt, Vahid Rostami, and Martin P. Nawrot

Venue: Frontiers in Neuroinformatics (2023)

Status: Published (publisher PDF)

DOI/URL: 10.3389/fninf.2023.941696

Supplementary information: See Appendix SI.1, p. 25.

Contribution statement: The study was designed by Felix J. Schmitt and Martin P. Nawrot. Felix J. Schmitt carried out the simulations and analyses. Vahid Rostami and Martin P. Nawrot supervised the project. Felix J. Schmitt and Martin P. Nawrot wrote the manuscript. All authors contributed to the article and approved the submitted version.

2.2 Metastable Neural Assemblies on a Wiring–Weight Continuum

Authors: Felix J. Schmitt, Franziska L. Müller, and Martin P. Nawrot

Venue: Manuscript under review (2026)

Status: Preprint available on bioRxiv; manuscript under review

DOI/URL: 10.64898/2026.03.16.712138

Supplementary information: See Appendix SI.2, p. 26.

Contribution statement: The study was designed by Felix J. Schmitt and Martin P. Nawrot. Felix J. Schmitt and Franziska L. Müller carried out the simulations and analyses. Felix J. Schmitt and Martin P. Nawrot supervised the project. Felix J. Schmitt and Martin P. Nawrot wrote the manuscript. All authors approved the submitted version.

2.3 Spiking attractor model of motor cortex explains modulation of neural and behavioral variability by prior target information

Authors: Vahid Rostami, Thomas Rost, Felix J. Schmitt, Sacha J. van Albada, Alexa Riehle, and Martin P. Nawrot

Venue: Nature Communications (2024)

Status: Published (publisher PDF)

DOI/URL: 10.1038/s41467-024-49889-4

Supplementary information: See Appendix SI.3, p. 27.

Contribution statement: The study was conceptualized by Vahid Rostami, Thomas Rost, Alexa Riehle, and Martin P. Nawrot. Experimental design and data acquisition were carried out by Alexa Riehle. Vahid Rostami, Thomas Rost, Felix J. Schmitt, and Martin P. Nawrot performed the data analysis, and Vahid Rostami, Thomas Rost, and Felix J. Schmitt developed the computational model. The original manuscript draft was prepared by Vahid Rostami and Martin P. Nawrot, and the manuscript was subsequently reviewed and edited by Vahid Rostami, Felix J. Schmitt, Sacha J. van Albada, Alexa Riehle, and Martin P. Nawrot. Funding was acquired by Sacha J. van Albada and Martin P. Nawrot. All authors approved the final version.

Chapter 3

Discussion

This thesis asked whether locally balanced E/I-clustered spiking networks provide a useful model class for metastable cortical decision dynamics. The three studies support a restrained conclusion. E/I-clustered spiking networks are computationally tractable, structurally robust, and capable of reproducing key neural and behavioral signatures of action selection under uncertainty. Realizing these advantages, however, requires systematic calibration against data. The first study established the computational basis for this by enabling efficient parameter sweeps, calibration, and robustness analysis for E/I-clustered spiking networks. This is preferable to relying on isolated hand-tuned examples because it allows the model behavior to be assessed across parameter space rather than at a few selected operating points. The second study showed that metastable dynamics persist across a continuum from structural to weight-based clustering. Correlation structure and the repertoire of population states, however, still depend on how clustering is implemented. The third study showed that E/I-clustered spiking networks reproduce uncertainty-dependent modulation of neural variability and reaction times during competitive action selection in motor cortex. The model does not address the full generation of movement. Its explanatory scope is limited to action selection during movement preparation.

3.1 Structural robustness and inferential limits

3.1.1 Structural degeneracy

The second study showed that metastable dynamics persist across a continuum of excitatory and inhibitory cluster implementations. This raises the question of whether a unique microscopic connectivity scheme is required, or whether structurally different circuits can support the same function.

Complex biological systems often exhibit hierarchical degeneracy, that is, the capacity of structurally different elements to realize the same function across multiple levels of organization (Tononi et al., 1999; Edelman and Gally, 2001; Albantakis et al., 2024). This property is closely related to robustness, adaptability, and evolvability. At the cellular level, neurons of the same class can differ in ion-channel conductances while preserving similar firing patterns (Goaillard and Marder, 2021). At the circuit level, different synaptic parameterizations can produce equivalent network dynamics (Prinz et al., 2004; Marder and Taylor, 2011). Behavioral solutions can also be degenerate. In the third study, both monkeys solved the task, but one anticipated the movement whereas the other largely ignored preparatory cues. RNNs show the same general property. In Jürgensen et al. (2024), multiple structurally distinct motifs were sufficient to explain second-order conditioning, demonstrating that the mapping from network structure to learned behavior can be many-to-one.

In the E/I-clustered network, this same form of degeneracy appears along the wiring–weight continuum. Here κ denotes the mixing parameter, with $\kappa = 0$ for purely structural clustering and $\kappa = 1$ for purely weight-based clustering. Different combinations of structural clustering and synaptic-weight clustering can produce the same metastable dynamics. These implementations are nevertheless not equivalent in every respect. They can differ in pairwise correlations, co-activation structure, robustness, and metabolic or technical cost, depending on the substrate (Deistler et al., 2022). For that reason, the second study does not imply that all implementations are interchangeable. It shows instead that the same metastable computation can arise from more than one microscopic realization.

This has a direct inferential consequence. From phenomenological data alone, the underlying circuit parameterization cannot in general be uniquely recovered. In that sense, the inverse problem is ill-posed: the mapping from structure to function is many-to-one, and the same observables may be generated by disparate circuit parameterizations (Marder and Taylor, 2011). If local cortical circuits implement an E/I-clustered network motif, the preferred implementation along the wiring–weight continuum may differ across brain areas or physical substrates because of differences in robustness requirements, developmental constraints, or energetic cost. Constraining this space more tightly would require additional structural measurements beyond neural activity alone. Fine-scale connectomic data, for example, could directly constrain the connectivity component of clustering and thereby narrow the range of admissible positions along the wiring–weight continuum, even if synaptic efficacy remains unobserved.

3.1.2 From robustness to identifiability

This structural degeneracy of the E/I-clustered network has direct consequences for how the calibration in the third study should be interpreted. That calibration is best understood as a point-fit analysis. By screening parameterizations against condition-dependent Fano factor (FF) modulation, preserved local coefficient of variation of inter-spike intervals (CV_2), firing-rate trajectories, and reaction times, the analysis showed that E/I-clustered network dynamics can reproduce the joint neural–behavioral phenomenology. Under the structural degeneracy established above, however, this falls short of identifiability. It does not establish whether the solution is unique, whether only parameter combinations rather than individual parameters are constrained, or whether structurally distinct implementations generate the same observables. Even without exact degeneracy, practical identifiability may remain poor when the available data are insensitive to some directions in parameter space. In that case, broad or multimodal posteriors are not fitting failures, but statements about what the experiment can and cannot determine.

The second study suggests how this situation could be improved. Switching statistics, dwell times, and variability quenching mainly constrain the dynamical regime, whereas pairwise correlations and co-activation structure are more diagnostic of structural implementation, despite the structural degeneracy along the κ axis. This argues for posterior calibration rather than further manual search. Simulation-based inference (SBI), a family of likelihood-free methods that approximate Bayesian posteriors over model parameters by training density estimators on simulated data, is well suited to this setting because the E/I-clustered spiking network model lacks a tractable likelihood. Rather than seeking a single best parameter vector, SBI asks which parameters or parameter combinations are identifiable given the data. It also indicates where the posterior remains broad and which predictions are stable across data-consistent realizations (Gonçalves et al., 2020; Cranmer et al., 2020; Boelts et al., 2025). For the E/I-clustered network, this would enable a more principled application to richer large-scale recording datasets by keeping parameter uncertainty explicit.

The MICrONS program provides a concrete example of how such constraints can be strengthened through multimodal data. Its core dataset combines dense calcium imaging with a co-registered electron-microscopic reconstruction of local connectivity (Bae et al., 2025). Beyond descriptive structure–function comparison, recent analyses use foundation-model and digital-twin approaches to jointly predict neural responses, anatomical cell types, and connectivity, thereby separating tuning into feature and spatial components (E. Y. Wang et al., 2025; Ding, Fahey, et al., 2025; Ding, Tran, et al., 2026). For model fitting, the important point is that functional organization is not a single target. Feature tuning, spatial organization, and connectivity constrain different parts of parameter space and should not be collapsed into a single activity-matching objective.

For the E/I-clustered network, at least three objective groups should therefore be distinguished. First, regime-level statistics such as firing-rate trajectories, switching structure, dwell times, and changes in FF. Second, variability and covariation statistics such as CV_2, FF, spike-count correlations, and reaction-time distributions. Third, structural priors such as adjacency constraints, cell-type-resolved parameter distributions, and connectomic constraints on connectivity and cell-type-specific organization. These terms should be normalized by measurement and simulation uncertainty rather than combined as an unweighted L_2 loss. Otherwise, the model can match mean rates or reaction times with the wrong source of variability or the wrong structural mechanism. A related result was

reported in the recent deep-learning-assisted cortical-circuit model of Ito et al. (2026). There, removing biological priors on synaptic efficacy distributions preserved activity fits while disrupting emergent connectivity rules. Activity-only fitting is therefore insufficient.

The most informative experiments are complementary rather than merely larger. For example, MICrONS-like connectomics constrains structural connectivity, but not synaptic efficacy, short-term dynamics, intrinsic excitability, or latent state. Large-scale calcium imaging constrains tuning and slow state dependence, but not millisecond delays or synchrony. For the E/I-clustered network, the most useful additions would be simultaneous within-session population recordings that remove the pseudopopulation problem, cell-type-resolved physiology that anchors intrinsic and synaptic priors, and perturbation datasets that test responses to intervention rather than passive activity alone. In that workflow, AI-based methods are not themselves the explanation. They are tools that make posterior inference, learned representations, and multi-objective calibration tractable.

Given that several implementations can support similar dynamics, the decisive test is no longer unique structural recovery, but explanatory reach: whether the same model class can account for behaviorally meaningful signatures in data. The most direct such test in the present work concerns decision dynamics during motor preparation under uncertainty.

3.2 Decision dynamics during motor preparation under uncertainty

The third study confronted the E/I-clustered network with condition-dependent variability quenching, preserved spiking irregularity, and reaction-time distributions during a delayed reach task under graded uncertainty. To assess what this comparison establishes, the following subsections examine the mechanism of variability quenching against the principal alternative account and then ask where the model's explanatory scope ends.

3.2.1 Mechanisms of variability quenching

Both stochastic SSNs and the E/I-clustered network account for the widespread reduction of trial-to-trial variability at stimulus onset (Churchland, B. M. Yu, et al., 2010; Hennequin, Ahmadian, et al., 2018). The key difference lies in the underlying mechanism. In the stochastic SSN, increasing stimulus drive strengthens effective recurrent interactions and, in particular, inhibitory restoring forces around a single stimulus-dependent stable state. Variability is therefore quenched around one attractor rather than by restricting switching among multiple coexisting states (Hennequin, Ahmadian, et al., 2018). In the E/I-clustered network, by contrast, variability quenching reflects a stimulus-dependent restriction of the accessible metastable state repertoire. In the task setting studied here, this predicts weaker variability quenching when more target clusters remain simultaneously viable during preparation. At the same time, the third study shows a stimulus-strength-dependent reduction in FF even when the cue projects to a single cluster. This indicates that, within the E/I-clustered network, variability quenching cannot be attributed solely to a reduction in the number of competing states, but also depends on stimulus-dependent changes in the local operating point.

A second point of comparison is within-trial spiking irregularity, which need not track trial-to-trial variability one-to-one in motor cortex (Riehle, Brochier, et al., 2018). The E/I-clustered network preserves approximately constant interspike interval (ISI) variability from spontaneous to evoked conditions by maintaining fluctuation-driven firing through tight local E/I balance, consistent with the recorded neural data. For SSN-like models, the picture is less clear. The stochastic SSN was introduced primarily to explain the suppression of shared and low-frequency variability around a single stable state (Hennequin, Ahmadian, et al., 2018). Related analyses of spiking E/I networks predict that internally generated variability and the coefficient of variation of interspike intervals can decrease for sufficiently strong inputs (Sanzeni et al., 2020). Recent mappings between SSN regimes and spiking networks, however, show that asynchronous irregular activity with nearly Poisson-like ISI statistics can also persist in biologically relevant operating regions (Ekelmans et al., 2023). It therefore remains unclear whether an SSN would preserve CV_2 under the specific conditions considered here, namely constant cue strength with graded uncertainty implemented through the number of potential targets.

This comparison is further limited by the chosen input model. In the present parametrization, cue information enters as a fixed-amplitude input to all excitatory neurons in each encoding cluster and therefore has a discrete input structure. Consequently, increasing the number of cued targets also increases the total external drive to the network. The observed modulation of FF with target number therefore reflects both a change in the number of viable metastable states and a change in total input strength. Broader or partially overlapping input tuning profiles could shift the balance between these contributions and might increase the role of gain-dependent variability quenching.

At the same time, the present results suggest a partial mechanistic overlap with the ISN framework. Beyond reducing the number of accessible metastable states, cue input may also quench variability by shifting the local operating point of clustered E/I populations. Consistent with this interpretation, the E/I-clustered network exhibits a broader and less steep effective input–output relation at the cluster level than the purely E-clustered network, while maintaining fluctuation-driven activity through tight local balance. This raises the possibility that variability quenching in the E/I-clustered network reflects both metastable state selection and an ISN-like stabilization mechanism.

3.2.2 Action selection as one component of motor-cortical decision dynamics

The preceding comparison concerned the preparatory epoch in isolation. However, even within motor areas, population dynamics differ across task phases and appear to be linked to the computations being performed (Shenoy et al., 2013). In motor cortex, preparatory activity occupies a subspace largely orthogonal to the subspace engaged during movement execution (Kaufman et al., 2014; Elsayed et al., 2016). While preparatory activity is often well described by fixed-point dynamics (Inagaki, Fontolan, et al., 2019), movement execution is more commonly associated with rotational or transient trajectories (Churchland, Cunningham, et al., 2012), though the rotational account has been challenged by traveling-wave descriptions (Kuzmina et al., 2024). Even the character of execution dynamics depends on the movement class: reaching and grasping engage qualitatively different population motifs within the same area (Suresh et al., 2020), and qualitatively different dynamics have been reported in supplementary motor area despite overlapping directional tuning at the single-neuron level (Lara et al., 2018). Importantly, these regimes are not rigidly sequential. Preparatory fixed-point states can be bypassed or compressed when no delay is imposed (Ames et al., 2014), and during rapid corrections, preparatory and execution-related dynamics can be expressed in parallel (Ames et al., 2019).

The model in the third study addresses only movement preparation, specifically the competitive selection among potential actions during the delay period. It does not provide a hypothesis for the neural mechanisms underlying commitment to a decision, nor for movement initiation and execution. Since the same neurons likely participate in these different dynamics, and no clear feedforward separation has so far been identified, the corresponding network motifs likely coexist within the same recurrent circuit. Theoretical work on motor control in recurrent networks offers a perspective on how such coexistence might be realized. Hennequin, Vogels, et al. (2014) showed that transient amplification in balanced excitatory–inhibitory networks can generate complex movement-like dynamics when controlled by appropriately structured inputs. Subsequent work by Schimel et al. (2024) and Kao et al. (2021) showed that motor preparation in trained RNNs emerges precisely when the network's autonomous dynamics are rich enough to require initial-condition control. In these models, recurrent connectivity provides a dynamical repertoire, loosely analogous to a reservoir, and preparation sets the initial conditions from which execution-period dynamics unfold.

An open question is whether an analogous separation of roles could also operate in the E/I-clustered network studied here. In that case, metastable attractor dynamics would govern target selection during preparation, whereas transient within-cluster dynamics would support execution after a cluster has been selected. Such an arrangement would require that individual clusters possess sufficiently rich internal dynamics to support movement-related trajectories when driven by appropriately structured descending input, a possibility that has not been tested for the present model class.

More broadly, the transition between dynamical regimes from attractor-like preparation to transient execution likely requires reconfiguration of the effective connectivity on a fast timescale (Bachschmid-Romano et al., 2023). Low-rank decompositions of trained RNNs show that such regime switches can be achieved through gain modulation that reshapes the low-dimensional dynamics without changing the anatomical connectivity (Dubreuil et al., 2022; Beiran et al., 2023). How

gain-modulated reconfiguration interacts with the E/I-clustered network studied here is an open problem, but neuromodulatory or thalamic gating provides a biologically plausible route (Kao et al., 2021).

While the model captures action selection during movement preparation, it does not by itself provide a mechanistic account of commitment, movement initiation, or movement execution. At the same time, its explanatory reach is not trivial: the inter-monkey differences in variability dynamics and reaction times were captured by a single parameter change. A further limitation concerns the model's dependence on an external behavioral readout. The RNN generates preparatory competition, but commitment is implemented by an additional leaky integrator with threshold. Reaction times and error rates are therefore properties of the combined recurrent-plus-decoder model, not of the recurrent circuit in isolation. These limitations concern the scope of the model's explanatory claims. A separate question is whether its architectural assumptions, including discrete clusters, point neurons, and noise-driven switching, should be taken literally as descriptions of cortical circuitry.

3.3 Biological plausibility and level of interpretation

Before examining the model's architectural assumptions, it is worth distinguishing the dynamical phenomenology from the mechanism proposed to explain it. With the rise of parallel spike-train recordings from cortical populations, hidden Markov models (HMMs) have been applied to single-trial activity in some areas to infer sequences of quasi-stationary states whose transitions relate to sensory processing and behavior (Abeles et al., 1995; Jones et al., 2007). Yet this should be interpreted first as a compact statistical description, not as direct evidence that the underlying circuit is a noise-driven multistable attractor system, because discrete latent states are already built into the HMM. Similar switching statistics do not necessarily imply noise-driven hopping between discrete attractors. For example, continuous deterministic dynamics in heteroclinic networks can generate semi-stable behavioral or neural states with apparently random transitions and controlled dwell-time statistics (Rabinovich and Varona, 2011; Morrison and Young, 2022). HMMs are therefore a valid and important tool for summarizing single-trial population activity, but they should not be understood as proof of the underlying network mechanism.

In the third study, the mechanistic value of the E/I-clustered network is not merely reproducing discrete-state phenomenology, but that the same architecture jointly explains state structure, preserved CV_2, condition-dependent FF modulation, firing-rate dynamics, and reaction times. An HMM can summarize such state sequences, but it cannot predict their covariation. The question is then how literally the main assumptions of this explanation should be taken, namely the network's scale, its discrete cluster boundaries, and its embedding in the larger cortical circuit.

3.3.1 A reduced local patch

The E/I-clustered network is best interpreted not as a literal cortical column, but as a reduced local patch of cortex. Although its neuron count overlaps with that of a small microcircuit, the model has neither explicit space nor laminar identity, and its clusters are defined only by recurrent coupling. The more defensible interpretation is therefore functional rather than anatomical: a downscaled balanced RNN chosen to preserve the preparatory dynamical regime, specifically metastable switching, irregular spiking, and variability quenching, without claiming to reproduce the full cellular and synaptic density of a tissue volume (Mountcastle, 1997; Horton and Adams, 2005; Potjans and Diesmann, 2014; Albada et al., 2015).

This distinction is important because downscaling has consequences. Full-density local patch models on the order of 1 mm^2 require roughly 10^5 neurons, whereas networks in the present size range preserve only selected statistics under scaling rules. The model should therefore not be read as the tissue directly sampled by an electrode. It is better understood as an effective mesoscale surrogate for the local dynamics that matter here. On that reading, the clusters represent functional subpopulations within a local motif, not anatomical modules with sharply defined borders.

3.3.2 Graded rather than discrete structure

Recent multimodal studies of cortical networks have begun to clarify the relationship between functional properties of single neurons and their structural embedding. The MICrONS consortium dataset, already referenced above for its implications for model fitting, provides the strongest current anatomical benchmark: a multimodal reconstruction of an approximately 1 mm^3 cortical volume from an 87-day-old mouse spanning primary and higher visual cortex, combining large-scale two-photon calcium imaging with serial section electron microscopy (Bae et al., 2025). Earlier studies had already shown that connectivity in mouse V1 reorganizes after eye opening from an initially random pattern toward one correlated with tuning similarity (Ko et al., 2013), and that the sparse strong synaptic connections providing the largest fraction of excitatory input arise from neurons with similar tuning (Cossell et al., 2015). The MICrONS dataset is dominated by excitatory pyramidal neurons (Pyr) and therefore provides the most direct current constraints on excitatory recurrent connectivity.

Ding, Fahey, et al. (2025) confirmed that neurons with similar response properties are preferentially connected and extended this result by showing that the same principle holds within and across layers, and even across areas including feedback pathways. Fine-scale synaptic connectivity was better predicted by similarity in feature tuning than by receptive field location, beyond the contribution of spatial proximity within the volume. At the same time, this like-to-like connectivity does not fully explain the observed similarity between connected neurons, which points to a higher-order functional organization beyond simple pairwise relationships. The resulting picture is therefore more graded than the E/I-clustered network assumed here and more closely resembles a continuous organization in feature space (Oldenburg et al., 2024).

Regarding inhibition, available evidence likewise argues against taking the abstract E/I-clustered network literally as a discretely partitioned connectivity motif. Early slice-mapping studies emphasized dense local inhibition, with nearby Pyr cells receiving input from many local interneurons (Fino and Yuste, 2011; Packer and Yuste, 2011; Karnani et al., 2014). More recent work suggests that cortical inhibition is structured within this broadly connected background. In the MICrONS reconstruction, inhibitory neurons formed overlapping groups that converged on selected excitatory subpopulations rather than one inhibitory population paired with one excitatory group (Schneider-Mizell et al., 2025). Cortical inhibition is also distributed across several interneuron classes with different subcellular targets and disinhibitory interactions, providing multiple routes for state-dependent modulation (Kepecs and Fishell, 2014; Pfeffer et al., 2013; Pi et al., 2013). Inhibitory connectivity is also not unstructured. In mouse V1, synaptic efficacies between parvalbumin-expressing (PV) interneurons and Pyr cells increase with response similarity even when connection probability remains comparatively dense (Znamenskiy et al., 2024). Inhibitory organization therefore appears aligned with excitatory organization, but in a graded and overlapping way rather than through discrete inhibitory clusters.

Distributed and partially overlapping feature representations, as observed in V1 and M1 and also in the data analyzed in the third study (Rickert et al., 2009), suggest that the underlying circuit structure is likely more graded than the discrete E/I-clustered network assumed here (Muir et al., 2017; Mackwood et al., 2021). More generally, parameterizing recurrent connectivity by neuron-specific latent variables is well established. The latent variable may encode physical location, yielding distance-dependent coupling (Rosenbaum and Doiron, 2014), or a preferred feature in an abstract tuning space, yielding structured connectivity and continuous attractor dynamics (Zhang, 1996). In this sense, the E/I-clustered network should be understood as a schematic, analytically tractable limiting case within a broader family of structured networks rather than as a literal anatomical model.

Because most of the detailed evidence comes from mouse visual cortex, it should be interpreted with some care. Visual cortex is a sensory area with strong feature-based organization, and it remains unclear whether the same principles apply to motor areas with a different computational role. The same uncertainty applies across species. In macaque motor cortex, for example, mixed selectivity and pronounced differences in population dynamics between motor preparation, initiation, and execution have been reported. The relevant latent variables may therefore be more closely related to task variables than to sensory features, but the corresponding feature space has not yet been identified. It is likewise unclear whether motor cortex follows an analogous similarity rule. In addition, the available evidence is based on a short temporal window for functional measurements and on a single structural snapshot from a continuously remodeling circuit, with dendritic spine turnover, synaptic efficacy changes, and representational drift over days to weeks. The biologically relevant question

is therefore not whether cortex literally contains static, non-overlapping E/I-clusters, but whether it transiently enters regimes in which excitatory and inhibitory structure are sufficiently co-organized to generate the same dynamical computation.

3.3.3 The motif's computational role in a distributed system

Reading the model as a local motif raises the question of how it relates to the distributed cortical activity observed experimentally. The E/I-clustered network receives topographically structured input and projects toward a downstream integrator, implicitly assuming that a single local circuit performs competitive selection and causally drives the decision. However, the observed signatures could equally reflect a decision computed upstream. In mouse anterolateral motor cortex (ALM), a premotor area, optogenetic perturbation during the delay period could switch discrete attractor-like states and lead to incorrect subsequent licking responses (Inagaki, Fontolan, et al., 2019), consistent with a causal contribution of local dynamics. Across areas, local attractor dynamics may also regulate which upstream signals are expressed downstream rather than directly generating motor commands (Finkelstein et al., 2021). Whether primate motor cortex harbors similarly causal attractor dynamics remains a possibility not tested in this thesis.

What is available are array recordings that constrain the spatial organization of local representations. Within local primary motor cortex (M1) patches, neurons with related movement features appear broadly intermingled across the recorded territory rather than spatially segregated (Vargas-Irwin et al., 2010), and directional tuning exhibits only weak local clustering during preparation, without evidence for a strict modular map (Ben-Shaul et al., 2003). At the same time, ordered spatiotemporal structure during movement initiation and execution spans millimeters of cortex (Rubino et al., 2006; Best et al., 2017; Riehle, Wirtssohn, et al., 2013). The experimental data in the third study comprise a pseudopopulation of 76 units recorded across multiple sessions. Across sessions, electrode placements inevitably sample different micro-territories and temporal snapshots, so the pseudopopulation conflates what may be distinct local instantiations of the E/I-clustered network.

This demands reinterpretation of what the model represents in the intact brain. Three scenarios are worth distinguishing. First, the motif may tile motor cortex as loosely coupled modules whose states partially align through shared input or lateral connectivity. Kim et al. (2023) showed in spiking networks fitted to ALM data that training only a subset of neurons was sufficient to reproduce task-related activity across untrained neurons through strong task-independent synapses. This suggests that attractor-like dynamics can be distributed well beyond the directly driven population. Brain-wide recordings during analogous memory-guided tasks confirm that choice-related activity is broadly distributed, though concentrated in specific circuit nodes (Chen et al., 2024; Steinmetz et al., 2019). Second, even within a single cortical area, the motif may be organized hierarchically: Daie et al. (2021) revealed that ALM contains recurrently connected modules that are spatially intermingled, each independently capable of maintaining selective persistent activity and weakly coupled with each other. Such modular substructure within one area is precisely the arrangement that, theoretically, gives rise to slow-switching assembly dynamics governed by the between-module connectivity graph (Schaub et al., 2015). If a similar architecture exists in primate motor cortex, metastable transitions at the local cluster level could nest within slower assembly sequences, generating multi-timescale dynamics. Third, following Finkelstein et al. (2021), each cortical site may act primarily as a local gate, with coordination arising through shared modulatory input rather than monolithic network computation.

Distinguishing these scenarios requires simultaneous recordings from identified neurons across millimeter-scale territories at single-spike resolution, combined with targeted causal perturbations. High-density probes sample along one axis, calcium imaging sacrifices temporal resolution, and multi-electrode arrays sample too sparsely. The E/I-clustered network should therefore be understood as a motif-level hypothesis: a sufficient local mechanism whose instantiation in the intact brain is likely more distributed, hierarchical, and functionally nuanced than any single-scale model can express.

3.3.4 From the local motif to larger-scale observables

The observation that movement-related activity extends across millimeters of cortex underscores why larger-scale modeling matters. A reduced local model is scientifically strongest when the mechanism it isolates can be embedded in wider circuitry and related to measurements above the single-neuron level. Computational models are especially useful here because they can connect hypotheses formulated at one spatial or temporal scale to observables at another. The relevant question is therefore not whether the present model is already a mesoscale model, but whether its clustered mechanism survives translation to larger circuits and richer signals.

A first positive result comes from the multi-area macaque visual-cortex model of Pronold et al. (2024), which embeds joint E/I-clustering into the full-density multi-area framework of Schmidt et al. (2018). That model was not designed to reproduce behavioral performance or task-related computations; its goal was to replicate statistical signatures of cortical activity, including firing-rate distributions and inter-area propagation. Clustering improved these activity statistics and reproduced stimulus-driven variability quenching. This supports an important but limited conclusion: the E/I-clustered network is at least embeddable in a distributed cortical architecture. It does not show that such clustering is the cortical mesoscale building block, and it does not remove the dependence on modeling choices specific to that scaffold.

Beyond spikes, extracellular recordings in motor cortex routinely capture local field potentials, which carry substantial information about behavioral state, movement parameters, and cortical dynamics (Rickert et al., 2009; Riehle, Wirtssohn, et al., 2013). If the E/I-clustered network is a building block of local cortical computation, it should leave signatures not only in spike statistics but also in the mesoscale signals that integrate synaptic activity across the local population. A point-neuron network does not directly generate a physical local field potential (LFP), because LFP recordings arise from transmembrane currents in spatially extended neurons and depend on morphology, synapse placement, laminar alignment, and volume conduction (Einevoll et al., 2013). Point-neuron models can nevertheless be linked to LFP once an explicit measurement model is added, for example by using weighted synaptic-current proxies or by replaying the spikes into multicompartment neurons in a forward model (Mazzoni et al., 2015; Hagen et al., 2016). Crucially, the simulated neurons are then not the literal physical generators of the recorded field. Rather, the spike network provides a compact description of the coordinated synaptic activity that, at the appropriate density and geometry, would give rise to an LFP.

For a model of the present scale, LFP predictions should be understood as qualitative constraints on synaptic coordination rather than as quantitative predictions of recorded field amplitudes or spatial profiles. This is also why LFP would add information rather than merely duplicate spike statistics: weak pairwise spike correlations can coexist with spatially coherent field signals, so the field constrains aspects of collective coordination that spikes alone leave underdetermined (Senk et al., 2024). If the clustered regime is the right local mechanism, it should leave a state-dependent signature in synaptic coordination and therefore in LFP structure. That makes LFP a useful next constraint, but one whose interpretation requires an explicit forward model bridging from the reduced network to the recorded field.

3.3.5 Candidate mechanisms for the formation of E/I-clustered networks

The preceding subsections examined how to interpret the model's architecture given current anatomical evidence. A complementary question is how such structure could arise in the first place. All three included studies assumed that the E/I-clustered network structure is already in place, but did not address how such structure is established. The third study assumed that neurons have organized into stable discrete clusters throughout the extensive training phase of the delayed center-out reach task. Given the graded connectomic picture, the biologically more plausible interpretation is not the literal formation of static, block-diagonal clusters, but the emergence of an E/I-clustered regime in which excitation and inhibition become sufficiently co-organized in feature space to support metastable dynamics. The mechanistic question can therefore be divided into the formation of strengthened cell assemblies, the co-tuning of inhibition, the respective roles of functional and structural plasticity, and the developmental constraints under which these processes unfold.

At the level of excitatory neurons, repeated co-activation can in principle create cell assemblies through Hebbian potentiation (Hebb, D. O., 1949), spike-timing-dependent plasticity (STDP), or rate-based rules such as the Bienenstock–Cooper–Munro (BCM) rule (Bienenstock et al., 1982), but these

are only stable if they are paired with compensatory processes to prevent runaway activity (Miehl et al., 2023). Modeling work has therefore consistently combined correlation-driven strengthening with homeostatic or heterosynaptic mechanisms, showing that assemblies reflecting previously experienced inputs can emerge and remain reactivatable by their own recurrent dynamics. Several studies are especially relevant here. Litwin-Kumar and Doiron (2014) showed that assemblies can form and remain stable under realistic synaptic plasticity and homeostatic control. Zenke et al. (2015) argued that stable memory formation requires several interacting compensatory mechanisms. Manz and Memmesheimer (2023) showed that a depression-dominant STDP rule is sufficient to maintain assemblies without additional compensatory mechanisms. X. Yang and Camera (2024) demonstrated that a fully local rule, explicitly reminiscent of BCM and belonging to the class of voltage-based STDP models, can generate clustered connectivity with metastable dynamics without imposing clustered structure by hand.

Matched inhibition is best understood as a co-learned part of the same process rather than as a static background correction. Inhibitory plasticity can stabilize firing rates and restore local E/I balance, but more recent work suggests that it can also shape structured assemblies. Vogels et al. (2011) established the canonical result that inhibitory plasticity can maintain balanced, asynchronous irregular activity in RNNs. Extending this idea, Mackwood et al. (2021) showed that stimulus-specific E/I assemblies require synergistic plasticity at both Pyr-to-PV and PV-to-Pyr synapses. This fits the experimental finding of Znamenskiy et al. (2024) that PV interneurons in mouse V1 are densely connected, yet the weights of reciprocal PV↔Pyr synapses are graded by response similarity. The inhibitory structure required by E/I-clustered networks is therefore most plausibly implemented as learned co-tuning in synaptic efficacies, preserving local balance while still allowing fluctuation-driven switching between metastable states.

It is also useful to distinguish functional from structural plasticity. Functional plasticity changes the strengths of existing synapses and may already be sufficient to move a circuit into an effectively clustered dynamical regime. Structural plasticity changes the graph itself through spine formation, elimination, and rewiring. Theoretical work shows that homeostatic structural plasticity can generate associative subnetworks from activity constraints alone (Gallinaro and Rotter, 2018; Gallinaro, Gašparović, et al., 2022), while experimental work shows that learning induces rapid spine formation and selective long-term stabilization of a subset of newly formed contacts (Xu et al., 2009; G. Yang et al., 2009). In the language of the wiring–weight continuum, functional plasticity can sharpen coupling within an existing scaffold, whereas structural plasticity can progressively reshape the scaffold itself. Furthermore, these processes can work together and restructure an initially functionally defined cluster into a more structurally defined one. On the wiring–weight continuum of the second study, this would correspond to a gradual shift from large κ toward smaller κ as learning consolidates initially weight-based clustering into structural clustering.

Finally, development and evolution likely specify biases and constraints rather than exact mature cluster membership. Preferential synapse formation among sister excitatory neurons, early lineage-dependent electrical coupling, and similar feature tuning among clonally related neurons all suggest that cortical circuits are not assembled from an initially uniform substrate, but from developmentally biased microcircuits that are later refined by experience (Y.-C. Yu, He, et al., 2012; Y.-C. Yu, Bultje, et al., 2009). The MICrONS reconstruction captures only a single snapshot of this ongoing developmental and experience-dependent process. Consistent with this view, recurrent connectivity becomes more feature-specific during postnatal visual experience (Ko et al., 2013). For assemblies on the scale of only a few hundred neurons, the most plausible picture is therefore not neuron-by-neuron genetic encoding, but the interaction of developmental priors with activity-dependent refinement of local motifs and common-neighbor structure (Perin et al., 2011). The E/I-clustered regime studied here is therefore best interpreted as a learned and developmentally biased operating point rather than a literal hard-wired anatomical partition, one that can shift through activity-dependent functional and structural remodeling (Hofer et al., 2009; Wilbrecht et al., 2010). However, even once established, this operating point must still be regulated online on the timescale of behavior.

3.3.6 Dynamic regulation of the metastable regime

An organism behaves on much shorter timescales than those of learning or development, and cortical dynamics are modulated on timescales from hundreds of milliseconds to minutes (McGinley

et al., 2015; Harris and Thiele, 2011) in response to arousal, attention, task difficulty, or neuromodulatory state. Metastable dynamics need to be tuned to the environmental and task-related timescales to be useful for computation. Discrimination is optimal when evidence is integrated on a timescale matched to the volatility of the generative process: in changing environments, ideal observers discount older observations at a rate set by the hazard of state changes, so overly short windows discard still-relevant evidence whereas overly long windows retain stale evidence after a change (Glaze et al., 2015; Radillo et al., 2017). Consistent with this principle, cortical population dynamics exhibit hallmarks of adaptive evidence accumulation in changing environments (Murphy et al., 2021), implying that the metastable operating point cannot be fully specified by fixed connectivity alone. Potential fast regulatory mechanisms, including gain modulation and cell-intrinsic adaptation, are therefore best treated as future extensions and are taken up in the Outlook.

In light of these considerations, the E/I-clustered network provides a sufficient mechanism for the joint neural–behavioral phenomenology observed in the third study. However, it does not establish necessity, causal identity, or control. The model constrains the space of explanations and generates falsifiable predictions, particularly regarding the relationship between structural implementation and pairwise correlation structure. Its inferential reach depends on whether the assumed connectivity is biologically plausible and whether the dynamics can be actively regulated.

3.4 Future directions

Posterior calibration and richer data constraints. Applying SBI to the third study's dataset would reveal parameter identifiability more clearly. Using the three objective groups introduced above, namely regime-level statistics, variability and covariation statistics, and structural priors, it could show which parameters are individually identifiable, which are constrained only in combination, and where the posterior remains broad. Two additions would sharpen these constraints. Simultaneous within-session population recordings would eliminate the need for pseudopopulation construction and provide access to single-trial co-variability. In addition, LFP would add a mesoscale observable sensitive to collective synaptic coordination, which pairwise spike statistics alone leave underdetermined.

Causal perturbation of the attractor landscape. If the local circuit implements attractor-based competition, transiently biasing one population during preparation should produce direction-specific behavioral effects. By contrast, a circuit that mainly relays an upstream decision should show at most transient disruption without systematic directional bias. A complementary prediction concerns basin structure: weak perturbations should produce relaxation on the timescale predicted by the mean-field analysis of the second study, whereas stronger perturbations should trigger state switches. Both predictions have precedent in mouse ALM (Inagaki, Fontolan, et al., 2019; Li, Daie, et al., 2016; Daie et al., 2021). Extending them to primate motor cortex and more than two alternatives remains challenging given the graded connectivity discussed above and the well-documented indirect effects of optogenetic perturbations (Li, Chen, et al., 2019; Inagaki, Chen, et al., 2022). Here, these predictions mainly serve to clarify which empirical observations would support a local attractor-based account.

Online regulation through spike frequency adaptation. Spike frequency adaptation (SFA) is a plausible fast regulatory mechanism. It is intrinsic to cortical neurons, operates on timescales from tens of milliseconds to seconds (Pozzorini et al., 2015), and can be modulated in excitatory and inhibitory populations by Acetylcholine (ACh) (McCormick et al., 1993; Colangelo et al., 2019). Preliminary simulations suggest that shifting adaptation between excitatory and inhibitory populations provides a fast online control knob for clustered dynamics (Figure 3.1). Excitatory-dominated adaptation flattens the effective landscape and facilitates state switches (Figure 3.1a,c1), whereas inhibitory-dominated adaptation stabilizes the currently active attractor and prolongs dwell times (Figure 3.1a,c2). The corresponding shifts in the psychometric curves in panel **b** suggest that discrimination performance may depend non-monotonically on adaptation strength. This parallels the inverted-U reported by Papadopoulos et al. (2026) for arousal-dependent modulation of clustering contrast. In an ISN regime, however, the two control channels are not independent, because both act through effective connectivity (Wyrick and Mazzucato, 2021). Testing this prediction would

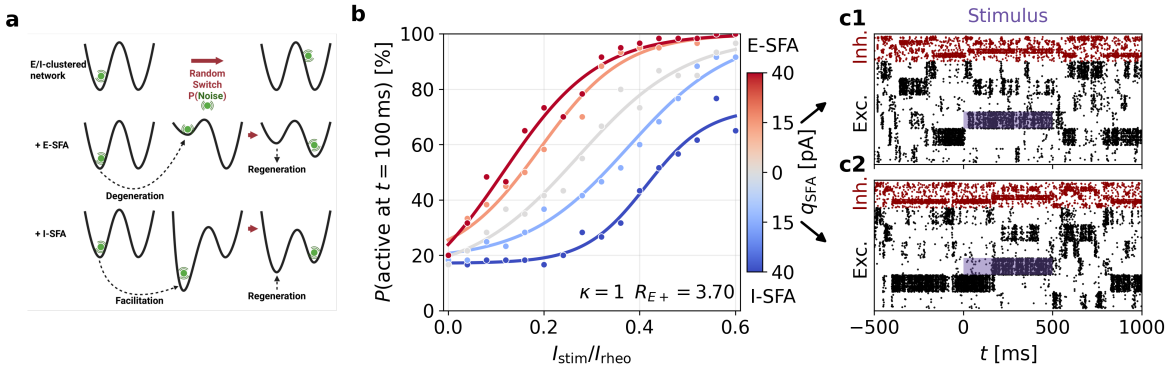

FIGURE 3.1: SFA provides a fast control channel for stimulus-driven switching in a clustered E/I network. Although excitatory and inhibitory adaptation can in principle be co-modulated, only pure E-SFA and pure I-SFA manipulations are shown here to isolate their opposing effects. Simulations used a 6-cluster network ($N_E = 1200$, $N_I = 300$) with DC background drive $E : 0.95\, I_{\text{rheo}}$ and $I : 1.10\, I_{\text{rheo}}$, where I_{rheo} denotes the population-specific rheobase current. In panel **b**, each curve is based on a single network realization and 60 trials at each of 16 stimulus strengths. Target clusters were pseudorandomized such that each of the six clusters was stimulated equally often and the resulting trial list was globally shuffled. Panels **c1,c2** show representative trials selected from separate 18-trial searches at $0.36\, I_{\text{rheo}}$. **(a)** Schematic: excitatory-dominated adaptation flattens the effective landscape and facilitates switching, whereas inhibitory-dominated adaptation stabilizes the current attractor. **(b)** Probability that the stimulated cluster is active 100 ms after stimulus onset as a function of stimulus strength. **(c1,c2)** Representative rasters for the arrow-marked small-adaptation conditions in **(b)**: small E-SFA (**c1**) and small I-SFA (**c2**), both evoked by a 500 ms stimulus of $0.36\, I_{\text{rheo}}$; in each case, the adapted population has $q_{\text{SFA}} = 15$ pA and $\tau_{\text{SFA}} = 150$ ms.

require access to the neuromodulatory state during task performance. Datasets combining neural recordings with direct imaging of cholinergic and noradrenergic axons (Reimer et al., 2016) or with pupil-derived arousal embeddings (Raut et al., 2025) would allow the inverted-U to be tested against condition-matched variability and discrimination performance, as demonstrated for auditory cortex by Papadopoulos et al. (2026). Whether an analogous relationship holds in motor cortex during action selection remains open.

Transfer to neuromorphic hardware. The BrainScaleS-2 accelerated neuromorphic system (Pehle et al., 2022) offers a platform to test whether metastable dynamics survive imprecisions qualitatively different from the κ-axis: limited weight resolution, device mismatch, and analog noise. As a first step toward future on-chip plasticity experiments, the E/I-clustered network was ported to BrainScaleS-2 at three points on the wiring–weight continuum ($\kappa = 0, 0.5$, and 1) using a shared base parameterization. All three configurations produced metastable assembly dynamics on hardware (Figure 3.2). This shows that the continuum identified in the second study is not just an artifact of idealized software simulation, but persists under device mismatch and limited weight resolution. Fixed indegree was necessary to maintain the tight E/I balance required for fluctuation-driven firing on the analog substrate. Silent neurons, arising from quenched variability across analog circuits, were more prevalent than in simulation, particularly toward the weight-clustered end of the continuum. These substrate constraints currently preclude the systematic robustness analyses performed in the first and second studies. Even so, successful emulation across κ provides the basis for the plasticity experiments outlined below. The system's 1000× acceleration opens a distinct scientific opportunity: one hour of emulation corresponds to roughly 40 days of biological time, placing long-term plasticity dynamics within practical reach. The on-chip plasticity processor could implement co-dependent E/I plasticity rules (Agnes and Vogels, 2024) to investigate whether such plasticity sustains stable cluster structure under ongoing metastable dynamics or instead drives reorganization or collapse. Realizing this program requires substantial implementation effort beyond the scope of the present work. Still, the proof-of-concept emulation across the κ continuum establishes the key prerequisite: a substrate on which metastable attractor dynamics are already present and can serve as the backdrop for plasticity. If successful, this line of investigation would bear directly on representational drift,

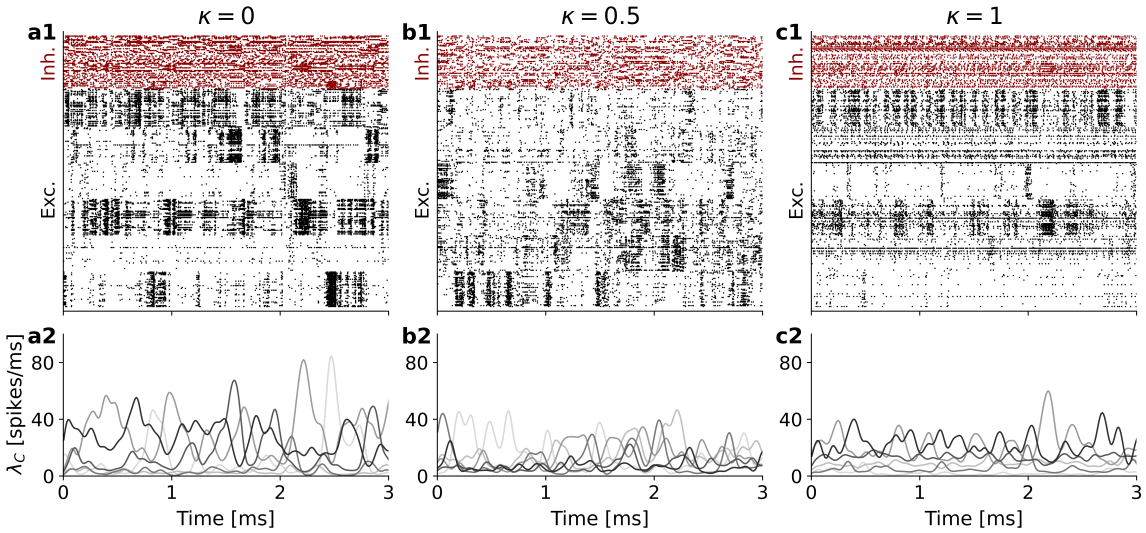

FIGURE 3.2: Neuromorphic reproduction of metastable switching across clustering mixtures on BrainScaleS-2, mirroring the software leaky integrate-and-fire (LIF) result (cf. Figure 5 in the second study) on accelerated analog hardware. All hardware realizations used six clusters with N_E=240, N_I=60, baseline connection probability p=0.15 with fixed indegree, R_{E+}=4.5, and R_j=0.65. **(a)** Structural clustering ($\kappa = 0$), **(b)** mixed clustering ($\kappa = 0.5$), and **(c)** weight clustering ($\kappa = 1$). Top panels **(a1–c1)** show raster plots (excitatory: black, inhibitory: red) spanning 3 ms wall-clock time (3 s biological equivalent at 1000× acceleration). Bottom panels **(a2–c2)** show per-cluster excitatory firing rates λ_C, estimated by Gaussian kernel convolution of spike trains ($\sigma = 50\,\mu$s, truncated at $\pm 3\sigma$), the hardware-time analog of the $\sigma = 50$ ms kernel used in the software simulations. All three regimes exhibit metastable switching consistent with the software simulations, confirming that the wiring–weight continuum translates to a physical substrate. Silent neurons from quenched analog variability are more prevalent toward the weight-clustered end. Addressing these substrate constraints and exploiting the system's on-chip plasticity processor for co-dependent E/I learning rules constitute the natural next steps enabled by this proof of concept.

testing the hypothesis that the relevant stability resides at the attractor-landscape level rather than at individual synapses (Deitch et al., 2021; Driscoll et al., 2022).

The E/I-clustered network studied here is not a complete account of cortical decision dynamics, but the tools, robustness results, and data constraints developed across the three studies provide a concrete foundation for pursuing these questions further.

SI.1 Efficient parameter calibration and real-time simulation of large-scale spiking neural networks with GeNN and NEST

Authors: Felix J. Schmitt, Vahid Rostami, and Martin P. Nawrot

Source: Frontiers in Neuroinformatics (2023), DOI: 10.3389/fninf.2023.941696

SI.2 Metastable Neural Assemblies on a Wiring–Weight Continuum

Authors: Felix J. Schmitt, Franziska L. Müller, and Martin P. Nawrot

Source: bioRxiv preprint / manuscript under review (2026), DOI: 10.64898/2026.03.16.712138

SI.3 Spiking attractor model of motor cortex explains modulation of neural and behavioral variability by prior target information

Authors: Vahid Rostami, Thomas Rost, Felix J. Schmitt, Sacha J. van Albada, Alexa Riehle, and Martin P. Nawrot

Source: Nature Communications (2024), DOI: 10.1038/s41467-024-49889-4

Bibliography

Abeles, M., Bergman, H., Gat, I., Meilijson, I., Seidemann, E., Tishby, N., and Vaadia, E. (1995). Cortical activity flips among quasi-stationary states. *Proceedings of the National Academy of Sciences* 92.19, 8616–8620.

Agnes, E. J. and Vogels, T. P. (2024). Co-dependent excitatory and inhibitory plasticity accounts for quick, stable and long-lasting memories in biological networks. *Nature Neuroscience* 27.5, 964–974.

Ahmadian, Y. and Miller, K. D. (2021). What is the dynamical regime of cerebral cortex? *Neuron* 109.21, 3373–3391.

Albada, S. J. v., Helias, M., and Diesmann, M. (2015). Scalability of Asynchronous Networks Is Limited by One-to-One Mapping between Effective Connectivity and Correlations. *PLOS Computational Biology* 11.9, e1004490.

Albantakis, L., Bernard, C., Brenner, N., Marder, E., and Narayanan, R. (2024). The Brain's Best Kept Secret Is Its Degenerate Structure. *Journal of Neuroscience* 44.40, e1339242024.

Ames, K. C., Ryu, S. I., and Shenoy, K. V. (2014). Neural Dynamics of Reaching following Incorrect or Absent Motor Preparation. *Neuron* 81.2, 438–451.

Ames, K. C., Ryu, S. I., and Shenoy, K. V. (2019). Simultaneous motor preparation and execution in a last-moment reach correction task. *Nature Communications* 10.1, 2718.

Amit, D. J. and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex* 7.3, 237–252.

Amit, D. J., Gutfreund, H., and Sompolinsky, H. (1985). Spin-glass models of neural networks. *Physical Review A* 32.2, 1007–1018.

Bachschmid-Romano, L., Hatsopoulos, N. G., and Brunel, N. (2023). Interplay between external inputs and recurrent dynamics during movement preparation and execution in a network model of motor cortex. *eLife* 12, e77690.

Bae, J. A., Baptiste, M., Baptiste, M. R., Bishop, C. A., Bodor, A. L., Brittain, D., Brooks, V., Buchanan, J., Bumbarger, D. J., Castro, M. A., Celii, B., Cobos, E., Collman, F., Costa, N. M. da, Danskin, B., Dorkenwald, S., Elabbady, L., Fahey, P. G., Fliss, T., Froudarakis, E., Gager, J., Gamlin, C., Gray-Roncal, W., Halageri, A., Hebditch, J., Jia, Z., Joyce, E., Ellis-Joyce, J., Jordan, C., Kapner, D., Kemnitz, N., Kinn, S., Kitchell, L. M., Koolman, S., Kuehner, K., Lee, K., Li, K., Lu, R., Macrina, T., Mahalingam, G., Matelsky, J., McReynolds, S., Miranda, E., Mitchell, E., Mondal, S. S., Moore, M., Mu, S., Muhammad, T., Nehoran, B., Neace, E., Ogedengbe, O., Papadopoulos, C., Papadopoulos, S., Patel, S., Vega, G. J. Y. P., Pitkow, X., Popovych, S., Ramos, A., Reid, R. C., Reimer, J., Rivlin, P. K., Rose, V., Sauter, Z. M., Schneider-Mizell, C. M., Seung, H. S., Silverman, B., Silversmith, W., Sterling, A., Sinz, F. H., Smith, C. L., Swanstrom, R., Suckow, S., Takeno, M., Tan, Z. H., Tolias, A. S., Torres, R., Turner, N. L., Walker, E. Y., Wang, T., Wanner, A., Wester, B. A., Williams, G., Williams, S., Willie, K., Willie, R., Wong, W., Wu, J., Xu, C., Yang, R., Yatsenko, D., Ye, F., Yin, W., Young, R., Yu, S.-c., Xenes, D., Zhang, C., and The MICrONS Consortium (2025). Functional connectomics spanning multiple areas of mouse visual cortex. *Nature* 640.8058, 435–447.

Barlow, H. B. (1972). Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology? *Perception* 1.4, 371–394.

Beiran, M., Meirhaeghe, N., Sohn, H., Jazayeri, M., and Ostojic, S. (2023). Parametric control of flexible timing through low-dimensional neural manifolds. *Neuron* 111.5, 739–753.e8.

Ben-Shaul, Y., Stark, E., Asher, I., Drori, R., Nadasdy, Z., and Abeles, M. (2003). Dynamical Organization of Directional Tuning in the Primate Premotor and Primary Motor Cortex. *Journal of Neurophysiology* 89.2, 1136–1142.

Best, M. D., Suminski, A. J., Takahashi, K., Brown, K. A., and Hatsopoulos, N. G. (2017). Spatio-Temporal Patterning in Primary Motor Cortex at Movement Onset. *Cerebral Cortex* 27.2, bhv327.

Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* 2.1, 32–48.

Boelts, J., Deistler, M., Gloeckler, M., Tejero-Cantero, Á., Lueckmann, J.-M., Moss, G., Steinbach, P., Moreau, T., Muratore, F., Linhart, J., Durkan, C., Vetter, J., Miller, B. K., Herold, M., Ziaeemehr, A., Pals, M., Gruner, T., Bischoff, S., Krouglova, N., Gao, R., Lappalainen, J. K., Mucsányi, B., Pei, F., Schulz, A., Stefanidi, Z., Rodrigues, P., Schröder, C., Zaid, F. A., Beck, J., Kapoor, J., Greenberg, D. S., Gonçalves, P. J., and Macke, J. H. (2025). sbi reloaded: a toolkit for simulation-based inference workflows. *Journal of Open Source Software* 10.108, 7754.

Bondanelli, G. and Ostojic, S. (2020). Coding with transient trajectories in recurrent neural networks. *PLOS Computational Biology* 16.2, e1007655.

Brinkman, B. A. W., Yan, H., Maffei, A., Park, I. M., Fontanini, A., Wang, J., and La Camera, G. (2022). Metastable dynamics of neural circuits and networks. *Applied Physics Reviews* 9.1, 011313.

Brunel, N. (2000). Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of Computational Neuroscience* 8.3, 183–208.

Chance, F. S., Abbott, L. F., and Reyes, A. D. (2002). Gain Modulation from Background Synaptic Input. *Neuron* 35.4, 773–782.

Chen, S., Liu, Y., Wang, Z. A., Colonell, J., Liu, L. D., Hou, H., Tien, N.-W., Wang, T., Harris, T., Druckmann, S., Li, N., and Svoboda, K. (2024). Brain-wide neural activity underlying memory-guided movement. *Cell* 187.3, 676–691.e16.

Christodoulou, G., Vogels, T. P., and Agnes, E. J. (2022). Regimes and mechanisms of transient amplification in abstract and biological neural networks. *PLOS Computational Biology* 18.8, e1010365.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., and Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature* 487.7405, 51–56.

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Santhanam, G., Sahani, M., and Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience* 13.3, 369–378.

Colangelo, C., Shichkova, P., Keller, D., Markram, H., and Ramaswamy, S. (2019). Cellular, Synaptic and Network Effects of Acetylcholine in the Neocortex. *Frontiers in Neural Circuits* 13, 24.

Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, R., Sader, E. N., Ko, H., Hofer, S. B., and Mrsic-Flogel, T. D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* 518.7539, 399–403.

Cranmer, K., Brehmer, J., and Louppe, G. (2020). The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences* 117.48, 30055–30062.

Cunningham, J. P. and Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience* 17.11, 1500–1509.

Daie, K., Svoboda, K., and Druckmann, S. (2021). Targeted photostimulation uncovers circuit motifs supporting short-term memory. *Nature Neuroscience* 24.2, 259–265.

Deco, G. and Hugues, E. (2012). Neural Network Mechanisms Underlying Stimulus Driven Variability Reduction. *PLOS Computational Biology* 8.3, 1–10.

Deistler, M., Macke, J. H., and Gonçalves, P. J. (2022). Energy-efficient network activity from disparate circuit parameters. *Proceedings of the National Academy of Sciences* 119.44, e2207632119.

Deitch, D., Rubin, A., and Ziv, Y. (2021). Representational drift in the mouse visual cortex. *Current Biology* 31.19, 4327–4339.e6.

Denève, S. and Machens, C. K. (2016). Efficient codes and balanced networks. *Nature Neuroscience* 19.3, 375–382.

Ding, Z., Tran, D., Ponder, K., Ding, Z., Froebe, R., Ntanavara, L., Fahey, P. G., Cobos, E., Baroni, L., Diamantaki, M., Wang, E. Y., Chang, A., Papadopoulos, S., Fu, J., Muhammad, T., Papadopoulos, C., Cadena, S. A., Evangelou, A., Willeke, K., Anselmi, F., Sanborn, S., Antolik, J., Froudarakis, E., Patel, S., Walker, E. Y., Reimer, J., Sinz, F. H., Ecker, A. S., Franke, K., Pitkow, X., and Tolias, A. S. (2026). Functional bipartite invariance in mouse primary visual cortex receptive fields. *Nature Neuroscience*.

Ding, Z., Fahey, P. G., Papadopoulos, S., Wang, E. Y., Celii, B., Papadopoulos, C., Chang, A., Kunin, A. B., Tran, D., Fu, J., Ding, Z., Patel, S., Ntanavara, L., Froebe, R., Ponder, K., Muhammad, T., Bae, J. A., Bodor, A. L., Brittain, D., Buchanan, J., Bumbarger, D. J., Castro, M. A., Cobos, E., Dorkenwald, S., Elabbady, L., Halageri, A., Jia, Z., Jordan, C., Kapner, D., Kemnitz, N., Kinn, S., Lee, K., Li, K., Lu, R., Macrina, T., Mahalingam, G., Mitchell, E., Mondal, S. S., Mu, S., Nehoran, B., Popovych, S., Schneider-Mizell, C. M., Silversmith, W., Takeno, M., Torres, R., Turner, N. L., Wong, W., Wu, J., Yin, W., Yu, S.-c., Yatsenko, D., Froudarakis, E., Sinz, F., Josić, K., Rosenbaum, R., Seung, H. S., Collman, F., Costa, N. M. da, Reid, R. C., Walker, E. Y., Pitkow, X., Reimer, J., and Tolias, A. S. (2025). Functional connectomics reveals general wiring rule in mouse visual cortex. *Nature* 640.8058, 459–469.

Douglas, R. J., Martin, K. A., and Whitteridge, D. (1989). A Canonical Microcircuit for Neocortex. *Neural Computation* 1.4, 480–488.

Driscoll, L. N., Duncker, L., and Harvey, C. D. (2022). Representational drift: Emerging theories for continual learning and experimental future directions. *Current Opinion in Neurobiology* 76, 102609.

Dubreuil, A., Valente, A., Beiran, M., Mastrogiuseppe, F., and Ostojic, S. (2022). The role of population structure in computations through neural dynamics. *Nature Neuroscience* 25.6, 783–794.

Edelman, G. M. and Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences* 98.24, 13763–13768.

Einevoll, G. T., Kayser, C., Logothetis, N. K., and Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience* 14.11, 770–785.

Ekelmans, P., Kraynyukova, N., and Tchumatchenko, T. (2023). Targeting operational regimes of interest in recurrent neural networks. *PLOS Computational Biology* 19.5, e1011097.

Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M., and Cunningham, J. P. (2016). Reorganization between preparatory and movement population responses in motor cortex. *Nature Communications* 7.1, 13239.

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience* 9.4, 292–303.

Finkelstein, A., Fontolan, L., Economo, M. N., Li, N., Romani, S., and Svoboda, K. (2021). Attractor dynamics gate cortical information flow during decision-making. *Nature Neuroscience* 24.6, 843–850.

Fino, E. and Yuste, R. (2011). Dense Inhibitory Connectivity in Neocortex. *Neuron* 69.6, 1188–1203.

Gallinaro, J. V., Gašparović, N., and Rotter, S. (2022). Homeostatic control of synaptic rewiring in recurrent networks induces the formation of stable memory engrams. *PLOS Computational Biology* 18.2, e1009836.

Gallinaro, J. V. and Rotter, S. (2018). Associative properties of structural plasticity based on firing rate homeostasis in recurrent neuronal networks. *Scientific Reports* 8.1, 3754.

Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal Population Coding of Movement Direction. *Science* 233.4771, 1416–1419.

Glaze, C. M., Kable, J. W., and Gold, J. I. (2015). Normative evidence accumulation in unpredictable environments. *eLife* 4, e08825.

Goaillard, J.-M. and Marder, E. (2021). Ion Channel Degeneracy, Variability, and Covariation in Neuron and Circuit Resilience. *Annual Review of Neuroscience* 44, 335–357.

Gonçalves, P. J., Lueckmann, J.-M., Deistler, M., Nonnenmacher, M., Öcal, K., Bassetto, G., Chintaluri, C., Podlaski, W. F., Haddad, S. A., Vogels, T. P., Greenberg, D. S., and Macke, J. H. (2020). Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife* 9, e56261.

Hagen, E., Dahmen, D., Stavrinou, M. L., Lindén, H., Tetzlaff, T., Albada, S. J. van, Grün, S., Diesmann, M., and Einevoll, G. T. (2016). Hybrid Scheme for Modeling Local Field Potentials from Point-Neuron Networks. *Cerebral Cortex* 26.12, 4461–4496.

Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature Reviews Neuroscience* 6.5, 399–407.

Harris, K. D. and Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nature Neuroscience* 18.2, 170–181.

Harris, K. D. and Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience* 12.9, 509–523.

Hebb, D. O. (1949). The Organization of Behavior: A Neuropsychological Theory. New York: John Wiley & Sons.

Hennequin, G., Ahmadian, Y., Rubin, D. B., Lengyel, M., and Miller, K. D. (2018). The Dynamical Regime of Sensory Cortex: Stable Dynamics around a Single Stimulus-Tuned Attractor Account for Patterns of Noise Variability. *Neuron* 98.4, 846–860.e5.

Hennequin, G., Vogels, T. P., and Gerstner, W. (2014). Optimal Control of Transient Dynamics in Balanced Networks Supports Generation of Complex Movements. *Neuron* 82.6, 1394–1406.

Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T., and Hübener, M. (2009). Experience leaves a lasting structural trace in cortical circuits. *Nature* 457.7227, 313–317.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* 79.8, 2554–2558.

Hopfield, J. J. and Tank, D. W. (1985). "Neural" computation of decisions in optimization problems. *Biological Cybernetics* 52.3, 141–152.

Horton, J. C. and Adams, D. L. (2005). The cortical column: a structure without a function. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360.1456, 837–862.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology* 160.1, 106–154.

Inagaki, H. K., Chen, S., Daie, K., Finkelstein, A., Fontolan, L., Romani, S., and Svoboda, K. (2022). Neural Algorithms and Circuits for Motor Planning. *Annual Review of Neuroscience* 45, 249–271.

Inagaki, H. K., Fontolan, L., Romani, S., and Svoboda, K. (2019). Discrete attractor dynamics underlies persistent activity in the frontal cortex. *Nature* 566.7743, 212–217.

Ito, S., Haufler, D., Fraile, J. G., Dai, K., Aman, J., Chen, G., Mirasso, C., Maass, W., and Arkhipov, A. (2026). Deep-learning-assisted simulation of a cortical circuit: integrating anatomy, physiology and function. *bioRxiv*, 2026.03.13.711751.

Jaeger, H. and Haas, H. (2004). Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. *Science* 304.5667, 78–80.

Jones, L. M., Fontanini, A., Sadacca, B. F., Miller, P., and Katz, D. B. (2007). Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proceedings of the National Academy of Sciences* 104.47, 18772–18777.

Jürgensen, A.-M., Schmitt, F. J., and Nawrot, M. P. (2024). Minimal circuit motifs for second-order conditioning in the insect mushroom body. *Frontiers in Physiology* 14, 1326307.

Kao, T.-C., Sadabadi, M. S., and Hennequin, G. (2021). Optimal anticipatory control as a theory of motor preparation: A thalamo-cortical circuit model. *Neuron* 109.9, 1567–1581.e12.

Karnani, M. M., Agetsuma, M., and Yuste, R. (2014). A blanket of inhibition: functional inferences from dense inhibitory connectivity. *Current Opinion in Neurobiology* 26, 96–102.

Kaufman, M. T., Churchland, M. M., Ryu, S. I., and Shenoy, K. V. (2014). Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience* 17.3, 440–448.

Kelso, J. A. S. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367.1591, 906–918.

Kepecs, A. and Fishell, G. (2014). Interneuron cell types are fit to function. *Nature* 505.7483, 318–326.

Khalil, H. K. (2002). Nonlinear systems. 3rd ed. Upper Saddle River, N.J: Prentice Hall. ISBN: 978-0-13-067389-3.

Khona, M. and Fiete, I. R. (2022). Attractor and integrator networks in the brain. *Nature Reviews Neuroscience* 23.12, 744–766.

Kim, C. M., Finkelstein, A., Chow, C. C., Svoboda, K., and Darshan, R. (2023). Distributing task-related neural activity across a cortical network through task-independent connections. *Nature Communications* 14.1, 2851.

Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S. B., and Mrsic-Flogel, T. D. (2013). The emergence of functional microcircuits in visual cortex. *Nature* 496.7443, 96–100.

Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7, 284–304.

Kuzmina, E., Kriukov, D., and Lebedev, M. (2024). Neuronal travelling waves explain rotational dynamics in experimental datasets and modelling. *Scientific Reports* 14.1, 3566.

La Camera, G., Fontanini, A., and Mazzucato, L. (2019). Cortical computations via metastable activity. *Current Opinion in Neurobiology* 58, 37–45.

Langdon, C., Genkin, M., and Engel, T. A. (2023). A unifying perspective on neural manifolds and circuits for cognition. *Nature Reviews Neuroscience* 24.6, 363–377.

Lara, A. H., Cunningham, J. P., and Churchland, M. M. (2018). Different population dynamics in the supplementary motor area and motor cortex during reaching. *Nature Communications* 9.1, 2754.

Levenstein, D., Alvarez, V. A., Amarasingham, A., Azab, H., Chen, Z. S., Gerkin, R. C., Hasenstaub, A., Iyer, R., Jolivet, R. B., Marzen, S., Monaco, J. D., Prinz, A. A., Quraishi, S., Santamaria, F., Shivkumar, S., Singh, M. F., Traub, R., Nadim, F., Rotstein, H. G., and Redish, A. D. (2023). On the Role of Theory and Modeling in Neuroscience. *Journal of Neuroscience* 43.7, 1074–1088.

Li, N., Chen, S., Guo, Z. V., Chen, H., Huo, Y., Inagaki, H. K., Chen, G., Davis, C., Hansel, D., Guo, C., and Svoboda, K. (2019). Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *eLife* 8, e48622.

Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. *Nature* 532.7600, 459–464.

Litwin-Kumar, A. and Doiron, B. (2012). Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neuroscience* 15.11, 1498–1505.

Litwin-Kumar, A. and Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature Communications* 5.1, 5319.

Maass, W., Natschläger, T., and Markram, H. (2002). Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations. *Neural Computation* 14.11, 2531–2560.

Mackwood, O., Naumann, L. B., and Sprekeler, H. (2021). Learning excitatory-inhibitory neuronal assemblies in recurrent networks. *eLife* 10, e59715.

Mante, V., Sussillo, D., Shenoy, K. V., and Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503.7474, 78–84.

Manz, P. and Memmesheimer, R.-M. (2023). Purely STDP-based assembly dynamics: Stability, learning, overlaps, drift and aging. *PLOS Computational Biology* 19.4, e1011006.

Marder, E. and Calabrese, R. L. (1996). Principles of rhythmic motor pattern generation. *Physiological Reviews* 76.3, 687–717.

Marder, E. and Taylor, A. L. (2011). Multiple models to capture the variability in biological neurons and networks. *Nature Neuroscience* 14.2, 133–138.

Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W. H. Freeman. ISBN: 0-7167-1284-9.

Mastrogiuseppe, F. and Ostojic, S. (2018). Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. *Neuron* 99.3, 609–623.e29.

Mazzoni, A., Lindén, H., Cuntz, H., Lansner, A., Panzeri, S., and Einevoll, G. T. (2015). Computing the Local Field Potential (LFP) from Integrate-and-Fire Network Models. *PLOS Computational Biology* 11.12, e1004584.

Mazzucato, L., Fontanini, A., and Camera, G. L. (2015). Dynamics of Multistable States during Ongoing and Evoked Cortical Activity. *Journal of Neuroscience* 35.21, 8214–8231.

McCormick, D. A., Wang, Z., and Huguenard, J. (1993). Neurotransmitter Control of Neocortical Neuronal Activity and Excitability. *Cerebral Cortex* 3.5, 387–398.

McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C. R., Tolias, A. S., Cardin, J. A., and McCormick, D. A. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron* 87.6, 1143–1161.

Miehl, C., Onasch, S., Festa, D., and Gjorgjieva, J. (2023). Formation and computational implications of assemblies in neural circuits. *The Journal of Physiology* 601.15, 3071–3090.

Morishima, M., Kobayashi, K., Kato, S., Kobayashi, K., and Kawaguchi, Y. (2017). Segregated Excitatory–Inhibitory Recurrent Subnetworks in Layer 5 of the Rat Frontal Cortex. *Cerebral Cortex* 27.12, 5846–5857.

Morrison, M. and Young, L.-S. (2022). Chaotic heteroclinic networks as models of switching behavior in biological systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 32.12, 123102.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain* 120.4, 701–722.

Muir, D. R., Molina-Luna, P., Roth, M. M., Helmchen, F., and Kampa, B. M. (2017). Specific excitatory connectivity for feature integration in mouse primary visual cortex. *PLOS Computational Biology* 13.12, e1005888.

Murphy, P. R., Wilming, N., Hernandez-Bocanegra, D. C., Prat-Ortega, G., and Donner, T. H. (2021). Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments. *Nature Neuroscience* 24.7, 987–997.

Najafi, F., Elsayed, G. F., Cao, R., Pnevmatikakis, E., Latham, P. E., Cunningham, J. P., and Churchland, A. K. (2020). Excitatory and Inhibitory Subnetworks Are Equally Selective during Decision-Making and Emerge Simultaneously during Learning. *Neuron* 105.1, 165–179.e8.

Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods* 169.2, 374–390.

Okun, M. and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neuroscience* 11.5, 535–537.

Oldenburg, I. A., Hendricks, W. D., Handy, G., Shamardani, K., Bounds, H. A., Doiron, B., and Adesnik, H. (2024). The logic of recurrent circuits in the primary visual cortex. *Nature Neuroscience* 27.1, 137–147.

Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., and Ferster, D. (2009). Inhibitory Stabilization of the Cortical Network Underlies Visual Surround Suppression. *Neuron* 62.4, 578–592.

Packer, A. M. and Yuste, R. (2011). Dense, Unspecific Connectivity of Neocortical Parvalbumin-Positive Interneurons: A Canonical Microcircuit for Inhibition? *Journal of Neuroscience* 31.37, 13260–13271.

Papadopoulos, L., Jo, S., Zumwalt, K., Wehr, M., Jaramillo, S., McCormick, D. A., and Mazzucato, L. (2026). Modulation of metastable ensemble dynamics explains the inverted-U relationship between tone discriminability and arousal in auditory cortex. *Neuron* 114.4, 740–758.e19.

Pehle, C., Billaudelle, S., Cramer, B., Kaiser, J., Schreiber, K., Stradmann, Y., Weis, J., Leibfried, A., Müller, E., and Schemmel, J. (2022). The BrainScaleS-2 Accelerated Neuromorphic System With Hybrid Plasticity. *Frontiers in Neuroscience* 16, 795876.

Perich, M. G., Narain, D., and Gallego, J. A. (2025). A neural manifold view of the brain. *Nature Neuroscience* 28.8, 1582–1597.

Perin, R., Berger, T. K., and Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences* 108.13, 5419–5424.

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature Neuroscience* 16.8, 1068–1076.

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J. I., Huang, Z. J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. *Nature* 503.7477, 521–524.

Potjans, T. C. and Diesmann, M. (2014). The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. *Cerebral Cortex* 24.3, 785–806.

Pozzorini, C., Mensi, S., Hagens, O., Naud, R., Koch, C., and Gerstner, W. (2015). Automated High-Throughput Characterization of Single Neurons by Means of Simplified Spiking Models. *PLOS Computational Biology* 11.6, e1004275.

Prinz, A. A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. *Nature Neuroscience* 7.12, 1345–1352.

Pronold, J., Morales-Gregorio, A., Rostami, V., and Albada, S. J. v. (2024). Cortical multi-area model with joint excitatory-inhibitory clusters accounts for spiking statistics, inter-area propagation, and variability dynamics. *bioRxiv*, 2024.01.30.577979.

Rabinovich, M. I., Huerta, R., Varona, P., and Afraimovich, V. S. (2008). Transient Cognitive Dynamics, Metastability, and Decision Making. *PLOS Computational Biology* 4.5, e1000072.

Rabinovich, M. I. and Varona, P. (2011). Robust Transient Dynamics and Brain Functions. *Frontiers in Computational Neuroscience* 5, 24.

Radillo, A. E., Veliz-Cuba, A., Josić, K., and Kilpatrick, Z. P. (2017). Evidence Accumulation and Change Rate Inference in Dynamic Environments. *Neural Computation* 29.6, 1561–1610.

Ramsauer, H., Schäfl, B., Lehner, J., Seidl, P., Widrich, M., Adler, T., Gruber, L., Holzleitner, M., Pavlović, M., Sandve, G. K., Greiff, V., Kreil, D., Kopp, M., Klambauer, G., Brandstetter, J., and Hochreiter, S. (2021). Hopfield Networks is All You Need. *arXiv*, 2008.02217.

Raut, R. V., Rosenthal, Z. P., Wang, X., Miao, H., Zhang, Z., Lee, J.-M., Raichle, M. E., Bauer, A. Q., Brunton, S. L., Brunton, B. W., and Kutz, J. N. (2025). Arousal as a universal embedding for spatiotemporal brain dynamics. *Nature* 647.8089, 454–461.

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., and Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications* 7.1, 13289.

Renart, A. and Machens, C. K. (2014). Variability in neural activity and behavior. *Current Opinion in Neurobiology* 25, 211–220.

Rickert, J., Riehle, A., Aertsen, A., Rotter, S., and Nawrot, M. P. (2009). Dynamic Encoding of Movement Direction in Motor Cortical Neurons. *Journal of Neuroscience* 29.44, 13870–13882.

Riehle, A., Brochier, T., Nawrot, M., and Grün, S. (2018). Behavioral Context Determines Network State and Variability Dynamics in Monkey Motor Cortex. *Frontiers in Neural Circuits* 12, 52.

Riehle, A., Wirtssohn, S., Gruen, S., and Brochier, T. (2013). Mapping the spatio-temporal structure of motor cortical LFP and spiking activities during reach-to-grasp movements. *Frontiers in Neural Circuits* 7, 48.

Rosenbaum, R. and Doiron, B. (2014). Balanced Networks of Spiking Neurons with Spatially Dependent Recurrent Connections. *Physical Review X* 4.2, 021039.

Rost, T., Deger, M., and Nawrot, M. P. (2018). Winnerless competition in clustered balanced networks: inhibitory assemblies do the trick. *Biological Cybernetics* 112.1, 81–98.

Rubino, D., Robbins, K. A., and Hatsopoulos, N. G. (2006). Propagating waves mediate information transfer in the motor cortex. *Nature Neuroscience* 9.12, 1549–1557.

Sanzeni, A., Histed, M. H., and Brunel, N. (2020). Response nonlinearities in networks of spiking neurons. *PLOS Computational Biology* 16.9, e1008165.

Saxena, S. and Cunningham, J. P. (2019). Towards the neural population doctrine. *Current Opinion in Neurobiology* 55, 103–111.

Schaub, M. T., Billeh, Y. N., Anastassiou, C. A., Koch, C., and Barahona, M. (2015). Emergence of Slow-Switching Assemblies in Structured Neuronal Networks. *PLOS Computational Biology* 11.7, e1004196.

Schimel, M., Kao, T.-C., and Hennequin, G. (2024). When and why does motor preparation arise in recurrent neural network models of motor control? *eLife* 12, RP89131.

Schmidt, M., Bakker, R., Shen, K., Bezgin, G., Diesmann, M., and Albada, S. J. v. (2018). A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. *PLOS Computational Biology* 14.10, e1006359.

Schmitt, F. J., Rostami, V., and Nawrot, M. P. (2023). Efficient parameter calibration and real-time simulation of large-scale spiking neural networks with GeNN and NEST. *Frontiers in Neuroinformatics* 17, 941696.

Schneider-Mizell, C. M., Bodor, A. L., Brittain, D., Buchanan, J., Bumbarger, D. J., Elabbady, L., Gamlin, C., Kapner, D., Kinn, S., Mahalingam, G., Seshamani, S., Suckow, S., Takeno, M., Torres, R., Yin, W., Dorkenwald, S., Bae, J. A., Castro, M. A., Halageri, A., Jia, Z., Jordan, C., Kemnitz, N., Lee, K., Li, K., Lu, R., Macrina, T., Mitchell, E., Mondal, S. S., Mu, S., Nehoran, B., Popovych, S., Silversmith, W., Turner, N. L., Wong, W., Wu, J., Reimer, J., Tolias, A. S., Seung, H. S., Reid, R. C., Collman, F., and Costa, N. M. da (2025). Inhibitory specificity from a connectomic census of mouse visual cortex. *Nature* 640.8058, 448–458.

Senk, J., Hagen, E., Albada, S. J. van, and Diesmann, M. (2024). Reconciliation of weak pairwise spike–train correlations and highly coherent local field potentials across space. *Cerebral Cortex* 34.10, bhae405.

Shenoy, K. V., Sahani, M., and Churchland, M. M. (2013). Cortical Control of Arm Movements: A Dynamical Systems Perspective. *Annual Review of Neuroscience* 36, 337–359.

Softky, W. and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of Neuroscience* 13.1, 334–350.

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLOS Biology* 3.3, e68.

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., and Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature* 576.7786, 266–273.

Suresh, A. K., Goodman, J. M., Okorokova, E. V., Kaufman, M., Hatsopoulos, N. G., and Bensmaia, S. J. (2020). Neural population dynamics in motor cortex are different for reach and grasp. *eLife* 9, e58848.

Taube, J. S. (1998). Head direction cells and the neurophysiological basis for a sense of direction. *Progress in Neurobiology* 55.3, 225–256.

Tononi, G., Sporns, O., and Edelman, G. M. (1999). Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences* 96.6, 3257–3262.

Tsodyks, M. V., Skaggs, W. E., Sejnowski, T. J., and McNaughton, B. L. (1997). Paradoxical Effects of External Modulation of Inhibitory Interneurons. *Journal of Neuroscience* 17.11, 4382–4388.

Turner-Evans, D. B., Jensen, K. T., Ali, S., Paterson, T., Sheridan, A., Ray, R. P., Wolff, T., Lauritzen, J. S., Rubin, G. M., Bock, D. D., and Jayaraman, V. (2020). The Neuroanatomical Ultrastructure and Function of a Biological Ring Attractor. *Neuron* 108.1, 145–163.e10.

Vargas-Irwin, C. E., Shakhnarovich, G., Yadollahpour, P., Mislow, J. M. K., Black, M. J., and Donoghue, J. P. (2010). Decoding Complete Reach and Grasp Actions from Local Primary Motor Cortex Populations. *Journal of Neuroscience* 30.29, 9659–9669.

Vegué, M., Perin, R., and Roxin, A. (2017). On the Structure of Cortical Microcircuits Inferred from Small Sample Sizes. *Journal of Neuroscience* 37.35, 8498–8510.

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., and Gerstner, W. (2011). Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. *Science* 334.6062, 1569–1573.

Vreeswijk, C. van and Sompolinsky, H. (1996). Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science* 274.5293, 1724–1726.

Vyas, S., Golub, M. D., Sussillo, D., and Shenoy, K. V. (2020). Computation Through Neural Population Dynamics. *Annual Review of Neuroscience* 43, 249–275.

Wang, E. Y., Fahey, P. G., Ding, Z., Papadopoulos, S., Ponder, K., Weis, M. A., Chang, A., Muhammad, T., Patel, S., Ding, Z., Tran, D., Fu, J., Schneider-Mizell, C. M., Costa, N. M. da, Reid, R. C., Collman, F., Costa, N. M. da, Franke, K., Ecker, A. S., Reimer, J., Pitkow, X., Sinz, F. H., Tolias, A. S., and MICrONS Consortium (2025). Foundation model of neural activity predicts response to new stimulus types. *Nature* 640.8058, 470–477.

Wang, X.-J. (2002). Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. *Neuron* 36.5, 955–968.

Wang, X.-J. (2008). Decision Making in Recurrent Neuronal Circuits. *Neuron* 60.2, 215–234.

Wehr, M. and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426.6965, 442–446.

Wilbrecht, L., Holtmaat, A., Wright, N., Fox, K., and Svoboda, K. (2010). Structural Plasticity Underlies Experience-Dependent Functional Plasticity of Cortical Circuits. *Journal of Neuroscience* 30.14, 4927–4932.

Wong, K.-F. and Wang, X.-J. (2006). A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. *Journal of Neuroscience* 26.4, 1314–1328.

Wyrick, D. and Mazzucato, L. (2021). State-Dependent Regulation of Cortical Processing Speed via Gain Modulation. *Journal of Neuroscience* 41.18, 3988–4005.

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., Jones, T., and Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* 462.7275, 915–919.

Yang, G., Pan, F., and Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature* 462.7275, 920–924.

Yang, X. and Camera, G. L. (2024). Co-existence of synaptic plasticity and metastable dynamics in a spiking model of cortical circuits. *PLOS Computational Biology* 20.7, e1012220.

Yu, Y.-C., Bultje, R. S., Wang, X., and Shi, S.-H. (2009). Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature* 458.7237, 501–504.

Yu, Y.-C., He, S., Chen, S., Fu, Y., Brown, K. N., Yao, X.-H., Ma, J., Gao, K. P., Sosinsky, G. E., Huang, K., and Shi, S.-H. (2012). Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. *Nature* 486.7401, 113–117.

Yuste, R., Cossart, R., and Yaksi, E. (2024). Neuronal ensembles: Building blocks of neural circuits. *Neuron* 112.6, 875–892.

Zenke, F., Agnes, E. J., and Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Communications* 6.1, 6922.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *Journal of Neuroscience* 16.6, 2112–2126.

Znamenskiy, P., Kim, M.-H., Muir, D. R., Iacaruso, M. F., Hofer, S. B., and Mrsic-Flogel, T. D. (2024). Functional specificity of recurrent inhibition in visual cortex. *Neuron* 112.6, 991–1000.e8.