

# Shaping dynamical neural computations using spatiotemporal constraints

Jason Z. Kim\*

*Department of Physics, Cornell University, Ithaca, NY 14853, USA*

Bart Larsen

*Department of Pediatrics, Masonic Institute for the Developing Brain, University of Minnesota*

Linden Parkes†

*Department of Psychiatry, Rutgers University, Piscataway, NJ 08854, USA*

(Dated: November 28, 2023)

Dynamics play a critical role in computation. The principled evolution of states over time enables both biological and artificial networks to represent and integrate information to make decisions. In the past few decades, significant multidisciplinary progress has been made in bridging the gap between how we understand biological *versus* artificial computation, including how insights gained from one can translate to the other. Research has revealed that neurobiology is a key determinant of brain network architecture, which gives rise to spatiotemporally constrained patterns of activity that underlie computation. Here, we discuss how neural systems use dynamics for computation, and claim that the biological constraints that shape brain networks may be leveraged to improve the implementation of artificial neural networks. To formalize this discussion, we consider a natural artificial analog of the brain that has been used extensively to model neural computation: the recurrent neural network (RNN). In both the brain and the RNN, we emphasize the common computational substrate atop which dynamics occur—the connectivity between neurons—and we explore the unique computational advantages offered by biophysical constraints such as resource efficiency, spatial embedding, and neurodevelopment.

Keywords: recurrent neural networks, dynamics, computation, spatial constraints, neurodevelopment

## I. INTRODUCTION

Dynamics have long underpinned computation. From the cycles of central pattern generators that support locomotion [1] to the networks of large-scale brain dynamics thought to regulate decision-making [2], it is clear that biological systems make ample use of their time-evolution to respond to their environment. Harnessing this dynamical computation, artificial recurrent neural networks (RNNs) have been trained to successfully perform the same computational tasks as humans [3, 4]. However, while inspired by the brain, training of RNNs is typically carried out in an unconstrained manner, leading to solutions that lack biophysical realism. Additionally, decades of neuroscience research has demonstrated the importance of biological constraints for achieving the brain’s unique structure and capabilities [5–8]. Here, we draw on literature from dynamical systems and neuroscience to discuss (i) how RNNs leverage dynamics to compute, and (ii) how biophysical constraints may shape this computation by guiding the formation of network structure. At the intersection of these goals exists an opportunity to study how biologically-constrained RNNs may yield more powerful and more interpretable

computational models.

To understand how biologically realistic neurons compute, there has been a long and rich history of modeling and interpreting neurobiological systems to leverage their computational capabilities. These quantitative models fall under the category of *dynamical systems*, whose evolution in time is determined by mathematical functions. At the scale of a single neuron, detailed circuit models of the ion channels that mediate membrane voltage have enabled quantitative understanding of signal propagation and computation in dendrites [9, 10]. At the scale of neural populations, mean-field models of excitatory and inhibitory neurons have enabled the study of neural circuits for biological sensing, imitation, and attention [11–13]. At the whole-brain level, both linear [14–17] and non-linear [18, 19] dynamical models have been used to simulate large-scale activity patterns, and have examined how those patterns spread across the brain’s white matter tracts. Across this broad range of systems, scales, and models, there exists a diversity of ways in which dynamics can be used for computation, as well as a crucial dependence of these dynamics on biophysical parameters.

More recently, with technological advances in deep learning, the study of neural computation has adopted a more functional direction that moves away from

---

\* Correspondence email address: jk2557@cornell.edu

† Correspondence email address: linden.parkes@rutgers.edu

biological realism. That is, rather than seeking a direct biophysical model [9, 11, 20, 21], RNNs posit a general dynamical system that is Turing complete [22], with parameters that are trained to solve computational tasks. We focus on RNNs because they have been used extensively to understand how general brain-like systems leverage dynamics to perform computation. Examples of this use include time-series prediction [23], source-separation [24], decision-making [3], odor classification [25], as well as concurrent performance of multiple cognitive tasks [26]. However, these insights and models often fail to translate into the real computational substrate of the brain—the neural architecture—because RNNs are trained without regard for biophysical constraints.

To merge biological realism with computational dynamics, we must first understand the physical embedding and constraints of the brain. In contrast to artificial RNNs, the brain is embedded within a circumscribed physical space [27], and its inter-connectivity is subject to limited metabolic resources [28]. This discrepancy makes it challenging to translate insights relating the structure, dynamics, and computation of biological brains to artificial RNNs. Neuroscience has studied these resource constraints for more than a century [29], suggesting that the brain is pressured to make efficient use of space, material, and time. That is, the brain must learn to communicate efficiently (time) while leveraging limited physical (space) as well as metabolic and cellular (material) resources. Critically, many of the brain’s topological features of connectivity and communication are thought to emerge as a consequence of navigating these pressures [5–8, 28]. These findings suggest that the brain’s resource constraints play a critical role in shaping its dynamic repertoire and computational capacity.

Here, our goal is to lay out promising new directions for improving the computational power and interpretability of RNN models of the brain. We posit that this goal will be achieved by placing biological constraints on RNNs that shape their structure and activity in systematic ways, which will in turn produce computationally improved dynamics. We focus on two aspects of RNN computational dynamics: the diversity of information that is represented by the neurons (expressivity), and the manipulation of low-dimensional internal representations (latent-spaces). In each section, we examine how biology shapes brain networks—with particular emphasis on the spatially-patterned macro-scale organizing principles of the cortex—and discuss how these constraints may be ported to RNNs to improve performance with interpretable structure and dynamics. Overall, we discuss how insights from biological and artificial computation can enrich each other towards a

new generation of biophysically realistic RNNs.

## II. THE RNN MODEL

To mathematically model the time-evolution of neural systems, we turn to dynamical systems which posit that the next state of a neural system can be written as a function of the current state and an input as

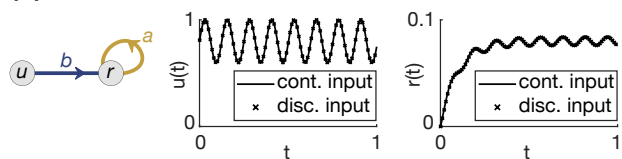
$$\mathbf{r}_{t+1} = f(\mathbf{r}_t, \mathbf{u}_t). \quad (1)$$

Here,  $\mathbf{r}_t \in \mathcal{R}^n$  is a vector of  $n$  neural activity states,  $\mathbf{u}_t \in \mathcal{R}^k$  is a vector of  $k$  inputs, and  $f$  is a function. As an example, let us consider a simple leaky integrator model with a single neuron which evolves according to

$$r_{t+1} = ar_t + bu_t, \quad (2)$$

where  $0 \leq a < 1$  and  $b$  are real numbers (Fig. 1A). As time evolves forward, the neuron state integrates the input  $bu_t$ , and the accumulated history of inputs decays at a rate set by  $a$ .

(A) One linear neuron



(B) Architecture of the RNN

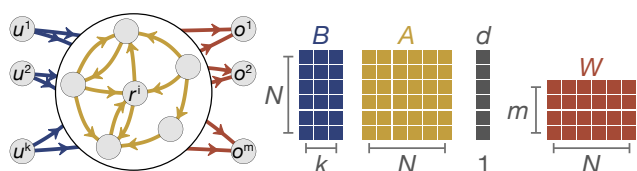


Figure 1. **One-dimensional dynamics and RNN architecture.** **A**, Schematic of a single neuron model of a leaky integrator with an input connection in blue, and a recurrent connection in gold (left), alongside examples of both a continuous and discretized input into the neuron (center), and the state of the input-driven neuron from the continuous (Eq. 5) and discretized (Eq. 2) dynamics (right). **B**, Schematic of an RNN with  $k$  inputs and  $m$  outputs, with input connections and matrix in blue, recurrent connections and matrix in gold, and output connections and matrix in maroon.

While Eq. 1 is written with  $t$  advancing in integer steps—thereby called a discrete-time dynamical system—many physical neural models evolve forward continuously in time as,

$$\frac{d}{dt}\mathbf{r}(t) = \hat{f}(\mathbf{r}(t), \mathbf{u}(t)). \quad (3)$$

We can approximate these continuous dynamics as discrete by evolving Eq. 3 in time using steps of  $\Delta t$  as,

$$\mathbf{r}_{t+\Delta t} = f(\mathbf{r}_t, \mathbf{u}_t) = \mathbf{r}_t + \int_t^{t+\Delta t} \hat{f}(\mathbf{r}(\tau), \mathbf{u}(\tau)) d\tau. \quad (4)$$

For example, the continuous-time version of the leaky integrator neuron is given by

$$\frac{d}{dt}r(t) = \hat{a}r(t) + \hat{b}u(t), \quad (5)$$

for  $\hat{a} \leq 0$  (Fig. 1A). In this case, the parameters of these two models can be interchanged through the transformation  $a = e^{\hat{a}\Delta t}$ ,  $b = (a - 1)\hat{b}/\hat{a}$ , but fundamental differences exist between continuous- and discrete-time systems [30]. Regardless of the system type, neural models make tradeoffs between complexity in the level of detail and tractability.

The RNN is a model that attempts to capture the biophysical quantity of the *interactions* between neurons through the connectivity matrix  $A$ . In tandem, the RNN simplifies the precise functional form of that interaction through the activation function  $f$ . In its most basic form, an RNN is a subset of dynamical systems (Eq. 3) that evolves in time as

$$\begin{aligned} \mathbf{r}_{t+1} &= f(A\mathbf{r}_t + B\mathbf{u}_t + \mathbf{d}), \\ \mathbf{o}_t &= g(\mathbf{r}_t). \end{aligned} \quad (6)$$

where  $A \in \mathcal{R}^{n \times n}$  is the connectivity matrix between neurons,  $B \in \mathcal{R}^{n \times k}$  is a matrix that linearly maps the inputs to the neurons,  $\mathbf{d} \in \mathcal{R}^n$  is a vector of bias terms, and  $f$  is an activation function (Fig. 1B). Rather than having  $f$  be a complex and biophysically motivated function, it is often approximated as a simple nonlinear function such as a sigmoid. The output of the RNN,  $\mathbf{o}_t$ , is usually taken to be some function  $g$  of the RNN state, and is often a linear output  $\mathbf{o}_t = W\mathbf{r}_t$ . Typically,  $A, B, \mathbf{d}$ , and  $W$  are treated as learnable parameters, some or all of which can be trained using a wide variety of methods [4, 31, 32].

We focus primarily on the computational role of the connectivity matrix  $A$ , as it dictates how the information in the RNN states is integrated as in the leaky integrator example (Eq. 2, 5). Despite its crucial importance in implementing computation, most uses of RNNs do not consider the biological pressures experienced by the brain while training RNN connectivity. In the following section, we describe how diversely the RNN states can *express* the inputs by leveraging  $A$ , and how biological processes and constraints reflect, mediate and accentuate this diversity.

### III. TUNING EXPRESSIVITY THROUGH REGULARIZED ACTIVITY AND NEUROMODULATION

When solving any computational problem, the expressivity of the language used is of crucial importance. The more expressive a language is, the greater the set of computations, formulae, and theorems comprise that language [33, 34]. For example, a programming language that supports conditional `if` statements can represent many more programs than an equivalent language that is without an `if` statement.

In the same way, neural networks can be viewed from the lens of expressivity. Specifically, we can ask: given arbitrary weights, what is the space of functions or dynamics that can be achieved? Previous work has demonstrated that shallow multi-layer perceptrons (MLPs) are universal function approximators [35, 36], and that RNNs are universal dynamics approximators [37]. However, even if a particular function or dynamical pattern is theoretically achievable, artificial neural networks—just like biological neural networks—must be trained from an initial condition. Hence, the study of expressivity extends beyond theoretical guarantees, and has been shown to rely heavily upon architectural features such as depth [38], neuron activation [39], and connectivity [40], which are crucial for explaining and engineering the success of modern-day neural networks.

In this section, we study three consequences of biological processes for expressivity. First, we introduce RNN expressivity as a richness of time history information about the inputs, and tie this richness to spatially-patterned temporal receptive fields in the brain that underpin information integration. Second, we study constraints on expressivity induced by resource constraints on neural activity as a putative learning mechanism. Finally, we explore the potential for neural networks to modulate their expressivity at short time-scales through neuromodulation. Together, the processes of the brain offer enticing and novel paradigms for training and constructing more expressive RNNs under biological constraints.

#### A. Expressivity as variable time-lagged integration of information

Expressivity of RNNs is intricately tied to the concept of *stability*: how quickly a perturbation to the RNN state decays [41]. If a perturbation decays very quickly, then the information contained in the perturbation cannot be used by the RNN for extended information processing. On the other hand, if the

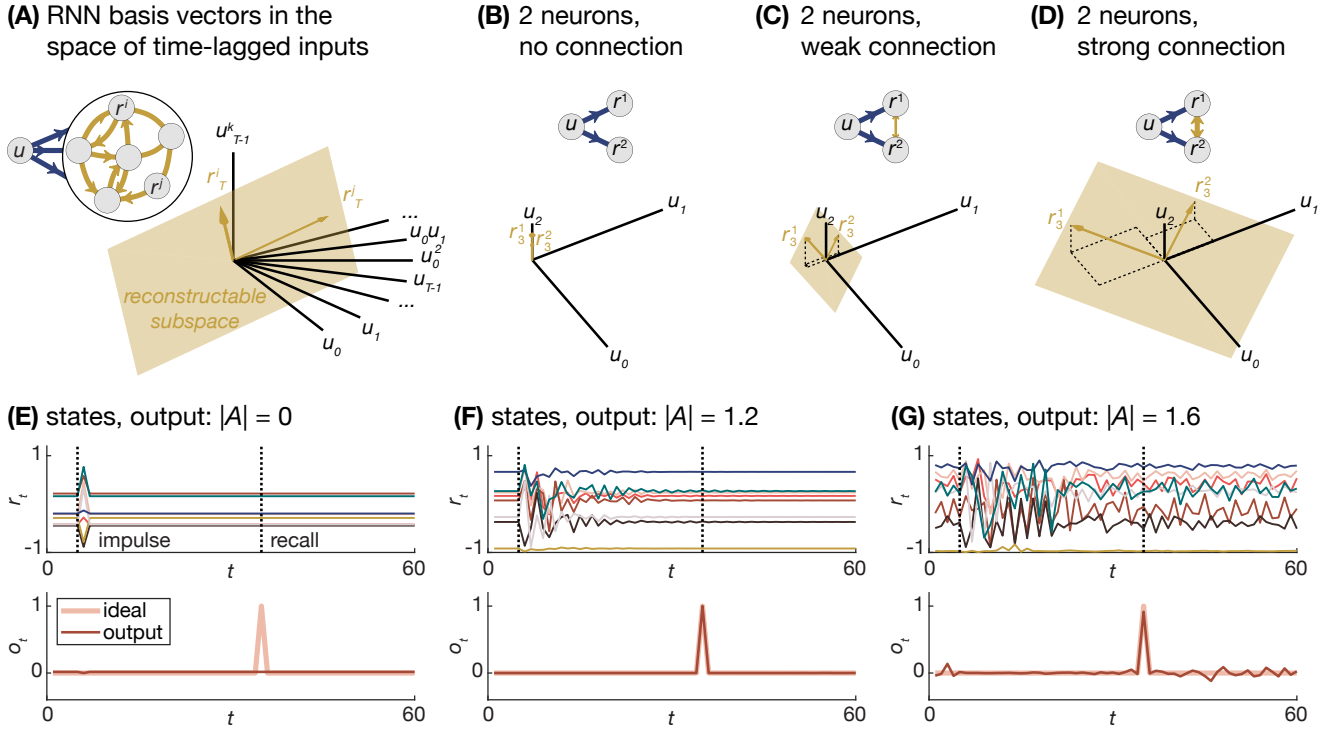


Figure 2. **Expressivity and the representational basis formed by RNNs.** **A** Schematic of an RNN with a single input, the basis vectors formed by neurons  $i$  and  $j$ , and the reconstructable subspace spanned by those vectors. **B** Schematic and reconstructable subspaces for a simple 2 neuron system with no recurrent connections, **C** weak recurrent connections, and **D** strong recurrent connections. **E** Example trajectories from a subset of 8 neurons of a 50 neuron system after receiving an impulse (top), alongside the ideal and true output of the RNN (bottom) after training a linear readout,  $o_t = W\mathbf{r}_t$ , to reproduce a time-lagged version of the impulse with  $|A| = 0$ , **F**  $|A| = 1.2$ , and **G**  $|A| = 1.6$ .

perturbation grows uncontrollably, then the temporal information in the inputs quickly becomes too complex to be represented by the finite number of neurons. As a result, an optimal amount of controlled stability should maximally preserve temporal information without saturating the RNN’s capacity.

We can quantify this intuition through a simple recursive substitution of Eq. 6,

$$\begin{aligned}
 \mathbf{r}_{t+1} &= f(A\mathbf{r}_t + B\mathbf{u}_t + \mathbf{d}) \\
 &= f(A(f(A\mathbf{r}_{t-1} + B\mathbf{u}_{t-1} + \mathbf{d}) + B\mathbf{u}_t + \mathbf{d})) \\
 &= \dots \\
 &= h(\mathbf{r}_0, \mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_t).
 \end{aligned} \tag{7}$$

Hence, in a noiseless system, the RNN state  $\mathbf{r}_{t+1}$  can be written as an explicit function  $h$  of the initial state,  $\mathbf{r}_0$ , and the full time history of the inputs,  $\mathbf{u}_\tau$ , mediated by the recursive application of  $A, B, \mathbf{d}$ , and the activation function  $f$ . The state of all neurons  $\mathbf{r}_{t+1}$  generates a *basis* for a subspace of the delay-embedded space of inputs  $\mathbf{u}_\tau$  (Fig. 2A), which means that the neuron states at time  $t + 1$  hold information about

the time history of the inputs mediated by  $A, B$ , and  $\mathbf{d}$  [42]. Thus, when computing an *output* using the neural states, we are implicitly computing an output using a basis of time-lagged input terms, where the connectivity defines the basis vectors, and therefore the reconstructable subspace (Fig. 2A). The more expressive this time-lagged basis, the greater the diversity of output functions which can be computed.

This expressivity is intimately tied to the connectivity matrix [43], and has been studied through many different lenses such as computation at the edge of chaos [40, 44, 45], criticality, and avalanches [46, 47]. The RNN’s stability is set by the specific activation function  $f$  and the connectivity matrix  $A$ . To gain intuition for this dependence, let us consider a simple linear 2 neuron system driven by one input. When there is no connectivity between the neurons, they store no time history of the inputs, and their state at time  $t = 3$  is purely a function of the input at time  $t = 2$  (Fig. 2B). When we add weak connections between the neurons, they begin to store some information about

the input at the previous time point  $t = 1$  (Fig. 2C). When we strengthen these connections, the neurons begin to store information from further back in time at  $t = 0$  (Fig. 2D).

To further develop this intuition for larger systems and a specific task, we consider a 50 neuron system whose connectivity is randomly initialized, and whose output is trained to recall an impulse from 30 time steps in the past. At the trivial limit of an RNN with no connectivity where  $|A| = 0$ , the recursion of Eq. 7 yields  $\mathbf{r}_{t+1} = f(B\mathbf{u}_t + \mathbf{d})$ , and we see that there is no time history of the input present in the RNN state. As a result, the RNN is unable to recall the input at a later point in time (Fig. 2E) [48]. As we increase the strength of connectivity, the RNN state stores more information about longer time lags of the input,  $\mathbf{u}_{t-\tau}$ , and is thus able to more accurately recall the input later in time (Fig. 2F). As the connectivity strength continues to further increase, the RNN state holds increasingly more time lags of the input until it saturates such that the number of neurons is smaller than the dimension of the space of time-lagged input functions, thereby forming an incomplete basis for that space (Fig. 2G).

This link between storing long time histories and expressivity is displayed prominently and spatially in the brain. Specifically, there exists a tight coupling between longer periods of temporal integration and higher-order computation, and this relationship varies systematically across the cortex [49]. At the macro-scale, cortical brain regions are thought to follow a dominant axis of variation that encodes a global processing hierarchy [50]. This gradient of brain organization is broadly referred to as the sensorimotor-association (S-A) axis [51, 52]. The S-A axis spans from primary cortices supporting sensation and movement at the bottom, to multimodal cortices supporting multisensory processing and integration in the middle, to transmodal association cortices supporting higher-order cognition at the top. The S-A axis is observed across multiple diverse features of brain structure and function [51] and is conserved across species [52, 53], demonstrating its evolutionary roots. Notably, as regions traverse up the S-A axis they undergo a progressive lengthening of their temporal receptive windows [49, 54]. Specifically, regions' intrinsic functional timescales vary over the S-A axis [55, 56] with regions at the top showing slower fluctuations reflecting longer temporal receptive windows. In turn, these longer windows are thought to enable greater accumulation and integration of information over time, facilitating higher-order cognition [49]. Conversely, regions at the bottom of the S-A axis show relatively fast dynamics, which is thought to underpin rapid

integration of recent sensory information [49].

This spatial patterning of receptive windows suggests that the brain—unlike naively constructed RNNs—distributes its computational expressivity systematically across the cortex, and research suggests that this may be critical for functional integration [49]. The above data suggest that RNNs too may benefit from spatially varying periods of temporal integration. Specifically, the stability of RNNs is typically only considered globally across the entire system, as the connectivity of many RNNs are initialized randomly. If the RNN is linear (i.e., if the activation function  $f$  is the identity matrix), then any ensuing dynamic instabilities are localized to linear subspaces, or *modes*, of neural activity [57]. However, if the RNN is nonlinear (e.g.,  $f = \tanh$ ), then instabilities bleed into other modes, making it difficult for the RNN to form a clean segregation of time-scales. Hence, varying periods of temporal integration could be achieved by specifying spatially varying penalties on neuronal timescales into the RNN cost functions. Such penalties may give rise to spatially segregated modules responsible for processing inputs at different time-scales.

## B. Learning by suppressing activity

While expressivity hinges on a careful balance of stable dynamics, biological neural networks are constrained by energy; more active neurons require more metabolic energy, which is a limited resource. As such, while artificial networks can maximize their expressivity through unconstrained backpropagation, the brain's capacity to learn is restricted by resource constraints; also, the extent to which the brain performs backpropagation remains unclear [58], which has motivated the machine learning community to consider more biologically-inspired optimization approaches. Hence, penalizing activity in RNNs should intuitively penalize computational capability through reduced expressivity. However, recent work has instead demonstrated important computational benefits of minimizing energy usage during training [59]. For example, Ali *et al.* [59] trained an RNN to predict sequences of handwritten digits, and examined how different optimization functions impacted model architecture and behavior. Specifically, Ali *et al.* [59] did not train their RNN to minimize prediction error through backpropagation. Instead, they trained their model to minimize absolute levels of neural activity prior to passing that activity through neurons' activation functions (here, ReLU). Such *preactivation minimization* is akin to selectively minimizing neurons' presynaptic inputs in biological networks. Critically, executing this cost function required no information

about task performance, and instead simply limited the RNN’s resources in a biologically plausible way.

Alongside good task performance, Ali *et al.* [59] observed dynamics in their RNN indicative of predictive coding. Predictive coding describes the hypothesis that the brain stores and updates expectations about its environment which it compares with incoming sensory evidence for those expectations [60]. Specifically, Ali *et al.* [59] observed activity patterns in their RNN suggestive of (i) selective self-inhibition in neurons receiving visual stimuli and (ii) prediction of future inputs in neurons not receiving visual stimuli. These results accord with hierarchically-organized predictive coding coupled to the S-A axis, wherein association cortices store predictions and via their distributed connectivity modulate activity in sensorimotor cortices [61–63]. Taken together, the findings of Ali *et al.* [59] indicate that while limiting the neurons’ activation might intuitively reduce their expressivity—thereby limiting their computational capability in RNNs—it can also serve as a unique mechanism for distributed learning that is more biophysically realistic than backpropagation.

### C. Dynamically Tuning Expressivity via Neuromodulation

The preceding sections discussed expressivity as an emergent property of a trained network that can be modified by placing certain constraints on RNN training. This static expressivity has been shown to be effective at performing a wide range of tasks, and is the basis of the success of *reservoir computing* [64]. Unlike in an RNN, the internal connectivity of the reservoir computer (RC) is not trained. Instead, only the output is trained, typically as a weighted sum of RC states [31]. Hence, RCs rely completely on the preexisting expressivity of their internal dynamics to generate a sufficiently expressive basis representation of their inputs. Because RCs can be trained without knowledge or modification of the internal system, a wide variety of physical systems have been explored as efficient RCs [65], including photonics [66], electrical circuits [67], hydrodynamic systems [24], and the brain [68]. However, expressivity in biological networks is not static, even in the presence of fixed weights. Instead it can be modulated dynamically over short time-scales and—similar to regions’ temporal receptive windows—this too is spatially patterned.

Previously, we discussed how the S-A axis tracks functional specialization and integration across the cortical mantle; cortical brain systems located at the bottom of the S-A axis are responsible for processing sensory/motor information while systems at the top of

the S-A axis are involved in processing higher-order cognition [50], and the brain’s connectivity allows for the hierarchical flow of information across these systems [69–71]. However, the functional roles of these different brain systems are not static. Instead the brain utilizes a complex array of neuromodulatory systems to actively reconfigure the brain’s dynamic repertoire [72]. In turn, this neuromodulation endows a relatively static network architecture (i.e., structural connectivity) with an increased capacity for functional flexibility. Reviewing all of the brain’s neuromodulatory mechanisms, and their effects on neural dynamics, is beyond the scope of this piece (see [72–75] for reviews), as these mechanisms comprise myriad cortico-cortical, cortico-subcortical, and subcortical-subcortical interactions. Here, we focus on a specific example that we believe is well positioned to be integrated into RNNs: the balance and modulation of cortical excitation and inhibition.

One fundamental neuromodulatory effect is that of dynamic changes to cortical excitation and inhibition. Neuron’s in the cortex receive a complex set of excitatory and inhibitory inputs, and the ratio between these inputs (E/I ratio) plays a critical role in coordinating an action potential. Following the S-A axis [51], the E/I ratio varies systematically across the cortex [76–78], leading to baseline differences in regions’ dynamics and computation [56, 79–81]. Moreover, incorporating regional variations to the E/I ratio into biophysical models has been shown to improve their fit to empirical functional data [80, 82], demonstrating that the E/I ratio shapes large-scale brain dynamics. However, unlike features of brain structure that track the S-A axis [51], regions’ baseline E/I ratio can be dynamically shifted via up- or downregulating the excitatory and inhibitory neurotransmitters of postsynaptic cells [72]. This regulation is achieved via multiple neurochemical pathways which can be driven exogenously—for example, via pharmaceutical agents [83] or chemogenetics [84, 85]—or endogenously, for example via the ascending noradrenergic arousal system (AAS) [72].

In dynamical systems, changes to neuronal excitation and inhibition are thought to engender population-level changes in neural gain (Fig. 3) [72]; the slope of a function that maps simulated neurons’ inputs to their outputs. By tuning the neural gain between coupled oscillators, Shine *et al.* [86] observed that increased gain lead to greater functional integration between neural populations. Critically, functional integration is thought to be an important computational property of the brain; in the human brain, functional integration fluctuates over short time scales [87] and facilitates cross-talk between the brain’s many functionally-specialized communities [88]. Thus, on-the-fly changes

to neurons’ E/I ratio facilitates a diverse range of dynamic behaviors [11]. This diversity allows brain function to flexibly decouple from its underlying structural architecture [89–92], which in turn supports a broader range of computations than would otherwise be possible.

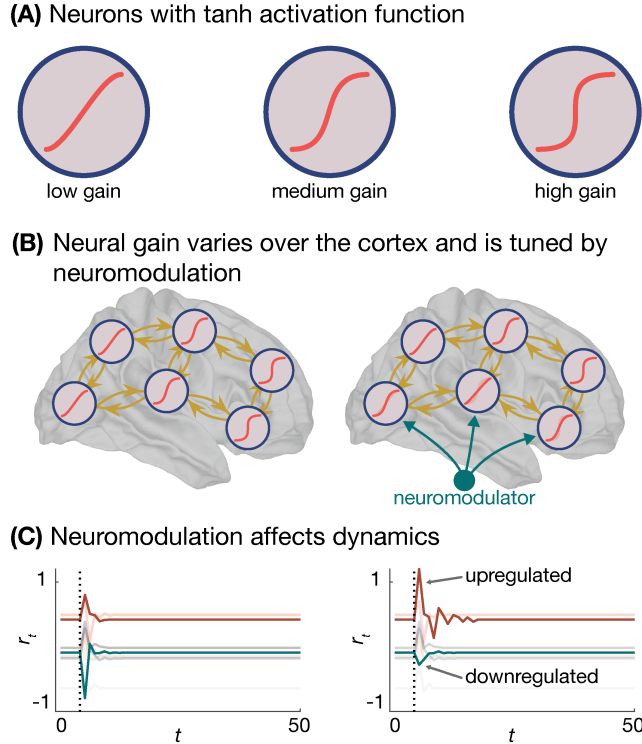


Figure 3. **Tuning neural gain in RNNs.** **A**, Neural gain can be dynamically tuned in a trained RNN by modifying the slope of neurons’ activation function. **B**, Neuroscientific research demonstrates that neural gain varies systematically across the cortex (left) and is tuned dynamically by neuromodulators (right). **C**, This neuromodulation can be used to change the dynamics of RNNs.

The above data suggests that modulation of regions’ E/I ratio gives rise to state-dependent dynamics that facilitate the brain’s computational expressivity. Recently, researchers have begun examining how E/I modulation might be instantiated in RNNs, with a particular focus on the aforementioned AAS. The AAS stems from the locus coeruleus, a small brainstem structure that provides diffuse noradrenergic projections spanning the cerebral cortex [72, 93]. These projections modulate neuronal excitability via the neurotransmitter noradrenaline, granting the AAS the capacity to modulate the E/I ratio. Drawing on this mechanism, Wainstein *et al.* [94] trained an RNN to perform a perceptual switching task, wherein

one visual stimuli (a plane) gradually morphed into another (a shark) and the RNN was tasked with reporting which stimuli it perceived at each time point. Once trained, Wainstein *et al.* [94] modified the slope of the artificial neurons’ activation function (i.e., the neural gain) and examined the corresponding change in perceptual switching. The authors observed that higher gain caused perceptual switches to occur earlier than expected, while lower gain caused the opposite. Additionally, Wainstein *et al.* [94] modeled the energy landscape of the RNN state-space and observed that increasing neurons’ gain flattened the landscape, allowing for easier state transitions (perceptual switches). Finally, the authors supported these modeling results with task-based fMRI data as well as pupillometry data, which is thought to be an indirect measure of noradrenaline-mediated arousal [95]. Together, the authors’ results demonstrate that a system’s computational function can be dynamically modulated in behaviorally meaningful ways, and that this reconfiguration may be underpinned by an internal capacity to regulate neural excitability. Critically, this dynamic reconfiguration unfolds on top of a static network architecture, wherein only neurons’ activation functions are tweaked while their trained weights are preserved.

The results of Wainstein *et al.* [94] demonstrate that the affects of neuromodulation can be introduced to RNNs, modifying their functional outcomes in behaviorally meaningful ways. However, as touched on above, the brain comprises multiple neuromodulatory systems that are capable of influencing regions’ excitation and inhibition, each of which subserve different functional goals [72] and each exhibit unique spatial patterning of their associated neurotransmitters and genes [77]. Future work examining how each of these neurotransmitter maps affect RNN behavior, across a diverse range of tasks, will be important to characterize how different neuromodulatory mechanisms influence expressivity. Indeed, other fields of dynamical systems (e.g. linear systems) have already begun pursuing these goals [96–98].

#### IV. COMPUTING WITH THE LATENT SPACES OF RNNs VIA CONSTRAINED CONNECTIVITY

While expressivity tells us what information about the input is contained in a specific state, it does not tell us about the computational meaning behind that state. Specifically, although Eq. 7 provides us with a map of how any input series  $\mathbf{u}_\tau$  is expressed as a specific neural state  $\mathbf{x}_{t+1}$ , the meaning of that state depends on the context of the problem being solved.

For example, while the dynamics of the transistors in a microprocessor can be known and simulated, the computational meaning of the transistor state depends on its internal, or *latent* representation [99].

As an illustration of latent representation, consider one of the fundamental memory elements of computers, the *set-reset latch* (SR-latch) [100], which simply remembers which of two inputs were pulsed most recently. A single nonlinear neuron with two inputs can be designed to mimic this behavior (Fig. 4A), where its state remains high if input  $u_1$  was last pulsed, and remains low if input  $u_2$  was last pulsed. Here, the state of the neuron is directly the output of a latch. Alternatively, this latch functionality can be defined in a *distributed* manner into a system of multiple neurons, where the high state is represented as some pattern of activity  $\mathbf{r}^*$ , the low state is represented as another pattern of activity  $\mathbf{r}^\dagger$ , and the input pulses transition the RNN state between these two (Fig. 4B). Here, no single neuron is responsible for the latch dynamics. Rather, these latent-space latch dynamics depend on the connectivity between neurons, as well as how that connectivity was formed by training. In this section, we discuss how RNNs represent and manipulate information in their latent space, and the consequence of biological constraints on these latent representations. Then, we draw on recent advances from the field of neurodevelopment to put forth new directions for studying biologically-constrained RNNs.

### A. Sparsity and attractor stability

Neural networks harness the power of internal, or *latent*, representations for computational tasks such as path integration [101], tracking [102], and spatial working memory [103]. In RNNs and dynamical systems, these latent representations are often referred to as *attractors*: sets of points,  $\mathcal{S} = \{\mathbf{s}_i\}$ , to which the dynamics evolve after a relatively long period of time. An early example of latent representations is associative memory in Hopfield networks [104], wherein a specific set of neural activity patterns,  $\mathcal{S} = \{\mathbf{s}_i\}$ , are stored as memories that could represent information such as an image. Specifically,  $\mathcal{S}$  are stored as *fixed-point attractors* [105] such that after stimulating neurons close to a specific pattern,  $\mathbf{x}_0 = \mathbf{s}_i + \epsilon$ , the neural states will evolve toward a stored memory,  $\mathbf{x}_{t \rightarrow \infty} \approx \mathbf{s}_i$ . In general, a fixed-point is a state  $\mathbf{x}^*$  such that

$$\mathbf{x}^* = f(\mathbf{x}^*). \quad (8)$$

The computational power of this Hopfield network is in using inputs to retrieve pre-defined information (i.e., *memories*) stored in attractors, and substantial work

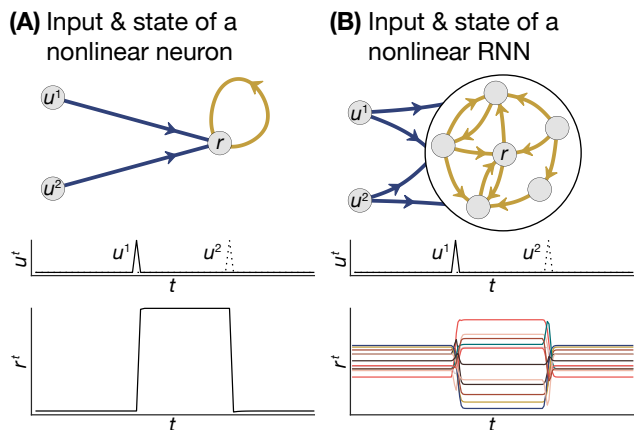


Figure 4. **Distributed latent-space representations.**

**A**, A nonlinear single-neuron model with two inputs (top), which has been designed to act as a memory circuit such that an impulse from input  $u_1$  or  $u_2$  will cause the state of the neuron to stay fixed at a high or low value, respectively (bottom). **B**, A nonlinear multi-neuron RNN which has been designed to act as a memory circuit (top), except that the RNN represents the states “high” and “low” as two stable fixed-points across all of its states, and these fixed-points are determined in a distributed manner *via* all of the connections between the neurons.

has gone into improving their computational capability [106] and biophysical realism [107]. Hence, fixed-point attractors are latent dynamical properties that can be harnessed for computation.

In addition to discrete memory states, RNNs can make use of the *geometry* of their attractors to form representations and make decisions. For example, *continuous-attractor neural networks* (CANNs) extend the concept of an attracting point to higher-dimensional manifolds, thereby forming attracting curves and surfaces such that the geometric position along these manifolds holds latent computational meaning. For example, the geometric trajectory of the neural network state along these manifolds can reflect a path traversed in real physical space by an agent, [101], the continuous tracking of a moving stimulus [102], and the recall of spatial location in the prefrontal cortex [103]. While the connectivity and dynamics of CANNs are precisely engineered to preserve translation-invariance along their structure, this continuum of attractors also emerges in trained RNN models [108], and even in models of the prefrontal cortex trained to integrate information given different contexts [3]. Hence, rather than the activity of one or a collection of specific neurons [109], it is the geometry of the attractor manifold that forms the internal representation of information in the RNN, and the RNN integrates external information



by moving its representation along that manifold [110].

Formation of attractors occurs as a consequence of a loss of energy in the system, which in turn results in the stabilization of dynamics [111]. This stabilization can be characterized using methods such as Lyapunov functions—energy-like quantities that monotonically decrease or dissipate throughout the dynamics [112, 113]—and Lyapunov exponents—the rate of convergence towards an attracting manifold [114]—to study existing systems. However, what is less clear is how connectivity can develop to improve the stabilization of these attractors. Intuitively, this energy dissipation usually takes the form of a loss in the neural *activity* (e.g. the “leaky” component of the leaky integrator in Eq. 2), whereas the parameters that can be learned in the RNN are the *connectivity*. Thus, the question becomes: how do we modify the RNN connectivity to achieve greater attractor stability?

A biologically-inspired optimization process that has proven useful for stabilizing RNN dynamics is *sparse-ness*. In order to minimize energy expenditure [111], the brain substantially prunes its connectivity [115] retaining only a sparse set of weights that are finely tuned to achieve its functional goals. In RNNs, inducing sparseness via weight pruning has been shown to provide several computational benefits. For example, Averbeck [116] trained RNNs with and without weight pruning to complete a working memory task. Compared to their unpruned counterparts, moderate amounts of pruning yielded RNNs that (i) exhibited better task performance; (ii) required fewer training epochs; (iii) had stronger connectivity weights; and (iv) were more resistant to task distractors [116]. Notably, regarding distractor resistance, pruned RNNs showed a smaller departure from their dynamic trajectories when they were perturbed by a distracting probe within the task. This result demonstrates that the sparse connectivity in the pruned RNNs strengthened their attractor basins (Fig. 5), making them more stable and resistant to undesired inputs.

### B. Deduction: Learning Problem Structure through Iterative Algorithms

While the geometry of dynamical attractors can be used for computational purposes, they also arise as the solutions to complex problems. For example, iterative methods are commonly used in optimization problems such as iterative refinement, [117], root-finding methods [118], and feasibility problems [119]. Critically, in these and many other examples, the solution is not learned in the typical deep learning sense (i.e., training), but rather emerges as an attractor to satisfy

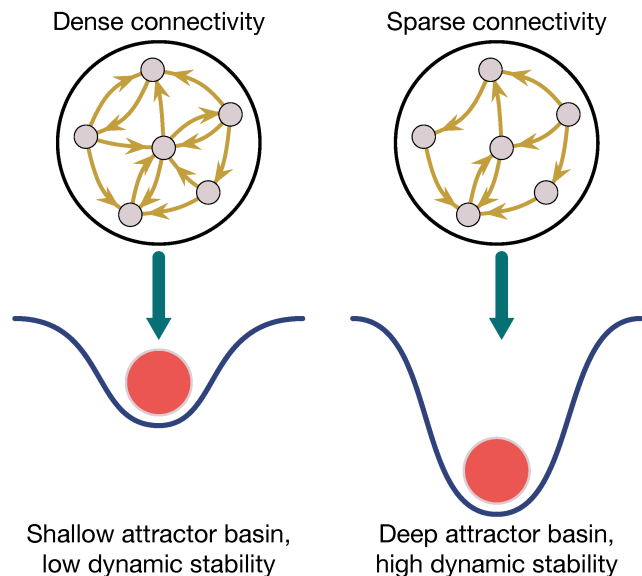


Figure 5. **Sparse connectivity leads to stable attractors.** Pruning connectivity weights in RNNs leads to sparse weight matrices that deepen the attractor basins of RNNs dynamics, making dynamics more stable.

the conditions of an iterative *algorithm*. In the same manner, RNNs need not only learn the attractor structure of specific input-output relations (i.e., through training), but have the potential to encode a specific algorithm in the iteration of the neural states, such that solutions to problems (i.e., tasks) emerge as attractors.

Biological neural networks demonstrate the ability to run iterations within their latent representations [120]. A prominent example is hippocampal replay, whereby hippocampal place cells will reactivate along the same sequence as in a prior navigation experience [121], even when the subject is not actively performing a navigation task. Another prominent example is dynamical inference, whereby neural activity in the dlPFC can reliably predict the future nonlinear trajectory of a ball, and RNN models which best replicate prediction behavior are trained on the sequence of the ball’s trajectory [122, 123]. Hence, RNNs are not only capable of learning attractor geometries, but also of learning and simulating the sequence of the problem structure, which may enable more generalizable solutions.

RNNs can be engineered to run iterative deductions through several means. One approach involves assigning to each neuron the state variable of an algorithm, and defining complex interaction dynamics such that the RNN state will settle on the solution as a stable attractor. For example, an RNN can be designed to solve  $k$ -satisfiability problems, which seek

an assignment of  $n$  Boolean variables that satisfy a set of  $c$  constraints, where each constraint places a condition on subsets of  $k$  Boolean variables (Fig. 6A) [124]. These RNNs evolve until the neural states find a solution [124]. Surprisingly, a wide variety of different dynamics and architectures can lead to different algorithms for solving the same satisfiability problem [125, 126], and other difficult optimization problems such as integer linear programming feasibility [127] or the  $n$ -queens problem [128]. Hence, algorithm variables can be directly represented by individual neurons, and the algorithm rules can be directly encoded in the connectivity and update rules  $f$  of RNNs.

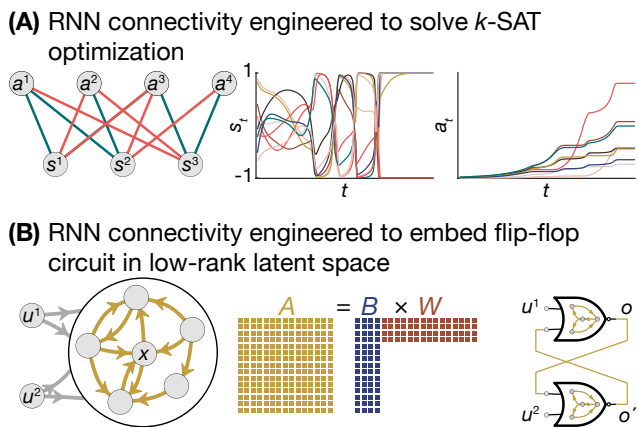


Figure 6. **Programming Algorithms into RNNs** **A**, Schematic of an RNN whose neurons  $s_i$  and  $a_j$  represent states in a  $k$ -SAT problem, and whose connectivity is defined such that the stable attracting manifold is guaranteed to be a solution (left). The dynamics of the RNN evolve in a complex manner as it searches and finds the assignment of states that satisfies the  $k$ -SAT problem (center, right). **B**, Schematic of the process for designing algorithms into RNNs, where the recurrent connectivity of the RNN (gold) can be constructed from a low-rank product of matrices  $(B, W)$ . Through this construction, the evolution of the RNN states concurrently evolves the state variables of the algorithm within a low-dimensional projection. This designed connectivity was used to program and simulate the SR-latch in Fig. 4B).

In addition to a direct, one-to-one encoding of problem structure as an RNN, we can also design algorithms into the latent spaces of RNNs [129]. Rather than ascribing each neuron a specific variable in an algorithm, we can embed an algorithm into the distributed connectivity of an RNN (Fig. 6B). One technique for this embedding is the *neural engineering framework* (NEF) which enables the design of iterations of the latent-space variables through the engineering of low-rank connectivity matrices

[130, 131]. Extensions enable the programming of iterations in pre-existing, higher-rank connectivity, and the ability to reverse-engineer representations from conventionally trained RNNs [42]. Other engineered architectures such as the differentiable neural computer [132] and the neural turing machine [133] emulate the structure of conventional computers using differentiable neural elements. Hence, RNNs have the capability to explicitly run complex algorithms in their latent spaces, which is crucial for generalizable computation; rather than learning individual solutions, we posit that training RNNs on solution sequences will enable them to learn and generalize problem-solving strategies [134], even into nonlinear, out-of-sample regimes [135].

While the above approaches may provide more generalizable RNN solutions than task-specific training, the added computational capabilities of engineered RNNs is accompanied by a further deviation of the corresponding connectivity from biology. Whether through the enforcing of low-rank connectivity [130, 136] or the segregation of memory and processing units in a neural von Neumann architecture [132], engineered neural connectivities lack many of the costs and constraints experienced by biological networks. A critical question then is how the brain formulates connectivities that permit sophisticated latent space computations while adhering to biological constraints. Prior work has demonstrated the capability of largely disordered RNNs to produce sequences that rely on recurrent connectivity [137], and the importance of the sequence of learning over many learning iterations—a curriculum—for RNN performance [138]. Hence, rather than forming low-rank structures *ab initio*, the brain defines its structure and dynamics through learning and plasticity on long time scales. We explore insights into the governing principles and computational advantages of this progression in neurodevelopment.

### C. Latent-space computational capability throughout neurodevelopment

Just as the topology of an RNN is sculpted over training epochs (Fig. 7A), the topology of the human connectome is sculpted throughout development (Fig. 7B). However, unlike the RNN, which may be trained in an unconstrained manner, neurodevelopment follows a carefully orchestrated and stereotyped program that unfolds dynamically across space and time. Specifically, cortical neurodevelopment is thought to spatially track the aforementioned S-A axis in a temporally staged manner [51, 139], and this staging is thought to underpin the emergence of cortical regions' functional specialization and inter-connectivity. Crucially, the asynchronous nature of this developmental program

is thought to underpin the sequential emergence of increasingly complex cognitive functions [140, 141], suggesting that neurodevelopment stages the brain’s acquisition of lower- and higher-order computational processes. Mechanistically, this program may be underpinned by windows of heightened neural plasticity that cascade up the S-A axis [139, 142] priming specific neural circuits at specific points in time for experience-dependent neural change.

Regions in the cortex are defined in part according to their laminar structure [143, 144], with different regions exhibiting variations in the number and size of their distinct layers, as well as different distributions of cells throughout those layers. Critically, cortical variations in cytoarchitecture conform to the S-A axis [145], and animal research demonstrates that this spatial patterning predicts regions’ extrinsic connectivity [146], including their strength [147], distance [147], and layer-wise projections [147, 148]. In humans, structural connectivity between regions at the bottom of the S-A axis refines relatively early in development, while connectivity at the top of the S-A axis does so later in development [71, 149–152]. Furthermore, recent work has shown that cytoarchitecture plays an important role in shaping how dynamics spread across the connectome throughout development [69]. Thus, the spatial patterning embedded in the S-A axis plays a key role in shaping connectome topology throughout development. But what about RNNs? Recent work by Achterberg *et al.* [153] regularized RNNs by using the Euclidean distance between regions to constrain training. The authors found that modularity [154] and small-worldness [155]—two complex topological features that are hallmarks of the human connectome [28]—emerged to a greater extent in these spatially-embedded RNNs compared to standard RNNs (see also recent work by Tanner *et al.* [156] for evidence of modularity in RNNs trained without spatial constraints). Additionally, this effect coincided with achieving higher out-of-sample task performance earlier in training compared to standard RNNs (though performance eventually converged; see their Figure 2A).

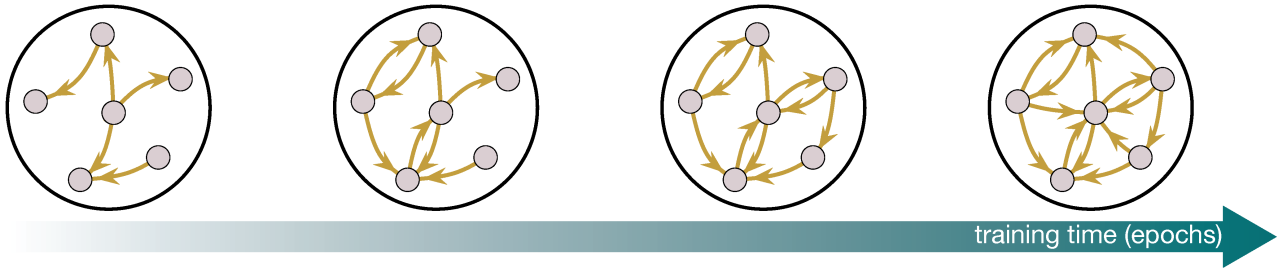
The results of Achterberg *et al.* [153] demonstrate that incorporating space-based inductive biases into RNN training causes them to converge on topological features observed in the human connectome. However, it remains unclear whether neurodevelopmentally-informed spatial constraints, like those embedded in the S-A axis [51], show similar effects. We posit that constraining RNNs using the S-A axis may outperform Euclidean distance-based spatial embedding, as the former is rooted in evolutionary programs of connectivity formation and functional specialization

[53]. Additionally, the spatial constraints deployed by Achterberg *et al.* [153] were static throughout training. As mentioned above, the S-A axis scaffolds connectome development in a temporally varying way [139], and incorporating this dynamic information will be critical to achieving realistic brain-like topology in RNNs. One approach would be to code spatially varying periods of heightened learning into RNNs, simulating traveling waves of heightened neural plasticity [139, 142]. Such an inductive bias could be achieved by including temporal cascades of weight training that flow bottom-up across the S-A axis.

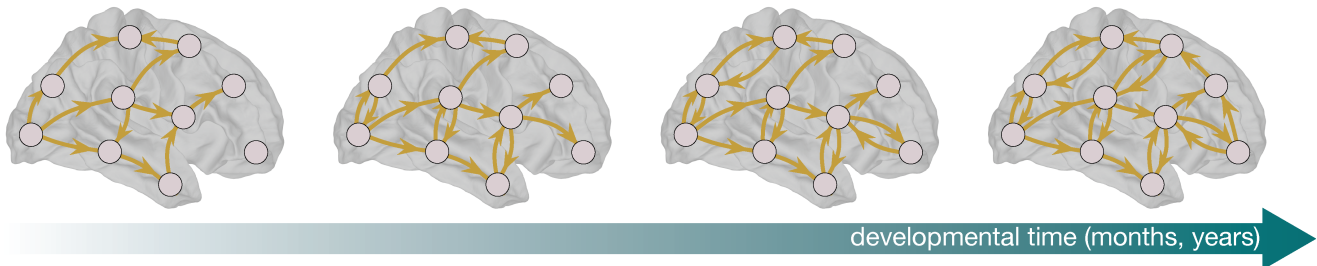
In addition to injecting spatially constrained inductive biases into RNNs, we can also directly assess the ability of the developing connectome to support latent-space computation. This analysis can be achieved through studying the *synchronization* between a given RNN and a particular latent attractor. Historically, synchronization has been shown to be crucial for computation in both biological cortical networks [157] and artificial RNNs [158], and is deeply related to consensus dynamics [159]. Intuitively, synchronization between two systems implies that both systems are evolving identically. This concept can be extended to *generalized synchronization*, which stipulates conditions under which a response system,  $\mathbf{r}_t$ , has synchronized in a general sense to a driving system,  $\mathbf{d}_t$  [160], such that rather than evolving identically such that  $\mathbf{r}_t = \mathbf{d}_t$  [161], the joint system has collapsed onto a function  $\phi$  of only the driver system such that  $\mathbf{r}_t = \phi(\mathbf{d}_t)$ . Under these conditions, the response system has *followed* the attractor structure of the driving system. If we choose the response system to be an RNN with an empirically-derived connectivity taken at a specific point in neurodevelopment, and the driver system to be a specific latent attractor dynamical system, then we can assess whether the RNN can follow the attractor structure of the driving system (Fig. 7C).

Of course, just because the RNN can synchronize to a particular latent-space attractor does not guarantee that it can maintain that attractor once the driving system is gone. In order for the RNN to *internalize* the attractor dynamics, theories of *invertible generalized synchronization* (IGS) stipulate conditions for which the attractor of the driver system can be invertibly reproduced and stabilized by the response system, and thus can be *learned* autonomously [162]. Hence, rather than modifying RNN connectivity to learn a latent space conditioned on spatial and resource constraints, IGS tests whether a given RNN connectivity that already obeys those spatial and resource constraints can stably generate a latent attractor (Fig. 7D). The IGS theory also indicates that the ability to internalize latent attractors from driving systems depends not only on the

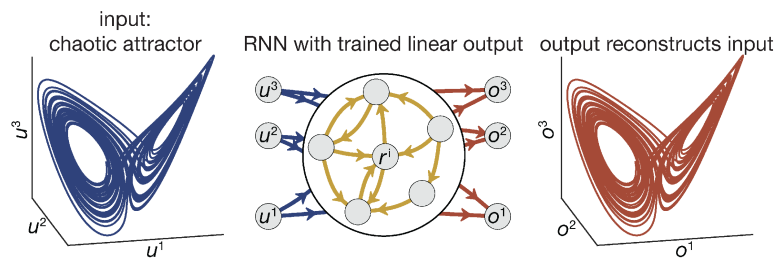
(A) RNNs comprise abstract nodes and edges, and undergo unconstrained training



(B) Brains comprise physically embedded nodes and edges, and undergo constrained development



(C) Generalized synchronization of RNN to attractor manifold



(D) Attractor stabilization via invertible generalized synchronization

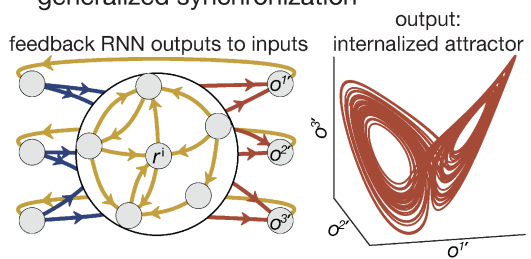


Figure 7. **Computational capabilities of RNNs and biological brains throughout development.** RNNs and brains are subject to different environments. **A**, RNNs comprise abstract nodes and edges, and their training is unconstrained. This allows RNNs to evolve toward any network architecture that optimally performs a given task. **B**, By contrast, brains comprise physically embedded nodes and edges, and the formations of their network architecture is tightly constrained by space and time. **C**, The extent to which pre-existing connectivities can *follow* latent representations can be captured through *generalized synchronization*, which measures how well an RNN that is driven by an attracting manifold (in this case, the chaotic Lorenz attractor) can reconstruct the driving signal from its states. **D**, The ability of pre-existing connectivities to *internalize* and generate latent representations is captured through *invertible generalized synchronization*, whereby an RNN that has learned output weights  $W$  that copy the attractor manifold can then *drive itself* by feeding those outputs back as inputs to autonomously generate the attractor *without any input*.

RNN's structure, but also the latent attractor. Thus, given RNNs of different structures throughout development, we may examine the IGS on their most-likely encountered driving signals corresponding to their position along the S-A axis.

## V. CONCLUSIONS

Neural systems compute using dynamics, and the dynamics of biological brains evolve atop the computational substrate of a spatially and resource constrained network. Here, we sought to jointly discuss advances in neuroscience and dynamical systems with a view to improving the computational power and interpretability of RNNs. Within the context of two computational capabilities—expressivity and latent-

space computing—we highlighted several avenues for future research that we believe will advance our understanding of dynamical computation. Through these avenues, we envision biologically interpretable, computationally improved RNN models of how the brain computes.

Central to this research program is the application of biophysical constraints on the computational dynamics of RNNs. While not exhaustive, the constraints discussed herein represent the diverse influences of neurobiology on network structure, and they have been shown to influence the emergence of complex behavior in humans. Furthermore, the influence of these constraints on RNN connectivity can be studied in combination. For example, the spatial-patterning of regions’ baseline E/I ratio emerges throughout development [139], indicating its connection to the S-A axis. Another example is to explore how to extend low-rank RNN design approaches to higher-rank connectivities that more closely match the spatial gradients of the S-A axis. Thus, these biophysical constraints provide fertile ground for future experimental, computational, and theoretical work into biologically-informed RNNs.

The methodological strategies for incorporating these constraints into computational RNN modeling are vast. Here, we have discussed several approaches: using known cortical structure and function as target connectivities of additional training constraints [51, 153], using resource constraints as an alternative learning mecha-

nism [32], dynamically altering connectivity with neuromodulation for increased expressivity [86], stabilizing latent attractors through pruning [116], engineering and modifying low-rank latent representations and algorithms [42, 130], probing the teachability of RNNs *via* synchronization [162], among many others. Hence, there is no one prescription that serves as a panacea for the rich problems that lie at the intersection of computation, dynamics, and neurobiology. Instead, we must continue developing diverse and creative approaches for maximizing the computational capabilities of neurodynamical models through the use of biological and developmental constraints.

## ACKNOWLEDGEMENTS

We gratefully acknowledge comments and conversations with Dr. Zhixin Lu, Dr. Harang Ju, Dr. Dale Zhou, Dr. Melody Lim, and Dr. Stephen Hanson. JZK was supported by the Bethe/KIC/Wilkins postdoctoral fellowship. BL was supported by the National Institute Of Mental Health of the National Institutes of Health under Award Number R00MH127293. LP was supported by the National Institute Of Mental Health of the National Institutes of Health under Award Number R00MH127296.

## REFERENCES

- 
- [1] Ijspeert, A. J. Central pattern generators for locomotion control in animals and robots: a review. *Neural networks* **21**, 642–653 (2008).
  - [2] Seeley, W. W. *et al.* Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience* **27**, 2349–2356 (2007).
  - [3] Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature* **503**, 78–84 (2013).
  - [4] Werbos, P. J. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE* **78**, 1550–1560 (1990).
  - [5] Betzel, R. F. *et al.* Generative models of the human connectome. *NeuroImage* **124**, 1054–1064 (2016).
  - [6] Oldham, S. & Fornito, A. The development of brain network hubs. *Developmental Cognitive Neuroscience* **36**, 100607 (2019).
  - [7] Oldham, S. *et al.* Modeling spatial, developmental, physiological, and topological constraints on human brain connectivity. *Science Advances* **8**, eabm6127 (2022).
  - [8] Akarca, D. *et al.* A generative network model of neurodevelopmental diversity in structural brain organization. *Nature Communications* **12**, 4216 (2021).
  - [9] Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology* **117**, 500 (1952).
  - [10] Petousakis, K.-E., Apostolopoulou, A. A. & Poirazi, P. The impact of hodgkin–huxley models on dendritic research. *The Journal of Physiology* (2022).
  - [11] Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal* **12**, 1–24 (1972).
  - [12] Pinto, D. J., Brumberg, J. C., Simons, D. J., Ermentrout, G. B. & Traub, R. A quantitative population model of whisker barrels: re-examining the wilson-cowan equations. *Journal of computational neuroscience* **3**, 247–264 (1996).
  - [13] Sadeghi, S., Mier, D., Gerchen, M. F., Schmidt, S. N. & Hass, J. Dynamic causal modeling for fmri with wilson-cowan-based neuronal equations. *Frontiers in Neuroscience* **14**, 593867 (2020).

- [14] Parkes, L. *et al.* Using network control theory to study the dynamics of the structural connectome. preprint, Neuroscience (2023).
- [15] Lynn, C. W. & Bassett, D. S. The physics of brain network structure, function and control. *Nature Reviews Physics* **1**, 318–332 (2019).
- [16] Seguin, C., Sporns, O. & Zalesky, A. Brain network communication: concepts, models and applications. *Nature Reviews Neuroscience* (2023).
- [17] Srivastava, P. *et al.* Models of communication and control for brain networks: distinctions, convergence, and future outlook. *Network Neuroscience* **4**, 1122–1159 (2020).
- [18] Breakspear, M. Dynamic models of large-scale brain activity. *Nature Neuroscience* **20**, 340–352 (2017).
- [19] Roberts, J. A. *et al.* Metastable brain waves. *Nature Communications* **10**, 1056 (2019).
- [20] FitzHugh, R. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal* **1**, 445–466 (1961).
- [21] McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* **5**, 115–133 (1943).
- [22] Chung, S. & Siegelmann, H. Turing completeness of bounded-precision recurrent neural networks. *Advances in Neural Information Processing Systems* **34**, 28431–28441 (2021).
- [23] Liu, Y., Gong, C., Yang, L. & Chen, Y. Dstp-rnn: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction. *Expert Systems with Applications* **143**, 113082 (2020).
- [24] Lu, Z., Kim, J. Z. & Bassett, D. S. Supervised chaotic source separation by a tank of water. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30** (2020).
- [25] Wang, P. Y., Sun, Y., Axel, R., Abbott, L. & Yang, G. R. Evolving the olfactory system with machine learning. *Neuron* **109**, 3879–3892.e5 (2021).
- [26] Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T. & Wang, X.-J. Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience* **22**, 297–306 (2019).
- [27] Roberts, J. A. *et al.* The contribution of geometry to the human connectome. *NeuroImage* **124**, 379–393 (2016).
- [28] Bullmore, E. & Sporns, O. The economy of brain network organization. *Nature Reviews Neuroscience* **13**, 336–349 (2012).
- [29] Ramón y Cajal, S. *Histology of the nervous system of man and vertebrates*. No. no. 6 in History of neuroscience (Oxford University Press, New York, 1995).
- [30] Dupont, E., Doucet, A. & Teh, Y. W. Augmented neural odes. *Advances in neural information processing systems* **32** (2019).
- [31] Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Computer science review* **3**, 127–149 (2009).
- [32] Ali, A., Ahmad, N., de Groot, E., van Gerven, M. A. J. & Kietzmann, T. C. Predictive coding is a consequence of energy efficiency in recurrent neural networks. *Patterns* **3** (2022).
- [33] Felleisen, M. On the expressive power of programming languages. *Science of computer programming* **17**, 35–75 (1991).
- [34] Sipser, M. Introduction to the theory of computation. *ACM Sigact News* **27**, 27–29 (1996).
- [35] Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural networks* **2**, 359–366 (1989).
- [36] Cybenko, G. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems* **2**, 303–314 (1989).
- [37] Schäfer, A. M. & Zimmermann, H. G. Recurrent neural networks are universal approximators. In *Artificial Neural Networks–ICANN 2006: 16th International Conference, Athens, Greece, September 10–14, 2006. Proceedings, Part I 16*, 632–640 (Springer, 2006).
- [38] Poole, B., Lahiri, S., Raghu, M., Sohl-Dickstein, J. & Ganguli, S. Exponential expressivity in deep neural networks through transient chaos. *Advances in neural information processing systems* **29** (2016).
- [39] Raghu, M., Poole, B., Kleinberg, J., Ganguli, S. & Sohl-Dickstein, J. On the expressive power of deep neural networks. In *international conference on machine learning*, 2847–2854 (PMLR, 2017).
- [40] Bertschinger, N. & Natschläger, T. Real-time computation at the edge of chaos in recurrent neural networks. *Neural computation* **16**, 1413–1436 (2004).
- [41] Sompolinsky, H., Crisanti, A. & Sommers, H.-J. Chaos in random neural networks. *Physical review letters* **61**, 259 (1988).
- [42] Kim, J. Z. & Bassett, D. S. A neural machine code and programming framework for the reservoir computer. *Nature Machine Intelligence* 1–9 (2023).
- [43] Rajan, K. & Abbott, L. F. Eigenvalue spectra of random matrices for neural networks. *Physical review letters* **97**, 188104 (2006).
- [44] Langton, C. G. Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: nonlinear phenomena* **42**, 12–37 (1990).
- [45] Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
- [46] Munoz, M. A. Colloquium: Criticality and dynamical scaling in living systems. *Reviews of Modern Physics* **90**, 031001 (2018).
- [47] Hochstetter, J. *et al.* Avalanches and edge-of-chaos learning in neuromorphic nanowire networks. *Nature Communications* **12**, 4008 (2021).
- [48] Ju, H., Kim, J. Z., Beggs, J. M. & Bassett, D. S. Network structure of cascading neural systems predicts stimulus propagation and recovery. *Journal of Neural Engineering* **17**, 056045 (2020).
- [49] Wolff, A. *et al.* Intrinsic neural timescales: temporal integration and segregation. *Trends in Cognitive Sciences* S1364661321002928 (2022).
- [50] Mesulam, M. Representation, inference, and transcendent encoding in neurocognitive networks of the human brain. *Annals of Neurology* **64**, 367–378 (2008).
- [51] Sydnor, V. J. *et al.* Neurodevelopment of the association cortices: Patterns, mechanisms, and implications

- for psychopathology. *Neuron* **109**, 2820–2846 (2021).
- [52] Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences* **113**, 12574–12579 (2016).
- [53] Xu, T. *et al.* Cross-species functional alignment reveals evolutionary hierarchy within the connectome. *NeuroImage* **223**, 117346 (2020).
- [54] Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A Hierarchy of Temporal Receptive Windows in Human Cortex. *Journal of Neuroscience* **28**, 2539–2550 (2008).
- [55] Gao, R., van den Brink, R. L., Pfeffer, T. & Voytek, B. Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *eLife* **9**, e61277 (2020).
- [56] Sydnor, V. J. *et al.* Intrinsic activity development unfolds along a sensorimotor–association cortical axis in youth. *Nature Neuroscience* **26**, 638–649 (2023).
- [57] Hespanha, J. P. *Linear systems theory* (Princeton university press, 2018).
- [58] Marblestone, A. H., Wayne, G. & Kording, K. P. Toward an Integration of Deep Learning and Neuroscience. *Frontiers in Computational Neuroscience* **10** (2016).
- [59] Ali, A., Ahmad, N., De Groot, E., Johannes Van Gerwen, M. A. & Kietzmann, T. C. Predictive coding is a consequence of energy efficiency in recurrent neural networks. *Patterns* **3**, 100639 (2022).
- [60] Millidge, B., Seth, A. & Buckley, C. L. Predictive Coding: a Theoretical and Experimental Review (2021). Publisher: arXiv Version Number: 4.
- [61] Bastos, A. M. *et al.* Canonical Microcircuits for Predictive Coding. *Neuron* **76**, 695–711 (2012).
- [62] Singer, W. Recurrent dynamics in the cerebral cortex: Integration of sensory evidence with stored knowledge. *Proceedings of the National Academy of Sciences* **118**, e2101043118 (2021).
- [63] Friston, K. Computational psychiatry: from synapses to sentience. *Molecular Psychiatry* (2022).
- [64] Jaeger, H. Tutorial on training recurrent neural networks, covering bppt, rtrl, ekf and the "echo state network" approach (2002).
- [65] Tanaka, G. *et al.* Recent advances in physical reservoir computing: A review. *Neural Networks* **115**, 100–123 (2019).
- [66] Van der Sande, G., Brunner, D. & Soriano, M. C. Advances in photonic reservoir computing. *Nanophotonics* **6**, 561–576 (2017).
- [67] Soriano, M. C. *et al.* Delay-based reservoir computing: noise effects in a combined analog and digital implementation. *IEEE transactions on neural networks and learning systems* **26**, 388–393 (2014).
- [68] Suárez, L. E., Richards, B. A., Lajoie, G. & Misić, B. Learning function from structure in neuromorphic networks. *Nature Machine Intelligence* (2021).
- [69] Parkes, L. *et al.* Asymmetric signaling across the hierarchy of cytoarchitecture within the human connectome. *Science Advances* **8**, eadd2185 (2022).
- [70] Pines, A. *et al.* Development of top-down cortical propagations in youth. *Neuron* S0896627323000387 (2023).
- [71] Baum, G. L. *et al.* Modular Segregation of Structural Brain Networks Supports the Development of Executive Function in Youth. *Current Biology* **27**, 1561–1572.e8 (2017).
- [72] Shine, J. M. *et al.* Computational models link cellular mechanisms of neuromodulation to large-scale neural dynamics. *Nature Neuroscience* **24**, 765–776 (2021).
- [73] Marder, E. & Thirumalai, V. Cellular, synaptic and network effects of neuromodulation. *Neural Networks* **15**, 479–493 (2002).
- [74] Bucher, D. & Goaillard, J.-M. Beyond faithful conduction: Short-term dynamics, neuromodulation, and long-term regulation of spike propagation in the axon. *Progress in Neurobiology* **94**, 307–346 (2011).
- [75] McCormick, D. A., Nestvogel, D. B. & He, B. J. Neuromodulation of Brain State and Behavior. *Annual Review of Neuroscience* **43**, 391–415 (2020).
- [76] Kim, Y. *et al.* Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. *Cell* **171**, 456–469.e22 (2017).
- [77] Burt, J. B. *et al.* Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature Neuroscience* **21**, 1251–1259 (2018).
- [78] Anderson, K. M. *et al.* Transcriptional and imaging-genetic association of cortical interneurons, brain function, and schizophrenia risk. *Nature Communications* **11**, 2889 (2020).
- [79] Deco, G. *et al.* How Local Excitation-Inhibition Ratio Impacts the Whole Brain Dynamics. *Journal of Neuroscience* **34**, 7886–7898 (2014).
- [80] Deco, G. *et al.* Dynamical consequences of regional heterogeneity in the brain’s transcriptional landscape. *Science Advances* **7**, eabf4752 (2021).
- [81] Gao, R., Peterson, E. J. & Voytek, B. Inferring synaptic excitation/inhibition balance from field potentials. *NeuroImage* **158**, 70–78 (2017).
- [82] Zhang, S. *et al.* In-vivo whole-cortex estimation of excitation-inhibition ratio indexes cortical maturation and cognitive ability in youth. preprint, bioRxiv (2023).
- [83] Larsen, B. *et al.* A developmental reduction of the excitation:inhibition ratio in association cortex during adolescence. *Science Advances* **8**, eabj8750 (2022).
- [84] Rocchi, F. *et al.* Increased fMRI connectivity upon chemogenetic inhibition of the mouse prefrontal cortex. *Nature Communications* **13**, 1056 (2022).
- [85] Markicevic, M. *et al.* Cortical Excitation:Inhibition Imbalance Causes Abnormal Brain Network Dynamics as Observed in Neurodevelopmental Disorders. *Cerebral Cortex* **30**, 4922–4937 (2020).
- [86] Shine, J. M., Aburn, M. J., Breakspear, M. & Poldrack, R. A. The modulation of neural gain facilitates a transition between functional segregation and integration in the brain. *eLife* **7**, e31130 (2018).
- [87] Shine, J. M. *et al.* Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nature Neuroscience* **22**, 289–296 (2019).

- [88] Bertolero, M. A., Yeo, B. T. T. & D’Esposito, M. The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences* **112** (2015).
- [89] Vázquez-Rodríguez, B. *et al.* Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences* **116**, 21219–21227 (2019).
- [90] Preti, M. G. & Van De Ville, D. Decoupling of brain function from structure reveals regional behavioral specialization in humans. *Nature Communications* **10**, 4747 (2019).
- [91] Mišić, B. *et al.* Network-Level Structure-Function Relationships in Human Neocortex. *Cerebral Cortex* **26**, 3285–3296 (2016).
- [92] Zamani Esfahlani, F., Faskowitz, J., Slack, J., Mišić, B. & Betzel, R. F. Local structure-function relationships in human brain networks across the lifespan. *Nature Communications* **13**, 2053 (2022).
- [93] Samuels, E. & Szabadi, E. Functional Neuroanatomy of the Noradrenergic Locus Coeruleus: Its Roles in the Regulation of Arousal and Autonomic Function Part II: Physiological and Pharmacological Manipulations and Pathological Alterations of Locus Coeruleus Activity in Humans. *Current Neuropharmacology* **6**, 254–285 (2008).
- [94] Wainstein, G. *et al.* Gain neuromodulation mediates perceptual switches: evidence from pupillometry, fMRI, and RNN Modelling. preprint, Research Square (2023).
- [95] Joshi, S. & Gold, J. I. Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences* **24**, 466–480 (2020).
- [96] Singleton, S. P. *et al.* Receptor-informed network control theory links LSD and psilocybin to a flattening of the brain’s control energy landscape. *Nature Communications* **13**, 5812 (2022).
- [97] Luppi, A. I. *et al.* Transitions between cognitive topographies: contributions of network structure, neuromodulation, and disease. preprint, bioRxiv (2023).
- [98] Aitken, K. & Mihalas, S. Neural population dynamics of computing with synaptic modulations. *Elife* **12**, e83035 (2023).
- [99] Jonas, E. & Kording, K. P. Could a neuroscientist understand a microprocessor? *PLoS computational biology* **13**, e1005268 (2017).
- [100] Torii, N. *et al.* Asic implementation of random number generators using sr latches and its evaluation. *EURASIP Journal on Information Security* **2016**, 1–12 (2016).
- [101] Samsonovich, A. & McNaughton, B. L. Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience* **17**, 5900–5920 (1997).
- [102] Fung, C. A., Wong, K. M. & Wu, S. A moving bump in a continuous manifold: a comprehensive study of the tracking dynamics of continuous attractor neural networks. *Neural Computation* **22**, 752–792 (2010).
- [103] Wimmer, K., Nykamp, D. Q., Constantinidis, C. & Compte, A. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature neuroscience* **17**, 431–439 (2014).
- [104] Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences* **79**, 2554–2558 (1982).
- [105] Strogatz, S. H. *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering* (CRC press, 2018).
- [106] Ramsauer, H. *et al.* Hopfield networks is all you need. *arXiv preprint arXiv:2008.02217* (2020).
- [107] Storkey, A. Increasing the capacity of a hopfield network without sacrificing functionality. In *Artificial Neural Networks—ICANN’97: 7th International Conference Lausanne, Switzerland, October 8–10, 1997 Proceedings 7*, 451–456 (Springer, 1997).
- [108] Smith, L. M., Kim, J. Z., Lu, Z. & Bassett, D. S. Learning continuous chaotic attractors with a reservoir computer. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **32** (2022).
- [109] Nichols, M. J. & Newsome, W. T. Middle temporal visual area microstimulation influences veridical judgments of motion direction. *Journal of Neuroscience* **22**, 9530–9540 (2002).
- [110] Jazayeri, M. & Ostojic, S. Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Current opinion in neurobiology* **70**, 113–120 (2021).
- [111] Howarth, C., Gleeson, P. & Attwell, D. Updated Energy Budgets for Neural Computation in the Neocortex and Cerebellum. *Journal of Cerebral Blood Flow & Metabolism* **32**, 1222–1232 (2012).
- [112] Giesl, P. & Hafstein, S. Review on computational methods for lyapunov functions. *Discrete and Continuous Dynamical Systems-B* **20**, 2291–2331 (2015).
- [113] Bellman, R. Vector lyapunov functions. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control* **1**, 32–34 (1962).
- [114] Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. Determining lyapunov exponents from a time series. *Physica D: nonlinear phenomena* **16**, 285–317 (1985).
- [115] Petanjek, Z. *et al.* Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences* **108**, 13281–13286 (2011).
- [116] Averbeck, B. B. Pruning recurrent neural networks replicates adolescent changes in working memory and reinforcement learning. *Proceedings of the National Academy of Sciences* **119**, e2121331119 (2022).
- [117] Moler, C. B. Iterative refinement in floating point. *Journal of the ACM (JACM)* **14**, 316–321 (1967).
- [118] Ypma, T. J. Historical development of the newton–raphson method. *SIAM review* **37**, 531–551 (1995).
- [119] Aragón Artacho, F. J., Campoy, R. & Tam, M. K. The douglas–rachford algorithm for convex and non-convex feasibility problems. *Mathematical Methods of Operations Research* **91**, 201–240 (2020).
- [120] Flesch, T., Juechems, K., Dumbalska, T., Saxe, A. & Summerfield, C. Orthogonal representations for robust context-dependent task performance in brains and neural networks. *Neuron* **110**, 1258–1270 (2022).



- [121] Gillespie, A. K. *et al.* Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron* **109**, 3149–3163 (2021).
- [122] Rajalingham, R., Sohn, H. & Jazayeri, M. Dynamic tracking of objects in the macaque dorsomedial frontal cortex. *bioRxiv* 2022–06 (2022).
- [123] Rajalingham, R., Piccato, A. & Jazayeri, M. Recurrent neural networks with explicit representation of dynamic latent variables can mimic behavioral patterns in a physical inference task. *Nature Communications* **13**, 5865 (2022).
- [124] Ercsey-Ravasz, M. & Toroczkai, Z. Optimization hardness as transient chaos in an analog approach to constraint satisfaction. *Nature Physics* **7**, 966–970 (2011).
- [125] Molnar, B., Toroczkai, Z. & Ercsey-Ravasz, M. Continuous-time neural networks without local traps for solving boolean satisfiability. In *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*, 1–6 (IEEE, 2012).
- [126] Yamashita, H., Suzuki, H., Toroczkai, Z. & Aihara, K. Bounded continuous-time satisfiability solver. In *International Symposium on Nonlinear Theory and its Applications (NOLTA2019)* (2019).
- [127] Li, C. & MacLennan, B. J. Continuous-time systems for solving 0-1 integer linear programming feasibility problems. *arXiv preprint arXiv:1905.04612* (2019).
- [128] Ding, Y., Li, Y., Xiao, M., Wang, Q. & Li, D. A high order neural network to solve n-queens problem. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 1–6 (IEEE, 2010).
- [129] Mastrogiuseppe, F. & Ostojic, S. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron* **99**, 609–623 (2018).
- [130] Eliasmith, C. & Anderson, C. H. *Neural engineering: Computation, representation, and dynamics in neurobiological systems* (MIT press, 2003).
- [131] DeWolf, T., Jaworski, P. & Eliasmith, C. Nengo and low-power ai hardware for robust, embedded neuro-robotics. *Frontiers in Neurobotics* **14**, 568359 (2020).
- [132] Graves, A. *et al.* Hybrid computing using a neural network with dynamic external memory. *Nature* **538**, 471–476 (2016).
- [133] Tkačik, J. & Kordik, P. Neural turing machine for sequential learning of human mobility patterns. In *2016 International joint conference on neural networks (IJCNN)*, 2790–2797 (IEEE, 2016).
- [134] Fawzi, A. *et al.* Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature* **610**, 47–53 (2022).
- [135] Kim, J. Z., Lu, Z., Nozari, E., Pappas, G. J. & Basset, D. S. Teaching recurrent neural networks to infer global temporal structure from local examples. *Nature Machine Intelligence* **3**, 316–323 (2021).
- [136] Valente, A., Ostojic, S. & Pillow, J. W. Probing the relationship between latent linear dynamical systems and low-rank recurrent neural network models. *Neural computation* **34**, 1871–1892 (2022).
- [137] Rajan, K., Harvey, C. D. & Tank, D. W. Recurrent network models of sequence generation and memory. *Neuron* **90**, 128–142 (2016).
- [138] Kepple, D., Engelken, R. & Rajan, K. Curriculum learning as a tool to uncover learning principles in the brain. In *International Conference on Learning Representations* (2022).
- [139] Larsen, B., Sydnor, V. J., Keller, A. S., Yeo, B. T. & Satterthwaite, T. D. A critical period plasticity framework for the sensorimotor–association axis of cortical neurodevelopment. *Trends in Neurosciences* S0166223623001674 (2023).
- [140] Kim, N. Y. & Kastner, S. A biased competition theory for the developmental cognitive neuroscience of visuospatial attention. *Current Opinion in Psychology* **29**, 219–228 (2019).
- [141] Tervo-Clemmens, B. *et al.* A Canonical Trajectory of Executive Function Maturation During the Transition from Adolescence to Adulthood. preprint, PsyArXiv (2022).
- [142] Larsen, B. & Luna, B. Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neuroscience & Biobehavioral Reviews* **94**, 179–195 (2018).
- [143] García-Cabezas, M. Á., Zikopoulos, B. & Barbas, H. The structural model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Structure and Function* **224**, 985–1008 (2019).
- [144] García-Cabezas, M. Á., Hacker, J. L. & Zikopoulos, B. A protocol for cortical type analysis of the human neocortex applied on histological samples, the atlas of von economo and koskinas, and magnetic resonance imaging. *Frontiers in Neuroanatomy* **14**, 576015 (2020).
- [145] Paquola, C. *et al.* Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLOS Biology* **17**, e3000284 (2019).
- [146] Barbas, H. General Cortical and Special Prefrontal Connections: Principles from Structure to Function. *Annual Review of Neuroscience* **38**, 269–289 (2015).
- [147] Markov, N. T. *et al.* Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology* **522**, 225–259 (2014).
- [148] Beul, S. F. Towards a "canonical" agranular cortical microcircuit. *Frontiers in Neuroanatomy* **8** (2015).
- [149] Huttenlocher, P. R., de Courten, C., Garey, L. J. & Van der Loos, H. Synaptogenesis in human visual cortex — evidence for synapse elimination during normal development. *Neuroscience Letters* **33**, 247–252 (1982).
- [150] Peter R., H. Synaptic density in human frontal cortex — Developmental changes and effects of aging. *Brain Research* **163**, 195–205 (1979).
- [151] Semple, B. D., Blomgren, K., Gimlin, K., Ferriero, D. M. & Noble-Haeusslein, L. J. Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Progress in Neurobiology* **106–107**, 1–16 (2013).
- [152] Buckner, R. L. & Krienen, F. M. The evolution of distributed association networks in the human brain. *Trends in Cognitive Sciences* **17**, 648–665 (2013).
- [153] Achterberg, J., Akarca, D., Strouse, D. J., Duncan, J. & Astle, D. E. Spatially embedded recurrent neural

- networks reveal widespread links between structural and functional neuroscience findings. *Nature Machine Intelligence* (2023). URL <https://www.nature.com/articles/s42256-023-00748-9>.
- [154] Sporns, O. & Betzel, R. F. Modular Brain Networks. *Annual Review of Psychology* **67**, 613–640 (2016).
- [155] Bassett, D. S. & Bullmore, E. T. Small-World Brain Networks Revisited. *The Neuroscientist* **23**, 499–516 (2017).
- [156] Tanner, J., L., S. M., Coletta, L., Gozzi, A. & Betzel, R. F. Functional connectivity modules in recurrent neural networks: function, origin and dynamics (2023). URL <https://arxiv.org/abs/2310.20601>. Publisher: arXiv Version Number: 1.
- [157] Fries, P. Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual review of neuroscience* **32**, 209–224 (2009).
- [158] Loebel, A. & Tsodyks, M. Computation by ensemble synchronization in recurrent networks with synaptic depression. *Journal of computational neuroscience* **13**, 111–124 (2002).
- [159] Li, Z., Duan, Z., Chen, G. & Huang, L. Consensus of multiagent systems and synchronization of complex networks: A unified viewpoint. *IEEE Transactions on Circuits and Systems I: Regular Papers* **57**, 213–224 (2009).
- [160] Rulkov, N. F., Sushchik, M. M., Tsimring, L. S. & Abarbanel, H. D. Generalized synchronization of chaos in directionally coupled chaotic systems. *Physical Review E* **51**, 980 (1995).
- [161] Pecora, L. M. & Carroll, T. L. Synchronization in chaotic systems. *Physical review letters* **64**, 821 (1990).
- [162] Lu, Z. & Bassett, D. S. Invertible generalized synchronization: A putative mechanism for implicit learning in neural systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30** (2020).