Distributed Representations Enable Robust Multi-Timescale Symbolic Computation in Neuromorphic Hardware

Author(s):
Cotteret, Madison; Greatorex, Hugh; Renner, Alpha; Chen, Junren; Neftci, Emre; Wu, Huaqiang; Indiveri, Giacomo; Ziegler, Martin; Chicca, Elisabetta

Publication date:
2025-01-13

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000716404

Originally published in:
arXiv, https://doi.org/10.48550/ARXIV.2405.01305

DISTRIBUTED REPRESENTATIONS ENABLE ROBUST MULTI-TIMESCALE SYMBOLIC COMPUTATION IN NEUROMORPHIC HARDWARE

Madison Cotteret[1,2,3,*] **Hugh Greatorex**[1,2], **Alpha Renner**[4], **Junren Chen**[5],
Emre Neftci[4], **Huaqiang Wu**[6], **Giacomo Indiveri**[5], **Martin Ziegler**[7], **Elisabetta Chicca**[1,2]

ABSTRACT

Programming recurrent spiking neural networks (RSNNs) to robustly perform multi-timescale computation remains a difficult challenge. To address this, we describe a single-shot weight learning scheme to embed robust multi-timescale dynamics into attractor-based RSNNs, by exploiting the properties of high-dimensional distributed representations. We embed finite state machines into the RSNN dynamics by superimposing a symmetric autoassociative weight matrix and asymmetric transition terms, which are each formed by the vector binding of an input and heteroassociative outer-products between states. Our approach is validated through simulations with highly non-ideal weights; an experimental closed-loop memristive hardware setup; and on Loihi 2, where it scales seamlessly to large state machines. This work introduces a scalable approach to embed robust symbolic computation through recurrent dynamics into neuromorphic hardware, without requiring parameter fine-tuning or significant platform-specific optimisation. Moreover, it demonstrates that distributed symbolic representations serve as a highly capable representation-invariant language for cognitive algorithms in neuromorphic hardware.

1 Introduction

Neuromorphic computing promises to match the efficiency of information processing in the brain by emulating biological neural and synaptic dynamics in hardware [1]. In contrast to more conventional artificial neural networks (ANNs), biological neurons have rich internal temporal dynamics and interact sparsely with unary events (spikes). Performing multi-timescale tasks – such as motor planning and execution – with biologically-plausible spiking neurons requires information to be retained and processed for periods much longer than the timescales available in individual neurons and synapses [2, 3]. Information must then be represented by the collective dynamics of recurrently-connected populations of neurons instead. Programming recurrent spiking neural networks (RSNNs) to perform long-timescale tasks with short-timescale neurons remains a formidable challenge however, due to the conflicting demands that the network should react quickly to input but otherwise be stable on long timescales [4].

In the face of these challenges, a possible approach would be to apply gradient-based iterative learning algorithms to train the network parameters to solve a particular task [5–7]. However, besides being computationally intensive and biologically implausible, such approaches are not well-suited to tasks with long timescales, due to the need to backpropagate gradients into the distant (potentially infinite) past [8, 9]. Furthermore, these approaches do not have

[1] Bio-Inspired Circuits and Systems (BICS) Lab, Zernike Institute for Advanced Materials, University of Groningen, Netherlands.
[2] Groningen Cognitive Systems and Materials Center (CogniGron), University of Groningen, Netherlands.
[3] Micro- and Nanoelectronic Systems (MNES), Technische Universität Ilmenau, Germany.
[4] Forschungszentrum Jülich, Germany.
[5] Institute of Neuroinformatics, University of Zürich and ETH Zürich, Switzerland.
[6] School of Integrated Circuits, Tsinghua University, Beijing, China.
[7] Energy Materials and Devices, Department of Materials Science, Kiel University, Germany.
[*] Corresponding author: m.cotteret@rug.nl

guarantees of stability and robustness, compared to those afforded by much simpler biologically-inspired models like the Hopfield attractor network [10, 11].

If the trained networks are to be deployed on mixed-signal neuromorphic hardware, one must usually perform calibrating chip-in-the-loop training, lest device-specific nonidealities absent during training lead to a catastrophic degradation in network performance upon deployment [6, 12–15]. To overcome these difficulties and to shift to more robust training procedures, it has been suggested that a more appropriate level of description for programming neuromorphic hardware may be at the level of distributed high-dimensional patterns of neuron activity, rather than individual neuron and synapse parameters [16]. With vector symbolic architectures (VSAs) – also known as hyperdimensional computing (HDC) – arbitrary symbolic data structures can be represented by high-dimensional random vectors, known as *hypervectors*, such that information is distributed across the entire vector [17–20]. These representations are then not dependent upon the precise functioning of individual components, which is a critical property for robust functioning in neuromorphic hardware and biology [21–24].

In this work, we show how distributed representations can be leveraged to program arbitrary state machines into RSNNs in a scalable and robust manner. For each state, we add an autoassociative outer product to the recurrent weight matrix to store it as a fixed-point attractor in the RSNN [10]. Additional heteroassociative outer product terms are then superimposed, which are each responsible for storing an input-triggered transition between attractor states. Neural state machines have already been shown to be important computational primitives for reproducing cognitive behaviours in neuromorphic agents [25], and have been used in the context of solving constraint satisfaction problems [26].

We first embed multiple state machines into simulated RSNNs with nonideal synaptic weights. To validate these simulations in neuromorphic hardware, we create a proof-of-concept implementation using a memristive crossbar setup run in a closed-loop with simulated neurons. Finally, we implement these networks on Intel's asynchronous digital neuromorphic research chip Loihi 2 [27], to demonstrate the algorithm's scalability.

This work demonstrates the capability of distributed symbolic representations as a functional abstraction layer for neuromorphic hardware [16]. It permits the description of complex RSNN behaviours from a high level while being invariant to the choice of underlying neural representation, and is highly robust to hardware-specific nonidealities. It is a step towards a general framework for interoperability of cognitive algorithms across varying neuromorphic hardware architectures, without requiring significant optimisation or fine-tuning upon deployment to each particular platform.

2 Results

State machines are perhaps the most comprehensive model of multi-timescale dynamics, as they react quickly to input but are otherwise stable on indeterminately-long timescales. They also have a clear utility for building more complex models of cognitive behaviours [25, 26, 28]. We focus on the task of embedding deterministic finite automata (DFAs) into RSNNs by expressing the DFA in terms of high-dimensional attractor states and transitions between them. For every state q in the finite set of states Q, we embed an associated fixed-point attractor state in the RSNN dynamics that is stable on long timescales. When input is given to the network, the RSNN should quickly transition between attractor states according to the desired state transition function $F : Q \times S \to Q$ in the DFA, where S is the set of all input symbols.

We generate independent hypervectors $\mathbf{q} \in \{0, 1\}^N$ to represent each DFA state q, and store them as attractors in the RSNN using a Hopfield-like outer-product learning rule [11, 29, 30]. The hypervectors follow a sparse block structure, i.e., the N elements are split up into M blocks of equal length L, and exactly one entry in each block is nonzero [31, 32].

We use sparse rather than dense vectors to benefit from the increased capacity in attractor networks [29, 30, 33] and distributed rather than localist representations for improved robustness properties [34]. We use a sparse block structure due to the simplicity of the required neural activation function, and compatibility with the VSA framework [32, 35]. Additionally, we move closer to a more plausible model of biological neural dynamics, where sparse representations enforced by local competitive mechanisms are abundant [36, 37].

For every state transition, we embed additional outer product terms such that if the correct subset of neurons is masked by external inhibition, the RSNN will transition between attractor states (see Methods). The terms are of the form $(\mathbf{q} - \mathbf{q}_0)(\mathbf{q}_0 \circ \overline{\mathbf{s}})^\intercal$ where $\overline{\mathbf{s}}$ is another (bipolar) hypervector, and "$\circ$" is a Hadamard product, acting as a hypervector binding operation. Without input to the network, these terms are pseudo-orthogonal to the network state, i.e. $\langle(\mathbf{q}_0 \circ \overline{\mathbf{s}}) \cdot \mathbf{q}_0\rangle = 0$ and so have negligible effect. When the network is masked by the correct (binary) hypervector $\mathbf{s}$ however (effectively an unbinding operation), they become similar to the network state, i.e. $(\mathbf{q}_0 \circ \overline{\mathbf{s}}) \cdot (\mathbf{q}_0 \wedge \mathbf{s}) = \frac{1}{2}M$ where

"$\wedge$" is component-wise AND, implementing the masking. This allows us to effectively modulate the strength of superimposed terms by masking neurons in the RSNN [22].

We then wish to program this weight matrix onto neuromorphic hardware to reliably and robustly realise the desired dynamics without requiring significant optimisation or adaptation to the target hardware. In contrast, deployment of RSNNs to neuromorphic hardware usually requires some degree of chip-in-the-loop training or post-deployment fine-tuning to ensure robustness to hardware-specific nonidealities [6, 12, 13, 38, 39]. This requirement severely limits the ability of neuromorphic hardware to be deployed at scale.

We demonstrate the reliability and generality of our approach by running experiments in three SNN environments. First in simulation, where we show the ideal functioning of the RSNN even with considerably nonideal weights. Second, a closed-loop memristive hardware setup with 64×64 resistive RAM (RRAM) devices acting as the synaptic weights, demonstrating the suitability of the approach for novel beyond-CMOS memristor-integrated neuromorphic hardware. Third, we demonstrate on Intel's neuromorphic research chip Loihi 2 that the approach scales up seamlessly to large state machines.

2.1 SNN simulation demonstrates robustness to noisy weights

The simulated RSNN consists of N leaky integrate-and-fire (LIF) neurons, with all-to-all synaptic connectivity, except between neurons in the same block, which have winner-take-all (WTA) connectivity between themselves (see Methods). This enforces that only one neuron in each block may spike at once, thereby restricting the neural activity to have a sparse block code structure.

Although the method generalises to arbitrary DFAs [22], we chose to implement a 23-state DFA which can perform modular division of binary numbers (Fig. 1), as it is a clear example of useful computation. Two inputs, s_0 and s_1, with associated random hypervectors $\mathbf{s}_0$ and $\mathbf{s}_1$, are input to the network to represent "0" and "1", respectively. A sequence of inputs thus corresponds to a dividend d in binary representation, and the final state corresponds to the result of the modular division $d \bmod 23$. The full state transition function is described by the transitions $q_n \xrightarrow{s_0} q_{(2n \bmod 23)}$ and $q_n \xrightarrow{s_1} q_{(2n+1 \bmod 23)}$ for all states q_n, $n \in \{0, \ldots, 22\}$. The weight matrix to implement this DFA was constructed as described in Methods. We then stochastically binarized the ideal weights and added independently-sampled Gaussian noise of standard deviation 0.5 to each binary weight value, such that the two weight distributions were highly overlapping. This demonstrates robustness to low-precision and noisy weights, inherent in mixed-signal neuromorphic hardware platforms. Despite these nonidealities, the RSNN performs the correct walk between attractor states for every input sequence (Fig. 1). The network is not reliant on a specific input timing scheme, as long as the inputs are given for long enough for a transition to take place (Fig. 2). A similar 300-state DFA with ideal weights (Fig. S3) was constructed in the same way, to demonstrate the seamless scaling to large state machines.

2.2 Memristive crossbar hardware implementation

To demonstrate that the approach is robust to nonidealities introduced by memristive devices, we run a closed-loop experiment with synapses on a memristive crossbar and neurons in simulation (Fig. 3). Whenever the simulated neurons spiked, the memristive crossbar was read by applying the binary vector of spiking neurons as a voltage input vector to the crossbar via the on-board DACs. The readout currents were read by the ADCs and then fed back into the simulation as postsynaptic currents. With this hybrid "computer in the loop" approach, we demonstrate the readiness of the algorithm for implementation on emerging hybrid SNN-memristor platforms [40].

We used an RRAM crossbar with 4096 devices. Due to this size constraint, we implemented smaller DFAs than in simulation. The weights were mapped to ternary values by thresholding, and then written to three conductance states on the RRAM devices. Although the device programming procedure in principle supports more weight states – limited by the precision of the ADCs (12-bit) – relaxation effects in the filamentary RRAM cause the conductance values to disperse in the milliseconds to seconds after programming, reducing the effective precision of the weights [41, 42]. After this dispersion, the RRAM devices store the synaptic weights in a non-volatile fashion [43]. We first implemented a DFA representing a 4-state counter. The RSNN performs the correct walk between attractor states, despite the fixed-pattern-noise and trial-to-trial weight nonidealities in the crossbar current readouts (Fig. 4). We then implemented another 4-state DFA but with two inputs, following an identical procedure. For every input sequence tested, the network performs the correct walk between attractor states (Fig. S4).

2.3 Large state machines on digital neuromorphic hardware

To demonstrate that the approach scales to large state machines on neuromorphic hardware and can be easily transferred to different types of neuromorphic systems, we implemented the RSNN on Intel's neuromorphic research chip

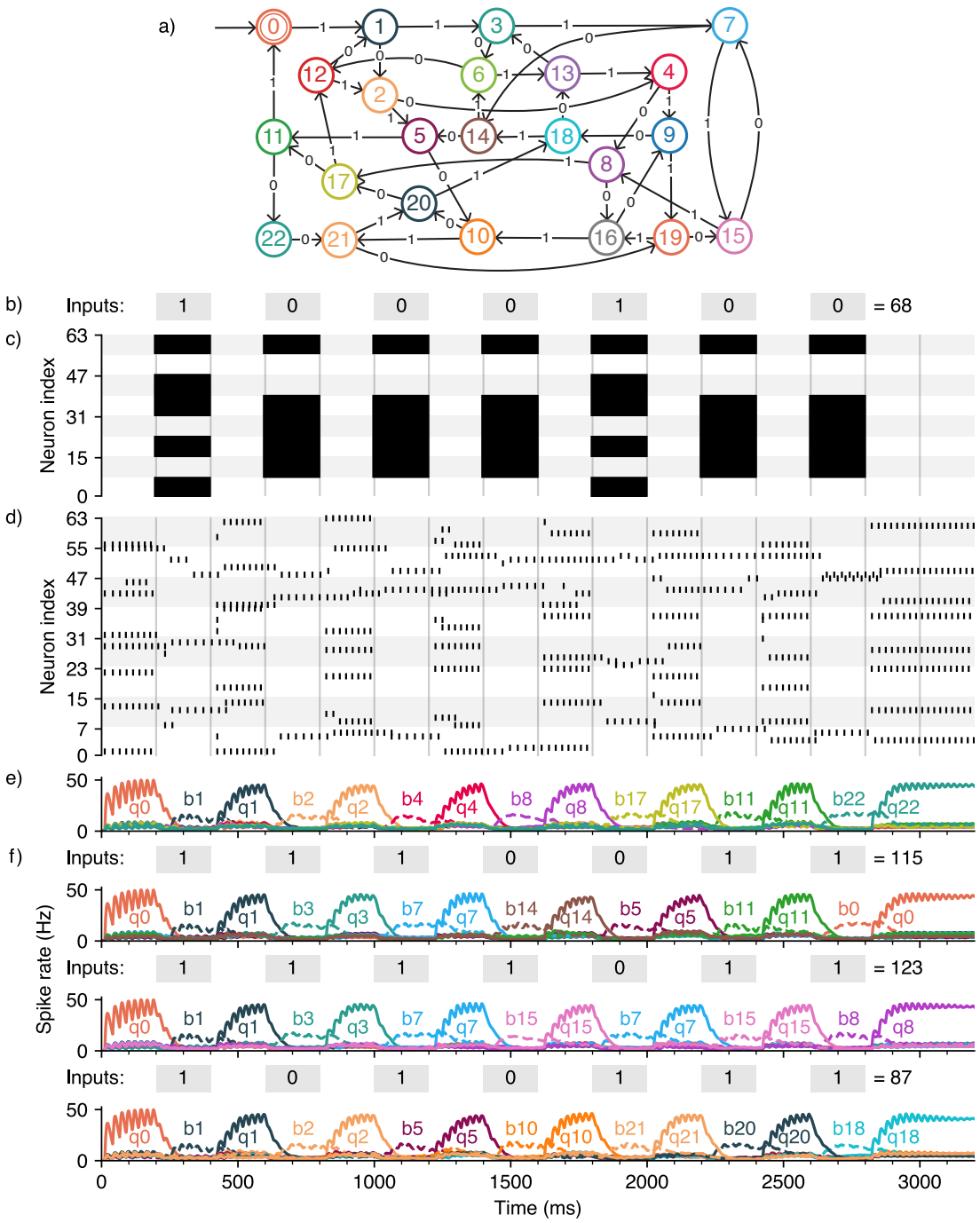

Fig. 1: An RSNN performing a walk on a 23-state DFA, using noisy 1-bit weights. **a)** The embedded DFA. If a binary number is input most-significant-bit first, the final state indicates the result of the input $\bmod\ 23$. **b)** The symbolic input to the RSNN. **c)** The input vector to the RSNN at any time, which each masks out half of the neurons therein. Each different input s corresponds to a different hypervector $\mathbf{s}$. Only the first 64 neurons are shown. **d)** A spike raster plot of activity within the RSNN. When the input to the network is constant, the network stabilises in an attractor state. When the input changes, some blocks are masked out or revealed, which may cause transitions to a new attractor state. The block-WTA mechanism ensures that only one neuron is persistently active in each block. **e)** The kernel-filtered mean firing rate of the neurons active in the attractor states $\mathbf{q}_0, \ldots, \mathbf{q}_{22}$, as well as the transition-facilitating bridge states $\mathbf{b}_0, \ldots, \mathbf{b}_{22}$. The bridge states are represented by dashed lines and coloured the same as their corresponding q attractor states. We here serially input the number 68 in binary format, and the RSNN correctly halts in the q_{22} state. **f)** The same RSNN is given different sequences of inputs, and the RSNN performs the correct walk between attractor states in all cases.

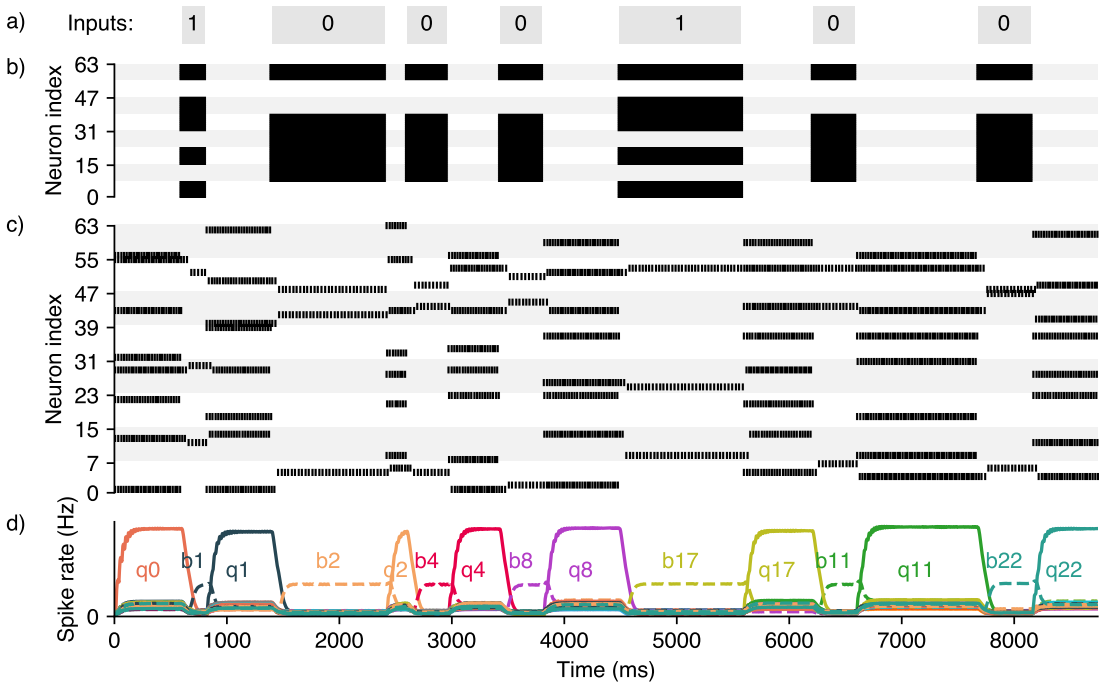

Fig. 2: The simulated RSNN with an irregular input timing scheme. **a)** The symbolic input to the network. Inputs were given for between 200 ms and 1000 ms. **b)** The corresponding input hypervector masking the neurons. **c)** A raster plot of spike activity in the RSNN. **d)** The mean firing rate of the neurons in each attractor state. The network is not reliant on an exact input timing scheme to perform the correct sequence of transitions. There is however still a minimum duration for which an input (or lack thereof) must be given, determined by the synaptic time constant.

Loihi 2. We implemented a functionally equivalent block-WTA mechanism with shunting inhibition synapses using a custom neuron microcode. This is required to give each neuron both a conventional integrative synapse and a shunting inhibition synapse. When a shunting synapse is stimulated, it resets the membrane potential of the target neuron to a subthreshold value (Fig. S5), implementing the within-block WTA and the block-wise masking of the RSNN by input. The integrative synapses implement the recurrent weight matrix as specified in Methods. We implemented the same 23-state DFA as in simulation (Fig. 1), now quantizing the ideal weights to the available 8-bit values. The network performs the correct walk between attractor states reliably for all given input sequences (Fig. 5).

2.4 Network capacity

Since the RSNNs used in this work are based upon a single-layer attractor network structure, we are limited by the theoretical memory capacity of such networks. That is, for a given network size N and block length L (determining the activation sparsity $1/L$), a maximum number of attractors can be embedded before unwanted cross-talk between patterns leads to a breakdown in attractor dynamics [11]. This limits the maximum size of DFA that can be embedded for a fixed N. For dense Hopfield networks, this results in the well-known linear capacity limit of $P < 0.14N$, where P is the number of storable patterns [10]. For sparse attractor networks, the capacity instead scales almost quadratically as $P \propto N^2/(\log N)^2$, provided the sparsity also scales appropriately [29, 30, 33]. We conducted non-spiking simulations to numerically investigate whether these scaling relations also hold with our approach, for both real ideal and stochastically binarised weights. We found that the capacity is proportional to $N^2/(\log N)^2$ in both cases, while the network capacity in the binary weight case is approximately $4\times$ smaller than with ideal weights (Fig. 6).

3 Discussion

It is widely believed that recurrent attractor dynamics are fundamental to the brain's ability to temporarily represent information in stable neuron firing patterns, without the need for components with intrinsically-long timescales [44], and despite the nonidealities present in biological systems [3, 45–48]. Networks that switch between discrete attractor

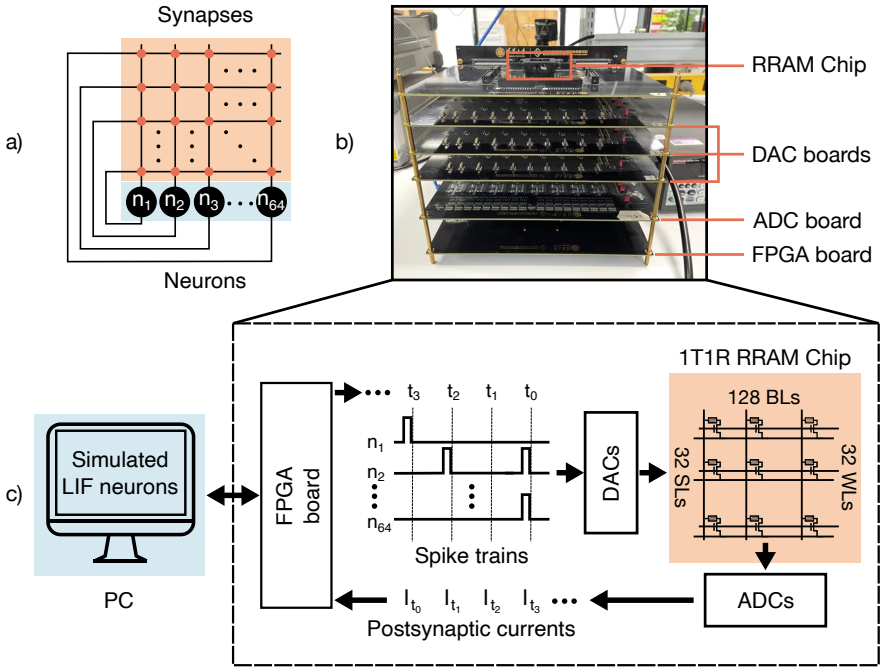

Fig. 3: The closed-loop experimental setup for running the RSNN using the 4K RRAM system. **a)** A simplified schematic of the 64-neuron RSNN. **b)** The physical RRAM measurement system, containing an FPGA board, three DAC boards, and one ADC board. **c)** The scheme of running the closed-loop experiment in which the LIF neurons are simulated on a PC. The FPGA board receives the control commands from the PC and operates all word-, source- and bit lines (WLs, SLs, BLs) of the RRAM chip in parallel. At each time step (t_n), the spikes generated by the neurons are sent to the synapse array on WLs, which are represented by voltage pulses. Meanwhile, a 0.2 V read voltage is applied to the SLs. The postsynaptic currents (I_{t_n}) produced by the RRAM cells accumulate on each BL, and the readout results are returned to the PC through the ADCs and FPGA.

states are used to explain neural recordings and are thought to underlie higher decision-making processes [28, 49–51]. These robust dynamics inspired the embedding of controlled attractor transition dynamics in RNNs, in particular in neuromorphic systems that mimic the brain's parallelism and asynchrony, but suffer from a similar inherent unreliability of single components as biological systems [52] (see [22] for a more comprehensive overview).

Distributed representations are scalable and robust

In previous efforts to embed attractor-based state machines into spiking neuromorphic hardware [25, 26, 53], each state was represented by a dedicated orthogonal (non-overlapping) population of neurons with a synaptic gating mechanism to trigger state transitions. Although this simplifies the state encoding, these localist state representations limit the robustness, flexibility, and scalability of the network in comparison to a distributed approach [34]. If each state is represented by a unique population of M neurons that are active for only one pattern, then a network of N neurons can represent at most N/M patterns. In contrast, sparse distributed single-layer attractor networks can store $N^2/(\log N)^2$ patterns, making more efficient use of the available synaptic resources [29, 30, 33]. Additionally, adding new states with localist representations requires adding or finding new neurons that are inactive for every other pattern. This is at odds with the mixed selectivity of neurons found throughout the brain [54].

We instead utilised the distributed representational framework of VSAs, in which high-dimensional random vectors can be composed to represent arbitrary data structures and algorithms [17]. The state-representing hypervectors thus have a nonzero probabilistic degree of overlap, but it is nonetheless possible to generate an exponential (in N) number of hypervectors for which significant overlap between any pair of hypervectors is unlikely (Section S1) [55, 56]. The sparsity of the hypervectors not only increases the network's memory capacity, but makes them inherently suitable for exploiting the efficiency of sparse activity in neuromorphic hardware [57].

The distributed representations also make the network highly robust to various nonidealities, such as noise, individual component failure and nonideal neuron dynamics. Their robustness to low-precision (quantized) and inaccurate synaptic weights is particularly relevant for memristive in-memory applications. This has prompted recent interest in combining VSAs with memristive crossbars [21, 58] and spiking neural networks [23, 59–61], where synaptic and

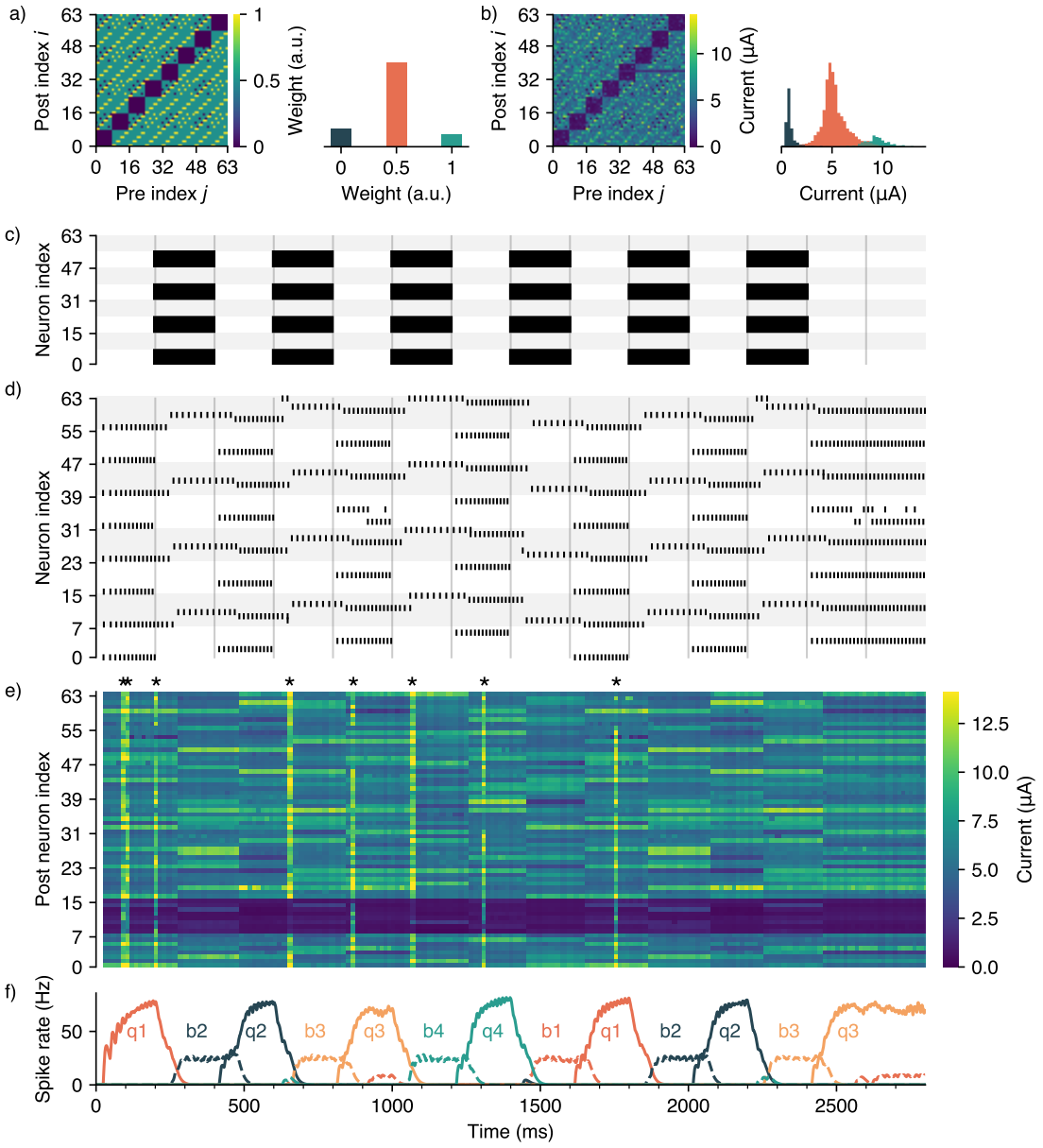

Fig. 4: A 4-state DFA embedded into an RSNN using a memristive crossbar with 64×64 RRAM devices as the synaptic weights. The DFA is described by $q_0 \xrightarrow{s} q_1 \xrightarrow{s} q_2 \xrightarrow{s} q_3 \xrightarrow{s} q_0$ for a single input s. **a)** The ternary weight matrix to be written to the RRAM crossbar, and a histogram of the values. **b)** Readout currents from each of the 64×64 RRAM devices after programming, and separate histograms for each of the ternary weight values. There is notable mismatch between the measured currents and the ideal values; a row of faulty devices giving almost no current; and devices giving anomalously large currents. **c)** The masking input to the network. **d)** A spike raster plot of the neurons within the RSNN. Due to the size constraints introduced by the crossbar, we chose the attractor hypervectors **q** to be orthogonal rather than random. **e)** Measured postsynaptic current readings from whenever a neuron in the second block spiked, chosen for the prominence of trial-to-trial current variation. The weights between neurons in the same block were programmed to the lowest weight, hence the horizontal band of low currents. At some times, multiple neurons fired within the same time step (labelled by $*$). **f)** The mean firing rates of the neurons in each attractor state. Despite the considerable nonidealities present, the RSNN performs the correct walk between attractor states.

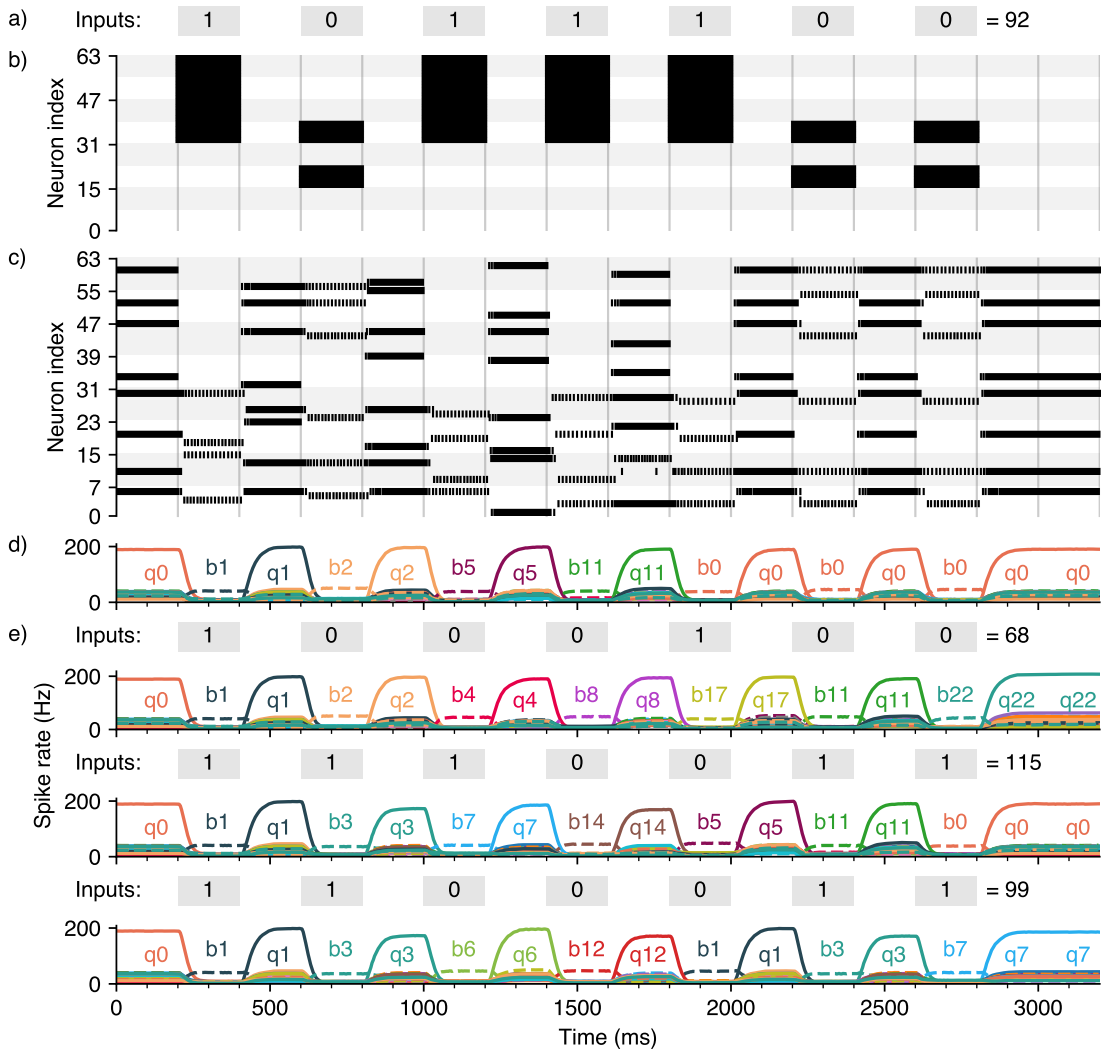

Fig. 5: The 23-state DFA on Intel's asynchronous digital neuromorphic research chip, Loihi 2. **a)** The input symbol to the network at any time, and **b)** the corresponding input vectors, which mask the network activity. **c)** A spike raster plot of the first 64 neurons. The shunting-inhibition WTA mechanism ensures that only one neuron in every block may spike at once (see Methods). **d)** Kernel firing-rate estimates of each q and b state, choosing arbitrarily that 1 time step on Loihi represents 1 ms. The sequence of inputs given corresponds to the binary representation of the number 92. The RSNN halts in the q_0 state, indicating correctly that 92 is divisible by 23. **e)** For all sequences of inputs, the correct walk between attractor states is performed.

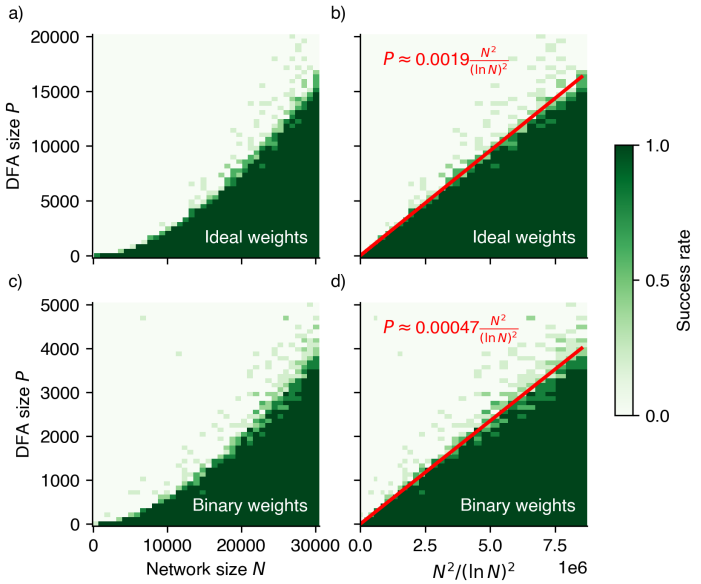

Fig. 6: Network capacity for varying network size N, in terms of the number of DFA states P that can reliably be emulated. Each DFA state has two outgoing edges, like in Fig. 1. **a)** A network with ideal weights displays an approximately quadratic capacity relation. **b)** The same data, but plotted against $N^2/(\log N)^2$, with a separation boundary obtained by a linear SVM. **c,d**) A network with stochastic binary weights displays the same scaling relation but reduced by a factor of ≈ 4. In both cases, the block length L was varied such that $L \propto N/\log N$, with reference values $L_0 = 8$, $N_0 = 2048$.

neural nonidealities are abundant. The results obtained from simulations and the closed-loop memristive crossbar setup confirm the system's reliable operation under hardware constraints, notably without requiring parameter fine-tuning, while the restriction of binary weights causes only a constant factor decrease in the network capacity.

Simultaneous auto- and heteroassociative memory

By exploiting the pseudo-orthogonality property of the hypervector binding operation, we were able to superimpose autoassociative and heteroassociative outer-product terms in the recurrent weight matrix without incurring considerable interference between them. Transitions are triggered by masking a subset of the neurons – constituting an effective unbinding operation – while in previous work, additional mechanisms were required to modulate between attractor and transition dynamical regimes [62–69]. Furthermore, the simplicity of this construction method means that the weight matrix could reasonably be learned via local Hebbian learning rules, such that states and transitions are learned and unlearned as needed to solve a particular task [70].

Fault-tolerant computation

Programming a certain DFA into an RSNN in one shot may be sufficient for some applications, however. In space, ionizing solar radiation causes unwanted bit-flips in systems using conventional charge-based memories, and so necessitates fault-tolerant computation schemes. Our system is intrinsically robust to such events since the attractor dynamics quickly reverse any erroneous flips in the neural state. Furthermore, RRAM-based memories are highly resistant to solar radiation, making memristive VSA-programmed attractor-based neuromorphic implementations of conventional algorithms potentially suitable for space and other high-energy applications [71].

The flexibility and robustness of the model enable a wide range of applications for neuromorphic hardware. It could be used to coordinate information flow in complex neuromorphic systems [53] or be directly incorporated into existing works leveraging VSA representations [72], such as online learning of robotics schemas [73] or general computation with distributed stack machines [16, 74].

Neuromorphic hardware abstraction

By utilizing distributed symbolic representations, we programmed an RSNN from an abstract higher level of description, which is separated from the exact details of the hardware. It relied on the existence of a fixed stereotypical

low-level neural motif (the block-WTA connectivity), but adjusted only the connections between neurons in different blocks, to easily program the same neural hardware for different purposes. This ability to abstract the function of lower layers and then build upon them relatively carefree is the foundation of conventional digital design approaches and bears resemblance to Marr's levels of analysis [16, 75].

The distributed representations of VSAs are especially suited as an abstraction layer to divorce high-level function from low-level neural dynamics, in part due to their invariance of the choice of hypervector representation [16, 17]. For example, in place of sparse block code representations, we could have instead used phasor representations [19, 76] or their sparse variants, which have an elegant connection to the periodic spiking of resonate-and-fire neurons [23, 59]. The same algorithm could be implemented on different neuromorphic hardware platforms, using whichever representation is best suited for the available neuron circuit. Neuromorphic computing would greatly benefit from a high-level neurally-inspired abstract programming language, which can be robustly implemented across different neuromorphic hardware platforms [16, 77, 78]. Although there have been many works towards such a unification [79–84], the majority of these focus on finding common ways to interface with different neuromorphic hardware platforms or to compose functional SNN modules, rather than on defining higher-level abstract primitives that are invariant of the neural representations on the hardware on which they are embedded.

This work demonstrates the capability of distributed representations for embedding computation through neural dynamics into neuromorphic hardware [16]. We embedded arbitrary state machines into RSNNs, as they are a clear example of multi-timescale computation: the switching between states is fast, but the autoassociative attractor dynamics ensure stability on long timescales [10, 11]. The extension of this approach to embedding general computational primitives beyond DFAs into neuromorphic RSNNs, such as continuous attractor networks [3], presents an exciting direction for future research. The fully-distributed representations ensure increasing robustness with increasing dimensionality, while the invariance to the choice of hypervector representation – many of which have elegant spiking implementations – may allow neuromorphic hardware to be programmed at a level that is independent of the underlying neural representations, but that can reliably and robustly be implemented on different neuromorphic hardware platforms without the need for fine-tuning.

4 Methods

4.1 VSA arithmetic

VSAs allow compositional data structures to be represented in a fully distributed manner, using high-dimensional random vectors, hypervectors, as the smallest units of representation [17–20]. They rely upon the fact that independently-generated hypervectors will very likely be approximately orthogonal to each other (the *pseudo-orthogonality* property) and so new hypervectors can easily be generated to represent new items. For most hypervector representations, the similarity between hypervectors is defined by the normalised inner product

$$\mathrm{sim}(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|_2 \|\mathbf{b}\|_2} \tag{1}$$

such that, for a set of independently-generated hypervectors $\{\mathbf{x}_\nu\}$ the pseudo-orthogonality property is then written as

$$\mathrm{sim}(\mathbf{x}_\nu, \mathbf{x}_\mu) \begin{cases} = 1 & \text{if} \quad \mu = \nu \\ \approx 0 & \text{otherwise} \end{cases} \tag{2}$$

Each VSA model is equipped with two operators to generate hypervectors which represent relations between other hypervectors, and thus also the symbols they represent. They are the superposition and binding operators, often corresponding simply to component-wise addition and multiplication respectively. The superposition operator produces a hypervector of the same dimensionality, which retains similarity to its operands. That is, if $\mathbf{a}$ and $\mathbf{b}$ are independent hypervectors, then $\mathrm{sim}(\mathbf{a} \oplus \mathbf{b}, \mathbf{a}) \approx \mathrm{sim}(\mathbf{a} \oplus \mathbf{b}, \mathbf{b}) \approx 1$ where "$\oplus$" represents a general superposition operator. Superpositions of hypervectors are thus used to efficiently represent sets of objects [55, 56].

The binding operator produces a hypervector that is dissimilar to both its operands. That is, if $\mathbf{a}$ and $\mathbf{b}$ are again independently-generated, then $\mathrm{sim}(\mathbf{a} \odot \mathbf{b}, \mathbf{a}) \approx \mathrm{sim}(\mathbf{a} \odot \mathbf{b}, \mathbf{b}) \approx 0$ where "$\odot$" is a general binding operation. Binding may be inverted by unbinding the same hypervector, i.e. $(\mathbf{a} \odot \mathbf{b}) \odot^{-1} \mathbf{b} = \mathbf{a}$. Superimposing pairs of bound hypervectors could then be used to create a hypervector representing a set of key-value pairs [18], for example, but more complex objects can similarly be represented such as graphs [17, 85], algorithms [74] or images [86].

In this work, we used two types of VSA models: dense bipolar hypervectors, also known as Multiply-Add-Permute (MAP) hypervectors, and Sparse Block Code (SBC) hypervectors [17]. MAP hypervectors are randomly sampled from

$\{-1, 1\}^N$, with approximately equal numbers of positive and negative entries. The binding operation for MAP hypervectors is the Hadamard product "$\circ$" (element-wise multiplication), which also serves as the inverse binding operation [20]. SBC hypervectors are binary-valued and have a sparse-block structure. That is, the N components are divided into M equally-sized blocks of length $L = N/M$, and we randomly choose exactly one component in each block to be 1-valued, with the rest 0 [31, 32]. By binding MAP and SBC hypervectors with a Hadamard product operation, we produce a bipolar sparse hypervector that is pseudo-orthogonal to the original SBC hypervector (see Section S4 for further discussion on the joint MAP-SBC VSA model). We can then superimpose weight matrices constructed using these hypervectors, each corresponding to a different network behaviour, without significant interference. Masking out some components constitutes an effective Hadamard unbinding operation. This operation allows that terms in the weight matrix bound to a MAP hypervector (and thus pseudo-orthogonal to all SBC hypervectors) may be unbound and align with the network state. As a result, state transitions in the RSNN can be triggered by selectively masking neurons with inhibitory input.

4.2 State machine embedding

A DFA is a finite state machine consisting of a set of states Q, an initial state $q_0 \in Q$, a finite set of input symbols S, and a state transition function $F : Q \times S \to Q$. In addition, a DFA denotes a subset $A \subseteq Q$ of its states as "accepting" states. A string of input symbols is then said to have been accepted by the DFA if it terminates in one of the acceptance states. As such, DFAs can be used to perform pattern matching – accepting or rejecting strings based on a predefined set of rules, regular expressions [87].

The DFA is embedded into the RSNN by first generating for each state $q \in Q$ an SBC hypervector $\mathbf{q} \in \{0, 1\}^N$ to represent it. For every input symbol $s \in S$, a dense binary hypervector $\mathbf{s} \in \{0, 1\}^N$ is generated, which, when input to the network, should trigger the correct transitions between attractor states. We restrict these hypervectors to be block-constant, with the same block length L as the SBC hypervectors, such that components in the same block have the same value. The 0 and 1 entries were chosen with equal probability. We will denote $\bar{\mathbf{s}} \in \{-1, 1\}^N$ as the equivalent bipolar representation of these hypervectors, given by $\bar{\mathbf{s}} = 2\mathbf{s} - 1$. For every state q, an additional SBC vector $\mathbf{b}$ is generated, which we refer to as its *bridge* state. They will be attractors of the RSNN dynamics only when there is input to the network. Transitions will be constructed such that, when we give input to the RSNN, it will not transition directly from a source state $\mathbf{q}$ to target state $\mathbf{q}'$, but rather will first transition from $\mathbf{q}$ to $\mathbf{b}'$, the bridge state belonging to the target state $\mathbf{q}'$. When the input is removed, the RSNN will finally flow from the $\mathbf{b}'$ to $\mathbf{q}'$ state, completing the transition. Splitting each transition up in this way ensures they are not dependent upon a specific input timing scheme (Fig. 2). The construction of the recurrent weight matrix $\mathbf{W} \in R^{N \times N}$ can be divided into three parts, first to embed the $\mathbf{q}$ states as attractors

$$\mathbf{W}_{\text{attr.}} = \sum_{q \in Q} (\mathbf{q} - f)(\mathbf{q} - f)^\mathsf{T} \tag{3}$$

where the coding level $f = 1/L$ is the fraction of nonzero elements in each $\mathbf{q}$ hypervector [29, 30]. We then embed the bridge states as

$$\mathbf{W}_{\text{brdg.}} = \sum_{q \in Q} \Big[(\mathbf{q} - f)(\mathbf{b} - f)^\mathsf{T} + \sum_{s \in S} (\mathbf{b} - \mathbf{q}) \big((\mathbf{b} - f) \circ \bar{\mathbf{s}} \big)^\mathsf{T} \Big] \tag{4}$$

where "$\circ$" is the Hadamard product binding operation. We are thus using an unorthodox mixture of the MAP [17, 20] and SBC [31, 32] VSA models (Section S4). These terms cause the bridge states $\mathbf{b}$ flow to their corresponding $\mathbf{q}$ states, except when a valid input is present, then they themselves become attractors of the RSNN's dynamics. To store transitions between states, outer product terms are then similarly added for each non-loop transition, the set of which we define as $E = \{(q, s, q') \text{ s.t. } q' = F(q, s),\ q' \neq q\}$. They are included as

$$\mathbf{W}_{\text{trans.}} = \sum_{\eta \in E} (\mathbf{b}' - \mathbf{q}) \big((\mathbf{q} - f) \circ \bar{\mathbf{s}} \big)^\mathsf{T} \tag{5}$$

where $\mathbf{q}$ is the attractor state for source state q for edge η, and $\mathbf{b}'$ the bridge state belonging to the target state $q' = F(q, s)$, as defined by E. The full weight matrix is then given by the superposition $\mathbf{W} = \mathbf{W}_{\text{attr.}} + \mathbf{W}_{\text{brdg.}} + \mathbf{W}_{\text{trans.}}$.

4.3 RSNN attractor dynamics

To understand how the described weight construction method results in the desired DFA being emulated, we coarsely model the RSNN by a neural state vector $\mathbf{z} \in \{0, 1\}^N$, indicating which neuron is active. We then consider the postsynaptic sum $\mathbf{h} = \mathbf{W}\mathbf{z}$, with and without input to the RSNN. It is instructive to work with the expectation of $\mathbf{h}$, where terms containing inner products between pseudo-orthogonal hypervectors can be discarded. We must, however,

keep in mind that for finite N, the postsynaptic sum will only approximate its expected value, due to the accumulation of nonzero inner products between realised hypervectors. See [55, 56] for a rigorous theoretical analysis. We first consider the dynamics when the network is in a DFA state, $\mathbf{z} = \mathbf{q}^\mu$. The expected postsynaptic sum $\langle \mathbf{h} \rangle$ is then given by

$$\begin{aligned} \langle \mathbf{h} \rangle &= \left\langle \mathbf{W}\mathbf{q}^\mu \right\rangle \\ &= \left\langle \mathbf{W}_{\text{attr.}}\mathbf{q}^\mu \right\rangle + \left\langle \mathbf{W}_{\text{brdg.}}\mathbf{q}^\mu \right\rangle + \left\langle \mathbf{W}_{\text{trans.}}\mathbf{q}^\mu \right\rangle \\ &\propto (\mathbf{q}^\mu - f) + \mathbf{0} + \mathbf{0} \\ &= \mathbf{q}^\mu - f \end{aligned} \tag{6}$$

such that the state $\mathbf{q}^\mu$ projects back to itself (Section S2). Neither $\mathbf{W}_{\text{brdge.}}$ nor $\mathbf{W}_{\text{trans.}}$ contribute to $\langle \mathbf{h} \rangle$, since they are constructed either from independent SBC hypervectors, or the same SBC hypervector $\mathbf{q}^\mu$ but bound to a bipolar hypervector $\bar{\mathbf{s}}$, both of which are pseudo-orthogonal to the state $\mathbf{q}^\mu$. Hence, in the large-N limit, and sufficiently close to the states $\mathbf{q}$, we can approximate the dynamics as being governed only by the $\mathbf{W}_{\text{attr.}}$ matrix. Since $\mathbf{W}_{\text{attr.}}$ is symmetric, this results in fixed-point attractor dynamics with an energy landscape consisting of stable minima at the states $\mathbf{q}$ (Section S3) [29, 30, 88].

4.4 RSNN transition dynamics

We now consider how $\langle \mathbf{h} \rangle$ changes when we apply an input $s^\lambda \in S$ to the RSNN, which we assume corresponds to a valid transition $\eta = (q^\mu, s^\lambda, q^{\mu\prime}) \in E$ from state q^μ to $q^{\mu\prime}$. Inputting s^λ means applying the corresponding hypervector $\mathbf{s}^\lambda \in \{0,1\}^N$ as a mask to the RSNN, such that all neurons in the blocks for which $\mathbf{s}^\lambda$ is zero are prevented from spiking. The expected postsynaptic sum is then given by

$$\begin{aligned} \langle \mathbf{h} \rangle &= \left\langle \mathbf{W}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \right\rangle \\ &= \left\langle \mathbf{W}_{\text{attr.}}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \right\rangle + \left\langle \mathbf{W}_{\text{brdg.}}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \right\rangle + \left\langle \mathbf{W}_{\text{trans.}}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \right\rangle \\ &\propto (\mathbf{q}^\mu - f) + \mathbf{0} + (\mathbf{b}^{\mu\prime} - \mathbf{q}^\mu) \\ &= \mathbf{b}^{\mu\prime} - f \end{aligned} \tag{7}$$

where $\wedge$ is a component-wise AND, implementing the masking operation (Section S2). Since the masked $\mathbf{q}^\mu$ state now does not project back to itself, but to $\mathbf{b}^{\mu\prime}$ (the bridge state belonging to the target $q^{\mu\prime}$ state), the RSNN transitions to the $\mathbf{b}^{\mu\prime}$ state. By applying $\mathbf{s}^\lambda$ as a mask to the RSNN, we caused heteroassociative terms to be projected out of the edge summation in $\mathbf{W}_{\text{trans.}}$, which push the RSNN from $\mathbf{q}^\mu$ to $\mathbf{b}^{\mu\prime}$.

A similar calculation can be done to show that the bridge states $\mathbf{b}$ are attractors only when the network is being masked by a valid input, i.e. $\langle \mathbf{W}(\mathbf{b}^{\mu\prime} \wedge \mathbf{s}) \rangle = \mathbf{b}^{\mu\prime} - f$. When the input is removed, however, this becomes $\langle \mathbf{W}\mathbf{b}^{\mu\prime} \rangle = \mathbf{q}^{\mu\prime} - f$, driving the RSNN towards the target state, completing the transition. By applying a sequence of inputs to the RSNN as masks (with pauses in between), we cause the RSNN to perform the desired sequence of transitions between attractor states. See Fig. S6 for a depiction of how this unfolds at the neural level of description. Including the bridge states in the state machine construction ensures that transitions are not dependent upon a specific input timing scheme, since an input must be given – and crucially then removed – to complete a transition (Fig. 2) [22]. If an input s were given that did not correspond to a valid transition for the current state q, no transition occurs.

Constructing attractor networks with transition dynamics by superimposing symmetric autoassociative terms and asymmetric heteroassociative terms, is not a new idea [62]. In previous works, to ensure that neither attractor nor transition dynamics dominated always, one had to either ensure a fine balance between the relative strengths of the two terms [63, 64, 89] or to introduce additional mechanisms to modulate their relative strengths, such as delayed synapses [65–67, 90] or short-term synaptic plasticity [68, 69]. By exploiting the properties of VSAs, we have achieved the same result without such mechanisms and without sacrificing robustness. By adding to $\mathbf{W}$ latent heteroassociative terms which were each bound to a bipolar hypervector $\bar{\mathbf{s}}$, in the absence of input, we could ignore their contribution to the RSNN dynamics, due to them being pseudo-orthogonal to the network state (Equation 6). However, when we mask the RSNN with the same $\mathbf{s}$ – an effective unbinding operation – these heteroassociative terms are revealed, and trigger the transition dynamics (Equation 7). Crucially, this is achieved without ever altering the synaptic weight matrix $\mathbf{W}$ during operation. We use the masking mechanism to realise an unbinding operation because it is not reliant on the synchronous arrival of inputs [22].

4.5 Spiking neuron model

We simulate leaky integrate-and-fire neurons (LIF), defined by

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\frac{1}{\tau_m}(u_i - u_{\text{rest}}) + \frac{1}{C}I_i(t) \tag{8}$$

where u_i is dynamical membrane voltage of the neuron i, τ_m and u_{rest} are the membrane time constant and rest potential respectively, I_i is the postsynaptic current to that neuron, and $C = 1\,\mathrm{F}$. When u exceeds the spiking threshold u_θ, the neuron emits a spike and u is quickly clamped to a subthreshold reset voltage (here $0\,\mathrm{mV}$) for an amount of time determined by the refractory period τ_{ref}. The postsynaptic currents $I_i(t)$ we modelled as second-order low-pass filters over spike inputs,

$$\begin{aligned} \tau_{\text{syn}}\frac{\mathrm{d}I_i}{\mathrm{d}t} &= -I_i + J_i(t) \\ \tau_{\text{syn}}\frac{\mathrm{d}J_i}{\mathrm{d}t} &= -J_i + \sum_{j=1}^{N} w_{ij} \sum_{\text{spks } k} \delta(t - t_j^k) \end{aligned} \tag{9}$$

where we introduce the synaptic time constant τ_{syn}, the Dirac delta function $\delta(\cdot)$, and we sum over the spike times t_j^k from all presynaptic neurons j, with synaptic weight w_{ij} measured in coulombs. This has solution

$$\begin{aligned} I_i(t) &= \sum_{j=1}^{N} w_{ij}\Big(K(t) \circledast \sum_k \delta(t - t_j^k)\Big)(t) \\ &= \big[\mathbf{W}\mathbf{z}(t)\big]_i \end{aligned} \tag{10}$$

where $K(t)$ is an exponentially-decaying alpha kernel with time constant τ_{syn}, $\circledast$ indicates temporal convolution, and $\mathbf{z}(t)$ defines the kernel-filtered spike activity. The current is continuous everywhere (in contrast to the more conventional first-order temporal filters [91]), which ensures that I_i is not affected by the precise timing of presynaptic spikes [92]. In the Loihi implementation, only a first-order synapse was used, so this is not strictly necessary, however.

To stabilise the network activity and ensure it always represents an SBC hypervector, we included local winner-take-all (WTA) connectivity between all neurons in the same block, as often done to enforce constraints in RSNNs [93]. It is implemented such that if any neuron spikes, then all neurons in the same block are forced into a refractory state (Fig. 7). Each block thus performs an argmax neural activation function on the vector of postsynaptic currents $\mathbf{I}(t) = \mathbf{W}\mathbf{z}(t)$ integrated between spike times. This behaviour is easily implementable in analog VLSI [94, 95]. The same mechanism is used to mask out a subset of neurons within the RSNN with external input. The firing rates $\mathbf{z}(t)$ and postsynaptic currents $\mathbf{I}(t)$ approximate the discrete dynamics of the network state $\mathbf{z}$ and the postsynaptic sum $\mathbf{h}$ respectively, as described in Section 4.3.

4.6 Spiking simulation

The neuron equations were simulated in the *Brian 2* SNN simulator [96], using the Euler ODE solver with a default time step of $0.05\,\mathrm{ms}$. The weight matrix to embed the chosen DFA was constructed as described in Methods, with $N = 2048$ neurons and blocks of length $L = 8$. These ideal weights were transformed to binary values, to match that biological synapses may, and neuromorphic synapses often do have very few bits of precision [97]. This binarisation was performed stochastically, via

$$w_{ij}^{\text{binary}} = \begin{cases} 1 \text{ w.p. } (1 + \exp[-\beta(w_{ij} - \langle w\rangle)/\sigma_w])^{-1} \\ 0 \text{ otherwise} \end{cases} \tag{11}$$

where $\langle w\rangle$ and σ_w are the mean and standard deviation of the ideal weight values respectively, and the steepness parameter β is heuristically set to $\beta = 2$. This is a common way to embed analog values into binary-valued memristive devices [98]. We then add noise to each weight value, to emulate that the distribution of the binary weight states may themselves be imperfect and even overlapping. The nonideal weights used in simulation are then given by

$$w_{ij}^{\text{nonideal}} = |w_{ij}^{\text{binary}} + \chi_{ij}| \tag{12}$$

where χ_{ij} are independent samples from a centered Gaussian distribution with standard deviation 0.5. Histograms of the ideal; binary; and nonideal weights are shown in Fig. S7. It is well known that attractor networks are robust to such nonidealities and that their effect is to reduce the number of attractors that can be reliably stored [11, 33, 99, 100].

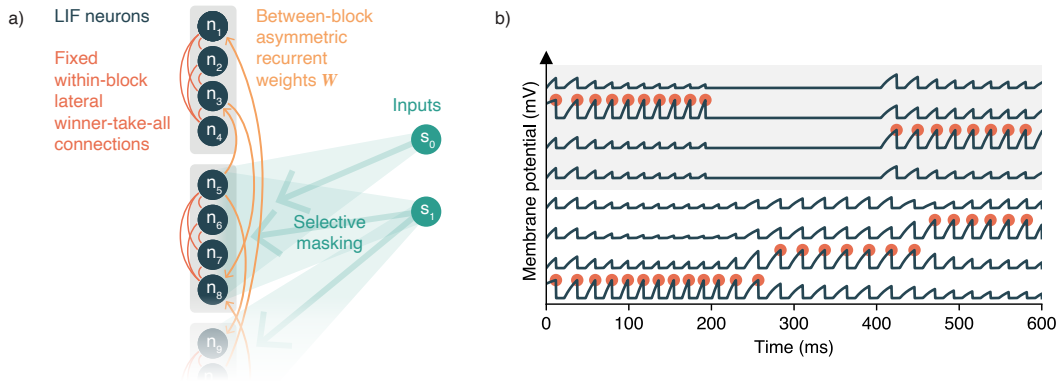

Fig. 7: **a)** RSNN network diagram. The neurons are split into blocks of equal length (here $L = 4$) with WTA connectivity between all neurons in the same block. The between-block weights $\mathbf{W}$ are programmed to embed the desired DFA in the RSNN dynamics. The connectivity from each input s is set according to its binary hypervector $\mathbf{s}$. **b)** Membrane potential traces $u_i(t)$ of the first 8 neurons within an RSNN. Whenever a neuron spikes (red dots), all neurons within the same block are forced into a refractory state. Thus only the neuron with the greatest postsynaptic current spikes, realising an argmax function in each block and confining the neural activity to sparse block codes. Between $200\,\mathrm{ms}$ and $400\,\mathrm{ms}$, the second block is masked by an input (either s_0 or s_1), which causes the network to switch between attractor states.

The RSNN is initialised in the q_0 state, and then a sequence of inputs s_0 or s_1 are given to the network for $200\,\mathrm{ms}$ at a time, with gaps in-between. We used this particular timing scheme for simplicity, but the proper functioning of the RSNN is not dependent upon the exact timing of inputs. Each input (and gap) may be given for arbitrarily long periods (Fig. 2), needing only to be longer than the attractor-switching time, which is of the order τ_{syn}.

The firing rate of each attractor state q is given by the dot product of its hypervector $\mathbf{q}$ and the vector of kernel-filtered spike activity, i.e.

$$m_q(t) = \frac{1}{M}\sum_{i=1}^{N}[\mathbf{q}]_i\big(K(t) \circledast \sum_{\text{spks } k}\delta(t - t_i^k)\big)(t) \tag{13}$$

where $m_q(t)$ is the firing rate of state q, and $K(t)$ is a normalised alpha kernel with time constant τ_{readout}. If the RSNN is exactly representing a state q, then $\langle m_q\rangle = \langle\nu\rangle$ where $\langle\nu\rangle$ is the average firing rate of the active neurons. Since the $\mathbf{q}$ states are not orthogonal, the expected firing rate of the other $q' \neq q$ attractor states $\langle m_{q'}\rangle = \langle\nu\rangle/L$ is nonzero. The firing rates of the bridge states $\mathbf{b}$ are likewise calculated.

4.7 RRAM crossbar experiment

The neuron equations were simulated in Python using Euler's method with the NumPy package. At each time step, if one or multiple neurons spiked, the memristive crossbar was read, and the measured currents were linearly scaled and then applied to the postsynaptic neurons in simulation. The experimental measurements were made using non-volatile 1T1R HfO_x-based RRAM chips integrating 32×128 devices. The hardware setup is shown in Fig. 3. An FPGA board in the measurement system supports the parallel operation of the WLs, BLs, and SLs of the RRAM chip.

With the 4096 RRAM devices, we could create a fully-connected RSNN with 64 neurons. In simulation and on Loihi, the dimensionality was large enough that we could rely upon the pseudo-orthogonality of randomly-generated hypervector representations. This is not reliably achieved for $N = 64$, and so we instead manually chose our sparse block code hypervectors $\mathbf{q}$ and $\mathbf{b}$ to be orthogonal. The ideal weight matrix $\mathbf{W}$ was constructed as described in Methods but with the f terms set to $f = 0$, to compensate for the fact that our attractor state vectors now have exactly 0 inner product. Since the RRAM devices supported ternary weight states (normalised to $0, 0.5, 1$), we mapped the ideal weights to ternary weights by applying two thresholds on either side of 0. The distributions of these ternary conductance states are shown in Fig. 4.

To store the 64×64 weight matrix on the physical RRAM chip, we partitioned it into two 32×64 matrices and then stored their concatenation on the available 32×128 devices (Fig. S8). To perform a current readout (whenever a neuron spiked), we input the binary vector of the first 32 neurons as a voltage via the DACs in parallel, and read the first 64 currents. Then, we input the binary vector of the latter 32 neurons' spike activity and added the two current vectors together on the PC. Therefore, all multiply-and-accumulate steps besides the last addition were performed in-memory.

4.8 Loihi implementation

The experiments were run on the Intel Loihi 2 Oheo Gulch research system using 1024 neurons (128 blocks of 8 neurons) on 95 neurocores. The code was written in Python using Intel's Lava package.

The network uses the regular current-based leaky integrate and fire neuron model on Loihi. However, different from the hard-coded model, each neuron has a conventional integrative synapse, with weights set by $\mathbf{W}$, and a shunting inhibitory synapse to implement the block-WTA and masking behaviour. To achieve the shunting inhibition synapse, we implemented the neuron using custom neuron microcode, a novel feature of Loihi 2. The microcode on Loihi 2 allows arbitrary instructions from a limited instruction set. To avoid having to create a neuron with 2 registers to store the input, we discriminate the integrative synapse's input and the shunting synapse's input by different weight ranges. The shunting synapse's weight was set to -2048, a value that is never reached by integrating regular input. So, in the microcode, before the regular neuron behaviour is computed, a comparison is made if the input value is below 1000. If it is, the membrane potential is reset to 0 (shunting) and 2048 is added to the input in order not to interfere with the integrative synapse.

The weight matrix was constructed as described in Section 4.2, and then quantized by linearly mapping the 4-sigma range of $\mathbf{W}$ to the available even integer weight values in the interval $[-254, 254]$. These recurrent weights were then written to the all-to-all recurrent synapse weights. Attractor transitions were triggered by stimulating the shunting inhibition synapses, with the input vectors $\mathbf{s}_0$ or $\mathbf{s}_1$ to input a 0 or 1, respectively. At the start of the experiment, the RSNN is initialised in the first attractor state q_0. Inputs are then given for 200 time steps each, with a 200 time step gap in between, wherein no input is given.

4.9 Capacity simulation

The RSNN was modelled as a single-layer fully-connected artificial RNN, with the discrete-time update rule

$$\mathbf{z}_{t+1} = \mathrm{bWTA}\Big[\mathbf{W} \cdot \big(\mathbf{z}_t \wedge \mathbf{i}_t\big)\Big] \tag{14}$$

where $\mathbf{z}_t \in \{0,1\}^N$ is a sparse block vector representing the network state on time step t, $\mathbf{W}$ is the recurrent weight matrix (see Section 4.2), bWTA is the block-winner-take-all activation function, and $\mathbf{i}_t \in \{0,1\}^N$ is the masking input to the network. If there is no input to the network, then $\mathbf{i}_t$ is a vector of all ones.

The network was tested with modular-division DFAs of varying sizes P, like in Fig. 1 for $P = 23$. Note that P does not directly correspond to the number of stored attractor states, since each DFA state actually stores two attractor states and, for these DFAs, two asymmetric transition terms. For various (P, N) pairs, we evaluated the network's ability to correctly emulate the DFA by giving sequences of 5 inputs (each lasting 10 time steps, with 10 time steps of no input between) to the network. A run was deemed to be successful if the hypervector $\mathbf{q}$ corresponding to the correct final DFA state had the greatest overlap with the network's final state. The block length L was varied such that the active number of neurons N/L scaled with $\log N$, as required to obtain the quadratic capacity relation [33]. Each (P, N) pair was tested 5 times with different hypervector instantiations.

References

[1] C. Mead, "Neuromorphic electronic systems", *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629–1636, 1990. DOI: 10.1109/5.58356.

[2] S. Ganguli, D. Huh, and H. Sompolinsky, "Memory traces in dynamical systems", *Proceedings of the National Academy of Sciences*, vol. 105, no. 48, pp. 18 970–18 975, 2008. DOI: 10.1073/pnas.0804451105.

[3] M. Khona and I. R. Fiete, "Attractor and integrator networks in the brain", *Nature Reviews Neuroscience*, vol. 23, no. 12, pp. 744–766, 2022. DOI: 10.1038/s41583-022-00642-0.

[4] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. DOI: 10.1162/neco.1997.9.8.1735.

[5] G. Bellec, F. Scherr, A. Subramoney, E. Hajek, D. Salaj, R. Legenstein, and W. Maass, "A solution to the learning dilemma for recurrent networks of spiking neurons", *Nature Communications*, vol. 11, no. 1, p. 3625, 2020. DOI: 10.1038/s41467-020-17236-y.

[6] B. Cramer, S. Billaudelle, S. Kanya, A. Leibfried, A. Grübl, V. Karasenko, C. Pehle, K. Schreiber, Y. Stradmann, J. Weis, J. Schemmel, and F. Zenke, "Surrogate gradients for analog neuromorphic computing", *Proceedings of the National Academy of Sciences*, vol. 119, no. 4, 2022. DOI: 10.1073/pnas.2109194119.

[7] E. O. Neftci, H. Mostafa, and F. Zenke, "Surrogate gradient learning in spiking neural networks", 2019. DOI: 10.48550/arXiv.1901.09948.

[8] J. Miller and M. Hardt, "Stable recurrent models," in *International Conference on Learning Representations*, 2019.

[9] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult", *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994. DOI: 10.1109/72.279181.

[10] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982. DOI: 10.1073/pnas.79.8.2554.

[11] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks*. 1989.

[12] T. Bohnstingl, A. Šurina, M. Fabre, Y. Demirağ, C. Frenkel, M. Payvand, G. Indiveri, and A. Pantazi, "Biologically-inspired training of spiking recurrent neural networks with neuromorphic hardware," in *2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2022, pp. 218–221. DOI: 10.1109/AICAS54282.2022.9869963.

[13] S. Schmitt, J. Klähn, G. Bellec, A. Grübl, M. Güttler, A. Hartel, *et al.*, "Neuromorphic hardware in the loop: Training a deep spiking network on the BrainScaleS wafer-scale system," in *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 2227–2234. DOI: 10.1109/IJCNN.2017.7966125.

[14] Y. Demirag, C. Frenkel, M. Payvand, and G. Indiveri, "Online training of spiking recurrent neural networks with phase-change memory synapses", 2021. DOI: 10.48550/arXiv.2108.01804.

[15] U. Cakal, Maryada, C. Wu, I. Ulusoy, and D. R. Muir, "Gradient-descent hardware-aware training and deployment for mixed-signal neuromorphic processors", *Neuromorphic Computing and Engineering*, vol. 4, no. 1, p. 014 011, 2024. DOI: 10.1088/2634-4386/ad2ec3.

[16] D. Kleyko, M. Davies, E. P. Frady, P. Kanerva, S. J. Kent, B. A. Olshausen, E. Osipov, J. M. Rabaey, D. A. Rachkovskij, A. Rahimi, and F. T. Sommer, "Vector symbolic architectures as a computing framework for emerging hardware", *Proceedings of the IEEE*, vol. 110, no. 10, pp. 1538–1571, 2022. DOI: 10.1109/JPROC.2022.3209104.

[17] D. Kleyko, D. A. Rachkovskij, E. Osipov, and A. Rahimi, "A survey on hyperdimensional computing aka vector symbolic architectures, part I: Models and data transformations", *ACM Computing Surveys*, vol. 55, no. 6, 130:1–130:40, 2022. DOI: 10.1145/3538531.

[18] P. Kanerva, "Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors", *Cognitive Computation*, vol. 1, no. 2, pp. 139–159, 2009. DOI: 10.1007/s12559-009-9009-8.

[19] T. A. Plate, "Holographic reduced representations", *IEEE Transactions on Neural Networks*, vol. 6, no. 3, pp. 623–641, 1995. DOI: 10.1109/72.377968.

[20] R. W. Gayler, "Multiplicative binding, representation operators & analogy", *Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences*, 1998.

[21] G. Karunaratne, M. Le Gallo, G. Cherubini, L. Benini, A. Rahimi, and A. Sebastian, "In-memory hyperdimensional computing", *Nature Electronics*, vol. 3, no. 6, pp. 327–337, 2020. DOI: 10.1038/s41928-020-0410-3.

[22] M. Cotteret, H. Greatorex, M. Ziegler, and E. Chicca, "Vector symbolic finite state machines in attractor neural networks", *Neural Computation*, vol. 36, no. 4, pp. 549–595, 2024. DOI: 10.1162/neco_a_01638.

[23] A. Renner, L. Supic, A. Danielescu, G. Indiveri, B. A. Olshausen, Y. Sandamirskaya, F. T. Sommer, and E. P. Frady, "Neuromorphic visual scene understanding with resonator networks", *Nature Machine Intelligence*, vol. 6, no. 6, pp. 641–652, 2024. DOI: 10.1038/s42256-024-00848-0.

[24] A. A. Faisal, L. P. J. Selen, and D. M. Wolpert, "Noise in the nervous system", *Nature Reviews Neuroscience*, vol. 9, no. 4, pp. 292–303, 2008. DOI: 10.1038/nrn2258.

[25] E. Neftci, J. Binas, U. Rutishauser, E. Chicca, G. Indiveri, and R. J. Douglas, "Synthesizing cognition in neuromorphic electronic systems", *Proceedings of the National Academy of Sciences*, vol. 110, no. 37, E3468–E3476, 2013. DOI: 10.1073/pnas.1212083110.

[26] D. Liang and G. Indiveri, "A neuromorphic computational primitive for robust context-dependent decision making and context-dependent stochastic computation", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 5, pp. 843–847, 2019. DOI: 10.1109/TCSII.2019.2907848.

[27] G. Orchard, E. P. Frady, D. B. D. Rubin, S. Sanborn, S. B. Shrestha, F. T. Sommer, and M. Davies, "Efficient neuromorphic signal processing with Loihi 2," in *2021 IEEE Workshop on Signal Processing Systems (SiPS)*, 2021, pp. 254–259. DOI: 10.1109/SiPS52927.2021.00053.

[28] P. Dayan, "Simple substrates for complex cognition", *Frontiers in Neuroscience*, vol. 2, p. 31, 2008. DOI: 10.3389/neuro.01.031.2008.

[29] S.-i. Amari, "Characteristics of sparsely encoded associative memory", *Neural Networks*, vol. 2, no. 6, pp. 451–457, 1989. DOI: 10.1016/0893-6080(89)90043-9.

[30] M. V. Tsodyks and M. V. Feigel'man, "The enhanced storage capacity in neural networks with low activity level", *Europhysics Letters (EPL)*, vol. 6, no. 2, pp. 101–105, 1988. DOI: 10.1209/0295-5075/6/2/002.

[31] M. Laiho, J. H. Poikonen, P. Kanerva, and E. Lehtonen, "High-dimensional computing with sparse vectors," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2015, pp. 1–4. DOI: 10.1109/BioCAS.2015.7348414.

[32] E. P. Frady, D. Kleyko, and F. T. Sommer, "Variable binding for sparse distributed representations: Theory and applications", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 5, pp. 2191–2204, 2023. DOI: 10.1109/TNNLS.2021.3105949.

[33] A. Knoblauch, G. Palm, and F. T. Sommer, "Memory capacities for synaptic and structural plasticity", *Neural Computation*, vol. 22, no. 2, pp. 289–341, 2010. DOI: 10.1162/neco.2009.08-07-588.

[34] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. 1986.

[35] A. Renner, Y. Sandamirskaya, F. Sommer, and E. P. Frady, "Sparse vector binding on spiking neuromorphic hardware using synaptic delays," in *Proceedings of the International Conference on Neuromorphic Systems 2022*, 2022, pp. 1–5.

[36] G. Palm, "Neural associative memories and sparse coding", *Neural Networks*, vol. 37, pp. 165–171, 2013. DOI: 10.1016/j.neunet.2012.08.013.

[37] C. J. Rozell, D. H. Johnson, R. G. Baraniuk, and B. A. Olshausen, "Sparse coding via thresholding and local competition in neural circuits", *Neural Computation*, vol. 20, no. 10, pp. 2526–2563, 2008. DOI: 10.1162/neco.2008.03-07-486.

[38] E. Neftci and G. Indiveri, "A device mismatch compensation method for VLSI neural networks," in *2010 Biomedical Circuits and Systems Conference (BioCAS)*, 2010, pp. 262–265. DOI: 10.1109/BIOCAS.2010.5709621.

[39] J. Büchel, D. Zendrikov, S. Solinas, G. Indiveri, and D. R. Muir, "Supervised training of spiking neural networks for robust deployment on mixed-signal neuromorphic processors", *Scientific Reports*, vol. 11, no. 1, p. 23 376, 2021. DOI: 10.1038/s41598-021-02779-x.

[40] H. Greatorex, O. Richter, M. Mastella, M. Cotteret, P. Klein, M. Fabre, *et al.*, "TEXEL: A neuromorphic processor with on-chip learning for beyond-CMOS device integration", 2024. DOI: 10.48550/arXiv.2410.15854.

[41] A. Fantini, G. Gorine, R. Degraeve, L. Goux, C. Y. Chen, A. Redolfi, S. Clima, A. Cabrini, G. Torelli, and M. Jurczak, "Intrinsic program instability in HfO2 RRAM and consequences on program algorithms," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 7.5.1–7.5.4. DOI: 10.1109/IEDM.2015.7409648.

[42] C. Wang, H. Wu, B. Gao, L. Dai, N. Deng, D. C. Sekar, Z. Lu, M. Kellam, G. Bronner, and H. Qian, "Relaxation effect in RRAM arrays: Demonstration and characteristics", *IEEE Electron Device Letters*, vol. 37, no. 2, pp. 182–185, 2016. DOI: 10.1109/LED.2015.2508034.

[43] J. Chen, S. Yang, H. Wu, G. Indiveri, and M. Payvand, "Scaling limits of memristor-based routers for asynchronous neuromorphic systems", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 71, no. 3, pp. 1576–1580, 2024. DOI: 10.1109/TCSII.2023.3343292.

[44] S. D'Agostino, F. Moro, T. Torchet, Y. Demirağ, L. Grenouillet, N. Castellani, G. Indiveri, E. Vianello, and M. Payvand, "DenRAM: Neuromorphic dendritic architecture with RRAM for efficient temporal processing with delays", *Nature Communications*, vol. 15, no. 1, p. 3446, 2024. DOI: 10.1038/s41467-024-47764-w.

[45] W. A. Little, "The existence of persistent states in the brain", *Mathematical Biosciences*, vol. 19, no. 1, pp. 101–120, 1974. DOI: 10.1016/0025-5564(74)90031-5.

[46] E. Rolls, "The mechanisms for pattern completion and pattern separation in the hippocampus", *Frontiers in systems neuroscience*, vol. 7, p. 74, 2013. DOI: 10.3389/fnsys.2013.00074.

[47] R. Chaudhuri and I. Fiete, "Computational principles of memory", *Nature Neuroscience*, vol. 19, no. 3, pp. 394–403, 2016. DOI: 10.1038/nn.4237.

[48] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, "Weak pairwise correlations imply strongly correlated network states in a neural population", *Nature*, vol. 440, no. 7087, pp. 1007–1012, 2006. DOI: 10.1038/nature04701.

[49] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome, "Context-dependent computation by recurrent dynamics in prefrontal cortex", *Nature*, vol. 503, no. 7474, pp. 78–84, 2013. DOI: 10.1038/nature12742.

[50] D. Sussillo and O. Barak, "Opening the black box: Low-dimensional dynamics in high-dimensional recurrent neural networks", *Neural Computation*, vol. 25, no. 3, pp. 626–649, 2013. DOI: 10.1162/NECO_a_00409.

[51] P. Miller, "Itinerancy between attractor states in neural systems", *Current opinion in neurobiology*, vol. 40, pp. 14–22, 2016. DOI: 10.1016/j.conb.2016.05.005.

[52] U. Rutishauser and R. Douglas, "State-dependent computation using coupled recurrent networks", *Neural computation*, vol. 21, pp. 478–509, 2009. DOI: 10.1162/neco.2008.03-08-734.

[53] S. Baumgartner, A. Renner, R. Kreiser, D. Liang, G. Indiveri, and Y. Sandamirskaya, "Visual pattern recognition with on on-chip learning: Towards a fully neuromorphic approach," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2020, pp. 1–5. DOI: 10.1109/ISCAS45731.2020.9180628.

[54] S. Fusi, E. Miller, and M. Rigotti, "Why neurons mix: High dimensionality for higher cognition", *Current Opinion in Neurobiology*, 2016. DOI: 10.1016/j.conb.2016.01.010.

[55] A. Thomas, S. Dasgupta, and T. Rosing, "A theoretical perspective on hyperdimensional computing", *Journal of Artificial Intelligence Research*, vol. 72, pp. 215–249, 2022. DOI: 10.1613/jair.1.12664.

[56] K. L. Clarkson, S. Ubaru, and E. Yang, "Capacity analysis of vector symbolic architectures", 2023. DOI: 10.48550/arXiv.2301.10352.

[57] K. Boahen, "Dendrocentric learning for synthetic intelligence", *Nature*, vol. 612, no. 7938, pp. 43–50, 2022. DOI: 10.1038/s41586-022-05340-6.

[58] H. E. Barkam, S. Yun, H. Chen, P. Gensler, A. Mema, A. Ding, G. Michelogiannakis, H. Amrouch, and M. Imani, "Reliable hyperdimensional reasoning on unreliable emerging technologies," in *2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, 2023, pp. 1–9. DOI: 10.1109/ICCAD57390.2023.10323935.

[59] E. P. Frady and F. T. Sommer, "Robust computation with rhythmic spike patterns", *Proceedings of the National Academy of Sciences*, vol. 116, no. 36, pp. 18 050–18 059, 2019. DOI: 10.1073/pnas.1902653116.

[60] Z. Zou, H. Alimohamadi, A. Zakeri, F. Imani, Y. Kim, M. H. Najafi, and M. Imani, "Memory-inspired spiking hyperdimensional network for robust online learning", *Scientific Reports*, vol. 12, no. 1, p. 7641, 2022. DOI: 10.1038/s41598-022-11073-3.

[61] J. Morris, H. W. Lui, K. Stewart, B. Khaleghi, A. Thomas, T. Marback, B. Aksanli, E. Neftci, and T. Rosing, "HyperSpike: Hyperdimensional computing for more efficient and robust spiking neural networks," in *2022 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2022, pp. 664–669. DOI: 10.23919/DATE54114.2022.9774644.

[62] D. Horn and M. Usher, "Neural networks with dynamical thresholds", *Physical Review A*, vol. 40, no. 2, pp. 1036–1044, 1989. DOI: 10.1103/PhysRevA.40.1036.

[63] F. T. Sommer and T. Wennekers, "Synfire chains with conductance-based neurons: Internal timing and coordination with timed input", *Neurocomputing*, vol. 65-66, pp. 449–454, 2005. DOI: 10.1016/j.neucom.2004.10.015.

[64] T. Kambara, Y. Kashimori, N. Usuba, and O. Hoshino, "Role of itinerancy among attractors as dynamical map in distributed coding scheme", *Neural Networks: The Official Journal of the International Neural Network Society*, vol. 10, no. 8, pp. 1375–1390, 1997. DOI: 10.1016/s0893-6080(97)00022-1.

[65] H. Gutfreund and M. Mezard, "Processing of temporal sequences in neural networks", *Physical Review Letters*, vol. 61, no. 2, pp. 235–238, 1988. DOI: 10.1103/PhysRevLett.61.235.

[66] D. J. Amit, "Neural networks counting chimes", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, no. 7, pp. 2141–2145, 1988. DOI: 10.1073/pnas.85.7.2141.

[67] M. F. J. Drossaers, "Hopfield models as nondeterministic finite-state machines," in *Proceedings of the 14th conference on computational linguistics - Volume 1*, 1992, pp. 113–119. DOI: 10.3115/992066.992087.

[68] B. Chen and P. Miller, "Attractor-state itinerancy in neural circuits with synaptic depression", *The Journal of Mathematical Neuroscience*, vol. 10, no. 1, p. 15, 2020. DOI: 10.1186/s13408-020-00093-w.

[69] P. Peretto and J. J. Niez, "Collective properties of neural networks," in *Disordered Systems and Biological Organization*, 1986, pp. 171–185.

[70] E. Osipov, D. Kleyko, and A. Legalov, "Associative synthesis of finite state automata model of a controlled object with hyperdimensional computing," in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, 2017, pp. 3276–3281. DOI: 10.1109/IECON.2017.8216554.

[71] M. Maestro-Izquierdo, M. B. González, P. Martín-Holgado, Y. Morilla, and F. Campabadal, "Gamma radiation effects on HfO2-based RRAM devices," in *2021 13th Spanish Conference on Electron Devices (CDE)*, 2021, pp. 23–26. DOI: 10.1109/CDE52135.2021.9455719.

[72] D. Kleyko, D. Rachkovskij, E. Osipov, and A. Rahimi, "A survey on hyperdimensional computing aka vector symbolic architectures, part II: Applications, cognitive models, and challenges", *ACM Computing Surveys*, vol. 55, no. 9, 175:1–175:52, 2023. DOI: 10.1145/3558000.

[73] P. Neubert, S. Schubert, and P. Protzel, "An introduction to hyperdimensional computing for robotics", *KI - Künstliche Intelligenz*, vol. 33, no. 4, pp. 319–330, 2019. DOI: 10.1007/s13218-019-00623-z.

[74] T. Yerxa, A. Anderson, and E. Weiss, "The hyperdimensional stack machine", *Cognitive Computing*, pp. 1–2, 2018.

[75] D. Marr, *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. 1982.

[76] A. J. Noest, "Phasor neural networks," in *Proceedings of the 1987 International Conference on Neural Information Processing Systems*, 1987, pp. 584–591.

[77] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, "Opportunities for neuromorphic computing algorithms and applications", *Nature Computational Science*, vol. 2, no. 1, pp. 10–19, 2022. DOI: 10.1038/s43588-021-00184-y.

[78] H. Jaeger, B. Noheda, and W. G. van der Wiel, "Toward a formal theory for computing machines made out of whatever physics offers", *Nature Communications*, vol. 14, no. 1, p. 4911, 2023. DOI: 10.1038/s41467-023-40533-1.

[79] F. Stefanini, E. O. Neftci, S. Sheik, and G. Indiveri, "PyNCS: A microkernel for high-level definition and configuration of neuromorphic electronic systems", *Frontiers in Neuroinformatics*, vol. 8, 2014. DOI: 10.3389/fninf.2014.00073.

[80] A. P. Davison, D. Brüderle, J. M. Eppler, J. Kremkow, E. Muller, D. Pecevski, L. Perrinet, and P. Yger, "PyNN: A common interface for neuronal network simulators", *Frontiers in Neuroinformatics*, vol. 2, 2009. DOI: 10.3389/neuro.11.011.2008.

[81] J. B. Aimone, W. Severa, and C. M. Vineyard, "Composing neural algorithms with Fugu," in *Proceedings of the International Conference on Neuromorphic Systems*, 2019, pp. 1–8. DOI: 10.1145/3354265.3354268.

[82] Y. Zhang, P. Qu, Y. Ji, W. Zhang, G. Gao, G. Wang, *et al.*, "A system hierarchy for brain-inspired computing", *Nature*, vol. 586, no. 7829, pp. 378–384, 2020. DOI: 10.1038/s41586-020-2782-y.

[83] C. Eliasmith, "A unified approach to building and controlling spiking attractor networks", *Neural Computation*, vol. 17, no. 6, pp. 1276–1314, 2005. DOI: 10.1162/0899766053630332.

[84] J. E. Pedersen, S. Abreu, M. Jobst, G. Lenz, V. Fra, F. C. Bauer, *et al.*, "Neuromorphic intermediate representation: A unified instruction set for interoperable brain-inspired computing", *Nature Communications*, vol. 15, no. 1, p. 8122, 2024. DOI: 10.1038/s41467-024-52259-9.

[85] P. Poduval, H. Alimohamadi, A. Zakeri, F. Imani, M. H. Najafi, T. Givargis, and M. Imani, "GrapHD: Graph-based hyperdimensional memorization for brain-like cognitive learning", *Frontiers in Neuroscience*, vol. 16, 2022. DOI: 10.3389/fnins.2022.757125.

[86] E. P. Frady, D. Kleyko, C. J. Kymn, B. A. Olshausen, and F. T. Sommer, "Computing on functions using randomized vector representations (in brief)," in *Neuro-Inspired Computational Elements Conference*, 2022, pp. 115–122. DOI: 10.1145/3517343.3522597.

[87] M. L. Minsky, *Computation: Finite and Infinite Machines*. 1967.

[88] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 5, pp. 815–826, 1983. DOI: 10.1109/TSMC.1983.6313075.

[89] J. Buhmann and K. Schulten, "Noise-driven temporal association in neural networks", *Europhysics Letters (EPL)*, vol. 4, no. 10, pp. 1205–1209, 1987. DOI: 10.1209/0295-5075/4/10/021.

[90] D. Kleinfeld, "Sequential state generation by model neural networks." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 24, pp. 9469–9473, 1986. DOI: 10.1073/pnas.83.24.9469.

[91] E. Chicca, F. Stefanini, C. Bartolozzi, and G. Indiveri, "Neuromorphic electronic circuits for building autonomous cognitive systems", *Proceedings of the IEEE*, vol. 102, no. 9, pp. 1367–1388, 2014. DOI: 10.1109/JPROC.2014.2313954.

[92] O. Richter, H. Greatorex, B. Hucko, M. Cotteret, W. Soares Girao, E. Janotte, M. Mastella, and E. Chicca, "A subthreshold second-order integration circuit for versatile synaptic alpha kernel and trace generation," in *Proceedings of the 2023 International Conference on Neuromorphic Systems*, 2023, pp. 1–4. DOI: 10.1145/3589737.3606008.

[93] H. Mostafa, L. K. Müller, and G. Indiveri, "Rhythmic inhibition allows neural networks to search for maximally consistent states", *Neural Computation*, vol. 27, pp. 2510–2547, 2015. DOI: 10.1162/neco_a_00785.

[94] M. Cotteret, O. Richter, M. Mastella, H. Greatorex, E. Janotte, W. S. Girão, M. Ziegler, and E. Chicca, "Robust spiking attractor networks with a hard winner-take-all neuron circuit," in *2023 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2023, pp. 1–5. DOI: 10.1109/ISCAS46773.2023.10181513.

[95] J. Abrahamsen, P. Hafliger, and T. Lande, "A time domain winner-take-all network of integrate-and-fire neurons," in *2004 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 5, 2004, pp. V–V. DOI: 10.1109/ISCAS.2004.1329537.

[96] M. Stimberg, R. Brette, and D. F. Goodman, "Brian 2, an intuitive and efficient neural simulator", *eLife*, vol. 8, e47314, 2019. DOI: 10.7554/eLife.47314.

[97] T. M. Bartol Jr, C. Bromer, J. Kinney, M. A. Chirillo, J. N. Bourne, K. M. Harris, and T. J. Sejnowski, "Nanoconnectomic upper bound on the variability of synaptic plasticity", *eLife*, vol. 4, e10778, 2015. DOI: 10.7554/eLife.10778.

[98] F. Zahari, E. Pérez, M. K. Mahadevaiah, H. Kohlstedt, C. Wenger, and M. Ziegler, "Analogue pattern recognition with stochastic switching binary CMOS-integrated memristive devices", *Scientific Reports*, vol. 10, no. 1, p. 14 450, 2020. DOI: 10.1038/s41598-020-71334-x.

[99] H. Sompolinsky, "The theory of neural networks: The Hebb rule and beyond," in *Heidelberg Colloquium on Glassy Dynamics*, 1987, pp. 485–527. DOI: 10.1007/BFb0057531.

[100] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins, "Non-holographic associative memory", *Nature*, vol. 222, no. 5197, pp. 960–962, 1969. DOI: 10.1038/222960a0.

Acknowledgements

Thanks to Friedrich Sommer, Denis Kleyko, Chris Kymn and Anthony Thomas for enlightening discussions and suggestions. This work has been supported by DFG projects NMVAC (432009531) and MemTDE (441959088). The authors would like to acknowledge the financial support of the CogniGron research center and the Ubbo Emmius Funds (Univ. of Groningen). We thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Hábrók high-performance computing cluster. We thank Intel for providing access to the Loihi 2 research platform and Philipp Stratmann for support with the microcode.

Author contributions

M.C. conceived the initial algorithm and its SNN implementation. H.G., A.R. implemented the algorithm on Loihi. H.W. supervised, and J.C., M.C., implemented the algorithm on the memristive crossbar setup. E.N., H.W., G.I., M.Z., E.C. supervised the experiments. All authors contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

S1 Sparse block code overlaps

We here derive a result, adapted from [1], that the number of sparse block code (SBC) hypervectors that can be randomly generated without accidentally incurring a significant degree of overlap between any pair of them, is exponential in dimension N, and so very large for reasonable values of N and L.

We generate SBC hypervectors of dimension N, with block length L, and number of blocks $M = N/L$. The coding level is then $f = 1/L$. The normalised similarity between SBC hypervectors is given by

$$\mathrm{sim}(\mathbf{a}, \mathbf{b}) = \frac{1}{M}\mathbf{a} \cdot \mathbf{b} \tag{1}$$

such that the the similarity between identical hypervectors $\mathrm{sim}(\mathbf{a}, \mathbf{a}) = 1$, and the expected similarity between independent hypervectors is $\langle \mathrm{sim}(\mathbf{a}, \mathbf{b}) \rangle = f < 1$.

We first provide a bound on the similarity between independent hypervectors. We can rewrite the similarity between SBC hypervectors as

$$\mathrm{sim}(\mathbf{a}, \mathbf{b}) = \frac{1}{M} \sum_{i=1}^{M} X_i \tag{2}$$

where $X_i \in \{0, 1\}$ denotes whether the same neuron in the i'th block is active in both hypervectors, with $\langle X_i \rangle = f$. By Hoeffding's inequality, this sum is bounded by

$$\mathbb{P}\big[\mathrm{sim}(\mathbf{a}, \mathbf{b}) - f \geq \theta\big] \leq \exp\big(-2M\theta^2\big) \tag{3}$$

where θ is a threshold of our choosing. This bound gives us assurances about how much the similarity between independent hypervectors can deviate from the expected value f. We can then set θ to represent the maximum acceptable similarity between independent hypervectors. Given that independent hypervectors have expected similarity f, and identical hypervectors have similarity 1, then a reasonable threshold might be halfway between them, $\theta = \frac{1}{2}(1 - f)$. For reasonable values of $N = 1000$, $L = 8$, the probability that a given pair of independent hypervectors exceeds θ is then less than 10^{-20}.

We however want a bound on the number K of SBC hypervectors that can be generated, for which *none* of the possible inner product pairs exceed θ. For K hypervectors, there are $\frac{1}{2}K(K-1) < \frac{1}{2}K^2$ pairs to check. We apply Boole's inequality, that the union probability of a set of events occurring $\mathbb{P}\big[\bigcup_i E_i\big] \leq \sum_i \mathbb{P}[E_i]$ for some countable events $\{E_i\}$. In our case, each event is whether a similarity exceeds θ, and so we have

$$\mathbb{P}[\text{any pair exceeds } \theta] \leq \delta = \sum_{j=1}^{\frac{1}{2}K^2} \exp\big(-2M\theta^2\big) = \frac{1}{2}K^2 \exp\big(-2M\theta^2\big) \tag{4}$$

where we have introduced $\delta \in [0, 1]$ as the acceptable probability of a collision, which we will set to $\delta \ll 1$. We can then rearrange for the number of hypervectors K,

$$K = \sqrt{2\delta} \exp\Big(M\theta^2\Big) = \sqrt{2\delta} \exp\Big(N\theta^2/L\Big) \tag{5}$$

and so the number of SBC hypervectors that can be generated before a collision becomes likely, is exponential in dimension N. For comparison, the number of orthogonal localist representations that can be generated is linear in N.

For reasonable values of $N = 1000$, $L = 8$, with an acceptable failure probability of $\delta = 0.0001$, then with probability greater than 0.9999, we can generate $K \approx 3 \times 10^9$ SBC hypervectors before a collision occurs. This is well in excess of the number of states that can be stored in a fully-connected attractor network, which scales at most with the number of synapses N^2.

S2 Postsynaptic sum calculations

We first calculate the mean and variance of the components of the postsynaptic sum $\mathbf{h}$, in the case that the network is already in state $\mathbf{q}^\mu \in Q$, and there is no input. We calculate the contribution to $\mathbf{h}$ due to the three superimposed matrices in $\mathbf{W}$ separately, for the sake of cleanliness. We will assume that the inner products between different

hypervectors are independent (while they are actually only pairwise independent), such that their variances combine linearly.

$$\begin{aligned}
[\mathbf{W}_{\text{attr.}}\mathbf{q}^\mu]_i &= \sum_\nu^Q (q_i^\nu - f)(\mathbf{q}^\nu - f)\cdot\mathbf{q}^\mu \\
&= M(1-f)(q_i^\mu - f) + \sum_{\nu\neq\mu}^Q \underbrace{(q_i^\nu - f)}_{\substack{\text{Mean = 0}\\ \text{Var. = } f(1-f)}} \underbrace{(\mathbf{q}^\nu - f)\cdot\mathbf{q}^\mu}_{\substack{\text{Mean = 0}\\ \text{Var. = } Mf(1-f)}} \\
&= M(1-f)(q_i^\mu - f) \pm f(1-f)\sqrt{(Q-1)M} \\
&\approx M(q_i^\mu - f) \pm f\sqrt{MQ}
\end{aligned} \tag{6}$$

where we are abusing notation somewhat by using Q to denote the number of states in the set Q (and will do the same with E and S), and on the last line we make the approximation that $Q \gg 1$ (there are many many DFA states) and $f \ll 1$ (our states are sparse). The corresponding calculation for the bridge matrix is

$$\begin{aligned}
[\mathbf{W}_{\text{brdg.}}\mathbf{q}^\mu]_i &= \sum_\nu^Q \Big[\underbrace{(q_i^\nu - f)}_{\substack{\text{Mean = 0}\\ \text{Var. = } f(1-f)}} \underbrace{(\mathbf{b}^\nu - f)\cdot\mathbf{q}^\mu}_{\substack{\text{Mean = 0}\\ \text{Var. = } Mf(1-f)}} + \sum_s^S \underbrace{(b_i^\nu - q_i^\nu)}_{\substack{\text{Mean = 0}\\ \text{Var. = } 2f(1-f)}} \underbrace{((\mathbf{b}^\nu - f)\circ\bar{\mathbf{s}})\cdot\mathbf{q}^\mu}_{\substack{\text{Mean = 0}\\ \text{Var. = } Mf(1-f)}} \Big] \\
&= 0 \pm f(1-f)\sqrt{MQ(1+2S)} \\
&\approx 0 \pm f\sqrt{MQ(1+2S)}
\end{aligned} \tag{7}$$

If we are constructing a DFA with many inputs (large S), but each state has incoming edges with only a few of these inputs, then the capacity of the network can be increased by here summing not over all inputs, but only over the unique inputs of the edges to the state q. The transition matrix calculation goes

$$\begin{aligned}
[\mathbf{W}_{\text{tran.}}\mathbf{q}^\mu]_i &= \sum_\eta^E (b'_i - q_i)((\mathbf{q} - f)\circ\bar{\mathbf{s}})\cdot\mathbf{q}^\mu \\
&= \sum_{\eta|q\neq q^\mu}^E \underbrace{(b'_i - q_i)}_{\substack{\text{Mean = 0}\\ \text{Var. = } 2f(1-f)}} \underbrace{((\mathbf{q} - f)\circ\bar{\mathbf{s}})\cdot\mathbf{q}^\mu}_{\substack{\text{Mean = 0}\\ \text{Var. = } Mf(1-f)}} + \sum_{\eta|q=q^\mu}^E \underbrace{(b'_i - q_i)}_{\substack{\text{Mean = 0}\\ \text{Var. = } 2f(1-f)}} \underbrace{((\mathbf{q}^\mu - f)\circ\bar{\mathbf{s}})\cdot\mathbf{q}^\mu}_{\substack{\text{Mean = 0}\\ \text{Var. = } M(1-f)^2}} \\
&= 0 \pm (1-f)\sqrt{2Mf^2(E-S) + 2MSf(1-f)} \\
&\approx 0 \pm \sqrt{2Mf^2E + 2MSf}
\end{aligned} \tag{8}$$

where we split the summation up into transitions whose source state is q^μ (and those with a different source state). We then assume the worse-case scenario that S of the E edges have q^μ as a source state. The final postsynaptic sum is given by

$$\begin{aligned}
[\mathbf{h}]_i = [\mathbf{W}\mathbf{q}^\mu]_i &\approx M(q_i^\mu - f) \pm \sqrt{M}\sqrt{2f^2Q(1+S) + 2f^2E + 2fS} \\
&\propto q_i^\mu - f \pm \sqrt{\frac{2}{N}}\sqrt{fQ(1+S) + fE + S}
\end{aligned} \tag{9}$$

from which we get $\langle \mathbf{W}\mathbf{q}^\mu \rangle \propto \mathbf{q}^\mu - f$.

We then calculate the postsynaptic sum when we are masking the network with an input s^λ, and we assume there exists an edge $\epsilon \in E$ with source state q^μ and input s^λ. Calculating $\mathbf{h}$ in the same way, we have

$$\begin{aligned}
[\mathbf{W}_{\text{attr.}}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)]_i &= \sum_\nu^Q (q_i^\nu - f)(\mathbf{q}^\nu - f) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \\
&= \frac{1}{2}M(1-f)(q_i^\mu - f) + \sum_{\nu \neq \mu}^Q \underbrace{(q_i^\nu - f)}_{\substack{\text{Mean = 0} \\ \text{Var. = } f(1-f)}} \underbrace{(\mathbf{q}^\nu - f) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)}_{\substack{\text{Mean = 0} \\ \text{Var. = } \frac{1}{2}Mf(1-f)}} \\
&= \frac{1}{2}M(1-f)(q_i^\mu - f) \pm f(1-f)\sqrt{(Q-1)\frac{1}{2}M} \\
&\approx \frac{1}{2}M(q_i^\mu - f) \pm f\sqrt{\frac{1}{2}MQ}
\end{aligned} \tag{10}$$

and for the bridge matrix

$$\begin{aligned}
[\mathbf{W}_{\text{brdg.}}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)]_i &= \sum_\nu^Q \Big[\underbrace{(q_i^\nu - f)}_{\substack{\text{Mean = 0} \\ \text{Var. = } f(1-f)}} \underbrace{(\mathbf{b}^\nu - f) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)}_{\substack{\text{Mean = 0} \\ \text{Var. = } \frac{1}{2}Mf(1-f)}} + \sum_s^S \underbrace{(b_i^\nu - q_i^\nu)}_{\substack{\text{Mean = 0} \\ \text{Var. = } 2f(1-f)}} \underbrace{((\mathbf{b}^\nu - f) \circ \bar{\mathbf{s}}) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)}_{\substack{\text{Mean = 0} \\ \text{Var. = } \frac{1}{2}Mf(1-f)}} \Big] \\
&= 0 \pm f(1-f)\sqrt{\frac{1}{2}MQ(1+2S)} \\
&\approx 0 \pm f\sqrt{\frac{1}{2}MQ(1+2S)}
\end{aligned} \tag{11}$$

and the transition matrix

$$\begin{aligned}
&[\mathbf{W}_{\text{tran.}}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)]_i = \sum_\eta^E (b_i' - q_i)((\mathbf{q} - f) \circ \bar{\mathbf{s}}) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \\
&= (1-f)\frac{M}{2}(b_i' - q_i^\mu) + \sum_{\substack{\eta \neq \epsilon \\ \text{and } q \neq q^\mu}}^E \underbrace{(b_i' - q_i)}_{\substack{\text{Mean = 0} \\ \text{Var. = } 2f(1-f)}} \underbrace{((\mathbf{q} - f) \circ \bar{\mathbf{s}}) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)}_{\substack{\text{Mean = 0} \\ \text{Var. = } \frac{M}{2}f(1-f)}} + \sum_{\substack{\eta \neq \epsilon \\ \text{and } q = q^\mu}}^E \underbrace{(b_i' - q_i)}_{\substack{\text{Mean = 0} \\ \text{Var. = } 2f(1-f)}} \underbrace{((\mathbf{q} - f) \circ \bar{\mathbf{s}}) \cdot (\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)}_{\substack{\text{Mean = 0} \\ \text{Var. = } \frac{M}{2}(1-f)^2}} \\
&= \frac{1}{2}M(1-f)(b_i' - q_i^\mu) \pm (1-f)\sqrt{M\big(f^2(E-S+1) + (S-1)f(1-f)\big)} \\
&\approx \frac{1}{2}M(b_i' - q_i^\mu) \pm \sqrt{M(f^2E + f(S-1))}
\end{aligned} \tag{12}$$

where we split the summation over all transitions E into the correct transition ϵ, incorrect transitions with different source states, and incorrect transitions with the same source state. We then assume the worst case where there are $S-1$ incorrect transitions with the same source state.

Combining the three again, we get

$$\begin{aligned}
[\mathbf{h}]_i = [\mathbf{W}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda)]_i &\approx \frac{M}{2}(b_i' - f) \pm \sqrt{M}\sqrt{f^2Q(1+S) + f^2E + f(S-1)} \\
&\propto (b_i' - f) \pm \frac{2}{\sqrt{N}}\sqrt{fQ(1+S) + fE + S - 1}
\end{aligned} \tag{13}$$

which in expectation gives us $\langle \mathbf{W}(\mathbf{q}^\mu \wedge \mathbf{s}^\lambda) \rangle \propto \mathbf{b}' - f$. The same procedure can be applied to calculate the mean and variance of the postsynaptic sum in all cases.

S3 Energy function

We consider a simplification of the network dynamics, where each block is updated asynchronously in discrete time steps, in which one block is picked at random, and the neuron with the greatest input spikes. The RSNN can then be

	Block-constant dense bipolar (bMAP)	Binary sparse block codes (SBC)	Bipolar sparse block codes (pSBC)
Block-constant dense bipolar (bMAP)	Hadamard ↓ bMAP	Hadamard ↓ pSBC	Hadamard ↓ pSBC
Binary sparse block codes (SBC)	Hadamard ↓ pSBC	LCC ↓ SBC	LCC ↓ pSBC
Bipolar sparse block codes (pSBC)	Hadamard ↓ pSBC	LCC ↓ pSBC	LCC ↓ pSBC

Binding operation ↓ Resulting hypervector type

Fig. S1: Suitable binding operations between hypervectors of different type, as well as the type of hypervector it produces.

coarsely described by a neural state vector $\mathbf{z}_t \in \{0,1\}^N$, indicating which neuron is active in each block on time step t. If we are in a regime where we can ignore the effects of the bridge and transition matrices, then we can make the substitution $\mathbf{W} = \mathbf{W}^{\text{attr.}}$. We can split the neural state vector $\mathbf{z}_t$ into the vector describing only the components in the updating block, $\mathbf{z}_t^b$, and the vector describing all the other $N - L$ components $\mathbf{z}_t^{\neg b}$, with $\mathbf{z}_t = \mathbf{z}_t^b + \mathbf{z}_t^{\neg b}$. If we define the energy E_t at time step t as $E_t = -\mathbf{z}_t^\intercal \mathbf{W} \mathbf{z}_t$, then the change in energy $\Delta E_{t+1} = E_{t+1} - E_t$ is given by

$$\Delta E_{t+1} = -(\mathbf{z}_{t+1}^b - \mathbf{z}_t^b)^\intercal \mathbf{W} \mathbf{z}_t^{\neg b} \tag{14}$$

where we have used that $\mathbf{W}$ is symmetric, and there are no interactions (weight is 0) between neurons in the same block [2]. Now, the block-WTA dynamics mean that each block effectively computes an argmax over its inputs, meaning that $\mathbf{z}_{t+1}^b$ maximises the inner product $\mathbf{z}_{t+1}^{b\intercal} \mathbf{W} \mathbf{z}_t^{\neg b}$, and so is the maximum possible value that $\mathbf{z}_t^{b\intercal} \mathbf{W} \mathbf{z}_t^{\neg b}$ can take. Thus, we have $\Delta E_t \leq 0$, i.e. the network dynamics descend the energy surface $E(\mathbf{z})$, which by construction

$$E(\mathbf{z}) = -\mathbf{z}^\intercal \mathbf{W}_{\text{attr.}} \mathbf{z} = -\sum_{q \in Q} \big(\mathbf{z} \cdot (\mathbf{q} - f)\big)^2 \tag{15}$$

consists of parabolic energy wells at our DFA states $\mathbf{q}$.

S4 Mixing SBC and MAP VSA models

In this work, we made simultaneous use of both the binary Sparse Block Code (SBC) [3, 4] and Multiply-Add-Permute (MAP) [5, 6] VSA models, with the latter model using dense bipolar hypervectors. Contrary to VSA orthodoxy, we utilised the MAP binding operation (Hadamard product, denoted "$\circ$") between both hypervector types in the construction of the recurrent weight matrices. We did not make use of the conventional binding operation with SBC hypervectors, blockwise Local Circular Convolution (LCC), denoted "$*_b$". Here, we briefly explore the system of hypervectors that arises when mixing the MAP and SBC models.

We used the Hadamard binding operation to bind dense bipolar MAP hypervectors to SBC hypervectors, as this produced *bipolar* SBC hypervectors which were pseudo-orthogonal to the binary SBC hypervectors representing the network state at any time. The dense bipolar MAP hypervectors used here needed to be block-constant: only N/L random samples of $\{-1, +1\}$ were drawn per hypervector, one for each block, and all components in the same block took this value.

In our experiments, this was required such that an effective unbinding of the MAP vectors could be performed by masking out entire blocks. Masking out only some neurons in each block (if the MAP vectors were not block-constant) conflicts with the winner-take-all mechanism led to the incorrect states being represented. For the joint MAP-SBC hypervector system, the MAP hypervectors must be block-constant to ensure the Hadamard and LCC binding operations are associative. Otherwise, with unconstrained MAP hypervectors, one would have

$$\big(\mathbf{MAP} \circ \mathbf{SBC}_1\big) *_b \mathbf{SBC}_2 \neq \mathbf{MAP} \circ \big(\mathbf{SBC}_1 *_b \mathbf{SBC}_2\big) \tag{16}$$

which is not a desirable property with current VSA approaches, not least of all because unbinding would then require information about the previous order of binding operations. For clarity, and to differentiate unconstrained dense bipolar hypervectors from block-constant dense bipolar hypervectors, we refer to the latter as *bMAP* hypervectors.

Hadamard binding a bMAP hypervector with an SBC hypervector produces a bipolar sparse block code hypervector, which we refer to as *pSBC* hypervectors, to differentiate them from the binary SBC hypervectors. The binding here must be Hadamard rather than LCC since LCC binding would produce a dense hypervector with all indices in each block having the same value, and so the SBC hypervector would not be at all recoverable. The pSBC hypervector produced by Hadamard binding bMAP and SBC hypervectors is then pseudo-orthogonal to the SBC vector, but not to the bMAP vector, the consequences of which we will shortly discuss. Independently-generated pSBC hypervectors are then pseudo-orthogonal to each-other (they have 0 expected inner product), can be bound together by LCC binding, and unbound using inverse-LCC binding. A summary of the suitable binding operations between the three hypervector types (bMAP, SBC, pSBC) is shown in Fig. S1.

pSBC hypervectors can thus be used in the same way as SBC hypervectors, with binding by LCC and superposition by element-wise addition, but with the added ease that independently-generated pSBC hypervectors have zero expected inner product.

The more interesting case is not to consider pSBC hypervectors that are independently generated, but rather which exist as a by-product of the simultaneous use of bMAP and SBC hypervectors. Since the bMAP hypervectors are constant-valued in each block, under LCC binding their effect amounts simply to a prefactor in each block, and so Hadmard binding commutes with LCC binding, i.e.

$$\begin{aligned}\mathbf{pSBC}_1 *_b \mathbf{pSBC}_2 &= \left(\mathbf{bMAP}_1 \circ \mathbf{SBC}_1\right) *_b \left(\mathbf{bMAP}_2 \circ \mathbf{SBC}_2\right) \\ &= \left(\mathbf{bMAP}_1 \circ \mathbf{bMAP}_2\right) \circ \left(\mathbf{SBC}_1 *_b \mathbf{SBC}_2\right) \\ &= \mathbf{bMAP}_3 \circ \mathbf{SBC}_3 \\ &= \mathbf{pSBC}_3\end{aligned} \tag{17}$$

Hence, LCC binding with pSBC hypervectors can be seen as actually performing two separate binding operations in the bMAP and SBC domains.

To exemplify the use of pSBC hypervectors and to empirically demonstrate viability as a VSA framework, we turn to the commonly used VSA analogical reasoning tutorial "What is the dollar of Mexico?" [7]. There, two hypervectors $\mathbf{usa}$ and $\mathbf{mex}$ represent information about the USA and Mexico respectively, by superimposing bound role-filler pairs created from the atomic hypervectors $\mathbf{cap}, \mathbf{cur}, \mathbf{wdc}, \mathbf{dol}, \mathbf{mxc}$ and $\mathbf{pes}$ representing "capital city", "currency", "Washington DC", "dollar", "Mexico City" and "peso" respectively [7].

S4.1 Case 1: All pSBC

We consider three cases. First, that all the atomic hypervectors are independently-generated pSBC hypervectors. The USA hypervector is then formed by LCC binding the corresponding role-filler hypervector pairs,

$$\mathbf{usa}_{\text{pSBC}} = \mathbf{cap}_{\text{pSBC}} *_b \mathbf{wdc}_{\text{pSBC}} + \mathbf{cur}_{\text{pSBC}} *_b \mathbf{dol}_{\text{pSBC}} \tag{18}$$

from which one can decode the currency used in the USA by unbinding the currency hypervector via inverse LCC: The only hypervector with significant overlap with $\mathbf{usa}_{\text{pSBC}} *_b^{-1} \mathbf{cur}_{\text{pSBC}}$ is $\mathbf{dol}_{\text{pSBC}}$, the correct answer (Fig. S2a). If $\mathbf{mex}_{\text{pSBC}}$ is constructed analagously to $\mathbf{usa}_{\text{pSBC}}$, the "What's the dollar of Mexico" example can be performed by calculating $\mathbf{mex}_{\text{pSBC}} *_b^{-1} \mathbf{usa}_{\text{pSBC}} *_b \mathbf{dol}_{\text{pSBC}}$, which is similar only to $\mathbf{pes}_{\text{pSBC}}$, as desired (Fig. S2d).

S4.2 Case 2: Roles SBC, fillers bMAP

In the second case, we consider that the roles (currency, capital) are represented by SBC hypervectors, and the fillers (Washington DC, dollar, ...) are represented by bMAP hypervectors. The USA hypervector is then constructed as

$$\mathbf{usa}_{\text{pSBC}} = \mathbf{cap}_{\text{SBC}} \circ \mathbf{wdc}_{\text{bMAP}} + \mathbf{cur}_{\text{SBC}} \circ \mathbf{dol}_{\text{bMAP}} \tag{19}$$

Decoding the currency of the USA is performed by unbinding the currency hypervector, again by inverse LCC. The resulting expression reads

$$\begin{aligned}\mathbf{usa}_{\text{pSBC}} *_b^{-1} \mathbf{cur}_{\text{SBC}} &= \left(\mathbf{cap}_{\text{SBC}} *_b^{-1} \mathbf{cur}_{\text{SBC}}\right) \circ \mathbf{wdc}_{\text{bMAP}} + \left(\mathbf{cur}_{\text{SBC}} *_b^{-1} \mathbf{cur}_{\text{SBC}}\right) \circ \mathbf{dol}_{\text{bMAP}} \\ &= \left(\mathbf{cap}_{\text{SBC}} *_b^{-1} \mathbf{cur}_{\text{SBC}}\right) \circ \mathbf{wdc}_{\text{bMAP}} + \mathbf{1}_{\text{SBC}} \circ \mathbf{dol}_{\text{bMAP}}\end{aligned} \tag{20}$$

where $\mathbf{1}_{\text{SBC}}$ is the identity hypervector under LCC, given by a hypervector where only the first element in each block is active. Hence, to correctly calculate similarities with the bMAP hypervectors, they should first be bound to the SBC identity vector. Results are shown in Fig. S2b. Note that, although the $\mathbf{dol}_{\text{bMAP}}$ hypervector correctly has the greatest

similarity, the $\mathbf{wdc}_{\text{bMAP}}$ hypervector also has a notable similarity, despite being bound to two SBC hypervectors. This is because the SBC hypervectors are not balanced, consisting only of non-negative values, and so the similarity with the $\mathbf{wdc}_{\text{bMAP}}$ is reduced by a factor of $1/L$, rather than to 0. A similar effect would be observed if another bound copy of $\mathbf{dol}_{\text{bMAP}}$ were superimposed, but bound to a different SBC hypervector.

Thus, unlike most VSA approaches, this binding method has produced a hypervector which retains similarity with one of its bound hypervectors, while still unambiguously coding for the binding of the two symbols. Although this raises the possibility of false positives, it opens an interesting avenue for future research exploring whether such properties could be used beneficially.

S4.3 Case 3: Roles bMAP, fillers SBC

In the third case, the roles are represented by bMAP hypervectors and the fillers by SBC hypervectors. To decode the currency of the USA, we unbind the currency hypervector $\mathbf{cur}_{\text{bMAP}}$, now via a Hadamard product, resulting in

$$\mathbf{usa}_{\text{pSBC}} \circ \mathbf{cur}_{\text{bMAP}} = \mathbf{cap}_{\text{bMAP}} \circ \mathbf{cur}_{\text{bMAP}} \circ \mathbf{wdc}_{\text{SBC}} + \mathbf{dol}_{\text{SBC}} \tag{21}$$

This expression is maximally similar to the $\mathbf{dol}_{\text{SBC}}$ as desired (Fig. S2c). However, due to the presence of an unbalanced SBC hypervector, this expression will also have a similarity of approximately $1/L$ with all the other SBC vectors.

Computing the full "What's the dollar of Mexico?" analogy as in [7] requires binding the USA to Mexico mapping to dollar $\mathbf{mex}_{\text{pSBC}} *_b^{-1} \mathbf{usa}_{\text{pSBC}} *_b \mathbf{dol}_{\text{SBC}}$, which reads:

$$\mathbf{mxc}_{\text{SBC}} *_b^{-1} \mathbf{wdc}_{\text{SBC}} *_b \mathbf{dol}_{\text{SBC}} + \mathbf{cap}_{\text{bMAP}} \circ \mathbf{cur}_{\text{bMAP}} \big(\mathbf{mxc}_{\text{SBC}} + \mathbf{pes}_{\text{SBC}} *_b^{-1} \mathbf{wdc}_{\text{SBC}} *_b \mathbf{dol}_{\text{SBC}}\big) + \mathbf{pes}_{\text{SBC}} \tag{22}$$

The first set of terms forms an SBC hypervector, that will have similarity $1/L$ with any of our SBC hypervectors. The second set of terms is a balanced pSBC hypervector, pseudo-orthogonal to all our atomic hypervectors. Since the only unbound term is the peso, this expression is maximally similar to $\mathbf{pes}_{\text{SBC}}$, as desired. However, it will have an overlap of $1 + 1/L$ due to the expected overlap with the first set of terms (Fig. S2f).

S4.4 Should multiple VSA models be explored?

When making simultaneous use of multiple VSA models, certain considerations must thus be taken into account. First, different hypervector types may affect what similarity structure is exposed/hidden through binding. Second, prior knowledge about the structure of a hypervector may be required to ascertain which binding operations are suitable. These requirements arguably detract from the self-consistency and simplistic elegance of VSAs that make them so attractive as a unified system of representation.

On the other hand, if one believes that the brain may utilise the principles of VSAs, it is nonetheless unlikely that a single common choice of representation would be used throughout. For this reason alone, it is worth considering how VSAs perform when multiple representation schemes are at play, and what additional algebraic structure is introduced. It should be explored further what general properties arise when multiple VSA models are integrated together, potentially operating on different parts of the hypervector or at different scales, and how this additional representational flexibility can be effectively utilised.

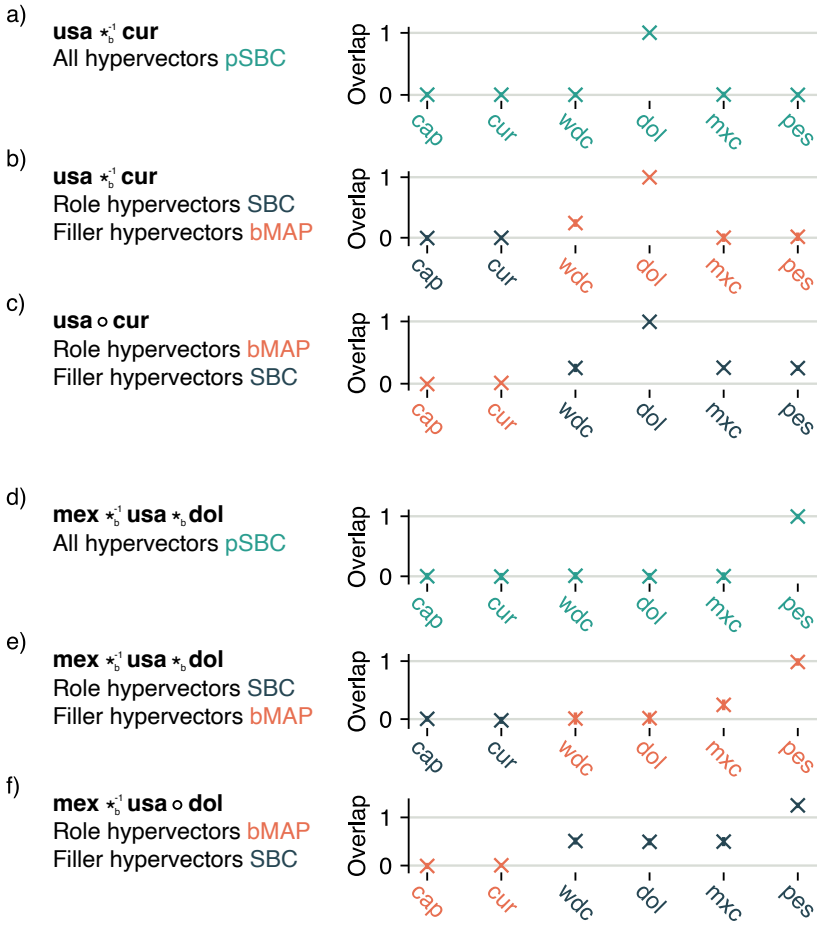

Fig. S2: Hypervector decoding with different combinations of SBC, bMAP, and pSBC hypervectors. **a)** Inner products between the atomic hypervectors and the hypervector formed by unbinding the currency hypervector from the USA hypervectors, if all atomic hypervectors are independently-generated pSBC hypervectors, or **b)** the role and filler hypervectors are SBC and bMAP hypervectors respectively, or **c)** vice versa. **d-f)** The same choices of hypervectors, put performing the "What's the dollar of Mexico?" decoding task [7]. Values used $N = 1024$, $L = 4$, 100 trials.

Parameter	Description	First simulation	Crossbar experiment	Big simulation	Loihi
N	Number of neurons	2048	64	2048	1024
L	Block length	8	8	16	8
$\langle \lvert w_{ij} \rvert \rangle$	Synaptic weight strength	$\approx 0.1\,\mathrm{mV\,F^{-1}}$	$\approx 4\,\mathrm{mV\,F^{-1}}$	$\approx 0.2\,\mathrm{mV\,F^{-1}}$	≈ 45
C	Membrane capacitance	$1\,\mathrm{F}$	$1\,\mathrm{F}$	$1\,\mathrm{F}$	-
τ_m	Membrane time constant	$20\,\mathrm{ms}$	$20\,\mathrm{ms}$	$20\,\mathrm{ms}$	10
u_θ	Spiking threshold	$20\,\mathrm{mV}$	$20\,\mathrm{mV}$	$20\,\mathrm{mV}$	1000
u_{rest}	Resting potential	$25\,\mathrm{mV}$	$25\,\mathrm{mV}$	$25\,\mathrm{mV}$	0
u_{reset}	Reset potential	$0\,\mathrm{mV}$	$0\,\mathrm{mV}$	$0\,\mathrm{mV}$	0
τ_{syn}	Synaptic time constant	$20\,\mathrm{ms}$	$20\,\mathrm{ms}$	$20\,\mathrm{ms}$	-
τ_{ref}	Refractory period	$10\,\mathrm{ms}$	$10\,\mathrm{ms}$	$10\,\mathrm{ms}$	0
τ_{readout}	Readout kernel timescale	$10\,\mathrm{ms}$	$10\,\mathrm{ms}$	$10\,\mathrm{ms}$	10

Table. 1: Parameters used in the simulation, closed-loop memristive crossbar and Loihi setups. The magnitude of the synaptic weights $\langle \lvert w_{ij} \rvert \rangle$ is the total amount of charge that each spike adds to a postsynaptic neuron's membrane.

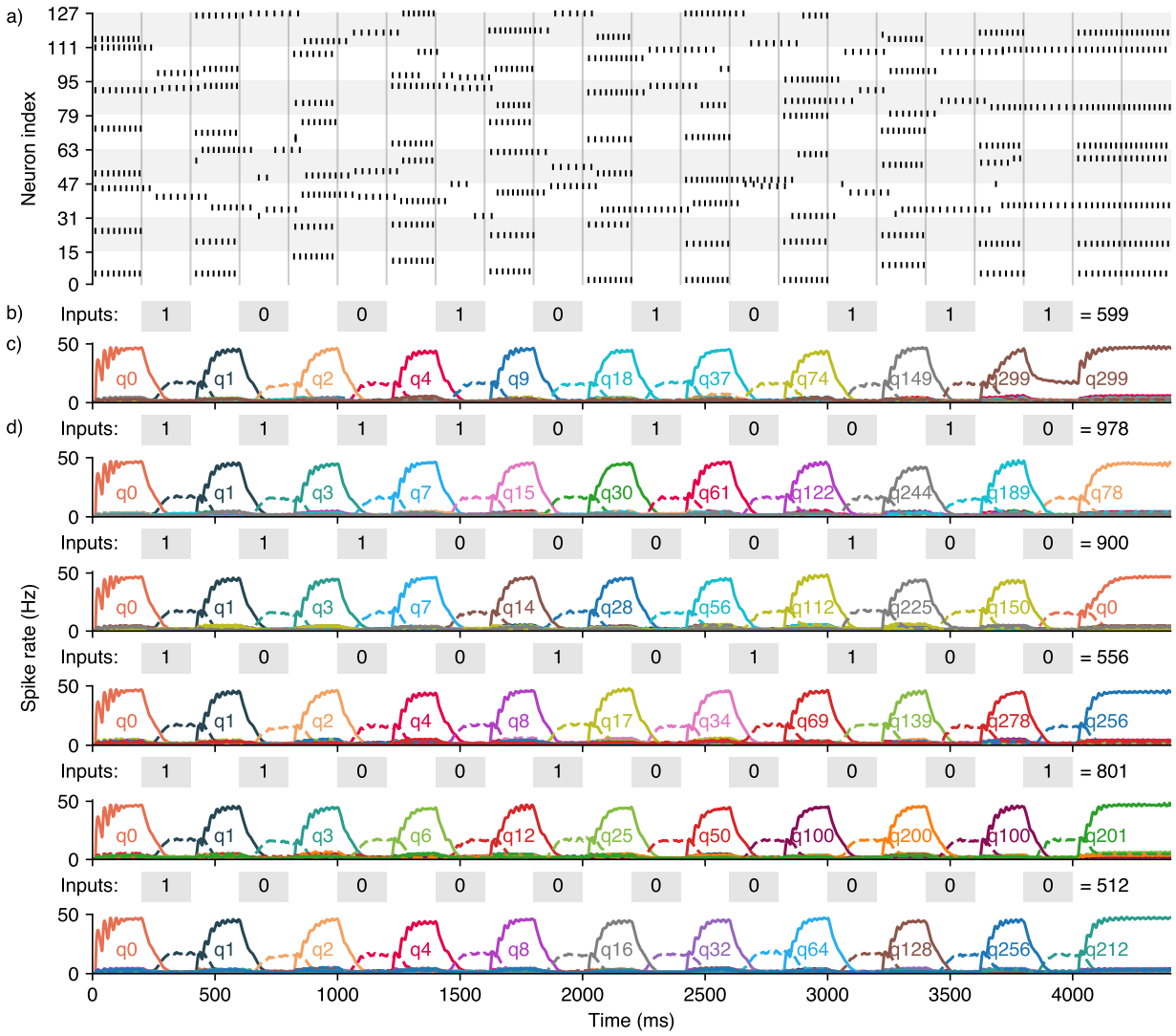

Fig. S3: A 300-state DFA embedded into an RSNN in simulation with ideal weights. The chosen DFA implements modular division of binary numbers by 300, generated in the same way as in Fig. 1. **a)** A raster plot of the first 64 neurons' spikes, with block length $L = 16$. **b)** The symbolic inputs to the RSNN. **c)** The mean firing rate of the neurons in each attractor state. **d)** Different input sequences are given to otherwise identical simulations. In all cases, the RSNN performs the correct walk between attractor states, with the final inhabited state indicating the result of the input binary number $\bmod 300$. By specifying our RSNN dynamics in the VSA framework, scaling up the algorithm to large state machines is seamless, and is limited only by the memory capacity of the attractor network.

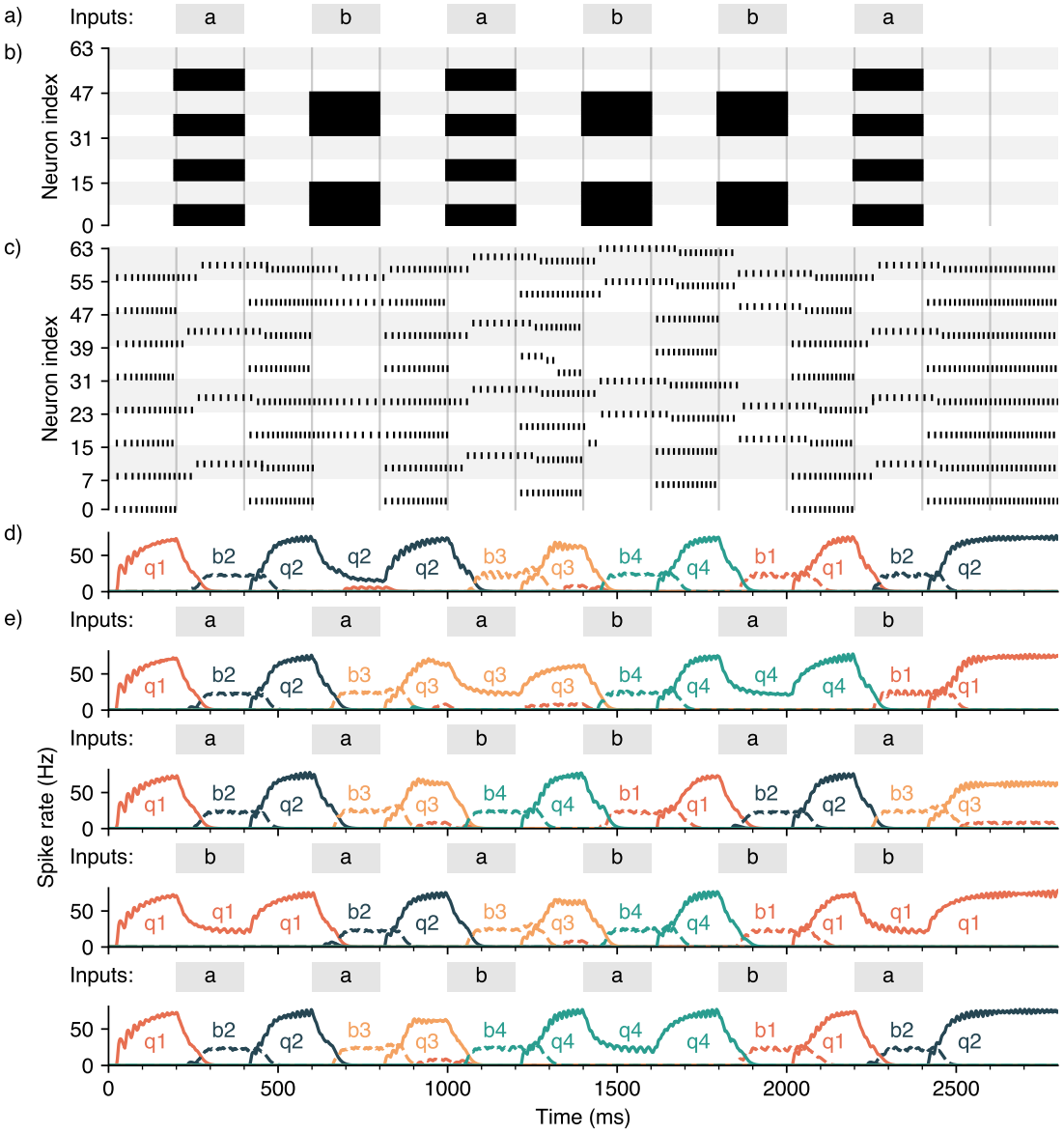

Fig. S4: The closed-loop memristive crossbar experiment, configured to embed a second DFA. The DFA is given by $q_0 \xrightarrow{a} q_1 \xrightarrow{a} q_2 \xrightarrow{b} q_3 \xrightarrow{b} q_0$. **a)** The symbolic input to the network. **b)** The masking input to the network, with a separate mask for a and b. Due to size constraints, the mask vectors were chosen to be orthogonal. **c)** The spiking activity of neurons within the RSNN. **d)** The mean firing rates of the neurons within each attractor state. **e)** The same experiment was performed with different sequences of inputs, which cause corresponding different walks between the states. The RSNN performs the correct walk between attractor states in all cases. Although there is erroneous neuron firing in some blocks (due to the nonideal device conductances), the distributed states ensure the overall dynamics are unaffected.

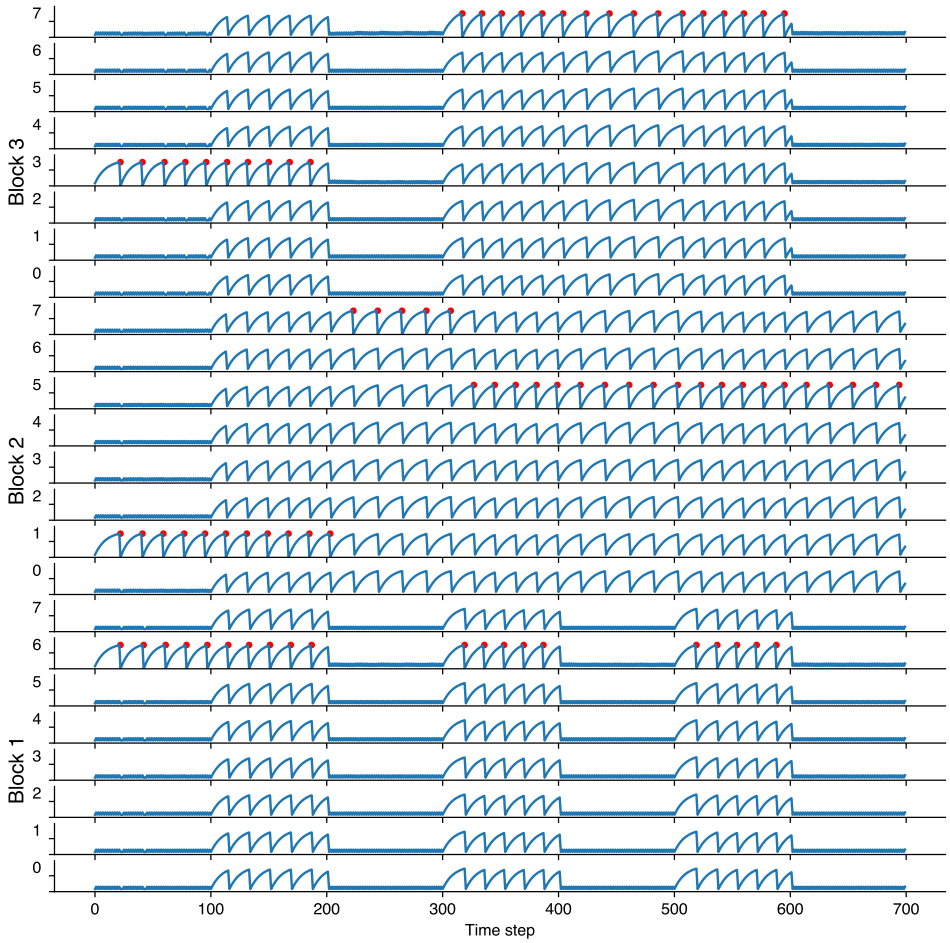

Fig. S5: The winner-take-all mechanism implemented on Loihi 2. Although every neuron is receiving enough excitatory input that it will spike if unhindered, the WTA mechanism - implemented using shunting inhibition - allows only the neuron with the greatest integrated current to spike. This is an easy way to stabilize the attractor dynamics and constrains the neural activity to sparse block codes.

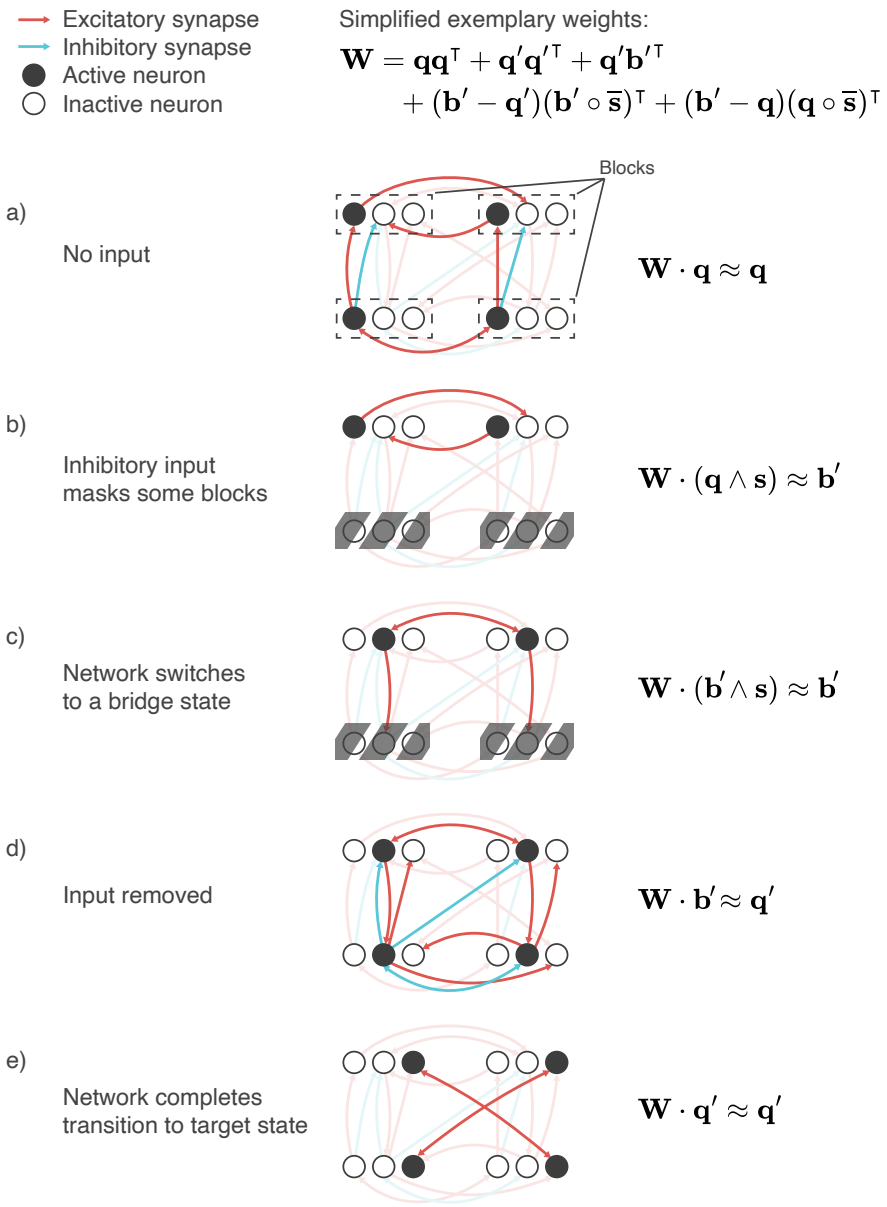

Fig. S6: Mechanistic description of how the constructed weight matrices result in attractor-transition dynamics. A simplified weight matrix is shown, which stores a transition from state $\mathbf{q}$ to $\mathbf{q}'$ through bridge state $\mathbf{b}'$. For simplicity, the $\mathbf{q}$, $\mathbf{b}'$ and $\mathbf{q}'$ hypervectors are here chosen as having the first, second and third neuron in each block ($L = 3$) be active respectively. From one panel to the next, the neuron with the greatest input in each block becomes active, which can be calculated by summing the number of +1 excitatory and -1 inhibitory synaptic inputs. **a)** The network starts in stable attractor state $\mathbf{q}$. **b)** Applying $\mathbf{s}$ as a network mask ($s_i = 0/1$ for the bottom/top neuron blocks) removes some blocks from the picture, such that $\mathbf{q}$ ceases to be a stable state. **c)** This causes the network to transition to the $\mathbf{b}'$ state, which is then stable. **d)** The eventual removal of the input causes the network to briefly inhabit the full, unmasked $\mathbf{b}'$ state, which is not stable. **e)** The network transitions to the $\mathbf{q}'$ state, completing the transition. The connectivity shown here does not correspond exactly to the weight matrix above, it has been chosen for visual simplicity. In general, the states will not be orthogonal but will be randomly-generated with expected overlap N/L^2.

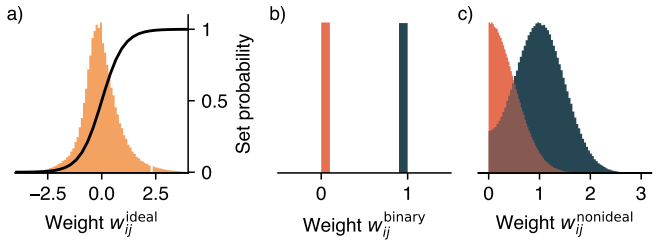

Fig. S7: Histograms of the recurrent weight values in simulation before and after nonidealities are applied. **a)** The ideal weights are shown in yellow, with a sigmoidal transfer function overlain. **b)** The ideal weights are stochastically binarised according to the sigmoidal set probability. **c)** The weights after adding independent Gaussian noise to each binarised weight value. The two histograms were each normalised to unit height. These nonidealities emulate that the synaptic weights are imprecise, unreliable and noisy. The weights w_{ij}^{nonideal} were used in the RSNN in Fig. 1.

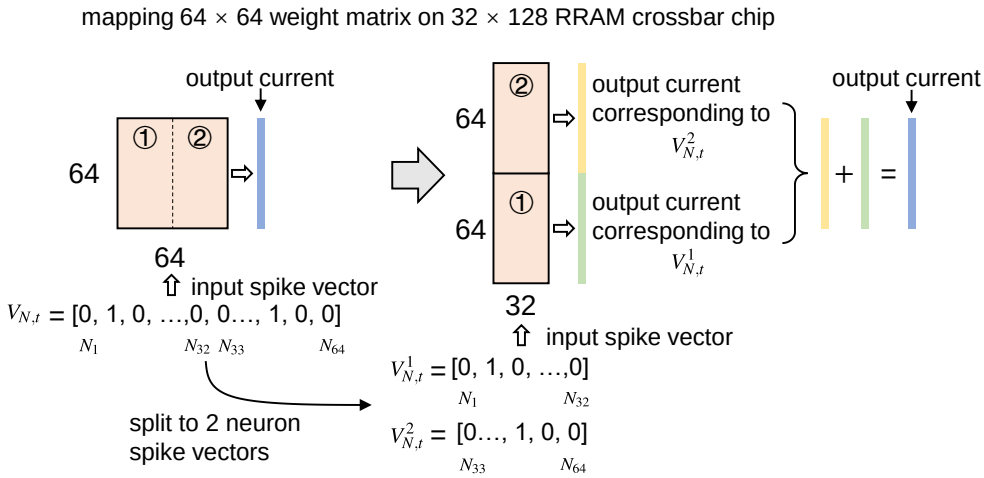

Fig. S8: The scheme to map the 64×64 weight matrix to the 32×128 RRAM crossbar chip. We split the size 64 neuron spike vector into two size 32 vectors. We then apply each vector separately as input to the RRAM crossbar one after the other and read the output currents as depicted. The sum of the two size 64 output current vectors is equivalent to applying a size 64 spike vector as input to a 64×64 crossbar array (shown left). The summation of the two output current vectors is performed in software and then fed back into the RSNN simulation as postsynaptic currents.

Supplementary references

[1] A. Thomas, S. Dasgupta, and T. Rosing, "A theoretical perspective on hyperdimensional computing", *Journal of Artificial Intelligence Research*, vol. 72, pp. 215–249, 2022. DOI: 10.1613/jair.1.12664.

[2] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 5, pp. 815–826, 1983. DOI: 10.1109/TSMC.1983.6313075.

[3] M. Laiho, J. H. Poikonen, P. Kanerva, and E. Lehtonen, "High-dimensional computing with sparse vectors," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2015, pp. 1–4. DOI: 10.1109/BioCAS.2015.7348414.

[4] E. P. Frady, D. Kleyko, and F. T. Sommer, "Variable binding for sparse distributed representations: Theory and applications", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 5, pp. 2191–2204, 2023. DOI: 10.1109/TNNLS.2021.3105949.

[5] R. W. Gayler, "Multiplicative binding, representation operators & analogy", *Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences*, 1998.

[6] D. Kleyko, D. A. Rachkovskij, E. Osipov, and A. Rahimi, "A survey on hyperdimensional computing aka vector symbolic architectures, part I: Models and data transformations", *ACM Computing Surveys*, vol. 55, no. 6, 130:1–130:40, 2022. DOI: 10.1145/3538531.

[7] P. Kanerva, "Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors", *Cognitive Computation*, vol. 1, no. 2, pp. 139–159, 2009. DOI: 10.1007/s12559-009-9009-8.