Article

Neuron

Recurrent Network Models of Sequence Generation and Memory

Highlights

- Sequences emerge in random networks by modifying a small fraction of their connections
- Analysis reveals new circuit mechanism for input-dependent sequence propagation
- Sequential activation may provide a dynamic mechanism for short-term memory

Authors

Kanaka Rajan, Christopher D. Harvey, David W. Tank

Correspondence

krajan@princeton.edu (K.R.), harvey@hms.harvard.edu (C.D.H.), dwtank@princeton.edu (D.W.T.)

In Brief

Rajan et al. show that neural sequences similar to those observed during memory-based decision-making tasks can be generated by minimally structured networks. Sequences may effectively mediate the short-term memory engaged in these tasks.

Rajan et al., 2016, Neuron *90*, 128–142
April 6, 2016
http://dx.doi.org/10.1016/j.neuron.2016.02.009

CellPress

CellPress

Neuron
Article

Recurrent Network Models of Sequence Generation and Memory

Kanaka Rajan,[1,*] Christopher D. Harvey,[2,*] and David W. Tank[3,*]
[1]Joseph Henry Laboratories of Physics and Lewis–Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
[2]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA
[3]Department of Molecular Biology and Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA
*Correspondence: krajan@princeton.edu (K.R.), harvey@hms.harvard.edu (C.D.H.), dwtank@princeton.edu (D.W.T.)
http://dx.doi.org/10.1016/j.neuron.2016.02.009

SUMMARY

Sequential activation of neurons is a common feature of network activity during a variety of behaviors, including working memory and decision making. Previous network models for sequences and memory emphasized specialized architectures in which a principled mechanism is pre-wired into their connectivity. Here we demonstrate that, starting from random connectivity and modifying a small fraction of connections, a largely disordered recurrent network can produce sequences and implement working memory efficiently. We use this process, called Partial In-Network Training (PINning), to model and match cellular resolution imaging data from the posterior parietal cortex during a virtual memory-guided two-alternative forced-choice task. Analysis of the connectivity reveals that sequences propagate by the cooperation between recurrent synaptic interactions and external inputs, rather than through feedforward or asymmetric connections. Together our results suggest that neural sequences may emerge through learning from largely unstructured network architectures.

INTRODUCTION

Sequential firing has emerged as a prominent motif of population activity in temporally structured behaviors, such as short-term memory and decision making. Neural sequences have been observed in many brain regions including the cortex (Luczak et al., 2007; Schwartz and Moran, 1999; Andersen et al., 2004; Pulvermüller and Shtyrov, 2009; Buonomano, 2003; Ikegaya et al., 2004; Tang et al., 2008; Seidemann et al., 1996; Fujisawa et al., 2008; Crowe et al., 2010; Harvey et al., 2012), hippocampus (Nádasdy et al., 1999; Louie and Wilson, 2001; Pastalkova et al., 2008; Davidson et al., 2009), basal ganglia (Barnes et al., 2005; Jin et al., 2009), cerebellum (Mauk and Buonomano, 2004), and area HVC of the songbird (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007). In all these instances, the observed sequences span a wide range of time durations, but individual neurons fire transiently only during a small portion of the full sequence. The ubiquity of neural sequences suggests that they are of widespread functional use and thus may be produced by general circuit-level mechanisms.

Sequences can be produced by highly structured neural circuits or by more generic circuits adapted through the learning of a specific task. Highly structured circuits of this type have a long history (Kleinfeld and Sompolinsky, 1989; Goldman, 2009), e.g., as synfire chain models (Hertz and Prügel-Bennett, 1996; Levy et al., 2001; Hermann et al., 1995; Fiete et al., 2010), in which excitation flows unidirectionally from one active neuron to the next along a chain of connected neurons, or as ring attractor models (Ben-Yishai et al., 1995; Zhang, 1996), in which increased (central) excitation between nearby neurons surrounded by long-range inhibition and asymmetric connectivity are responsible for time-ordered neural activity. These models typically require imposing a task-specific mechanism (e.g., for sequences) into their connectivity, producing specialized networks. Neural circuits are highly adaptive and involved in a wide variety of tasks, and sequential activity often emerges through learning of a task and retains significant variability. It is therefore unlikely for highly structured approaches to produce models with flexible circuitry or to generate dynamics with the temporal complexity needed to recapitulate experimental data.

In contrast, random networks interconnected with excitatory and inhibitory connections in a balanced state (Sompolinsky et al., 1988), rather than being specifically designed for one single task, have been modified by training to perform a variety of tasks (Buonomano and Merzenich, 1995; Buonomano, 2005; Williams and Zipser, 1989; Pearlmutter, 1989; Jaeger and Haas, 2004; Maass et al., 2002, 2007; Sussillo and Abbott, 2009; Jaeger, 2003). Here we built on these lines of research and asked whether a general implementation using relatively unstructured random networks could create sequential dynamics resembling experimental data. We used data from sequences observed in the posterior parietal cortex (PPC) of mice trained to perform a two-alternative forced-choice (2AFC) task in a virtual reality environment (Harvey et al., 2012), and we also constructed models that extrapolated beyond these experimental data.

To address how much network structure is required for sequences like those observed in the recordings, we introduced a new modeling framework called Partial In-Network Training or PINning. In this scheme, any desired fraction of the initially random connections within the network can be modified by a synaptic change algorithm, enabling us to explore the full range of network architectures between completely random and fully

structured. Using networks constructed by PINning, we first demonstrated that sequences resembling the PPC data were most consistent with minimally structured circuitry, with small amounts of structured connectivity to support sequential activity patterns embedded in a much larger fraction of unstructured connections. Next we investigated the circuit mechanism of sequence generation in such largely random networks containing some learned structure. Finally, we determined the role sequences play in short-term memory, e.g., by storing information during the delay period about whether a left or right turn was indicated early in the 2AFC task (Harvey et al., 2012). Going beyond models meant to reproduce the experimental data, we analyzed multiple sequences initiated by different sensory cues and computed the capacity of this form of short-term memory.

RESULTS

Sequences from Highly Structured or Random Networks Do Not Match PPC Data

Our study is based on networks of rate-based model neurons, in which the outputs of individual neurons are firing rates and the units are interconnected through excitatory and inhibitory synapses of various strengths (Experimental Procedures). To interpret the outputs of the rate networks in terms of experimental data, we extracted firing rates from the calcium fluorescence signals recorded in the PPC using two complementary deconvolution methods (Figures 1C and 1D; see also Figure S1 and Experimental Procedures). We defined two measures to compare the rates from the model to the rates extracted from data. The first, *bVar*, measures the stereotypy of the data or the network output by quantifying the variance explained by the translation of an activity profile with an invariant shape (Figure 1G; Experimental Procedures). The second metric, *pVar*, measures the percentage variance of the PPC data (Harvey et al., 2012) captured by the different network outputs, and it is useful for tracking network performance across different parameters (Figure 1H; Experimental Procedures). *bVar* and *pVar* are used throughout this paper (Figures 1, 2, 5, and 6).

Many models have suggested that highly structured synaptic connectivity, i.e., containing ring-like or chain-like interactions, is responsible for neural sequences (Ben-Yishai et al., 1995; Zhang, 1996; Hertz and Prügel-Bennett, 1996; Levy et al., 2001; Hermann et al., 1995; Fiete et al., 2010). We therefore first asked how much of the variance in long-duration neural sequences observed experimentally, e.g., in the PPC (Harvey et al., 2012), measured as *bVar*, was consistent with a bump of activity moving across the network, as expected from highly structured connectivity. The stereotypy of PPC sequences, which were averaged over hundreds of trials and pooled across different animals, was found to be quite small (*bVar* = 40% for Figures 1D and 2A). *bVar* was lower in both single-trial data (10%–15% for the data in Figure S13 of Harvey et al., 2012) and trial-averaged data from a single mouse (15% for the data in Figure 2c of Harvey et al., 2012). The small fraction of the variance explained by a moving bump (low *bVar*), combined with a weak relationship between the activity of a neuron and anatomical location in the PPC (Figure 5d in Harvey et al., 2012), motivated us to consider network architectures with disordered connectivity composed of a balanced set of excitatory and inhibitory weights drawn independently from a random distribution (Experimental Procedures).

In random network models, when excitation and inhibition are balanced on average, the ongoing dynamics have been shown to be chaotic (Sompolinsky et al., 1988); however, the presence of external stimuli can channel the ongoing dynamics in these networks by suppressing their chaos (Molgedey et al., 1992; Bertschinger and Natschläger, 2004; Rajan et al., 2010, 2011). In experiments, strong inputs have also been shown to reduce the Fano factor and trial-to-trial variability associated with spontaneous activity (Churchland et al., 2010; White et al., 2012). Thus, we asked whether PPC-like sequences could be constructed either from the spontaneous activity or the input-driven dynamics of random networks. We simulated a random network of rate-based neurons operating in a chaotic regime (Figure 1A; Experimental Procedures; N = 437, network size chosen to match the dataset under consideration). The individual firing rates were normalized by the maximum over the trial duration (10.5 s, Figure 1C) and sorted in ascending order of their times of center-of-mass (t_{COM}), matching the procedures applied to the experimental data (Harvey et al., 2012). Although the resulting ordered spontaneous activity was sequential (not shown, but similar to Figure 1B), the level of extra-sequential background (*bVar* = 5% ± 2% and *pVar* = 0.15% ± 0.1%) was higher than data. Sparsifying this background activity by increasing the threshold of the sigmoidal activation function increased *bVar* to a maximum of 22%, still considerably smaller than the data value of 40% (Figure S4).

Next we added time-varying inputs to the random network to represent the visual stimuli in the virtual environment (right panel of Figure 1A; Experimental Procedures). The sequence obtained by normalizing and sorting the rates from the input-driven random network did not match the PPC data (*bVar* = 10% ± 2% and *pVar* = 0.2% ± 0.1%; Figures 1B, left panel of 1H, and S4).

External inputs and disordered connectivity were insufficient to evoke sequences resembling the data (Figure 1B), which are more structured and temporally constrained (Figures 1C, 1D, and 2A compared to Figure 1B). Therefore, sequences like those observed during timing and memory experiments (Harvey et al., 2012) are unlikely to be an inherent property of completely random networks. Furthermore, since neural sequences arise during the learning of various experimental tasks, we asked whether initially disordered networks could also be modified by training to produce realistic sequences.

Temporally Constrained Neural Sequences Emerge with Synaptic Modification

To construct networks that match the activity observed in the PPC, we developed a training scheme, Partial In-Network Training or PINning, in which a user-defined fraction of synapses were modified (Experimental Procedures). In this synaptic modification scheme, the inputs to individual model neurons were compared directly with target functions or templates derived from the experimental data (Fisher et al., 2013), on both left and right correct-choice outcome trials from 437 trial-averaged neurons, pooled across six mice, during a 2AFC task (Figures 2A and 5D). During training, the internal synaptic weights in the connectivity matrix of the recurrent network were modified using a variant of the recursive least-squares

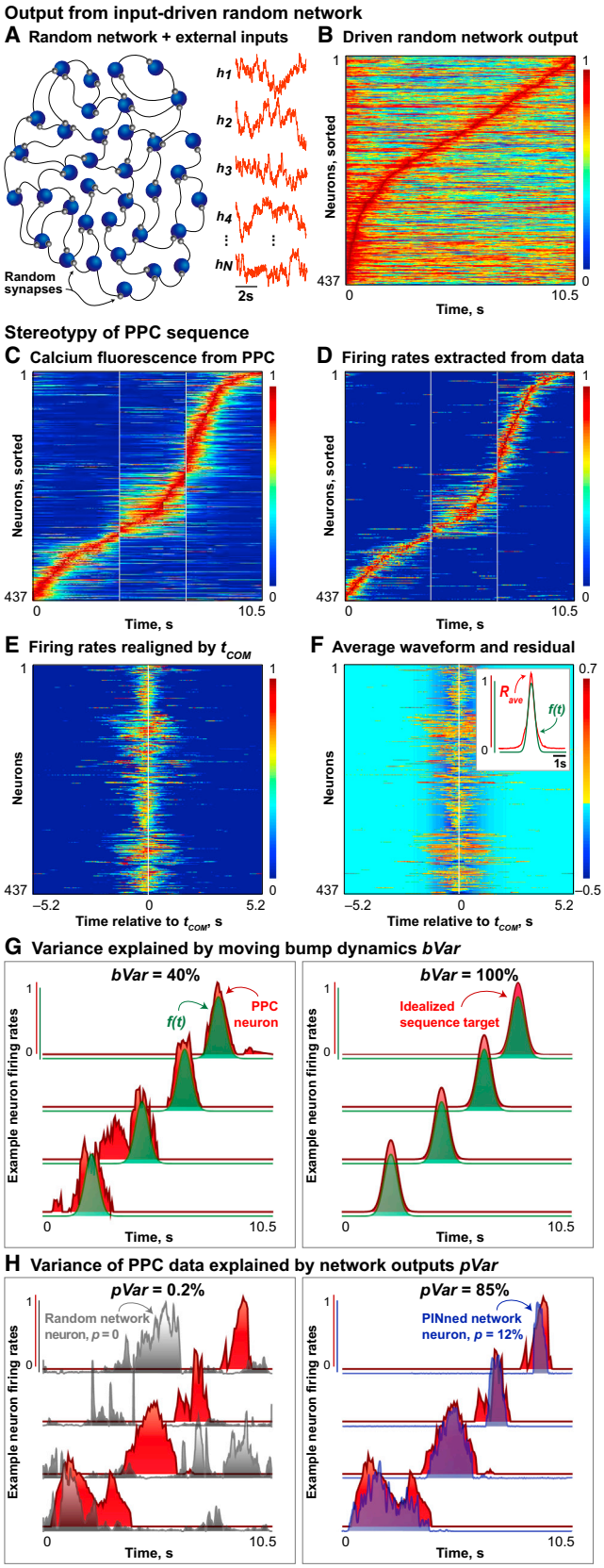


Figure 1. Stereotypy of PPC Sequences and Random Network Output

(A) Schematic of a randomly connected network of rate neurons (blue) operating in a spontaneously active regime and a few example irregular inputs (orange traces) are shown here. Random synapses are depicted in gray. Note that our networks were at least as big as the size of the data under consideration (Harvey et al., 2012) and were typically all-to-all connected, but only a fraction of these neurons and their interconnections are depicted in these schematics.

(B) When the individual firing rates of the random (untrained) network driven by external inputs in (A) were normalized by the maximum per neuron and sorted by their times of center-of-mass (t_{COM}), the activity appeared ordered. However, this sequence contained large amounts of background activity relative to the PPC data (*bVar* = 12%, *pVar* = 0.2%).

(C) Calcium fluorescence (i.e., normalized $\Delta F/F$) data were collected from 437 trial-averaged PPC neurons during a 10.5 s-long 2AFC experiment from both left and right correct-choice outcomes, pooled from six mice. Vertical gray lines indicate the time points corresponding to the different epochs of the task: the cue period ending at 4 s, the delay period ending at 7 s, and the turn period concluding at the end of the trial, at 10.5 s.

(D) Normalized firing rates extracted from (C) using deconvolution methods are shown here (see also Figure S1).

(E) The firing rates from the 437 neurons shown in (D) were realigned by their t_{COM} and plotted here.

(F) (Inset) A typical waveform (R_{ave} in red), obtained by averaging the realigned rates shown in (E) over neurons, and a Gaussian curve with mean = 0 and variance = 0.3 (*f(t)*, green) that best fit the neuron-averaged waveform are plotted here. (Main panel) Residual activity not explained by translations of the best fit to R_{ave}, *f(t)*, is shown here.

(G) Variance in the population activity explained by translations of the best fit to R_{ave}, *f(t)*, is a measure of the stereotypy (*bVar*). Left panel shows the normalized firing rates from four example PPC neurons (red) from (D) and the curves *f(t)* for each neuron (green); *bVar* = 40% for these data. Right panel shows the normalized firing rates of four model neurons (red) from a network generating an idealized sequence (see also Figure S2A) and the corresponding curves *f(t)* for each (green); *bVar* = 100% here.

(H) Variance of the PPC data explained by the outputs of different PINned networks is given by *pVar* and illustrated with four example neurons here. The left panel shows the normalized firing rates of four example PPC neurons (red) picked from (D) and four model neuron outputs from a random network driven by time-varying external inputs (gray, network schematized in A) with no training (p = 0). For this example, *pVar* = 0.2%. The right panel shows the same PPC neurons as in the left panel in red, along with four model neuron outputs from a sparsely PINned network with p = 12% plastic synapses. For this example, *pVar* = 85%.

(RLS) or first-order reduced and controlled error (FORCE) learning rule (Haykin, 2002; Sussillo and Abbott, 2009) until the network rates matched the target functions (Experimental Procedures). Crucially, the learning rule was applied only to all the synapses of a randomly selected and often small fraction of neurons, such that only a fraction p ($pN^2 << N^2$) of the total number of synapses was modified (plastic synapses, depicted in orange in Figures 2B, 5A, and 6A). While every neuron had a target function, only the outgoing synapses from a subset of neurons (i.e., from pN chosen neurons) were subject to the learning rule (Figure 2B; Experimental Procedures). The remaining elements of the synaptic matrix remained unmodified and in their random state (random synapses, depicted in gray in Figures 1A, 2B, 5A, and 6A). The PINning method therefore provided a way to span the entire spectrum of possible neural architectures, from disordered networks with random connections (p = 0 for Figures 1A and 1B), through networks with partially structured connectivity (p < 25% for Figures 2, 5, and 6), to networks containing entirely trained connections (p = 100%, Figures 3D–3F).

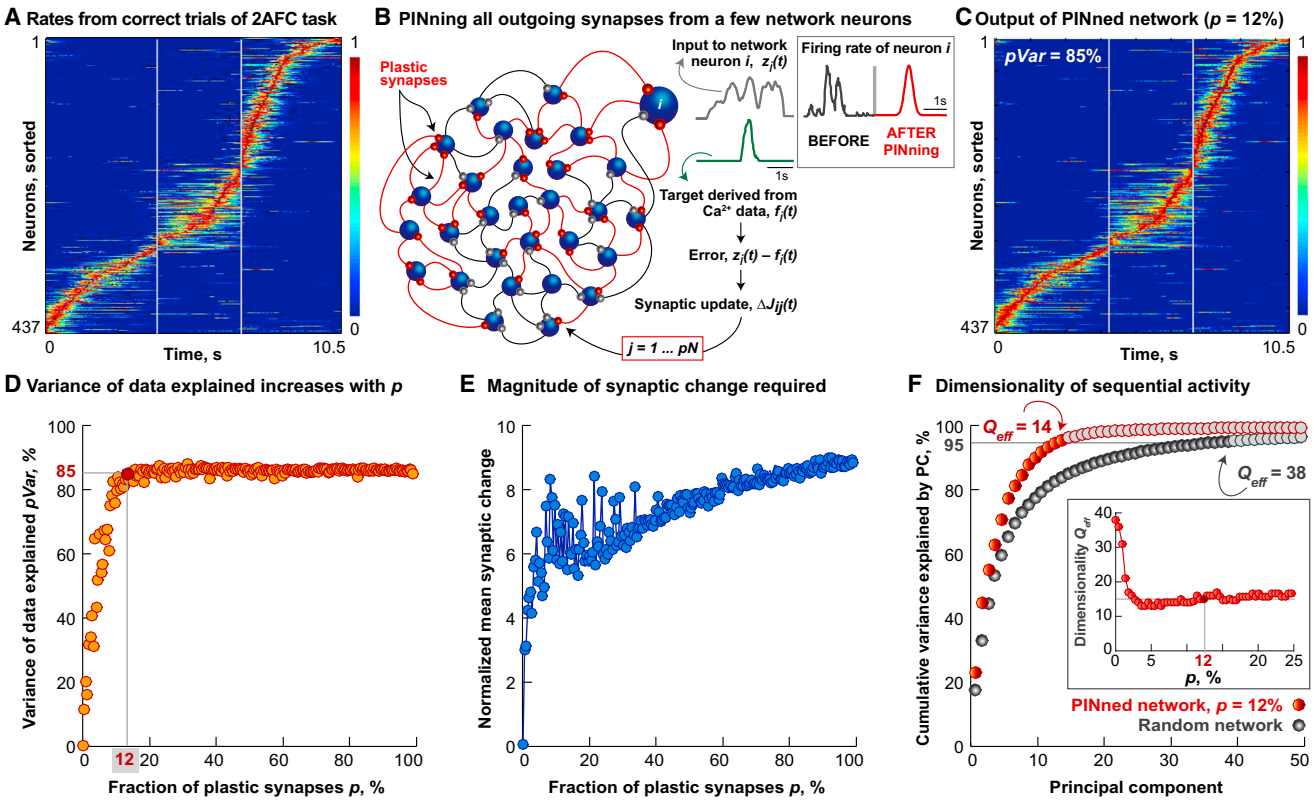


Figure 2. Partial In-Network Training Matches PPC-like Sequences

(A) This is identical to Figure 1C.

(B) Schematic of the activity-based modification scheme we call Partial In-Network Training or PINning is shown here. Only the synaptic weights carrying inputs from a small and randomly selected subset, controlled by the fraction p of the $N = 437$ rate neurons (blue), were modified (plastic synapses, depicted in orange) at every time step by an amount proportional to the difference between the input to the respective neuron at that time step ($z_i(t)$, plotted in gray) and a target waveform ($f_i(t)$, plotted in green), the presynaptic firing rate r_j, and the inverse cross-correlation matrix of the firing rates of PINned neurons (denoted by the matrix $\boldsymbol{P}$; see Experimental Procedures). Here the target functions for PINning were extracted from the rates shown in (A).

(C) Normalized activity from the network with $p = 12\%$ plastic synapses (red circle in D). This activity was temporally constrained, had a relatively small amount of extra-sequential activity (*bVar* = 40% for C and A), and showed a good match with data (*pVar* = 85%).

(D) Effect of increasing the PINning fraction, p, in a network producing a single PPC-like sequence as in (A), is shown here. *pVar* increased from 0 for random networks with no plasticity (i.e., $p = 0$) to *pVar* = 50% for $p = 8\%$ (not shown) and asymptoted at *pVar* = 85% as $p \geq 12\%$ (highlighted by red circle, outputs in C).

(E) The total magnitude of synaptic change to the connectivity as a function of p is shown here for the matrix, $J_{PINned,\ 12\%}$. The normalized mean synaptic change grew from a factor of ~7 for sparsely PINned networks ($p = 12\%$) to ~9 for fully PINned networks ($p = 100\%$), producing the PPC-like sequence. This means that individual synapses changed more in small-p networks, but the total change across the synaptic matrix was smaller.

(F) Dimensionality of the sequence is computed by plotting the cumulative variance explained by the different principal components (PCs) of the 437-neuron PINned network generating the PPC-like sequence (orange circles) with $p = 12\%$ and *pVar* = 85%, relative to a random network ($p = 0$, gray circles). When $p = 12\%$, $Q_{eff} = 14$, and smaller than the 38 for the random untrained network. Inset shows Q_{eff} (depicted in red circles) of the manifold of the overall network activity as p increased.

We first applied PINning to a network with templates obtained from a single PPC-like sequence (Figures 2A–2C), using a range of values for the fraction of plastic synapses, p. Modification of only a small percentage ($p = 12\%$) of the connections in a disordered network was sufficient for its sequential outputs to become temporally constrained (*bVar* = 40%) and to match the PPC data with high fidelity (*pVar* = 85%, Figure 2C; see Experimental Procedures for cross-validation analysis). Although our networks were typically as large as the size of the experimental dataset ($N = 437$ for Figures 1 and 2 and $N = 569$ for Figure 5), our results are consistent both for larger N networks and for networks in which non-targeted neurons are included to simulate the effect of unobserved but active neurons present in the data (Figure S3). The dependence of *pVar* on p is shown in Figure 2D.

Figure 2D quantifies the amount of structure needed for generating sequences in terms of the relative fraction of synapses modified from their initially random values, but what is the overall magnitude of synaptic change required? As shown in Figure 2E, we found that, although the individual synapses changed more in sparsely PINned (small p) networks, the total amount of change across the synaptic connectivity matrix was smaller (other implications are described subsequently).

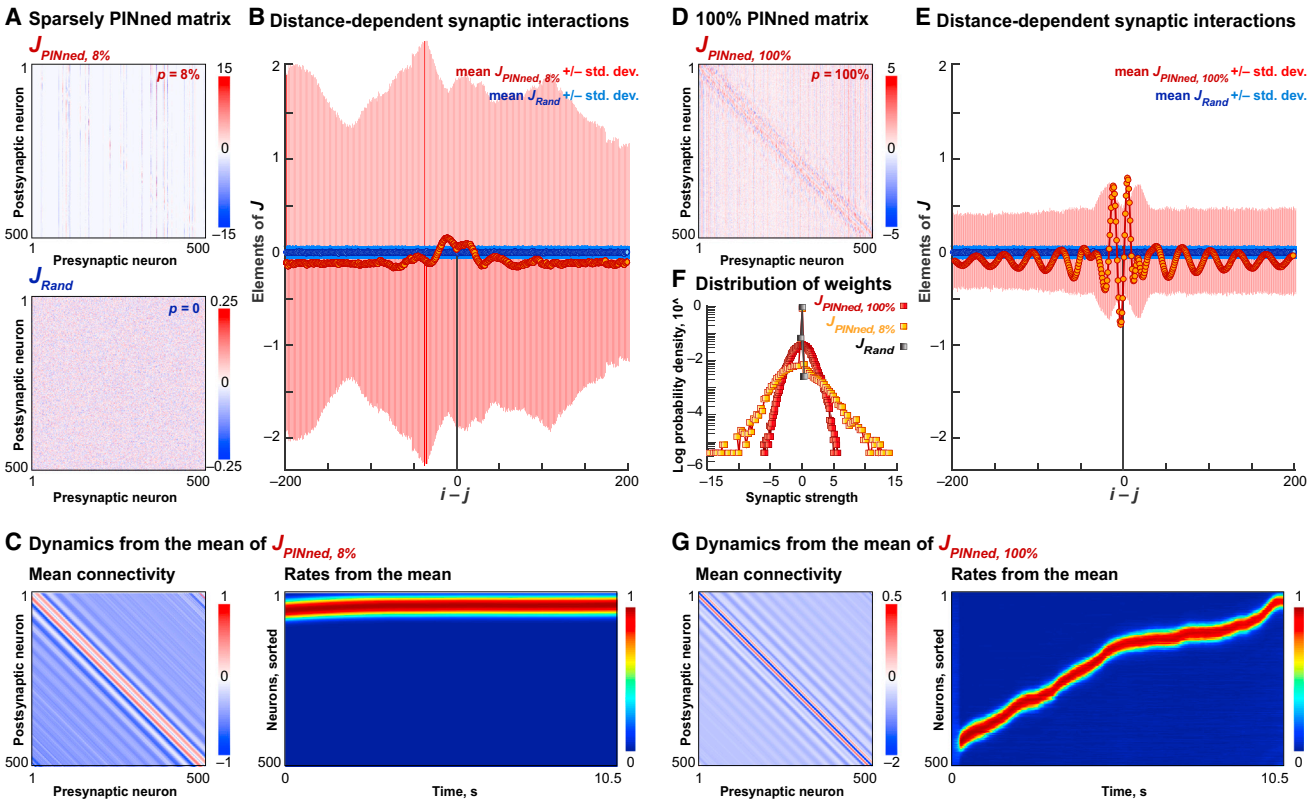


Figure 3. Properties of PINned Connectivity Matrices

(A) Synaptic connectivity matrix of a 500-neuron network with $p = 8\%$ producing an idealized sequence with $pVar = 92\%$ (highlighted by a red circle in Figure S2B), denoted by $\boldsymbol{J_{PINned,\ 8\%}}$, and that of the randomly initialized network with $p = 0$, denoted by $\boldsymbol{J_{Rand}}$, are shown here. Color bars indicate in pseudocolor the magnitudes of synaptic strengths after PINning.

(B) Influence of neurons away from the sequentially active neurons was estimated by computing the mean (circles) and the standard deviation (lines) of the elements of $\boldsymbol{J_{Rand}}$ (in blue) and $\boldsymbol{J_{PINned,\ 8\%}}$ (means in orange, standard deviations in the red lines) in successive off-diagonal stripes away from the principal diagonal. These quantities are plotted as a function of the inter-neuronal distance, $i - j$. In units of $i - j$, 0 corresponds to the principal diagonal or self-interactions, and the positive and the negative terms are the successive interaction magnitudes a distance $i - j$ away from the primary sequential neurons.

(C) Dynamics from the band averages of $\boldsymbol{J_{PINned,\ 8\%}}$ are shown here. The left panel is a synthetic matrix generated by replacing the elements of $\boldsymbol{J_{PINned,\ 8\%}}$ by their means (orange circles in B). The normalized activity from a network with this synthetic connectivity is shown on the right. Although there was a localized bump around $i - j = 0$ and long-range inhibition, these features led to fixed-point activity (right panel).

(D) Same as (A), except the connectivity matrix from a fully PINned ($p = 100\%$) network, denoted by $\boldsymbol{J_{PINned,\ 100\%}}$, is shown.

(E) Same as (B), except comparing $\boldsymbol{J_{PINned,\ 100\%}}$ and $\boldsymbol{J_{Rand}}$. Band averages (orange circles) were bigger and more asymmetric compared to those for $\boldsymbol{J_{PINned,\ 8\%}}$. Notably, these band averages were also negative for $i - j = 0$ and in the neighborhood of 0.

(F) Logarithm of the probability density of the elements of $\boldsymbol{J_{Rand}}$ (gray squares), $\boldsymbol{J_{PINned,\ 8\%}}$ (yellow squares), and, for comparison, $\boldsymbol{J_{PINned,\ 100\%}}$ (red squares) is shown here.

(G) Same as (C), except showing the firing rates from the mean $\boldsymbol{J_{PINned,\ 100\%}}$. In contrast with $\boldsymbol{J_{PINned,\ 8\%}}$, the band averages of $\boldsymbol{J_{PINned,\ 100\%}}$ could be sufficient to evoke a Gaussian bump qualitatively similar to the moving bump produced by a ring attractor model, since its movement is driven by the asymmetry in the mean connectivity.

To uncover how the sequential dynamics are distributed across the population of active neurons in PINned networks, we used principal component analysis (PCA) (e.g., Rajan et al., 2011; Sussillo, 2014; and references therein). For an untrained random network (here with $N = 437$ and $p = 0$) operating in a spontaneously active regime (Experimental Procedures), the top 38 principal components accounted for 95% of the total variance (therefore, the effective dimensionality, $Q_{eff} = 38$; gray circles in Figure 2F). In comparison, Q_{eff} of the data in Figure 1D is 24. In the network with $p = 12\%$ with outputs matching the PPC data, the dimensionality was lower ($Q_{eff} = 14$; orange circles in Figure 2F). Q_{eff} asymptoted around 12 dimensions for higher p values (inset of Figure 2F). The circuit dynamics are therefore higher dimensional for the output of the PINned network than for a sinusoidal bump attractor, but lower than the data.

Circuit Mechanism for Sequential Activation through PINning

To develop a simplified prototype for further investigation of the mechanisms of sequential activation, we created a synthetic sequence idealized from data, as long as the duration of the PPC sequence (10.5 s, Figure 2A). We first generated a Gaussian curve, *f(t)* (green curve in inset of Figure 1F and left panel of Figure S2A) that best fit the average extracted from the t_{COM}-aligned

PPC data (R_{ave}, red curve in inset of Figure 1F and left panel of Figure S2A). The curve, *f(t)*, was translated uniformly over a 10.5-s period to derive a set of target functions (N = 500, right panel of Figure S2A, *bVar* = 100%). Increasing fractions, p, of the initially random connectivity matrix were trained by PINning with the target functions of the idealized sequence. As before, *pVar* increased with p, reached *pVar* = 92% at p = 8% (red circle in Figure S2B), asymptoting at $p \sim 10\%$. The plasticity required for producing the idealized sequence was therefore smaller than the p = 12% required for the PPC-like sequence (Figure 2D), due to the lack of the idiosyncrasies present in the data, e.g., irregularities in the temporal ordering of individual transients and background activity off the sequence.

Two features are critical for the production of sequential activity. The first is the formation of a subpopulation of active neurons (bump), maintained by excitation between co-active neurons and restricted by inhibition. The second is an asymmetry in the synaptic inputs from neurons ahead in the sequence and those behind, needed to make the bump move. We therefore looked for these two features in PINned networks by examining both the structure of their connectivity matrices and the synaptic currents into individual neurons.

In a classic moving bump/ring attractor network, the synaptic connectivity is only a function of the distance between pairs of network neurons in the sequence, $i-j$, with distance corresponding to how close neurons appear in the ordered sequence (for our analysis of the connectivity in PINned networks, we also ordered the neurons in a similar manner). Furthermore, the structure that sustains, constrains, and moves the bump is all contained in the connectivity matrix, localized in $|i-j|$, and asymmetric. We looked for similar structure in the trained networks.

We considered three synaptic matrices interconnecting a population of model neurons, N = 500, before PINning (randomly initialized matrix with p = 0, denoted by $\boldsymbol{J_{Rand}}$), after sparse PINning ($\boldsymbol{J_{PINned,\ 8\%}}$), and after full PINning ($\boldsymbol{J_{PINned,\ 100\%}}$, built as a useful comparison). To analyze how the synaptic strengths in these three matrices varied with $i-j$, we first computed the means and the standard deviations of the diagonals and the means and the standard deviations of successive off-diagonal bands moving away from the diagonals, i.e., $i-j$ = constant (Experimental Procedures). These band averages and the fluctuations were then plotted as a function of the interneuronal distance, $i-j$ (Figures 3B and 3E). Next we generated "synthetic" interaction matrices, in which all the elements along each diagonal were replaced by their band averages and fluctuations, respectively. Finally, these synthetic matrices were used in networks of rate-based neurons and driven by the same inputs as the original PINned networks (Experimental Procedures).

In the p = 100% model, the band averages of $\boldsymbol{J_{PINned,\ 100\%}}$ formed a localized and asymmetric profile (orange circles in Figure 3E) that is qualitatively consistent with moving bump/ring attractor dynamics (Ben-Yishai et al., 1995; Zhang, 1996). In contrast, the band averages for $\boldsymbol{J_{PINned,\ 8\%}}$ (orange circles in Figure 3B) exhibited a localized zone of excitation for small values of $i-j$ that was symmetric, a significant inhibitory self-interactive feature at $i-j=0$, and diffuse flanking inhibition for larger values of $i-j$. This is reminiscent of the features expected in the stationary bump models (Ben-Yishai et al., 1995). Furthermore, neither the band averages of $\boldsymbol{J_{PINned,\ 8\%}}$ (shown in Figure 3C) nor the fluctuations (not shown) were, by themselves, sufficient to produce moving sequences similar to the full matrix $\boldsymbol{J_{PINned,\ 8\%}}$. Instead, the outputs of the synthetic networks built from the components of the band-averaged $\boldsymbol{J_{PINned,\ 8\%}}$ were stationary bumps (right panel of Figure 3C). In this case, what causes the bump to move?

We considered the fluctuations around the band averages of the sparsely PINned connectivity matrix, $\boldsymbol{J_{PINned,\ 8\%}}$. As expected (see color bars in Figure 3A and lines in Figures 3B and 3E), the fluctuations around the band averages of $\boldsymbol{J_{PINned,\ 8\%}}$ (red lines in Figure 3B) were much larger and more structured than those of $\boldsymbol{J_{Rand}}$ (small blue lines in Figure 3B). To uncover the mechanistic role of these fluctuations, we examined the input to each neuron produced by the sum of the fluctuations of $\boldsymbol{J_{PINned,\ 8\%}}$ and the external input (Figure 4). We realigned the sums of the fluctuations and the external inputs for all the neurons in the network by their t_{COM} and then averaged over neurons (e.g., see Experimental Procedures). This yielded an aligned population average (bottom right panel in Figure 4) that clearly revealed the asymmetry responsible for the movement of the bump across the network. Therefore, in the presence of external inputs that are constantly changing in time, the mean synaptic interactions do not have to be asymmetric, as observed for $\boldsymbol{J_{PINned,\ 8\%}}$ (Figure 3B). Instead, the variations in the fluctuations of $\boldsymbol{J_{PINned,\ 8\%}}$ (i.e., after the mean has been subtracted) and the external inputs create the asymmetry that moves the bump along. It is difficult to visualize this asymmetry at an individual neuron level because of fluctuation, necessitating this type of population-level measure. While the mean synaptic interactions in sparsely PINned networks cause the formation of the localized bump of excitation, it is the non-trivial interaction of the fluctuations in these synaptic interactions with the external inputs that causes the bump to move across the network. Therefore, this is a novel circuit mechanism for non-autonomous sequence propagation in a network.

Additionally, we looked at other PINning-induced trends in the elements of $\boldsymbol{J_{PINned,\ 8\%}}$ more directly. For all the three connectivity matrices of sequential networks constructed by PINning, there was a substantial increase in the magnitudes spanned by the synaptic weights as p decreased (compare the color bars in Figures 3A and 3D). These magnitude increases were manifested in PINned networks with different p values in different ways (Figure 3F). The initial weight matrix $\boldsymbol{J_{Rand}}$ had 0 mean (by design), 0.005 variance (of order $1/N$, Experimental Procedures), 0 skewness, and 0 kurtosis. The partially structured matrix $\boldsymbol{J_{PINned,\ 8\%}}$ had a mean of –0.1, variance of 2.2, skewness at –2, and kurtosis of 30, all of which were indicative of a probability distribution that was asymmetric about 0 with heavy tails from a small number of strong weights. This corresponds to a network in which the large sequence-facilitating synaptic changes come from a small fraction of the weights, as suggested experimentally (Song et al., 2005). In $\boldsymbol{J_{PINned,\ 8\%}}$, the ratio of the size of the largest synaptic weight to the size of the typical is ~20. If we assume the typical synapse corresponded to a postsynaptic potential (PSP) of 0.05 mV, then the large synapses had a 1 mV PSP, which is within the range of experimental data (Song et al., 2005). For comparison purposes, the connectivity for a fully structured network, $\boldsymbol{J_{PINned,\ 100\%}}$, had a mean of 0, variance of 0.7, skewness of –0.02, and kurtosis of 0.2, corresponding to a

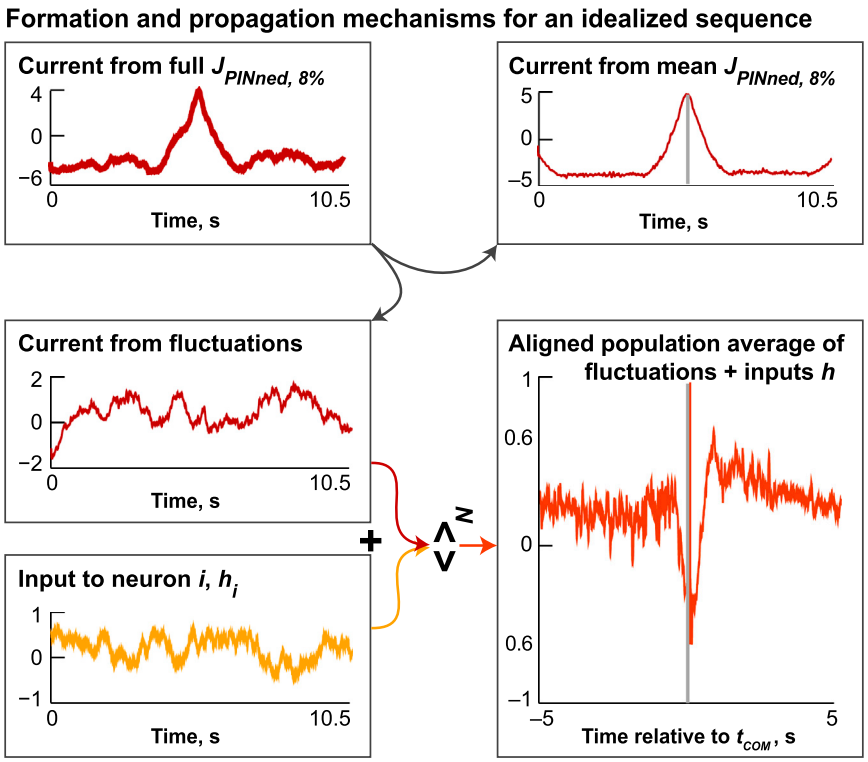


Figure 4. Mechanisms for Formation and Propagation of Sequence

The currents from the matrix $J_{PINned,\ 8\%}$ (red trace in the top left panel) and from its components are examined here. Band averages of $J_{PINned,\ 8\%}$ caused the bump to form, as shown in the plot of the current from mean $J_{PINned,\ 8\%}$ to one neuron in the network (red trace in the top panel on the right; see also Figure 3C). The cooperation of the fluctuations (whose currents are plotted in red in the middle panel on the left) with the external inputs (in yellow in the bottom left) caused the bump to move. We demonstrated this by considering the currents from the fluctuations around mean $J_{PINned,\ 8\%}$ for all the neurons in the network combined with the currents from the external inputs. The summed currents were realigned to the t_{COM} of the bump and then averaged over neurons. The resulting curve, an aligned population average of the sum of the fluctuations and external inputs to the network, is plotted in orange in the bottom panel on the right and revealed the asymmetry responsible for the movement of the bump across the network.

matrix in which the synaptic changes responsible for sequences were numerous and distributed throughout the network.

Finally, we determined that synaptic connectivity matrices obtained by PINning were fairly sensitive to small amounts of structural noise, i.e., perturbations in the matrix $J_{PINned,\ 8\%}$. However, when stochastic noise (Experimental Procedures) was used during training, slightly more robust networks were obtained (Figure S6E).

Delayed Paired Association and Working Memory Can Be Implemented through Sequences in PINned Networks

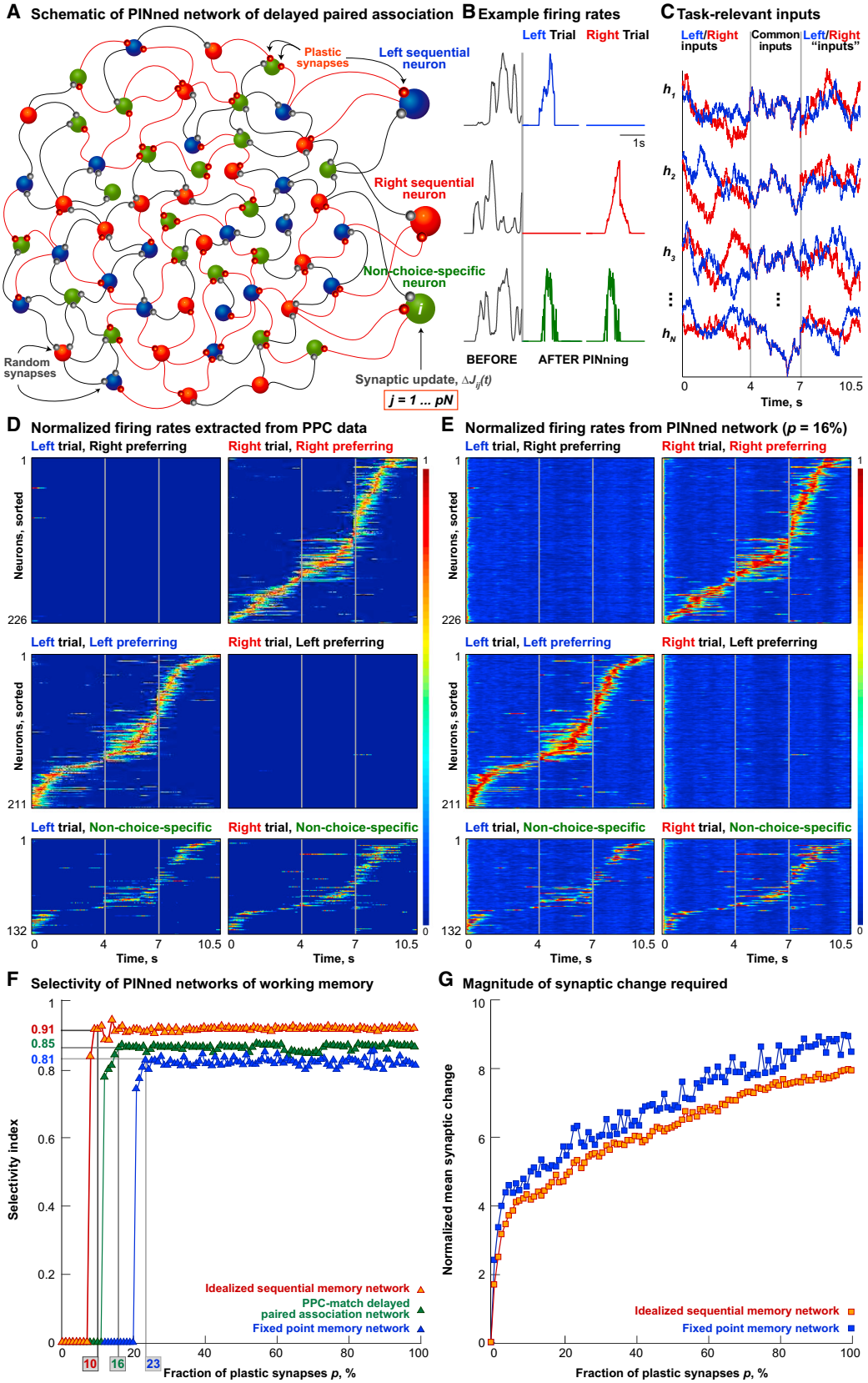
Delayed paired association (DPA) tasks, such as the 2AFC task from Harvey et al. (2012), engage working memory because the mouse must remember the identity of the cue, or cue-associated motor response, as it runs through the T-maze. Therefore, in addition to being behavioral paradigms that exhibit sequential activity (Harvey et al., 2012), DPA tasks are useful for exploring the different neural correlates of short-term memory (Gold and Shadlen, 2007; Brunton et al., 2013; Hanks et al., 2015; Amit, 1995; Amit and Brunel, 1995; Hansel and Mato, 2001; Hopfield and Tank, 1985; Shadlen and Newsome, 2001; Harvey et al., 2012). By showing that the partially structured networks we constructed by PINning could accomplish a 2AFC task, here we argued that sequences could mediate an alternative form of short-term memory.

During the first third of the 2AFC experiment (Harvey et al., 2012), the mouse received either a left or a right visual cue, and during the last third, it had two different experiences depending on whether it made a left turn or a right turn. Therefore, we modeled the cue period and the turn period of the maze by two different inputs (blue and red traces in Figure 5C). In the middle third, the delay period, when the left- or right-specific visual cues were off, the mouse ran through a section of the maze that was visually identical for both types of trials. In our simulation of this task, the inputs to individual network neurons coalesced into the same time-varying waveform during the delay periods of both the left and the right trials (purple traces in Figure 5C). The correct execution of the task therefore depended on the network generating more than one sequence to maintain the memory of the cues separately, even when the sensory inputs were identical.

A network with only p = 16% plastic synapses generated outputs that were consistent with the data (Figures 5D and 5E, *pVar* = 85%, *bVar* = 40%, compare with Figure 2c in Harvey et al., 2012; here N = 569, with 211 network neurons selected at random to activate in the left trial condition, depicted in blue in Figure 5A; 226 to fire in the right sequence, red in Figure 5A; and the remaining 132 to fire in the same order in both left and right sequences, non-choice-specific neurons, depicted in green in Figure 5A). This network retained the memory of cue identity by silencing the left-preferring neurons during the delay period of a right trial and the right-preferring neurons during a left trial, and generating sequences with the active neurons. Non-choice-specific neurons were sequentially active in the same order in both trials, like real no-preference PPC neurons observed experimentally (Figures 5D and 5E; also Figure 7b in Harvey et al., 2012).

Comparison with Fixed-Point Memory Networks

Are sequences a comparable alternative to fixed-point models for storing memories? We compared two types of sequential memory networks with a fixed-point memory network (Figure 5F). As before, different fractions of plastic synapses, controlled by p,



(legend on next page)

were embedded by PINning against different target functions (Experimental Procedures). These targets represented the values of a variable being stored and were chosen for the dynamical mechanism by which memory is implemented: idealized for the sequential memory network (orange in Figure 5F, based on Figure S2A), firing rates extracted from PPC data (Harvey et al., 2012) for the PPC-like DPA network (green in Figure 5F, outputs from the PPC-like DPA network with p = 16% plasticity are shown in Figure 5E), and constant valued targets for the fixed-point-based memory network (blue in Figure 5F).

To compare the task performance of these three types of memory networks, we computed a selectivity index (Experimental Procedures, similar to Figure 4 in Harvey et al., 2012). We found that the network exhibiting long-duration population dynamics and memory activity through idealized sequences (orange triangles in Figure 5F) had a selectivity of 0.91, when only 10% of its synapses were modified by PINning (p = 10%). In comparison, the PPC-like DPA network (Figure 5E) needed p = 16% of its synapses to be structured to match data (Harvey et al., 2012) and to achieve a selectivity of 0.85 (green triangles in Figure 5F). Both sequential memory networks performed comparably with the fixed-point network in terms of their selectivity-p relationship. The fixed-point memory network achieved an asymptotic selectivity of 0.81 for p = 23%. The magnitude of synaptic change required (Experimental Procedures) was also comparable between sequential and fixed-point memory networks, suggesting that sequences may be a viable alternative to fixed points as a mechanism for storing memories.

During experiments, the fraction of trials on which the mouse made a mistake was about 15%–20% (accuracy of the performance of mice was found to be 83% ± 9% correct, Harvey et al., 2012). We interpret errors as arising from trials in which delay period activity failed to retain the identity of the cue, leading to chance performance when the animal made a turn. Given a 50% probability of turning in the correct direction by chance, this implies that the cue identity was forgotten on 30%–40% of trials. Adding noise to our model, we could reproduce this level of delay period forgetting at a noise amplitude of $\eta_c = 0.4–0.5$ (Figure S6B). It should be noted that this value is in a region where the model shows a fairly abrupt decrease in performance.

Capacity of Sequential Memory Networks

We have discussed one specific instantiation of DPA through a sequence-based memory mechanism, the 2AFC task. We next extended the same basic PINning framework to ask whether PINned networks could accomplish memory-based tasks that require the activation of more than two non-interfering sequences. We also computed the capacity of such multi-sequential networks (denoted by N_s) as a function of the network size (N), PINning fraction (p), fraction of non-choice-specific neurons ($N_{Non\text{-}choice\text{-}specific}/N$), and temporal sparseness (fraction of neurons active at any instant, N_{Active}/N, Experimental Procedures).

We adjusted the PINning parameters for a type of memory task that is mediated by multiple non-interfering sequences (Figure 6). Different fractions of synaptic weights in an initially random network were trained by PINning to match different targets (Figure 6A), non-overlapping sets of Gaussian curves (similar to Figure S2A), each evenly spaced and collectively spanning 8 s. The width of these waveforms, N_{Active}/N, was varied as a parameter that controlled the sparseness of the sequence (Experimental Procedures). Because the turn period was omitted here for clarity, the modeled multi-sequential tasks were 8 s long. Similar to DPA (Figure 5), each neuron received a different filtered white noise input for each cue during the cue period (0–4 s); but, during the delay period (4–8 s), they

Figure 5. DPA in PINned Networks of Working Memory

(A) Schematic of a sparsely PINned network implementing DPA through a 2AFC task (Harvey et al., 2012). The targets for PINning were firing rates extracted from Ca^{2+} imaging data from left-preferring PPC cells (schematized in blue), right-preferring cells (schematized in red), and cells with no choice preference (in green, called non-choice-specific neurons). As in Figure 2B, the learning rule was applied only to p% of the synapses.

(B) Example single-neuron firing rates, normalized to the maximum, before (gray) and after PINning (blue trace for left-preferring, red trace for right-preferring, and green trace for non-choice-specific neurons) are shown here.

(C) A few example task-specific inputs (h_i for $i = 1, 2, 3, \ldots, N$) are shown here. Each neuron got a different irregular, spatially delocalized, filtered white noise input, but received the same one on every trial. Left trial inputs are in blue and the right trial ones, in red.

(D) Normalized firing rates extracted from trial-averaged Ca^{2+} imaging data collected in the PPC during a 2AFC task are shown here. Spike trains were extracted by deconvolution (see Experimental Procedures) from mean calcium fluorescence traces for the 437 choice-specific and 132 non-choice-specific, task-modulated cells (one cell per row) imaged on preferred and opposite trials (Harvey et al. 2012). These firing rates were used to extract the target functions for PINning (schematized in A). Traces were normalized to the peak of the mean firing rate of each neuron on preferred trials and sorted by the t_{COM}. Vertical gray lines indicate the time points corresponding to the different epochs of the task: the cue period ending at 4 s, the delay period ending at 7 s, and the turn period concluding at the end of the trial, at 10.5 s.

(E) The outputs of the 569-neuron network with p = 16% plastic synapses, sorted by t_{COM} and normalized by the peak of the output of each neuron, showed a match with the data (D). For this network, $pVar$ = 85%.

(F) Selectivity index, shown here, was computed as the ratio of the difference and the sum of the mean activities of preferred neurons at the 10-s time point during preferred trials and the mean activities of preferred neurons during opposite trials (see Experimental Procedures). Task performance of three different PINned networks of working memory as a function of the fraction of plastic synapses, p—an idealized sequential memory model (orange triangles, N = 500), a model that exhibits PPC-like dynamics (green triangles, N = 569, using rates from D), and a fixed-point memory network (blue triangles, N = 500)—are shown here. The network that exhibited long-duration population dynamics and memory activity through idealized sequences (orange triangles) had selectivity = 0.91, when p = 10% of its synapses were plastic; the PPC-like network needed p = 16% plastic synapses for selectivity = 0.85, and the fixed-point network needed p = 23% of its synapses to be plastic for selectivity = 0.81.

(G) Magnitude of synaptic change (computed as for Figure 2E) for the three networks shown in (F) is plotted here. The idealized sequential (orange squares) and the fixed-point memory network (blue squares) required comparable amounts of mean synaptic change to execute the DPA task, growing from ~3 to ~9 for p = 10%–100%. Across the matrix, the total amount of synaptic change required was smaller, even though individual synapses changed more in the sparsely PINned case. The PPC-like memory network is omitted here for clarity.

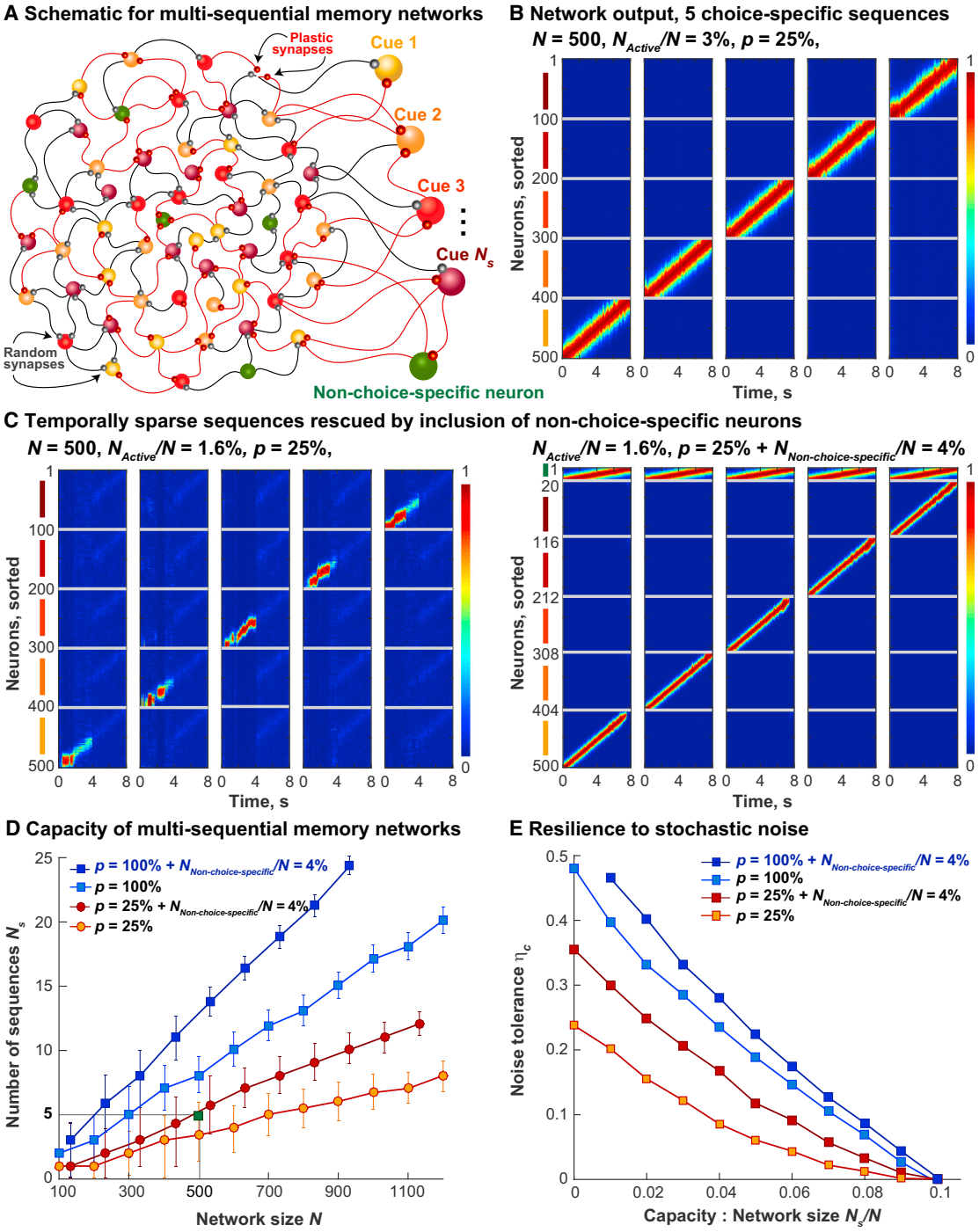


Figure 6. Capacity of Multi-sequential Memory Networks

(A) Schematic of a multi-sequential memory network (N = 500) is shown here. N/N_s neurons were assigned to each sequence that we wanted the network to simultaneously produce. Here N_s is the capacity. Only p% of the synapses were plastic. The target functions were identical to Figure S2A, however, their widths, denoted by N_{Active}/N, were varied as a task parameter controlling how many network neurons were active at any instant.

(B) The normalized firing rates of the 500-neuron network with p = 25% and N_{Active}/N = 3% are shown here. The memory task (the cue periods and the delay periods only) was correctly executed through five choice-specific sequences; during the delay period (4–8 s), neurons fired in a sequence only on trials of the same type as their cue preference.

(C) The left panel shows the same network as (B) failing to perform the task correctly when the widths of the targets were halved (N_{Active}/N = 1.6%), sparsifying the sequences in time. Memory of the cue identity was not maintained during the delay period. The right panel shows the result of including a small number of

(legend continued on next page)

coalesced to a common cue-invariant waveform, different for each neuron. In the most general case, we assigned N/N_s neurons to each sequence. Here N_s was the number of memories (capacity), which was also the number of trial types. Once again, only p% of the synapses in the network were plastic.

A correctly executed multi-sequential task is one in which, during the delay period, network neurons fire in a sequence only on trials of the same type as their cue preference and are silent during other trials. A network of 500 neurons with p = 25% plastic synapses performed such a task easily, generating N_s = 5 sequences with delay period memory of the appropriate cue identity with temporal sparseness of N_{Active}/N = 3% (Figure 6B).

The temporal sparseness was found to be a crucial factor that determined how well a network performed a multi-sequential memory task (Experimental Procedures). When sequences were forced to be sparser than a minimum (e.g., by using narrower target waveforms, for this 8 s-long task, when N_{Active}/N < 1.6%), the network failed. Although there was sequential activation, the memory of the five separately memorable cues was not maintained across the delay period. Surprisingly, this failure could be rescued by adding a small number of non-choice-specific neurons ($N_{Non\text{-}choice\text{-}specific}/N$ = 4%) that fired in the same order in all five trial types (Figure 6C). In other words, the $N - N_{Non\text{-}choice\text{-}specific}$ network neurons (<N) successfully executed the task using five sparse choice-specific sequences and one non-choice-specific sequence. The capacity of the network for producing multiple sparse sequences increased when non-choice-specific neurons were added to stabilize them, without needing a concomitant increase in the synaptic modification. Therefore, we predict that non-choice-specific neurons, also seen in the PPC (Harvey et al., 2012), may function as a kind of ''conveyor belt'' of working memory, providing recurrent synaptic current to sequences that may be too sparse to sustain themselves otherwise in a non-overlapping manner.

Finally, we focused on the memory capacity and noise tolerance properties. The presence of non-choice-specific neurons increased the capacity of networks to store memories if implemented through sequences (Figure 6C). Up to a constant factor, the capacity, the maximum of the fraction N_s/N, scaled in proportion to the fraction of plastic synapses, p, and the network size, N, and inversely with sparseness, N_{Active}/N (Figure 6D). The slope of this capacity-to-network size relationship increased when non-choice-specific neurons were included, because they enabled the network to carry sparser sequences. We also found that networks with a bigger fraction of plastic synaptic connections and networks containing non-choice-specific neurons were more stable against stochastic perturbations (Figure 6E). Once the amplitude of added stochastic noise exceeded the maximum tolerance (denoted by η_c, Experimental Procedures), however, non-choice-specific neurons were not effective at repairing the memory capacity (not shown). Non-choice-specific neurons could not rescue inadequately PINned schemes either (small p, not shown).

DISCUSSION

In this paper, we used and extended earlier work based on liquid-state (Maass et al., 2002) and echo-state machines (Jaeger, 2003), which has shown that a basic balanced-state network with random recurrent connections can act as a general purpose dynamical reservoir, the modes of which can be harnessed for performing different tasks by means of feedback loops. These models typically compute the output of a network through weighted readout vectors (Buonomano and Merzenich, 1995; Maass et al., 2002; Jaeger, 2003; Jaeger and Haas, 2004) and feed the output back into the network as an additional current, leaving the synaptic weights within the dynamics-producing network unchanged in their randomly initialized configuration. The result was that networks generating chaotic spontaneous activity prior to learning (Sompolinsky et al., 1988) produced a variety of regular non-chaotic outputs that matched the imposed target functions after learning (Sussillo and Abbott, 2009). The key to making these ideas work is that the feedback carrying the error during the learning process forces the network into a state in which it produces less variable responses (Molgedey et al., 1992; Bertschinger and Natschläger, 2004; Rajan et al., 2010, 2011). A compelling example, FORCE learning (Sussillo and Abbott, 2009) has been implemented as part of several modeling studies (Laje and Buonomano, 2013; Sussillo, 2014 and references therein), but it had a few limitations. Specifically, the algorithm in Sussillo and Abbott (2009) included a feedback network separate from the one generating the trajectories, with aplastic connections, and learning was restricted to the readout weights. The approach we developed here, called PINning, avoids these issues while producing sequences resembling experimental data (Results). During PINning, only a fraction of the recurrent synapses in an initially random network are modified in a task-specific manner and by a biologically reasonable amount, leaving the majority of the network heterogeneously wired. Furthermore, the fraction of plastic connections is a tunable parameter that controls the contribution of the structured portion of the network relative to the random part, allowing these models to interpolate between highly structured and completely

non-choice-specific neurons ($N_{Non\text{-}choice\text{-}specific}/N$ = 4%) that fired in the same order in trials of all five types. Including non-choice-specific neurons restored the memory of cue identity during the delay period, without increasing p.

(D) Capacity of multi-sequential memory networks, N_s, as a function of network size, N, for different values of p is shown here. Mean values are indicated by orange circles for p = 25% PINned networks, red circles for p = 25% + $N_{Non\text{-}choice\text{-}specific}/N$ = 4%, light blue squares for p = 100%, and dark blue squares for p = 100% + $N_{Non\text{-}choice\text{-}specific}/N$ = 4% neurons. PINned networks containing additional non-choice-specific neurons had temporally sparse target functions with N_{Active}/N = 1.6%, the rest had N_{Active}/N = 3%. Error bars were calculated over five different random instantiations each and decreased under the following conditions: as network size was increased, as fraction of plastic synapses p was increased, and moderately, with the inclusion of non-choice-specific neurons. The green square highlights the 500-neuron network whose normalized outputs are plotted in (B).

(E) Resilience of different multi-sequential networks, computed as the critical amount of stochastic noise (denoted here by η_c) tolerated before the memory task fails, is shown here. Tolerance, plotted here as a function of the ratio of multi-sequential capacity and network size, N_s/N, decreased as p was lowered and as capacity increased, although it fell slower with the inclusion of non-choice-specific neurons. Only mean values are shown here for clarity.

random architectures until we find the point that best matches the relevant experimental data (a similar point is made in Barak et al., 2013). PINning is not the most general way to restrict learning to a limited number of synapses, but it allows us to do so without losing the efficiency of the learning algorithm.

To illustrate the applicability of such partially structured networks, we used data recorded from the PPC of mice performing a working memory-based decision-making task in a virtual reality environment (Harvey et al., 2012). It might be unlikely that an association cortex such as the PPC evolved specialized circuitry solely for sequence generation, because the PPC mediates other complex temporal tasks, e.g., from working memory and decision making to evidence accumulation and navigation (Shadlen and Newsome, 2001; Gold and Shadlen, 2007; Freedman and Assad, 2011; Snyder et al., 1997; Andersen and Cui, 2009; Bisley and Goldberg, 2003; McNaughton et al., 1994; Nitz, 2006; Whitlock et al., 2008; Calton and Taube, 2009; Hanks et al., 2015). We also computed a measure of stereotypy from PPC data (*bVar*, Figures 1C–1G) and found it to be much lower than if the PPC were to generate sequences based on highly specialized chain-like or ring-like connections or if it read out sequential activity from a structured upstream region. We therefore started with random recurrent networks, imposed a small amount of structure in their connectivity through PINning, and duplicated many features of the experimental data (Harvey et al., 2012), e.g., choice-specific neural sequences and delay period memory of cue identity (Results).

We analyzed the structural features in the synaptic connectivity matrices of sparsely PINned networks, concluding that the probability distribution of the synaptic strengths is heavy tailed due to the presence of a small percentage of strong interaction terms (Figure 3F). There is experimental evidence that there might be a small fraction of very strong synapses embedded within a milieu of relatively weak synapses in the cortex (Song et al., 2005), most recently from the primary visual cortex (Cossell et al., 2015). Synaptic distributions measured in slices have been shown to have long tails (Song et al., 2005), and the experimental result in Cossell et al. (2015) has demonstrated that rather than occurring at random, these strong synapses significantly contribute to network tuning by preferentially interconnecting neurons with similar orientation preferences. These strong synapses may be the plastic synapses that are induced by PINning in our scheme. The model also exhibited some structural noise sensitivity (Figure S6E), and this was not completely removed by training in the presence of noise. It is possible that dynamic mechanisms of ongoing plasticity in neural circuits could enhance stability to structural fluctuations.

In this paper, the circuit mechanisms underlying both the *formation* of a localized bump of excitation in the connectivity and the manner in which the bump of excitation *propagates* across the network were elucidated separately (Figures 3 and 4). The circuit mechanism for the propagation of the sequence was found to be non-autonomous, relying on a complex interplay between the recurrent connections and the external inputs. The mechanism for propagation was therefore distinct from the standard moving bump/ring attractor model (Ben-Yishai et al., 1995; Zhang, 1996), but it had similarities to the models developed in Fiete et al. (2010) and Hopfield (2015) (explored further in Figure S5). In particular, the model in Fiete et al. (2010) generates highly stereotyped sequences similar to area HVC by initializing as a recurrently connected network and subsequently using spike timing-dependent plasticity, heterosynaptic competition, and correlated external inputs to learn sequential activity. A critical difference between the model in Fiete et al. (2010) and our approach lies in the circuitry underlying sequences: their learning rule results in synaptic chain-like connectivity.

In this paper, we suggest an alternative hypothesis that sequences might be a more general and effective dynamical form of working memory (Results), making the prediction that sequences may be increasingly observed in many working memory-based tasks (also suggested in Harvey et al., 2012). This contrasts with previous models that relied on fixed-point attractors to retain information and exhibited sustained activity (Amit, 1992, 1995; Amit and Brunel, 1995; Hansel and Mato, 2001; Hopfield and Tank, 1985). We computed the capacity of sequential memory networks for storing more than two memories (Results) by extending the sparse PINning approach developed above, and we interpreted this capacity as the "computational bandwidth" of a general purpose circuit to perform different timing-based computations.

The term pre-wired is used in this paper to mean a scheme in which a principled mechanism for executing a certain task is first assumed and then incorporated into the network circuitry, e.g., a moving bump architecture (Ben-Yishai et al., 1995; Zhang, 1996) or a synfire chain (Hertz and Prügel-Bennett, 1996; Levy et al., 2001; Hermann et al., 1995; Fiete et al., 2010). In contrast, the models built and described in this paper are constructed without bias or assumptions. If a moving bump architecture (Ben-Yishai et al., 1995; Zhang, 1996) had been assumed at the beginning and the network pre-wired accordingly, we would of course have uncovered it through the analysis of the synaptic connectivity matrix (similar to Results; Figures 3, 4, and S5; see also Experimental Procedures). However, by starting with an initially random configuration and learning a small amount of structure, we found an alternative mechanism for input-dependent sequence propagation (Figure 4). We would not have encountered this mechanism for the non-autonomous movement of the bump by pre-wiring a different mechanism into the connectivity of the model network. While the models constructed here are indeed trained to perform the task, the fact that they are unbiased means that the opportunity was present to uncover mechanisms that were not thought of a priori.

EXPERIMENTAL PROCEDURES

Network Elements

We considered a network of N fully interconnected neurons described by a standard firing-rate model, where $N = 437$ for the PPC-like sequence in Figures 1D and 2A, $N = 500$ for the single idealized sequence in Figure S2 and the multi-sequential memory task in Figure 6, and $N = 569$ for the 2AFC task in Figure 5 (we generally built networks of the same size as the experimental dataset we were trying to model, however, the results obtained remain applicable to larger networks, Figure S3). For mathematical details of the network, see the Supplemental Experimental Procedures.

Design of External Inputs

To represent sensory (visual and proprioceptive) stimuli innervating the PPC, the external inputs to the neurons in the network were made from

filtered and spatially delocalized white noise that was frozen (repeated from trial to trial, Figure 1A). For motivation of this choice and for mathematical details of the frozen noise input, see the Supplemental Experimental Procedures. In addition, we also tested the resilience of the memory networks we built (Figure 6E) to injected stochastic noise (see the Supplemental Experimental Procedures). We defined resilience or noise tolerance as the critical amplitude of this stochastic noise, denoted by η_c, at which the delay period memory failed and the selectivity dropped to 0 (Figure 6E; see also Figure S6).

Extracting Target Functions from Calcium Imaging Data

To derive the target functions for our activity-dependent synaptic modification scheme termed PINning, we converted the calcium fluorescence traces from the PPC data (Harvey et al., 2012) into firing rates using two complementary methods. See the Supplemental Experimental Procedures for details of both deconvolution methods (see also Figure S1).

Synaptic Modification Rule for PINning

During PINning, the inputs of individual network neurons were compared directly with the target functions extracted from data to compute a set of error functions. During learning, the subset of plastic internal weights in the connectivity matrix of the random network, denoted by p, underwent modification at a rate proportional to the error, the presynaptic firing rate of each neuron, and a $pN \times pN$ matrix, $\boldsymbol{P}$, that tracked the rate fluctuations across the network. For further details of the learning rule, see the Supplemental Experimental Procedures.

Dimensionality of Network Activity, Q_{eff}

We defined the effective dimensionality of the activity, Q_{eff}, as the number of principal components that captured 95% of the variance in the activity (Figure 2F), computed as in the Supplemental Experimental Procedures.

Stereotypy of Sequence, *bVar*

bVar quantified the variance of the data or the network output explained by the translation of an activity profile with an invariant shape. See the Supplemental Experimental Procedures for details.

Percentage Variance of Data Explained by Model, *pVar*

We quantified the match between the experimental data or the set of target functions and the outputs of the model by *pVar*, the amount of variance in the data captured by the model. Details of how this was computed are in the Supplemental Experimental Procedures.

Magnitude of Synaptic Change

In Figures 2E and 5G, we computed the magnitude of the synaptic change required to implement a single PPC-like sequence, an idealized sequence, and three types of memory tasks. In combination with the fraction of plastic synapses in the PINned network, p, this metric characterizes the amount of structure imposed in an initially random network to produce the desired temporally structured dynamics. See also the Supplemental Experimental Procedures.

Selectivity Index for Memory Task

In Figure 5F, a selectivity index was computed (similar to Figure 4 in Harvey et al., 2012) to assess the performance of different PINned networks at maintaining memories during the delay period of DPA tasks. This metric is based on the ratio of the difference and the sum of the mean activities of preferred neurons at the end of the delay period ($\sim$10 s for Figure 5, after Harvey et al., 2012 and $\sim$7 s for Figure 6) during preferred trials and the mean activities of preferred neurons during opposite trials. This was computed as described in the Supplemental Experimental Procedures.

Temporal Sparseness of Sequences, N_{Active}/N

The temporal sparseness of a sequence is defined as the fraction of neurons active at any instant during the sequence, denoted here by N_{Active}/N. This was computed as in the Supplemental Experimental Procedures.

Analyzing the Structure of PINned Synaptic Connectivity Matrices

This is pertinent to the Results (Figures 3 and 4), in which we quantified how the synaptic strength varies with distance between pairs of network neurons in connectivity space, $i - j$. These were computed as described in the Supplemental Experimental Procedures. See Figure 3B for the sparsely PINned matrix, $\boldsymbol{J}_{PINned,\ 8\%}$, relative to the randomly initialized matrix, $\boldsymbol{J}_{Rand}$, and Figure 3E for the fully PINned matrix, $\boldsymbol{J}_{PINned,\ 100\%}$. The same analysis also was used for Figure S5.

Cross-Validation Analysis

This pertains to quantifying how well PINning-based models did and is described in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.02.009.

AUTHOR CONTRIBUTIONS

Conceptualization, K.R., C.D.H., and D.W.T.; Methodology, K.R.; Formal Analysis, K.R.; Investigation, K.R.; Data Curation, C.D.H.; Writing – Original Draft, K.R.; Writing – Review and Editing, K.R., C.D.H., and D.W.T.; Resources, D.W.T.; Supervision, D.W.T.

ACKNOWLEDGMENTS

The authors thank Larry Abbott for providing guidance and critiques throughout this project; Eftychios Pnevmatikakis and Liam Paninski for their deconvolution algorithm (Pnevmatikakis et al., 2016; Vogelstein et al., 2010); and Dmitriy Aronov, Bill Bialek, Selmaan Chettih, Cristina Domnisoru, Tim Hanks, and Matthias Minderer for comments. This work was supported by the NIH (D.W.T., R01-MH083686, RC1-NS068148, and 1U01NS090541-01; C.D.H., R01-MH107620 and R01-NS089521), a grant from the Simons Collaboration on the Global Brain (D.W.T.), a National Alliance for Research on Schizophrenia and Depression (NARSAD) Young Investigator Award from the Brain & Behavior Research Foundation (K.R.), a fellowship from the Helen Hay Whitney Foundation (C.D.H.), the New York Stem Cell Foundation (C.D.H.), and a Burroughs Wellcome Fund Career Award at the Scientific Interface (C.D.H.). C.D.H. is a New York Stem Cell Foundation-Robertson Investigator and a Searle Scholar.

Received: August 22, 2015
Revised: December 3, 2015
Accepted: February 2, 2016
Published: March 10, 2016

REFERENCES

Amit, D.J. (1992). Modeling Brain Function: The World of Attractor Neural Networks (Cambridge University Press).

Amit, D.J. (1995). The Hebbian paradigm reintegrated: local reverberations as internal representations. Behav. Brain Sci. *18*, 617.

Amit, D.J., and Brunel, N. (1995). Learning internal representations in an attractor neural network with analogue neurons. Network *6*, 359–388.

Andersen, R.A., and Cui, H. (2009). Intention, action planning, and decision making in parietal-frontal circuits. Neuron *63*, 568–583.

Andersen, R.A., Burdick, J.W., Musallam, S., Pesaran, B., and Cham, J.G. (2004). Cognitive neural prosthetics. Trends Cogn. Sci. *8*, 486–493.

Barak, O., Sussillo, D., Romo, R., Tsodyks, M., and Abbott, L.F. (2013). From fixed points to chaos: three models of delayed discrimination. Prog. Neurobiol. *103*, 214–222.

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Ben-Yishai, R., Bar-Or, R.L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. Proc. Natl. Acad. Sci. USA *92*, 3844–3848.

Bertschinger, N., and Natschläger, T. (2004). Real-time computation at the edge of chaos in recurrent neural networks. Neural Comput. *16*, 1413–1436.

Bisley, J.W., and Goldberg, M.E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. Science *299*, 81–86.

Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and humans can optimally accumulate evidence for decision-making. Science *340*, 95–98.

Buonomano, D.V. (2003). Timing of neural responses in cortical organotypic slices. Proc. Natl. Acad. Sci. USA *100*, 4897–4902.

Buonomano, D.V. (2005). A learning rule for the emergence of stable dynamics and timing in recurrent networks. J. Neurophysiol. *94*, 2275–2283.

Buonomano, D.V., and Merzenich, M.M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. Science *267*, 1028–1030.

Calton, J.L., and Taube, J.S. (2009). Where am I and how will I get there from here? A role for posterior parietal cortex in the integration of spatial information and route planning. Neurobiol. Learn. Mem. *91*, 186–196.

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. *13*, 369–378.

Cossell, L., Iacaruso, M.F., Muir, D.R., Houlton, R., Sader, E.N., Ko, H., Hofer, S.B., and Mrsic-Flogel, T.D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex. Nature *518*, 399–403.

Crowe, D.A., Averbeck, B.B., and Chafee, M.V. (2010). Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. J. Neurosci. *30*, 11640–11653.

Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Hippocampal replay of extended experience. Neuron *63*, 497–507.

Fiete, I.R., Senn, W., Wang, C.Z., and Hahnloser, R.H.R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. Neuron *65*, 563–576.

Fisher, D., Olasagasti, I., Tank, D.W., Aksay, E.R., and Goldman, M.S. (2013). A modeling framework for deriving the structural and functional architecture of a short-term memory microcircuit. Neuron *79*, 987–1000.

Freedman, D.J., and Assad, J.A. (2011). A proposed common neural mechanism for categorization and perceptual decisions. Nat. Neurosci. *14*, 143–146.

Fujisawa, S., Amarasingham, A., Harrison, M.T., and Buzsáki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. Nat. Neurosci. *11*, 823–833.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Goldman, M.S. (2009). Memory without feedback in a neural network. Neuron *61*, 621–634.

Hahnloser, R.H.R., Kozhevnikov, A.A., and Fee, M.S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature *419*, 65–70.

Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., and Brody, C.D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature *520*, 220–223.

Hansel, D., and Mato, G. (2001). Existence and stability of persistent states in large neuronal networks. Phys. Rev. Lett. *86*, 4175–4178.

Harvey, C.D., Coen, P., and Tank, D.W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature *484*, 62–68.

Haykin, S. (2002). Adaptive Filter Theory (Prentice Hall).

Hermann, M., Hertz, J.A., and Prügel-Bennett, A. (1995). Analysis of synfire chains. Network *6*, 403–414.

Hertz, J., and Prügel-Bennett, A. (1996). Learning short synfire chains by self-organization. Network *7*, 357–363.

Hopfield, J.J. (2015). Understanding emergent dynamics: Using a collective activity coordinate of a neural network to recognize time-varying patterns. Neural Comput. *27*, 2011–2038.

Hopfield, J.J., and Tank, D.W. (1985). "Neural" computation of decisions in optimization problems. Biol. Cybern. *52*, 141–152.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., and Yuste, R. (2004). Synfire chains and cortical songs: temporal modules of cortical activity. Science *304*, 559–564.

Jaeger, H. (2003). Adaptive nonlinear system identification with echo state networks. In Advances in Neural Information Processing Systems 15 (NIPS 2002), S. Becker, S. Thrun, and K. Obermayer, eds. (MIT Press), pp. 593–600.

Jaeger, H., and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science *304*, 78–80.

Jin, D.Z., Fujii, N., and Graybiel, A.M. (2009). Neural representation of time in cortico-basal ganglia circuits. Proc. Natl. Acad. Sci. USA *106*, 19156–19161.

Kleinfeld, D., and Sompolinsky, H. (1989). Associative network models for central pattern generators. In Methods in Neuronal Modeling, C. Koch and I. Segev, eds. (MIT Press), pp. 195–246.

Kozhevnikov, A.A., and Fee, M.S. (2007). Singing-related activity of identified HVC neurons in the zebra finch. J. Neurophysiol. *97*, 4271–4283.

Laje, R., and Buonomano, D.V. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat. Neurosci. *16*, 925–933.

Levy, N., Horn, D., Meilijson, I., and Ruppin, E. (2001). Distributed synchrony in a cell assembly of spiking neurons. Neural Netw. *14*, 815–824.

Louie, K., and Wilson, M.A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. Neuron *29*, 145–156.

Luczak, A., Barthó, P., Marguet, S.L., Buzsáki, G., and Harris, K.D. (2007). Sequential structure of neocortical spontaneous activity in vivo. Proc. Natl. Acad. Sci. USA *104*, 347–352.

Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Comput. *14*, 2531–2560.

Maass, W., Joshi, P., and Sontag, E.D. (2007). Computational aspects of feedback in neural circuits. PLoS Comput. Biol. *3*, e165.

Mauk, M.D., and Buonomano, D.V. (2004). The neural basis of temporal processing. Annu. Rev. Neurosci. *27*, 307–340.

McNaughton, B.L., Mizumori, S.J., Barnes, C.A., Leonard, B.J., Marquis, M., and Green, E.J. (1994). Cortical representation of motion during unrestrained spatial navigation in the rat. Cereb. Cortex *4*, 27–39.

Molgedey, L., Schuchhardt, J., and Schuster, H.G. (1992). Suppressing chaos in neural networks by noise. Phys. Rev. Lett. *69*, 3717–3719.

Nádasdy, Z., Hirase, H., Czurkó, A., Csicsvari, J., and Buzsáki, G. (1999). Replay and time compression of recurring spike sequences in the hippocampus. J. Neurosci. *19*, 9497–9507.

Nitz, D.A. (2006). Tracking route progression in the posterior parietal cortex. Neuron *49*, 747–756.

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. Science *321*, 1322–1327.

Pearlmutter, B.A. (1989). Learning state space trajectories in recurrent neural networks. Neural Comput. *1*, 263–269.

Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., et al. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. Neuron *89*, 285–299.

Pulvermüller, F., and Shtyrov, Y. (2009). Spatiotemporal signatures of large-scale synfire chains for speech processing as revealed by MEG. Cereb. Cortex *19*, 79–88.

Rajan, K., Abbott, L.F., and Sompolinsky, H. (2010). Stimulus-dependent suppression of chaos in recurrent neural networks. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *82*, 011903.

Rajan, K., Abbott, L.F., and Sompolinsky, H. (2011). Inferring stimulus selectivity from the spatial structure of neural network dynamics. In Advances in Neural Information Processing Systems 23 (NIPS 2010), J.D. Lafferty, C.K.I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, eds. (Neural Information Processing Systems Foundation).

Schwartz, A.B., and Moran, D.W. (1999). Motor cortical activity during drawing movements: population representation during lemniscate tracing. J. Neurophysiol. *82*, 2705–2718.

Seidemann, E., Meilijson, I., Abeles, M., Bergman, H., and Vaadia, E. (1996). Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. J. Neurosci. *16*, 752–768.

Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. *86*, 1916–1936.

Snyder, L.H., Batista, A.P., and Andersen, R.A. (1997). Coding of intention in the posterior parietal cortex. Nature *386*, 167–170.

Sompolinsky, H., Crisanti, A., and Sommers, H.J. (1988). Chaos in random neural networks. Phys. Rev. Lett. *61*, 259–262.

Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. *3*, e68.

Sussillo, D. (2014). Neural circuits as computational dynamical systems. Curr. Opin. Neurobiol. *25*, 156–163.

Sussillo, D., and Abbott, L.F. (2009). Generating coherent patterns of activity from chaotic neural networks. Neuron *63*, 544–557.

Tang, A., Jackson, D., Hobbs, J., Chen, W., Smith, J.L., Patel, H., Prieto, A., Petrusca, D., Grivich, M.I., Sher, A., et al. (2008). A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. J. Neurosci. *28*, 505–518.

Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., and Paninski, L. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. J. Neurophysiol. *104*, 3691–3704.

White, B., Abbott, L.F., and Fiser, J. (2012). Suppression of cortical neural variability is stimulus- and state-dependent. J. Neurophysiol. *108*, 2383–2392.

Whitlock, J.R., Sutherland, R.J., Witter, M.P., Moser, M.B., and Moser, E.I. (2008). Navigating from hippocampus to parietal cortex. Proc. Natl. Acad. Sci. USA *105*, 14755–14762.

Williams, R.J., and Zipser, D. (1989). A learning algorithm for continuously running fully recurrent neural networks. Neural Comput. *1*, 270–280.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. J. Neurosci. *16*, 2112–2126.

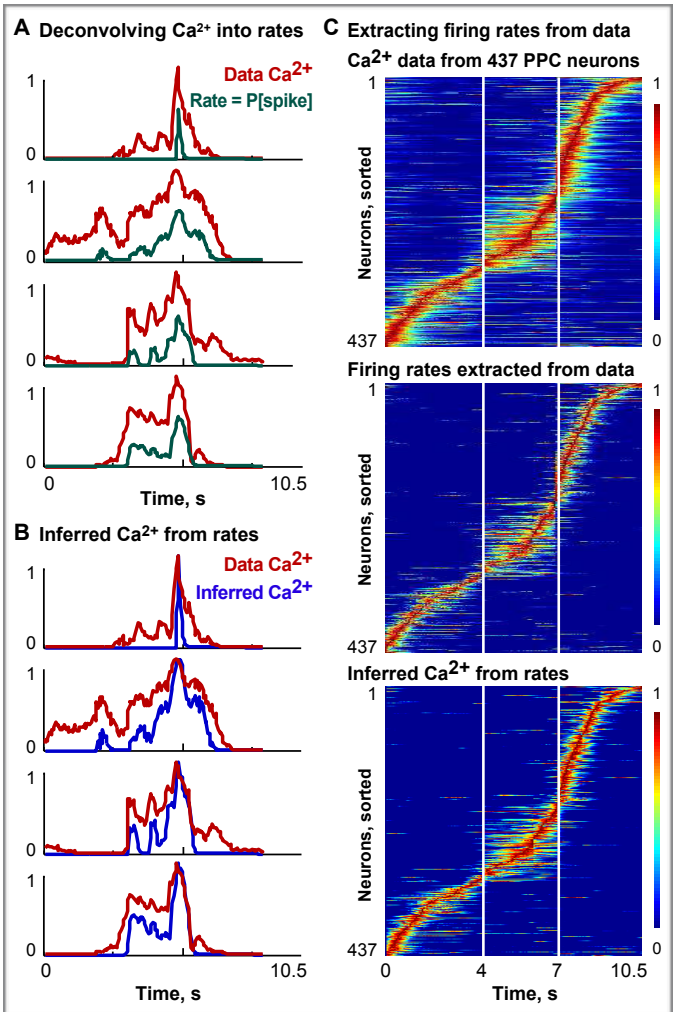
Neuron, Volume *90*

Supplemental Information

Recurrent Network Models of Sequence Generation and Memory

Kanaka Rajan, Christopher D. Harvey, and David W. Tank

SUPPLEMENTAL INFORMATION



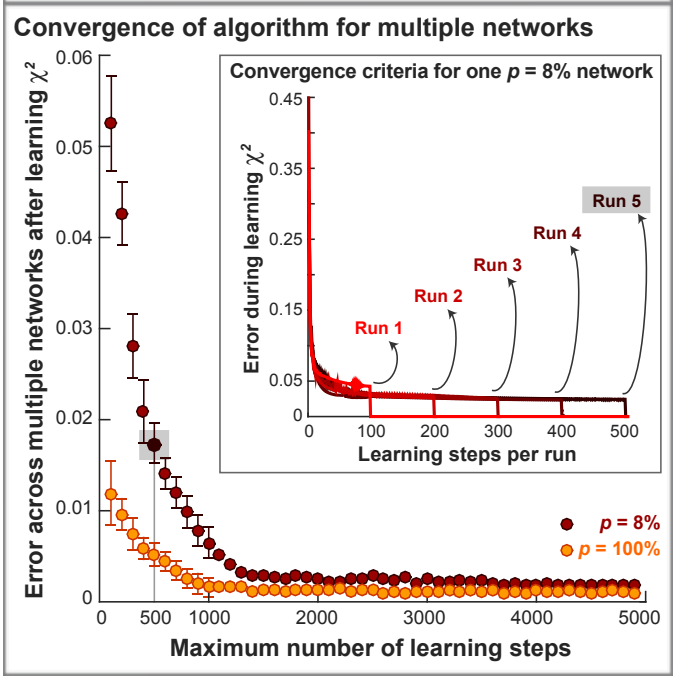


Figure S1 (related to Figure 1): Top Panel, Extracting Target Functions from Ca^{2+} Fluorescence Data by Deconvolution and Bottom Panel (D), Convergence of the PINning Algorithm

A. The use of deconvolution algorithms for the extraction of firing rates is illustrated here. A few example PPC neurons showing the Ca^{2+} fluorescence signals in red (replotted from [Harvey, Coen & Tank, 2012]) and their extracted firing rates (normalized to a maximum of 1) in green. These rates were obtained using both methods described in Experimental Procedures (see also Supplemental Experimental Procedures S7.3). The implementation of the Bayesian algorithm (method #2 from Supplemental Experimental Procedures S7.3) yields spike times along with the statistical confidence of a spike arriving at that particular time point (P[spike]). However, since the frame rate of the imaging experiments in [Harvey, Coen & Tank, 2012] was relatively slow (64ms per frame), we used these "probabilistic" spike trains as a normalized firing rate estimate. To verify, we first deconvolved single trial Ca^{2+} data, and then performed an average over all the single trial P[spike] estimates to obtain the smooth firing rates for each of the 437 sequential neurons that we show in the middle panel of (C).

B. To see how accurate the firing rate estimates extracted from the data were, we re-convolved the extracted firing rates of the example units in (A) through a difference of exponentials with a rise time of 52ms and decay time of 384ms (see Supplemental Experimental Procedures S7.3). The PPC Ca^{2+} data are plotted in red and the inferred Ca^{2+} are plotted in blue here.

C. Top panel shows the original Ca^{2+} fluorescence signals from the PPC (similar to Figure 2A of main text, adapted from Figure 2c in [Harvey, Coen & Tank, 2012]). Middle panel shows the firing rates (from 0 to a maximum of 6Hz, normalized by the peak to 1) extracted from these data by using deconvolution methods. Bottom panel shows the inferred Ca^{2+} signals obtained by convolving these extracted firing rates by a Ca^{2+} impulse response function as explained for (B) (see also Experimental Procedures 3 and Supplemental Experimental Procedures S7.3).

D. Bottom Panel shows the convergence of the PINning algorithm. The convergence of our PINning algorithm (Experimental Procedures 4, see also Supplemental Experimental Procedures S4) is shown here for multiple sequential networks generating sequential outputs (each, similar to Supplemental Figure S2A). We plot the χ^2 error between the target functions and the outputs of several PINned networks as a function of the number of learning steps for two values of $p - p = 8\%$ plastic synapses (mean values in the red circles, means computed over 5 instantiations each) and $p = 100\%$ plastic synapses (mean values in the yellow circles, means computed over 5 instantiations each) for comparison purposes. When the χ^2 error drops below 0.02, which for both sparsely and fully PINned networks occurs before the 500th learning step, we terminate the learning and simulate the network with the PINned connectivity matrix (denoted as $\boldsymbol{J}_{PINned,\,p\%}$ in general) for an additional 50 steps before the program graphs the network outputs (firing rates, inferred calcium to compare with data, statistics of $\boldsymbol{J}_{PINned,\,p\%}$, etc.). The point highlighted in the gray square corresponds to a PINned network with $p = 8\%$ that ran for 500 learning steps, at the end of which, the χ^2 error was 0.018. Additionally, this network had a *pVar* (Experimental Procedures 7, see also Supplemental Experimental Procedures S7.7) of 92%. **Inset** shows the speed of convergence for runs of different lengths for the $p = 8\%$ PINned network. Run #5 corresponds to the example network highlighted in the gray square.

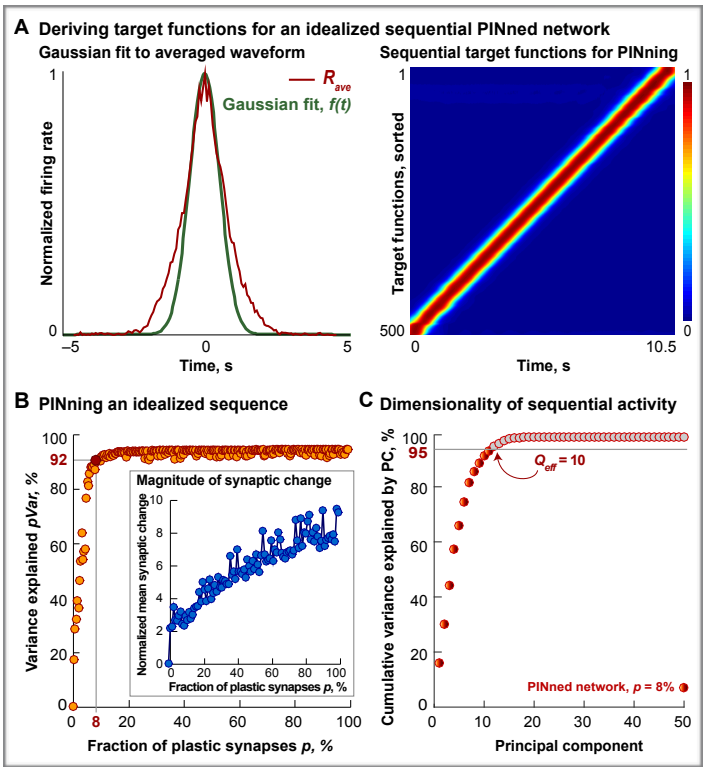


Figure S2 (related to Figures 3 and 6): Idealized Sequence Generating PINned Network

A. Left panel shows a Gaussian with mean = 0 and variance = 0.3, denoted by *f(t)* (green trace), that best fits the neuron-averaged waveform (red trace, R_{ave}, identical to the inset of Figure 1F). This waveform, *f(t)* is used to generate the target functions (right panel) for a network of 500 rate-based model neurons PINned to produce an idealized sequence of population activity (*bVar* = 100%).

B. Effect of increasing the fraction of plastic synapses, *p*, in a network constructed by PINning to produce the single idealized sequence is shown here. *pVar* (Experimental Procedures 7, see also Supplemental Experimental Procedures S7.7) plotted as a function of *p,* increases from 0 for a random unmodified network (*p* = 0) network and plateaus at *pVar* = ~92% for and above *p* = 8%. The sequence-facilitating properties of the connectivity matrix in the *p* = 8% network highlighted in red are analyzed in Figures 3 and 4 of the main text. **Inset** shows the magnitude of synaptic change required to generate an idealized sequence as a function of *p*, computed as in Figure 2E (Experimental Procedures 8, see also Supplemental Experimental Procedures S7.8). The overall magnitude grows from a factor of ~3 for sparsely PINned networks (*p* = 8%) to between 8 and 9 for fully PINned networks (*p* = 100%) producing a single idealized sequence that match the targets in (A). As explained in the main text, although the individual synapses change more in sparsely PINned (small *p*) networks, the total amount of change across the synaptic connectivity matrix is smaller.

C. Dimensionality of sequential activity is computed (as in Figure 2F, Experimental Procedures 5, see also Supplemental Experimental Procedures S7.5) for the 500-neuron PINned network generating an idealized sequence (orange circles) with *p* = 8% and *pVar* = 92%. Of the 500 possible PCs that can capture the total variability in the activity of this 500-neuron network, 10 PCs account for over 95% of the variance when *p* = 8%, i.e., Q_{eff} = 10.

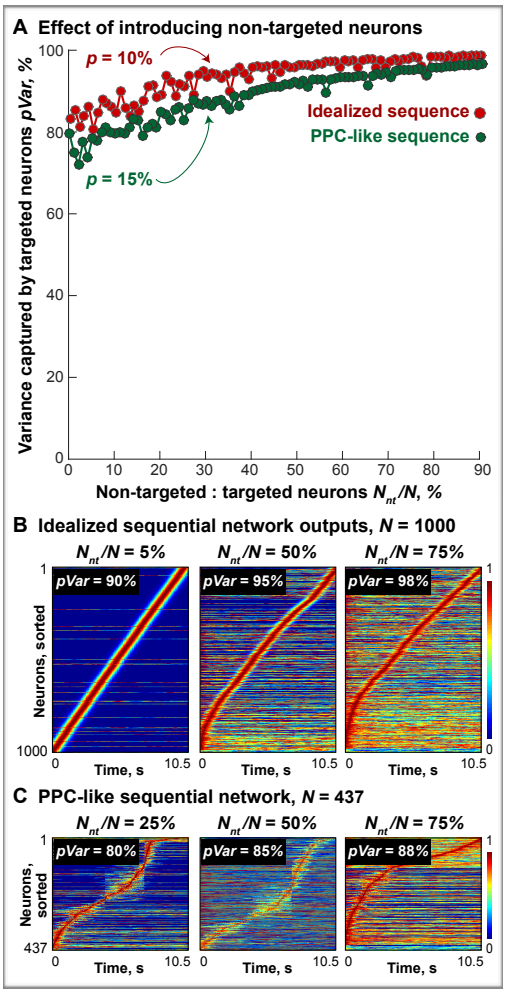


Figure S3 (related to Figures 1, 2 and 5): Simulating Unobserved Neurons by Including Non-targeted Neurons in PINned Networks

A. Variance of the target functions captured by the outputs of the network neurons that have been PINned, *pVar* (evaluated as described in Experimental Procedures 7, see also Supplemental Experimental Procedures S7.7) is plotted here as a function of the ratio of non-targeted to targeted neurons in the network, denoted by N_{nt}/N. Overall, *pVar* does not decrease appreciably when the relative number of untrained neurons introduced into the PINned network is increased. Interestingly however, *pVar* improves slightly for sparsely PINned (p = 10% networks in the red circles for an idealized sequence-generating network and p = 15% networks in green for a PPC-like sequence) when untrained neurons are introduced. This improvement gets smaller as p increases (not shown). The inclusion of non-targeted neurons in the networks constructed by PINning simulates the effect of unobserved but active neurons that may exist in the experimental data and might influence neural activity. It should be noted, however, that these additional neurons, however, might add irregularity to the sorted outputs of the full network (including both targeted and non-targeted neurons) and reduce the stereotypy of the overall outputs (indicated by a decrease in the *bVar* computed over the full network, not shown). This effect is independent of p.

B. Example network outputs are shown here for 3 values of N_{nt}/N. for a network generating an idealized sequence similar to Supplemental Figure 4. *pVar* = 90% for N_{nt}/N = 5%, 95% for N_{nt}/N = 50% and 98% for N_{nt}/N = 75%. The sequences become noisier overall as more randomly fluctuating untrained neurons are introduced, but the percent variance of the targets captured by the PINned neurons remains largely unaffected, even showing a slight improvement.

C. Same as panel (B), except for a 437-neuron network constructed by PINning to generate a PPC-like sequence similar to Figure 2C with different fractions of neurons left untrained. *pVar* = 80% for N_{nt}/N = 25%, 85% for N_{nt}/N = 50% and 88% for N_{nt}/N = 75%, confirming the same general trend as above.

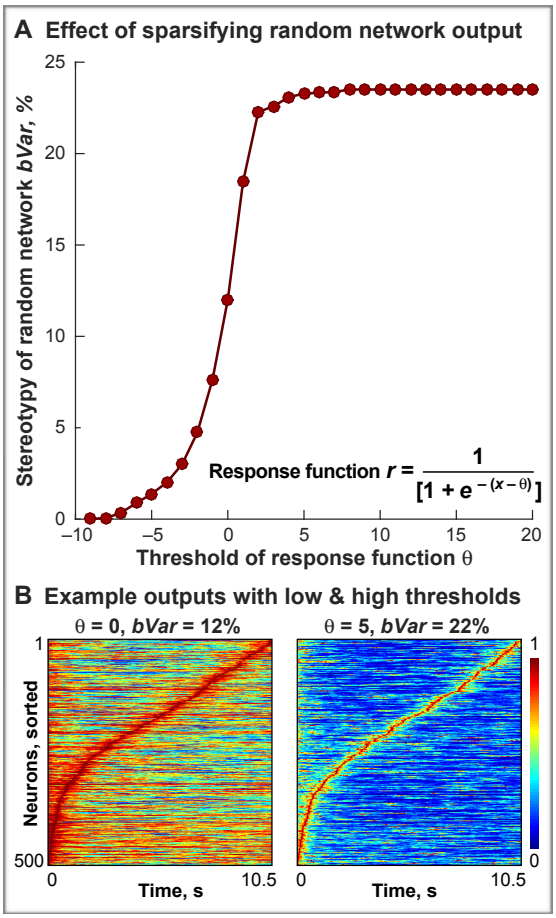


Figure S4 (related to Figure 1): Changing the Threshold of the Response Function and Sparsifying the Random Network Output

A. Effect of sparsifying the outputs of random networks of model neurons constructed with different thresholds, denoted by θ, is shown here. Stereotypy of the sequences made by sorting the outputs of these random networks, *bVar*, increases from 0 to 12% for $\theta = 0$ networks used as initial configurations throughout the paper, and saturates at 22% for networks with threshold values of $\theta > 2$.

B. Example outputs from two random networks, one with $\theta = 0$ (left, with *bVar* = 12%, this is identical to Figure 1B) and one with $\theta = 5$ (right, with *bVar* = 22%) are shown here. Firing rates are normalized by the t_{COM} and sorted to yield the sequences shown here.

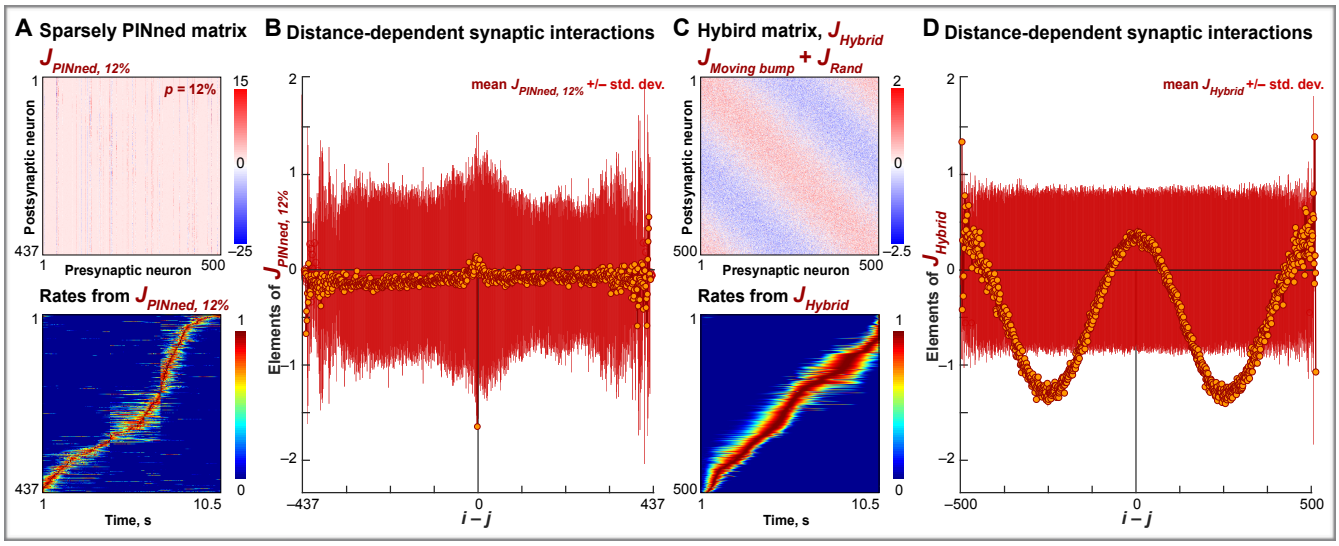


Figure S5 (related to Figure 3): Comparison of the Sparsely PINned Matrix, $J_{PINned,\ 12\%}$ and the Additive Hybrid Connectivity Matrix of the Form, $J_{Hybrid} = J_{Moving\ bump} + J_{Rand}$

A. Synaptic connectivity matrix of a 437-neuron network with p = 12% that produces a PPC-like sequence with $pVar$ = 85%, denoted by $J_{PINned,\ 12\%}$, and the firing rates obtained (identical to Figure 2C, also highlighted by the red circle in Figure 2D) are shown here.

B. Influence of neurons away from the sequentially active neurons is estimated by computing the mean (circles) and the standard deviation (lines) of the elements of J_{Rand} (in blue) and $J_{PINned,\ 8\%}$ (in orange) in successive off-diagonal "stripes" away from the principal diagonal (as described in Experimental Procedures 11 and analogous to Figure 3 in the main text). These quantities are plotted as a function of the "inter-neuron distance", $i - j$. In units of $i - j$, 0 corresponds to the principal diagonal or self-interactions, and the positive and the negative terms are the successive interaction magnitudes of neurons a distance $i - j$ away from the primary sequential neurons.

C. Same as panel (A), except for the synaptic connectivity matrix from an additive hybrid of the form, $J_{Hybrid} = J_{Moving\ bump} + J_{Rand}$ is shown here, where, J_{Rand}, is a random matrix similar to the one shown in the lower panel of Figure 3A and $J_{Moving\ bump}$ is the connectivity for a moving bump model [Yishai, Bar-Or & Sompolinsky, 1995]. The hybrid matrix contains a structured and a random part, and is constructed by the addition of a moving bump connectivity matrix [Yishai, Bar-Or & Sompolinsky, 1995], $J_{\text{Moving bump}} = -J_0 + J_2[\cos(\phi_i - \phi_j)] + 0.06 \times [\sin(\phi_i - \phi_j)]$, $\phi = 0,...,\pi$) and a random matrix, J_{Rand}, similar to the one used to initialize PINning (lower panel of Figure 3A). Mathematically, the hybrid is of the form, $J_{\text{Hybrid}} = \frac{1}{N}[A_B J_{\text{Moving bump}}] + \frac{1}{\sqrt{N}}[A_R J_{\text{Rand}}]$, where A_B is the relative amplitude of the structured or moving bump part, scaled by network size, N, and A_R is the relative amplitude of the random part of the N-neuron hybrid network, scaled by $\sqrt{N}$. The lower panel shows the firing rates from the additive hybrid network whose connectivity is given by J_{Hybrid}. $pVar$ for the output of this hybrid network is only 1%, however, its stereotypy, $bVar$ = 92%.

D. Same as (B), except for J_{Hybrid}. Band-averages (orange circles) are bigger and more asymmetric compared to those for $J_{PINned,\ 12\%}$. Notably, these band-averages are positive for $i - j = 0$ and in the neighborhood of 0. Fluctuations around the band-averages (red lines) for J_{Hybrid} are less structured than those for $J_{PINned,\ 12\%}$ and result from J_{Rand}.

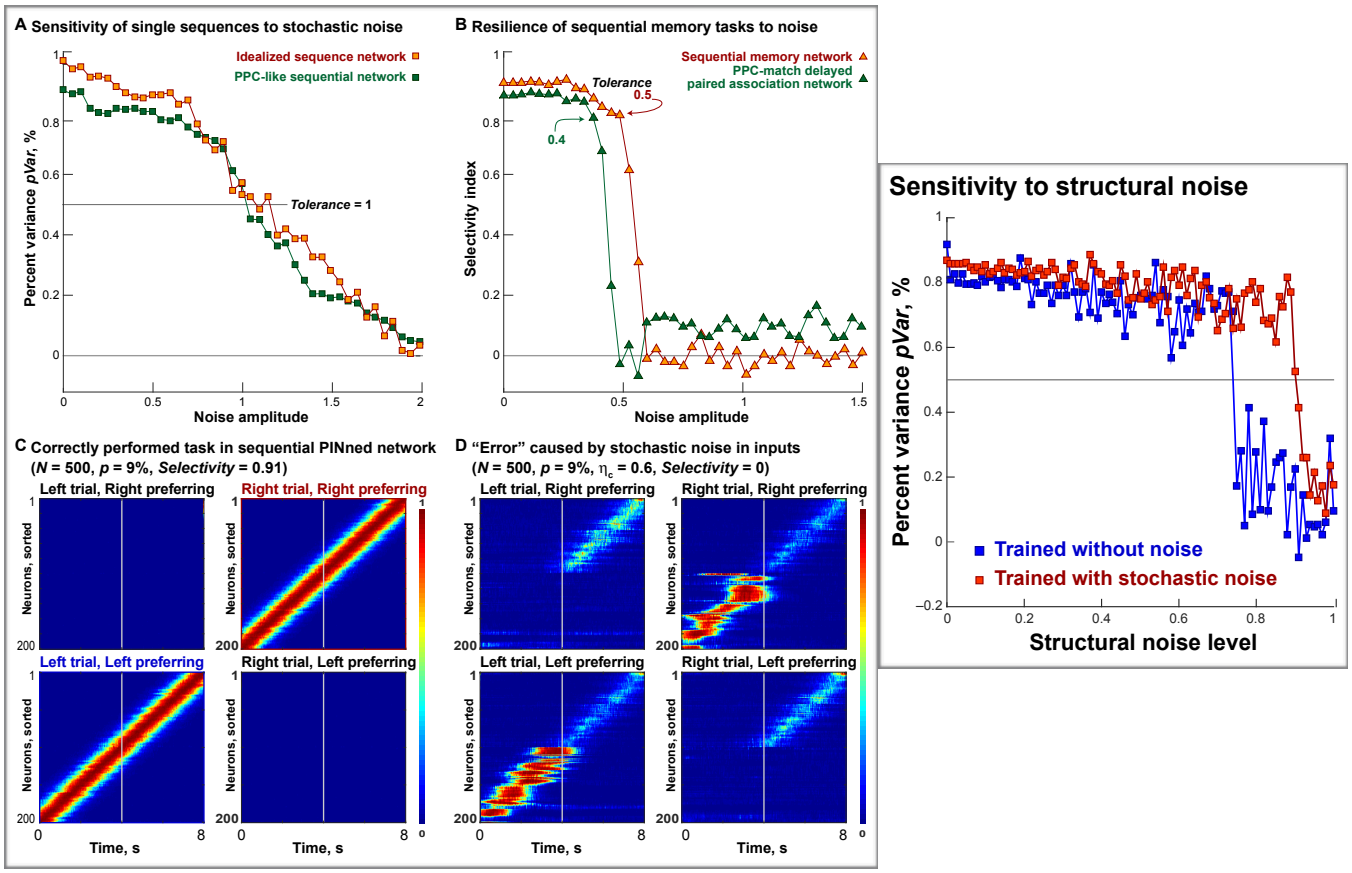


Figure S6 (related to Figures 3 and 6): Left Panel, Robustness of PINned Networks to Stochastic Noise in Inputs and Right Panel, Sensitivity to Structural Noise

A. Stochastic noise, η is injected as an additional current to test whether, and how much, the neural sequences learned through PINning are stable against perturbations (as described in Experimental Procedures 2, see also Supplemental Experimental Procedures S7.2), and the results are shown here. Percent variance of the target functions explained by the network outputs, *pVar*, drops as amplitude of the injected noise is increased. The maximum tolerance is indicated by the gray line and for single sequences, we define it as the amplitude of noise at which *pVar* drops below 50% and denote it by η_c. For the idealized sequence and for the PPC-like sequence, $\eta_c = 1$.

B. Noise tolerance of memory networks that implement working memory through an idealized sequence (orange triangles) and through PPC-like sequences (green triangles, identical to the network shown in Figure 5) is plotted here. Selectivity of both networks drops at different values of η, indicating the maximum resilience or noise-tolerance of each, denoted by η_c – the sequential memory network has a maximum tolerance of $\eta_c = 0.5$ while the PPC-match memory network has a maximum noise tolerance of $\eta_c = 0.4$.

C. Correctly performed delayed paired association task with delay period memory of the cue identity implemented through two idealized sequences (each similar to Supplemental Figure S2A) in a network of 500 neurons with $p = 9\%$ plastic synapses. For the outputs shown here, the turn period is omitted for clarity, and the delay period starts at the 5s time-point in the 10.5s-long task. The performance of this network, quantified by the selectivity index (Experimental Procedures 9) is 0.91.

D. The same network from panel (D) fails to perform the task (selectivity = 0) because the high levels of stochastic noise present in the inputs ($\eta = 0.6$, here) quenches delay period memory.

E. Right Panel shows sensitivity of PINned networks to structural noise in the synaptic connectivity matrix. Structural noise (described by $\text{level} \times \frac{\mathcal{N}(0,1)}{\sqrt{N}}$, where N = 500 here) is added to the connections in the sparsely PINned connectivity matrix, $\boldsymbol{J_{PINned,\ 8\%}}$ to test whether, and by how much, the synaptic connections are finely tuned. *pVar*, computed as in Experimental Procedures 9, is plotted as a function of structural noise amplitude, for sequential networks obtained by PINning in the presence of stochastic noise in the inputs (red squares) and for networks trained without any noise (blue squares). Gray line is at *pVar* = 50%, the noise-free-PINned network has a tolerance (noise amplitude at which *pVar* drops below 50%) of 0.6, and the network trained with noise, 0.8. Training in the presence of stochastic noise therefore leads to slightly more robust networks, although the drop in *pVar* with the addition of structural noise does not fully recover with training noise.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES S7: Detailed Experimental Procedures

7.1. Network elements (see also Experimental Procedures 1 in main text)

We consider a network of *N* fully interconnected neurons described by a standard firing rate model. Each model neuron is characterized by an activation variable, x_i for *i = 1, 2, ... N*, where, *N* = 437 for the PPC-like sequence in Figures 1D and 2A, *N* = 500 for the single idealized sequence in Supplemental Figure S2 and the multi-sequential memory task in Figure 6, and *N* = 569 for the 2AFC task in Figure 5 (we generally build networks of the same size as the experimental dataset we are trying to model, however the results obtained remain applicable to larger networks, see for example, Supplemental Figure S3), and a nonlinear response function, $\phi(x) = \frac{1}{1+e^{[-(x-\theta)]}}$. This function ensures that the firing rates, $r_i = \phi(x_i)$, go from a minimum of 0 to a maximum at 1. Adjusting θ allows us to set the firing rate at rest, $x = 0$, to some convenient and biologically realistic background firing rate, while retaining a maximum gradient at $x = \theta$. We use $\theta = 0$ (but see also Supplemental Figure S4).

We introduce a recurrent synaptic weight matrix ***J*** with element J_{ij} representing the strength of the connection from presynaptic neuron *j* to postsynaptic neuron *i* (schematic in Figure 1A) The individual synaptic weights are initially chosen independently and randomly from a Gaussian distribution with mean and variance given by $\langle J_{ij} \rangle_J = 0$ and $\langle J_{ij} \rangle_J^2 = g^2 / N$, and are either held fixed or modifiable, depending on the fraction of plastic synapses *p* that can change by applying a learning algorithm (Experimental Procedures 4).

The activation variable for each network neuron x_i is determined by,

$$\tau \frac{dx_i}{dt} = -x_i + \sum_j^N J_{ij}\phi(x_j) + h_i .$$

In the above equation, τ = 10ms is the time constant of each unit in the network and the control parameter *g* determines whether ($g > 1$) or not ($g < 1$) the network produces spontaneous activity with non-trivial dynamics [Sompolinsky, Crisanti & Sommers, 1988; Rajan, Abbott & Sompolinsky, 2010; Rajan, Abbott & Sompolinsky, 2011]. We use *g* values between 1.2 and 1.5 for the networks in this paper, so that the randomly initialized network generates chaotic spontaneous activity prior to activity-dependent modification (gray trace in the schematic in Figure 2B), but produces a variety of regular non-chaotic outputs that match the imposed target functions afterward (red trace in the schematic in Figure 2B, see also results in Figures 2, 5 and 6, see also Experimental Procedures 4 later). The network equations are integrated using Euler method with an integration time step, *dt* = 1ms. h_i is the external input to the unit *i*.

7.2. Design of External Inputs (see also Experimental Procedures 2 in main text)

During the course of the real two-alternative forced-choice (2AFC) experiment [Harvey, Coen & Tank, 2012], as the mouse runs through the virtual environment, the different patterns projected onto the walls of the maze (colored dots, stripes, pillars, hatches, etc.) translate into time-dependent visual inputs arriving at the PPC. Therefore, to represent sensory (visual and proprioceptive) stimuli innervating the PPC neurons, the external inputs to the neurons in the network, denoted by $h(t)$, are made from filtered and spatially delocalized white noise that is frozen (repeated from trial to trial), using the equation, $\tau_{WN} \frac{dh}{dt} = -h(t) + h_0\eta(t)$, where η is a random variable drawn from a Gaussian distribution with 0 mean and unit variance, and the parameters h_0 and τ_{WN} control the scale of these inputs and their correlation time, respectively. We use $h_0 = 1$ and $\tau_{WN} = 1$ s. There are as many different inputs as there are model neurons in the network, with individual model neurons receiving the same input on every simulated trial. A few example inputs are shown in the right panel of Figure 1A.

In addition to the frozen noise *h*, which acts as external inputs to these networks, described above, we also test the resilience of the memory networks we built (Figure 6E) to injected stochastic noise. This stochastic injected noise varies randomly (i.e., is a Gaussian random variable between 0 and 1, drawn from a zero mean and unit variance distribution) and

independently at every time step. The diffusion constant of the white noise is given by $A_\eta^2 / 2\tau$, where the amplitude is A_η^2 and τ is the time constant of the network units (we use 10ms, as detailed in Experimental Procedures 1). We define "Resilience" or "Noise Tolerance" as the critical amplitude of this stochastic noise, denoted by η_c, at which the delay period memory fails and the Selectivity Index of the memory network drops to 0 (Figure 6E, see also Supplemental Figure S6).

7.3. Extracting Target Functions From Calcium Imaging Data (see also Experimental Procedures 3 in main text)

To derive the target functions for our activity-dependent synaptic modification scheme termed Partial In-Network Training or PINning, we convert the calcium fluorescence traces from PPC recordings [Harvey, Coen & Tank, 2012] into firing rates using two complementary methods. We find that for this dataset, the firing rates estimated by the two methods agree quite well (Supplemental Figure S1).

The first method is based on the assumption that the calcium impulse response function, which is a difference of exponentials ($\mathrm{K} \propto \mathrm{e}^{-t/384} - e^{-t/52}$, with a rise time of 52ms and a decay time of 384ms [Tian et al, 2009; Harvey, Coen & Tank, 2012]), is approximated by an alpha function of the form $\mathrm{K} \propto \mathrm{te}^{-t/\tau_{\mathrm{Ca}}}$, where there is only a single (approximate) time constant for the filter, τ_{Ca} = 200ms. According to this assumption, the scaled firing rate *s* and calcium concentration, $[Ca^{2+}]$ are related by,

$$\tau_{Ca}\frac{dCa_i}{dt} = -Ca_i(t) + x_i(t) \quad \text{and} \quad \tau\frac{dx_i}{dt} = -x_i(t) + s_i(t),$$

where, x*(t)* is an auxiliary variable.The inverse of the above model is obtained by taking a derivative of the calcium data, writing,

$$x_i(t) = Ca_i(t) + \tau_{Ca}\frac{dCa_i}{dt} \quad \text{and} \quad s_i(t) = x_i(t) + \tau\frac{dx_i}{dt}.$$

Once we have *s(t),* we rectify it and choose a smoothing time constant τ_R for the firing rate we need to compute. Finally, integrating the equation, $\tau_R\frac{dR_i}{dt} = -R_i(t) + s_i(t)$, and normalizing by the maximum gives us an estimate for the firing rates extracted from the calcium data, denoted by *R* (Supplemental Figure S1).

The second method is a fast Bayesian deconvolution algorithm [Pnevmatikakis et al, 2016, available online at https://github.com/epnev/continuous_time_ca_sampler] that infers spike trains from calcium fluorescence data. The inputs to this algorithm are the rise time (52ms) and the decay time (384ms) of the calcium impulse response function [Tian et al, 2009; Harvey, Coen & Tank, 2012] and a noise parameter [Pnevmatikakis et al, 2016]. Typically, if the frame rate for acquiring the calcium images is low enough (the data in [Harvey, Coen & Tank, 2012] are imaged at 64ms per frame), the outputs from this algorithm can be interpreted as a normalized firing rate. To verify the accuracy of the firing rate outputs obtained from trial-averaged calcium data (for example, from Figure 2c in [Harvey, Coen & Tank, 2012]), we smoothed the spike trains we got from the above method for each trial separately through a Gaussian of the form $\sum_i e^{\frac{-(t-t_i)}{2\tau_R^2}}$, normalized by $\sqrt{2\pi}\times\tau_R$, and then averaged over single trials to get trial-averaged firing rates (this smoothing and renormalization procedure has also been recommended for faster imaging times [Pnevmatikakis et al, 2016]).

Once the values of τ_R and τ_{Ca} are determined that make the results obtained by both deconvolution methods consistent (we used τ_R = 100ms and τ_{Ca} = 384ms), we used the firing rates extracted as target functions for PINning through the transform, $f_i(t) = \ln\left[\frac{R_i(t)}{1-R_i(t)}\right]$. The above expression is obtained by solving the activation function relating input current to firing rate of model neurons, $R_i(t) = \frac{1}{1+e^{-f_i(t)}}$, since the goal of PINning is to match the input to neuron *i*, say, denoted

by $z_i(t)$ to its target function, denoted by $f_i(t)$. Finally, to verify our estimates, we re-convolved (Supplemental Figure S1) the output firing rates from the network neurons with a difference of exponentials using a rise time of 52ms and a decay time of 384ms.

7.4. Synaptic Modification Rule For PINning (see also Experimental Procedures 4 in main text)

During PINning, the inputs of individual network neurons are compared directly with the target functions to compute a set of error functions, i.e., $e_i(t) = z_i(t) - f_i(t)$, for *i = 1, 2, ... N.* Individual neuron inputs are expressed as $z_i(t) = \sum_j J_{ij} r_j(t)$, where $r_j(t)$ is the firing rate of the j^{th} or the presynaptic neuron.

During learning, the subset of plastic internal weights in the connectivity matrix ***J*** of the random recurrent network, denoted by the fraction *p*, undergo modification at a rate proportional to the error term, the presynaptic firing rate of each neuron, r_j and a *pN x pN* matrix, ***P*** (with elements P_{ij}) that keeps track of the rate fluctuations across the network at every time step. Here, *p* is the fraction of neurons whose outgoing synaptic weights are plastic; since this is a fully connected network, this is also the fraction of plastic synapses in the network. Mathematically, $P_{ij} = < r_i r_j >^{-1}$, the inverse cross-correlation matrix of the firing rates of the network neurons (P_{ij} is computed for all *i* but is restricted to *j = 1, 2, 3,...,pN*). The basic algorithm is schematized in Figure 2B. At time *t*, for *i = 1, 2, ... N* neurons, the learning rule is simply that the elements of the matrix ***J*** are moved from their values at a time step Δt earlier through $J_{ij}(t) = J_{ij}(t-1) + \Delta J_{ij}(t)$. Here, the synaptic update term, according to the RLS/FORCE procedure [Haykins, 2002; Sussillo & Abbott, 2009] (since other methods for training recurrent networks, such as backpropagation would be too laborious for our purposes) follows, $\Delta J_{ij}(t) = c[z_i(t) - f_i(t)] \sum_k P_{jk}(t) r_k(t)$, where the above update term is restricted to the *p*% of plastic synapses in the network, which are indexed by *j* and *k* in the above expression. While c can be thought of as an effective learning rate, it is given by the formula, $c = \frac{1}{1 + r'(t)\mathbf{P}(t)r(t)}$. The only free parameter in the learning rule is ***P(0)*** (but the value to which it is set is not critical [Sussillo & Abbott, 2009]). When there are multiple sequences (such as in Figures 5 and 6), we choose *pN* synapses that are plastic and we use those same synapses for all the sequences.

The matrix ***P*** is generally not explicitly calculated but rather updated according to the rule, $\mathbf{P}(t) = \mathbf{P}(t-1) - \frac{\mathbf{P}(t-1)r(t)r'(t)\mathbf{P}(t-1)}{1 + r'(t)\mathbf{P}(t-1)r(t)}$ in matrix notation, which includes a regularizer [Haykins, 2002]. In our scheme, all indices in the above expression are restricted to the neurons with plastic synapses in the network. The algorithm requires the matrix ***P*** to be initialized to the identity matrix times a factor that controls the overall learning rate, i.e., $P(0) = \alpha \times \mathrm{I}$, and in practice, values from 1 to 10 times the overall amplitude of the external inputs (denoted by h_0 in Experimental Procedures 2) driving the network are effective (other values are explored in [Sussillo & Abbott, 2009]).

For numerically simulating the PINned networks whose sequential outputs are shown in Figures 2C, 5 and 6, the integration time step used is *dt* = 1ms (as described in Experimental Procedures 1 and 2, we use Euler method for integration). The learning occurs at every time step for the *p*% of pre-synaptic neurons with plastic outgoing synaptic weights. Starting from a random initial state (Experimental Procedures 1), we first run the program for 500 learning steps, which include both the network dynamics and the PINning algorithm, and then an additional 50 steps with only the network dynamics after the learning has been terminated (convergence metrics below, see also Supplemental Figure S1D). A "step" is defined as one run of the program for the duration of the relevant trial, denoted by *T*. Each step is equivalent to *T* = 10500 time points (10.5s) in Figures 2C and 5E, and 8000 time points (8s) in Figure 6B.

On a standard laptop computer, the first 500 such steps for a 500-unit rate-based network producing a 10s-long sequence (such as in Supplemental Figure S2A) take approximately 8 minutes to complete in realtime and the following 50, about 30 seconds in realtime (about 1s/step for the 10s-long single sequence example, scaling linearly with network size *N*, total duration or length of the trial *T*, and number of sequences produced.)

The convergence of the PINning algorithm was assayed as follows: (a) By directly comparing the outputs with the data (as in the case of Figures 2A–C and 5D–E) or the set of target functions used (Supplemental Figure S2A) and (b) By calculating and following the χ^2 -squared error between the network rates and the targets, both during PINning (inset of Supplemental Figure S1D) and at the end of the simulation (Supplemental Figure S1D). The performance of the PINning algorithm was assayed by computing the percent variance of the data or the targets captured by the sequential outputs of different networks (*pVar*, see Experimental Procedures 7, see also Figure 1H) or by computing the Selectivity Index of the memory network (Selectivity, see Experimental Procedures 9, see also Figure 5F).

7.5. Dimensionality Of Network Activity (Q_{eff}) (see also Experimental Procedures 5 in main text)

We use state space analysis based on PCA (see for example, [Rajan, Abbott & Sompolinsky, 2011; Sussillo, 2014] and references therein) to describe the instantaneous network state by diagonalizing the equal-time cross-correlation matrix of network firing rates given by, $Q_{ij} = \langle (r_i(t) - < r_i >)(r_j(t) - < r_j >) \rangle$, where <> denotes a time average. The eigenvalues of this matrix expressed as a fraction of their sum indicate the distribution of variances across different orthogonal directions in the activity trajectory. We define the effective dimensionality of the activity, Q_{eff}, as the number of principal components that capture 95% of the variance in the dynamics (Figure 2F).

7.6. Stereotypy Of Sequence (*bVar*), % (see also Experimental Procedures 6 in main text)

bVar quantifies the variance of the data or the network output that is explained by the translation along the sequence of an activity profile with an invariant shape. For example, for Figure 1C–G, we extracted an aggregate waveform, denoted by R_{ave} (red trace in Figures 1D and the left panel of Supplemental Figure S2A), by averaging the t_{COM}-realigned firing rates extracted from trial-averaged PPC data collected during a 2AFC task [Harvey, Coen & Tank, 2012]. Undoing the t_{COM} shift, we can write this function for the aggregate or "typical" bump-like waveform as $R_{\text{ave}}(t - t_i)$. The amount of variability in the data that is explained by the moving bump, $R_{\text{ave}}(t - t_i)$ is given by a measure we call *bVar*.

$bVar = \left[1 - \frac{\langle R_i(t) - R_{\text{ave}}(t - t_i) \rangle^2}{\langle R_i(t) - \bar{R}(t) \rangle^2} \right]$, where $<> = \sum_i^N \sum_t^T$. In the above expressions, *R*'s denote the firing rates extracted from calcium data (Experimental Procedures 3); $R_i(t)$ is the firing rate of the i^{th} PPC neuron at time *t* and $\bar{R}(t)$ is the average over neurons. The total duration,*T* is 10.5s in Figures 1, 2 and 5, and 8s in Figure 6.

In some ways, *bVar* is similar to the ridge-to-background ratio computed during the analysis of experimental data for measuring the level of background activity (see for example, Supplemental Figure 14 in [Harvey, Coen & Tank, 2012]); however, *bVar* additionally quantifies the stereotypy of the shape of the transient produced by individual neurons.

7.7. Percent Variance Of Data Explained By Model (*pVar*), % (see also Experimental Procedures 7 in main text)

We quantify the match between the experimental data or the set of target functions, and the outputs of the model by the amount of variance of the data that is captured by the model, $pVar = \left[1 - \frac{\langle D_i(t) - r_i(t) \rangle^2}{\langle D_i(t) - \bar{D}(t) \rangle^2} \right]$, which is one minus the ratio of the Frobenius norm of the difference between the data and the outputs of the network, and the variance of the data. The data referred to here, denoted by *D*, is trial-averaged data, such as from Figures 1D, 2A and 5D.

7.8. Magnitude Of Synaptic Change (see also Experimental Procedures 8 in main text)

In Figure 2E and in Figure 5G, we compute the magnitude of the synaptic change required to implement a single PPC-like sequence, an idealized sequence and three memory tasks, respectively. In combination with the fraction of plastic synapses in the PINned network, *p*, this metric characterizes the amount of structure that needs to be imposed in an initially random network to produce the desired temporally structured dynamics. This is calculated as, normalized mean synaptic

change = $\frac{\sum_{ij} |J_{ij}^{\text{PINned, p\%}} - J_{ij}^{\text{Rand}}|}{\sum_{ij} |J_{ij}^{\text{Rand}}|}$, where, following the same general notation as in the main text, $\boldsymbol{J_{PINned,\,p\%}}$ denotes the connectivity matrix of the PINned network constructed with p% plastic synapses and $\boldsymbol{J_{Rand}}$ denotes the initial random connectivity matrix ($p = 0$).

7.9. Selectivity Index For Memory Task (see also Experimental Procedures 9 in main text)

In Figures 5F, a Selectivity Index is computed (similar to Figure 4 in [Harvey, Coen & Tank, 2012]) to assess the performance of different PINned networks at maintaining cue-specific memories during the delay period of delayed paired association tasks. This metric is based on the ratio of the difference and the sum of the mean activities of preferred neurons at the end of the delay period during preferred trials, and the mean activities of preferred neurons during opposite trials. We compute Selectivity Index as, $\frac{1}{2}\left[\frac{< r >_{r.t}^{r.n} - < r >_{l.t}^{r.n}}{< r >_{r.t}^{r.n} + < r >_{l.t}^{r.n}} + \frac{< r >_{l.t}^{l.n} - < r >_{r.t}^{l.n}}{< r >_{l.t}^{l.n} + < r >_{l.t}^{l.n}}\right]$, where the notation is as follows:

$< r >_{\text{r.t}}^{\text{r.n}} = \frac{1}{N_{\text{right pref neurons}}} \sum_{i}^{\text{right pref neurons}} r_i$ and $< r >_{\text{r.t}}^{\text{l.n}} = \frac{1}{N_{\text{left pref neurons}}} \sum_{i}^{\text{left pref neurons}} r_i$ are the average firing rates of right-preferring and left-preferring model neurons on right trials;

$< r >_{\text{l.t}}^{\text{r.n}} = \frac{1}{N_{\text{right pref neurons}}} \sum_{i}^{\text{right pref neurons}} r_i$ and $< r >_{\text{l.t}}^{\text{l.n}} = \frac{1}{N_{\text{left pref neurons}}} \sum_{i}^{\text{left pref neurons}} r_i$ are the average firing rates of right-preferring and left-preferring model neurons on left trials of the simulated task. The end of the delay period is at approximately 10s for the network in Figure 5 (after [Harvey, Coen & Tank, 2012]) and at ~7s time point for the network in Figure 6.

7.10. Temporal Sparseness Of Sequences ($\boldsymbol{N_{Active}/N}$) (see also Experimental Procedures 10 in main text)

The temporal sparseness of a sequence is defined as the fraction of neurons active at any instant during the sequence, the fraction, N_{Active}/N. To compute this, first, the normalized firing rate from each model neuron in the network or from data [Harvey, Coen & Tank, 2012], denoted by $R_i(t)$, is realigned by the center-of-mass, $i_{COM}(t)$, given by, $i_{\text{COM}}(t) = \sum_i i * R(i,t) / \sum_i R(i,t)$, where i is the neuron number and t is time. The realigned rates are then averaged over time to obtain $<R>_{\text{time}}$, i.e., $\langle R(i) \rangle_{\text{time}} = \langle R(i - i_{\text{COM}}(t), t) \rangle_{\text{time}}$ after using a circular shift rule to undo the i_{COM}-shift. The standard deviation of the best Gaussian that fits this curve $<R>_{\text{time}}$ is the number, N_{Active}, and the ratio of N_{Active} to the network size, N, yields the temporal sparseness of the sequence, N_{Active}/N. For the data in Figure 2A, for example, N_{Active}/N = 3% (i.e., 16 neurons out of a total of 437). Practically speaking, decreasing this fraction makes the sequence narrower and at a critical value of sparseness (N_{Active}/N = 1.6% in Figure 6), there is not enough current in the network to propagate the sequence. However, up to a point, making the sequences sparser increases the capacity of a network for carrying multiple non-interfering sequences (i.e., sequences with delay period memory of cue identity), without demanding an increase in either the fraction of plastic synapses, p, or in the overall magnitude of synaptic change.

7.11. Analyzing the Structure of PINned Synaptic Connectivity Matrices (see also Experimental Procedures 11 in main text)

This is pertinent to Section 3 (Figures 3 and 4), in which we quantify how the synaptic strength varies with the "distance" between pairs of network neurons in connectivity space, $i - j$, in PINned sequential networks. We first compute the means and the standard deviations of the principal diagonals, i.e., $\sum_{i=j}^{N} \frac{J_{ij}}{N} \rightarrow \sum_{i=1}^{N} \frac{J_{ii}}{N}$, in the 3 connectivity matrices under consideration here – $\boldsymbol{J_{PINned,\,8\%}}$, $\boldsymbol{J_{Rand}}$, and just for comparison purposes, $\boldsymbol{J_{PINned,\,100\%}}$. Then, we compute the means and the

standard deviations of successive off-diagonal "stripes" moving away from the principal diagonal, i.e., $\sum_{i-j=a} \frac{J_{ij}}{N} \rightarrow \sum_{i=a+1}^{N} \frac{J_{i(i-a)}}{N}$, for the same three matrices. These are plotted in Figures 3B for the sparsely PINned matrix, $\boldsymbol{J_{PINned,\ 8\%}}$, relative to the randomly initialized matrix, $\boldsymbol{J_{Rand}}$, and in Figure 3E, for the fully PINned matrix constructed for comparison purposes, $\boldsymbol{J_{PINned,\ 100\%}}$. The same analysis is also used to compare different partially structured matrices in Supplemental Figure S5.

7.12: Cross-Validation Analysis (see also Experimental Procedures 12 in main text)

We first divided the data from 436 neurons in Figure 1D into two separate "synthetic" sequences by assigning the even-numbered cells to one (let's call this Sequence A, containing 218 neurons) and the odd-numbered cells to another sequence (say, Sequence B, also with 218 neurons). Then we constructed a PINning-based network with 218 model neurons, exactly like we described in the main text, using data from Sequence A as target functions for PINning. Next, we computed the percent variance of the data in Sequence A and the data in Sequence B, denoted by $pVar_{Test\ A,\ Train\ A}$ and $pVar_{Test\ B,\ Train\ A}$, respectively, that are captured by the outputs of the PINned network (Experimental Procedures 7). Following a similar procedure, we also computed $pVar_{Test\ A,\ Seq\ B}$ and $pVar_{Test\ B,\ Seq\ B}$, after PINning a second network against target functions derived from Sequence B.

We obtained the following estimates for all fractions of plastic synapses, p :

$pVar_{Test\ A,\ Train\ A} = 91 \pm 2\%$ $\qquad$ $pVar_{Test\ B,\ Train\ A} = 45 \pm 2\%$

$pVar_{Test\ A,\ Train\ B} = 46 \pm 2\%$ $\qquad$ $pVar_{Test\ B,\ Train\ B} = 90 \pm 2\%$

For comparison purposes, a random network such as the one in Figure 1B only captures a tiny amount of the variability of data (in this notation, $pVar_{Test\ Data,\ Random\ Network} = 0.2\%$, see also, right panel of Figure 1H). Additionally, the data from one set only accounts for 49% of the variance of the data of the other set, i.e., $pVar = 49\%$. Thus the model does almost as well as it possibly could.

REFERENCE

Tian L, Hires SA, Mao T, Huber D, Chiappe ME, Chalasani SH, Petreanu L, Akerboom J, McKinney SA, Schreiter ER, Bargmann CI, Jayaraman V, Svoboda K, Looger LL (2009) Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. Nat Methods. 6(12): 875-881.