



Published in final edited form as:

Neuron. 2016 April 6; 90(1): 128–142. doi:10.1016/j.neuron.2016.02.009.

Recurrent Network models of sequence generation and memory

Kanaka Rajan, Christopher D Harvey, and David W Tank

Joseph Henry Laboratories of Physics and Lewis–Sigler Institute for Integrative Genomics,
Princeton University, Princeton, NJ.

Department of Neurobiology, Harvard Medical School, Boston, MA.

Department of Molecular Biology and Princeton Neuroscience Institute, Princeton University,
Princeton, NJ.

Kanaka Rajan: krajan@princeton.edu; Christopher D Harvey: harvey@hms.harvard.edu; David W Tank:
dwtank@princeton.edu

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 [Harvey, Coen and Tank, 2012]. 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, Burdick, Musallam, Pesaran and Cham, 2004; Pulvermutter and Shtyrov, 2009; Buonomano, 2003; Ikegaya, Aaron, Cossart, Aronov, Lampl, Ferster and Yuste, 2004; Tang et al, 2008; Seidemann, Meilijson, Abeles, Bergman and Vaadia, 1996; Fujisawa, Amarasingham, Harrison and Buzsaki 2008; Crowe, Averbek

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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. and Supervision, D.W.T.

and Chafee, 2010; Harvey, Coen and Tank, 2012], hippocampus [Nadasdy, Hirase, Czurko, Csicsvari and Buzsaki, 1999; Louie and Wilson, 2001; Pastalkova, Itskov, Amarasingham and Buzsaki, 2008; Davidson, Kloosterman and Wilson, 2009], basal ganglia [Barnes, Kubota, Hu, Jin and Graybiel, 2005; Jin, Fuji and Graybiel, 2009], cerebellum [Mauk and Buonomano, 2004], and area HVC of the songbird [Hahnloser, Kozhevnikov and Fee, 2002; Kozhevnikov and Fee, 2007]. 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 Prugel-Bennett, 1996; Levy et al, 2001; Hermann, Hertz and Prugel-Bennet, 1995; Fiete, Senn, Wang and Hahnloser, 2010], in which excitation flows unidirectionally from one active neuron to the next along a chain of connected neurons, or as ring attractor models [Yishai, Bar-Or and Sompolinsky, 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, Crisanti and Sommers, 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; Pearlmuter, 1989; Jaeger and Haas, 2004; Maass, Joshi and Sontag, 2007; Sussillo and Abbott, 2009; Maass, Natschlagler and Markram, 2002; 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 the 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, Coen and Tank, 2012], and 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 are 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, Coen and Tank, 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

1. 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 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 extract firing rates from the calcium fluorescence signals recorded in the PPC using two complementary deconvolution methods (Figures 1C and D; see also Figure S1 and Experimental Procedures). We define 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 and Experimental Procedures). The second metric, $pVar$, measures the percent variance of the PPC data [Harvey, Coen and Tank, 2012] captured by the different network outputs and is useful for tracking network performance across different parameters (Figure 1H and Experimental Procedures). $bVar$ and $pVar$ are used throughout this paper (Figures 1–6).

Many models have suggested that highly structured synaptic connectivity, i.e., containing ring-like or chain-like interactions, is responsible for neural sequences [Yishai, Bar-Or and Sompolinsky, 1995; Zhang, 1996; Hertz and Prugel-Bennett, 1996; Levy et al, 2001; Hermann, Hertz and Prugel-Bennet, 1995; Fiete, Senn, Wang and Hahnloser, 2010]. We therefore first asked how much of the variance in long-duration neural sequences seen experimentally, e.g., in the PPC [Harvey, Coen and Tank, 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 was found to be quite small – $bVar = 40\%$ for Figures 1D and 2A, which were averaged over hundreds of trials and pooled across different animals. $bVar$ was lower in both single trial data (10–15% for data in Supplemental Figure 13 from [Harvey, Coen and Tank, 2012]) and trial-averaged data from a single mouse (15% for the data in Figure 2c from [Harvey, Coen and Tank, 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, Coen and Tank, 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, Crisanti and Sommers, 1988]; however, the presence of external stimuli can channel the ongoing dynamics in these networks by suppressing their chaos [Molgedey, Schuchhardt and Schuster, 1992; Bertschinger and Natschlager, 2004; Rajan, Abbott and Sompolinsky, 2010; Rajan, Abbott and Sompolinsky, 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, Abbott and Fiser, 2011]. 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 (10.5s, Figure 1C) and sorted in ascending order of their time of center-of-mass (t_{COM}), matching the procedures applied to experimental data [Harvey, Coen and Tank, 2012]. Although the resulting ordered spontaneous activity was sequential (not shown, but similar to Figure 1B), the level of extra-sequential background ($bVar = 5 \pm 2\%$ 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 data at 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, $bVar = 10 \pm 2\%$ and $pVar = 0.2 + 0.1\%$ (Figure 1B, left panel of Figure 1H and Figure 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–D and 2A, compared to Figure 1B). Therefore, sequences like those observed during timing and memory experiments [Harvey, Coen and Tank, 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 be modified by training to produce realistic sequences.

2. Temporally Constrained Neural Sequences Emerge With Synaptic Modification

To construct networks that match the activity seen in the PPC, we developed a training scheme, Partial In-Network Training (PINning), in which a user-defined fraction of synapses were modified (Experimental Procedures). For synaptic modification, the inputs to individual model neurons were compared directly with target functions or templates derived from the data [Fisher, Olasagasti, Tank, Aksay and Goldman, 2013], on both left and right correct-choice outcome trials from 437 trial-averaged neurons, pooled across 6 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 (RLS) or first-order reduced and controlled error (FORCE) learning rule [Haykins, 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 ($pN^2 \ll N^2$) 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, also 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 B), through networks with partially structured connectivity ($p < 25\%$ for Figures 2, 5 and 6) to networks containing entirely trained connections ($p = 100\%$, Figure 3D–F).

We first applied PINning to a network with templates obtained from a single PPC-like sequence (Figures 2A–C) 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 (Figure 1F) 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). Although our networks were typically as large as the size of the experimental dataset ($N = 437$ for Figures 1–2 and 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 in Section 3).

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, Abbott and Sompolinsky, 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.

3. 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 (10.5s, 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 10.5s 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, $pVar$ increases with p , reaches $pVar=92\%$ at $p=8\%$ (red circle in Figure S2B), asymptoting at $p \simeq 10\%$. 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 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 ring/bump 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 order 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 3 connectivity matrices interconnecting a population of model neurons, $N=500$ – before PINning (randomly initialized matrix, denoted by \mathbf{J}_{Rand}), after sparse PINning ($\mathbf{J}_{\text{PINned}, 8\%}$), and after full PINning ($\mathbf{J}_{\text{PINned}, 100\%}$, built as a useful comparison). To analyze how the synaptic strengths in these 3 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., $i-j = \text{constant}$ (Experimental Procedures). These band-averages and the fluctuations were 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 $\mathbf{J}_{\text{PINned}, 100\%}$ formed a localized and asymmetric profile (orange circles in Figure 3E) that is qualitatively similar to moving ring attractor dynamics [Yishai, Bar-Or and Sompolinsky, 1995; Zhang, 1996]. In contrast, the band-averages for $\mathbf{J}_{\text{PINned}, 8\%}$ (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 [Yishai, Bar-Or and Sompolinsky, 1995]. Furthermore, neither the band-averages of $\mathbf{J}_{PINned, 8\%}$ (shown in Figure 3C) nor the fluctuations (not shown) were, by themselves, sufficient to produce moving sequences similar to the full matrix $\mathbf{J}_{PINned, 8\%}$. Instead, the outputs of the synthetic networks built from the components of the band-averaged $\mathbf{J}_{PINned, 8\%}$ 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 matrices, $\mathbf{J}_{PINned, 8\%}$. As expected (see color-bars in Figure 3A, lines in Figures 3B and 3F), the fluctuations around the band-averages of $\mathbf{J}_{PINned, 8\%}$ (red lines in Figure 3B) were much larger and more structured than \mathbf{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 $\mathbf{J}_{PINned, 8\%}$ and the external input. 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 (for 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 $\mathbf{J}_{PINned, 8\%}$ (Figure 3B). Instead, the variations in the fluctuations of $\mathbf{J}_{PINned, 8\%}$ (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 $\mathbf{J}_{PINned, 8\%}$ more directly. For all the 3 connectivity matrices of sequential networks, there was a substantial increase in the magnitudes spanned by the synaptic weights as p decreased (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, \mathbf{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, $\mathbf{J}_{PINned, 8\%}$ 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 $\mathbf{J}_{PINned, 8\%}$, 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 post-synaptic potentiation (PSP) of 0.05mV, then the "large" synapses had a 1mV PSP, which is within the range of experimental data [Song et al, 2005]. For comparison purposes, the connectivity for a fully structured network, $\mathbf{J}_{PINned, 100\%}$, had a mean = 0, variance = 0.7, skewness = -0.02 ,

and kurtosis = 0.2, corresponding to a case 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 are fairly sensitive to small amounts of structural noise, i.e., perturbations in the matrix J^{PINned} , 8%. However, when stochastic noise (Experimental Procedures) is used during training, slightly more robust networks were obtained (Figure S6E).

4. Delayed Paired Association And Working Memory Can be Implemented Through Sequences in PINned Networks

Delayed Paired Association (DPA) tasks, such as the two-alternative forced-choice (2AFC) task from [Harvey, Coen and Tank, 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, Coen and Tank, 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, Coen and Tank, 2012]. By showing that the partially structured network 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, Coen and Tank, 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. In the middle third, the delay period, when the left- or right-specific visual cues were off (blue and red traces in Figure 5C), 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 > 1 sequences to maintain the memory of the cue 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 E, $pVar = 85\%$, $bVar = 40\%$, compare with Figure 2c in [Harvey, Coen and Tank, 2012], here, $N = 569$, 211 network neurons selected at random to activate in the left trial condition, 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, 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 seen experimentally (Figure 5E, also Figure 7b in [Harvey, Coen and Tank, 2012]).

5. 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 , 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, Coen and Tank, 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 3 types of memory networks, we computed a selectivity index (Experimental Procedures, similar to Figure 4 in [Harvey, Coen and Tank, 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 = 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, Coen and Tank, 2012] and to achieve a selectivity = 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 makes a mistake is about 15–20% (accuracy of the performance of mice was found to $83 \pm 9\%$ correct [Harvey, Coen and Tank, 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 makes a turn. Given a 50% probability of turning in the correct direction by chance, this implies that the cue identity is forgotten on 30–40% of trials. Adding noise to our model, we can 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.

6. 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 required the activation of >2 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 ($N_{Non-choice-specific}/N$), and temporal sparseness (fraction of neurons active at any instant, N_{Active}/N , Experimental Procedures).

We adjusted the PINning parameters for a type of a 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 8s. 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 is omitted here for clarity, multi-sequential tasks are 8s-long here. Similar to DPA (Figure 5), each neuron received a different filtered white noise input for each cue during the cue period (0–4s), but during the delay (4–8s), they 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 is the number of memories (capacity), which is 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 8s-long task, when $N_{Active}/N < 1.6\%$) the network failed. Although there was sequential activation, the memory of the 5 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 ($N_{Non-choice-specific}/N = 4\%$) that fired in the same order in all 5 trial types (Figure 6C). In other words, the $N - N_{Non-choice-specific}$ network neurons ($< N$) successfully executed the task using 5 sparse choice-specific sequences and 1 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, Coen and Tank, 2012], may function as a “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, Natschlagler and Markram, 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, Natschlagler and Markram, 2002; Jaeger, 2003; Jaeger and Haass, 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, Crisanti and Sommers, 1988] produced a variety of regular non-chaotic outputs that match 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, Schuchhardt and Schuster, 1992; Bertschinger and Natschlagler, 2004; Rajan, Abbott and Sompolinsky, 2010; Rajan, Abbott and Sompolinsky, 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 readout weights. The approach we developed here, called Partial In-Network Training or PINning, avoids these issues, while producing sequences resembling data (Sections 1 and 2). 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 random architectures until we find the point that best matches the relevant data (a similar point 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 posterior parietal cortex (PPC) of mice performing a working memory-based decision making task in a virtual reality environment [Harvey, Coen and Tank, 2012]. It is 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., working memory and decision making to evidence accumulation and navigation [Shadlen and Newsome, 2001; Gold and Shadlen, 2007; Freedman and Assad, 2011; Snyder, Batista and

Anderson, 1997; Anderson and Cui, 2009; Bisley and Goldberg, 2003; McNaughton, 1994; Nitz, 2006; Whitlock, Sutherland, Witter, Moser and Moser, 2008; Carlton and Taube, 2009; Hanks et al, 2015]. We also computed a measure of stereotypy from PPC data (*bVar*, Figure 1C–G) and found it to be much lower than if the PPC were to generate sequences based on 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, Coen and Tank, 2012], e.g., choice-specific neural sequences and delay-period memory of cue identity.

We analyzed the structural features in the synaptic connectivity matrix 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 sea of relatively weak synapses in the cortex [Song et al, 2005], most recently, from the primary visual cortex [Cossell, Iacaruso, Muir, Houlton, Sader, Ko, Hofer and Mrsc-Flogel, 2015]. Synaptic distributions measured in slices have been shown to have long tails [Song and Nelson, 2005], and the experimental result in [Cossell, Iacaruso, Muir, Houlton, Sader, Ko, Hofer and Mrsc-Flogel, 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 exhibits some structural noise sensitivity (Figure S6E), and this is not completely removed by training in the presence of noise. It is possible that dynamic mechanisms of ongoing plasticity 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 (Figure 3). 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 [Yishai, Bar-Or and Sompolinsky, 1995; Zhang, 1996], but had similarities to the models developed in [Fiete, Senn, Wang and Hahnloser, 2010; Hopfield, 2015] (explored in Figure S5). In particular, the model in [Fiete, Senn, Wang and Hahnloser, 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 [Fiete, Senn, Wang and Hahnloser, 2010] and our approach lies in the circuitry underlying sequences – their learning rule results in synaptic chain-like connectivity.

We suggest an alternative hypothesis that sequences might be a more general and effective dynamical form of working memory (Sections 4 and 5), making the prediction that sequences may be increasingly observed in many working memory-based tasks (also suggested in [Harvey, Coen and Tank, 2012]). This contrasts with previous models that relied on fixed point attractors to retain information and exhibited sustained activity [Amit,

1992; Amit, 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 (Section 6) by extending the sparse PINning approach developed in Section 4 and 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 [Yishai, Bar-Or and Sompolinsky, 1995; Zhang, 1996], or a synfire chain [Hertz and Prugel-Bennett, 1996; Levy et al, 2001; Hermann, Hertz and Prugel-Bennet, 1995; Fiete, Senn, Wang and Hahnloser, 2010]. In contrast, the models built and described in this paper are constructed without bias or assumptions. If a moving bump architecture [Yishai, Bar-Or and Sompolinsky, 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 Section 3, 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

(see also Supplemental Experimental Procedures S7)

1. Network elements

We consider 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 build networks of the same size as the experimental dataset we are trying to model, however the results obtained remain applicable to larger networks, Figure S3). For mathematical details of the network, see Experimental Procedures S7.1.

2. Design of External Inputs

To represent sensory (visual and proprioceptive) stimuli innervating the PPC, the external inputs to the neurons in the network are made from filtered and spatially delocalized white noise that is frozen (repeated from trial to trial, Figure 1A). For motivating this choice and for details of the frozen noise input, see Experimental Procedures S7.2. In addition, we also test the resilience of the memory networks we built (Figure 6E) to injected stochastic noise (see Experimental Procedures S7.2). 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 drops to 0 (Figure 6E, see also Figure S6).

3. Extracting Target Functions From Calcium Imaging Data

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 the PPC data [Harvey, Coen and Tank, 2012] into firing rates using two complementary methods. See Experimental Procedures S7.3 for details of both deconvolution methods (see also Figure S1).

4. Synaptic Modification Rule For PINning

During PINning, the inputs of individual network neurons are 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 , undergoes modification at a rate proportional to the error, the presynaptic firing rate of each neuron, and a $pN \times pN$ matrix, P that tracks the rate fluctuations across the network. For further details of the learning rule, see Experimental Procedures S7.4.

5. Dimensionality Of Network Activity (Q_{eff})

We define the effective dimensionality of the activity, Q_{eff} , as the number of principal components that capture 95% of the variance in the activity (Figure 2F), computed as in Experimental Procedures S7.5.

6. Stereotypy Of Sequence ($bVar$), %

$bVar$ quantifies the variance of the data or the network output explained by the translation of an activity profile with an invariant shape. See Experimental Procedures S7.6 for details.

7. Percent Variance Of Data Explained By Model ($pVar$), %

We quantify 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 is computed are in Experimental Procedures S7.7.

8. Magnitude Of Synaptic Change

In Figures 2E and 5G, we compute the magnitude of the synaptic change required to implement a single PPC-like sequence, an idealized sequence and three memory tasks. In combination with the fraction of plastic synapses in the PINned network, p , this metric characterizes the amount of structure that is imposed in an initially random network to produce the desired temporally structured dynamics. See also Experimental Procedures S7.8.

9. Selectivity Index For Memory Task

In Figure 5F, a Selectivity Index is computed (similar to Figure 4 in [Harvey, Coen and Tank, 2012]) to assess the performance of different PINned networks at maintaining 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 (~10s for Figure 5 (after [Harvey, Coen and Tank, 2012]) and ~7s for Figure 6) during preferred trials, and the mean activities of preferred neurons during opposite trials. This is computed as described in Experimental Procedures S7.9.

10. 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, N_{Active}/N . This is computed as in Experimental Procedures S7.10.

11. Analyzing the Structure of PINned Synaptic Connectivity Matrices

This is pertinent to Section 3 (Figures 3 and 4), in which we quantify how the synaptic strength varies with “distance” between pairs of network neurons in connectivity space, $i - j$. These are computed as described in Experimental Procedures S7.11. See Figure 3B for the sparsely PINned matrix, $J_{PINned, 8\%}$, relative to the randomly initialized matrix, J_{Rand} , and Figure 3E, for the fully PINned matrix, $J_{PINned, 100\%}$. The same analysis is also used for Figure S5.

12. Cross-validation Analysis

This pertains to quantifying how well PINning-based models do and is described in Experimental Procedures S7.12.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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, 2014; 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 (DWT: R01-MH083686; RC1-NS068148; 1U01NS090541-01 and CDH: R01-MH107620; R01-NS089521), a grant from the Simons Collaboration on the Global Brain (DWT), a Young Investigator Award from NARSAD/Brain and Behavior Foundation (KR), a fellowship from the Helen Hay Whitney Foundation (CDH), the New York Stem Cell Foundation (CDH), and a Burroughs Wellcome Fund Career Award at the Scientific Interface (CDH). CDH is a New York Stem Cell Foundation-Robertson Investigator and a Searle Scholar.

REFERENCES

- Amit, DJ. Modeling Brain Function: The World of Attractor Neural Networks. Cambridge University Press; 1992.
- Amit DJ. The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behav Brain Sci.* 1995; 18:617.
- Amit DJ, Brunel N. Learning internal representations in an attractor neural network with analogue neurons. *Network.* 1995; 6:359–388.
- Andersen RA, Cui H. Intention, action planning, and decision making in parietal-frontal circuits. *Neuron.* 2009; 63:568–583. [PubMed: 19755101]
- Andersen RA, Burdick JW, Musallam S, Pesaran B, Cham JG. Cognitive neural prosthetics. *Trends Cogn Sci.* 2004; 8:486–493. [PubMed: 15491902]
- Barak O, Sussillo D, Romo R, Tsodyks M, Abbott LF. From fixed points to chaos: three models of delayed discrimination. *Progress in neurobiology.* 2013; 103:214–222.
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature.* 2005; 437:1158–1161. [PubMed: 16237445]

- Ben-Yishai R, Bar-Or RL, Sompolinsky H. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci*. 1995; 92:3844–3848. [PubMed: 7731993]
- Bertschinger N, Natschläger TN. Real-time computation at the edge of chaos in recurrent neural networks. *Neural Computation*. 2004; 16(7):1413–1436. [PubMed: 15165396]
- Bisley JW, Goldberg ME. Neuronal activity in the lateral intraparietal area and spatial attention. *Science*. 2003; 299:81–86. [PubMed: 12511644]
- Bisley JW, Goldberg ME. Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci*. 2010; 33:1–21. [PubMed: 20192813]
- Brunton BW, Botvinick MM, Brody CD. Rats and Humans can Optimally Accumulate Evidence for Decision-making. *Science*. 2013; 340:95–98. [PubMed: 23559254]
- Buonomano DV. Timing of neural responses in cortical organotypic slices. *Proc Natl Acad Sci USA*. 2003; 100:4897–4902. [PubMed: 12668762]
- Buonomano DV, Merzenich MM. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*. 1995; 267:1028–1030. [PubMed: 7863330]
- Buonomano DV. A learning rule for the emergence of stable dynamics and timing in recurrent networks. *J Neurophysiol*. 2005; 94(4):2275–2283. [PubMed: 16160088]
- Calton JL, Taube JS. 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*. 2009; 91:186–196. [PubMed: 18929674]
- Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci*. 2010; 13(3):369–378. [PubMed: 20173745]
- Cossell L, Iacaruso MF, Muir DR, Houlton R, Sader EN, Ko H, Hofer SB, Mrsic-Flogel TD. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*. 2015; 518:399–405. [PubMed: 25652823]
- Crowe DA, Averbach BB, Chafee MV. Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *J Neurosci*. 2010; 30:11640–11653. [PubMed: 20810885]
- Davidson TJ, Kloosterman F, Wilson MA. Hippocampal replay of extended experience. *Neuron*. 2009; 63:497–507. [PubMed: 19709631]
- Fiete IR, Senn W, Wang C, Hahnloser RHR. Spike time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron*. 2010; 65(4):563–576. [PubMed: 20188660]
- Fisher D, Olasagasti I, Tank DW, Aksay ER, Goldman MS. A modeling framework for deriving the structural and functional architecture of a short-term memory microcircuit. *Neuron*. 2013; 79(5):987–1000. [PubMed: 24012010]
- Freedman DJ, Assad JA. A proposed common neural mechanism for categorization and perceptual decisions. *Nat Neurosci*. 2011; 14:143–146. [PubMed: 21270782]
- Fujisawa S, Amarasingham A, Harrison MT, Buzsaki G. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nature Neurosci*. 2008; 11:823–833. [PubMed: 18516033]
- Gold JI, Shadlen MN. The neural basis of decision making. *Annu Rev Neurosci*. 2007; 30:535–574. [PubMed: 17600525]
- Goldman MS. Memory without feedback in a neural network. *Neuron*. 2009; 61:621–634. [PubMed: 19249281]
- Hahnloser RHR, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*. 2002; 419:65–70. [PubMed: 12214232]
- Hanks TD, Kopec CD, Brunton BW, Duan CA, Erlich JC, Brody CD. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*. 2015; 520:220–223. [PubMed: 25600270]
- Hansel D, Mato G. Existence and stability of persistent states in large neuronal networks. *Phys Rev Lett*. 2001; 10:4175–4178. [PubMed: 11328124]

- Harvey CD, Coen P, Tank DW. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*. 2012; 14(7392):62–68. 484. [PubMed: 22419153]
- Haykin, S. *Adaptive Filter Theory*. Upper Saddle River, NJ: Prentice Hall; 2002.
- Hermann M, Hertz J, Prugel-Bennet A. Analysis of synfire chains. *Network: Computation Neural Systems*. 1995; 6:403–414.
- Hertz J, Prügel-Bennett A. Learning short synfire chains by self-organization. *Network*. 1996; 7(2): 357–363. [PubMed: 16754397]
- Hopfield JJ, Tank DW. “Neural” computation of decisions in optimization problems. *Biol Cyber*. 1985; 55:141–146.
- Hopfield JJ. Understanding emergent dynamics: Using a collective activity coordinate of a neural network to recognize time-varying patterns. *Neural Comp*. 2015; 27(10):2011–2038.
- Ikegaya Y, Aaron G, Cossart R, Aronov D, Lampl I, Ferster D, Yuste R. Synfire chains and cortical songs: temporal modules of cortical activity. *Science*. 2004; 304:559–564. [PubMed: 15105494]
- Jaeger, H. Adaptive nonlinear system identification with echo state networks. In: Becker, S.; Thrun, S.; Obermayer, K., editors. *Advances in Neural Information Processing Systems*. Vol. 15. Cambridge MA: MIT Press; 2003. p. 593-600.
- Jaeger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science*. 2004; 304:78–80. [PubMed: 15064413]
- Jin DZ, Fuji N, Graybiel AM. Neural representation of time in cortico-basal ganglia circuits. *Proc Natl Acad Sci USA*. 2009; 106:19156–19561. [PubMed: 19850874]
- Kleinfeld, D.; Sompolinsky, H. Associative network models for central pattern generators. In: Koch; Segev, editors. *Methods in Neuronal Modeling*. Cambridge, MA: MIT Press; 1989. p. 195-246.
- Kozhevnikov AA, Fee MS. Singing-related activity of identified HVC neurons in the zebra finch. *J Neurophysiol*. 2007; 97:4271–4283. [PubMed: 17182906]
- Laje R, Buonomano DV. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat. Neurosci*. 2013; 16:925–933. [PubMed: 23708144]
- Levy N, Horn D, Meilijson I, Ruppin E. Distributed synchrony in a cell assembly of spiking neurons. *Neural Netw*. 2001; 14(6–7):815–824. [PubMed: 11665773]
- Louie K, Wilson MA. Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*. 2001; 29:145–156. [PubMed: 11182087]
- Luczak A, Bartho P, Marguet SL, Buzsaki G, Harris KD. Sequential structure of neocortical spontaneous activity in vivo. *Proc Nat Acad Sci USA*. 2007; 104(1):347–352. [PubMed: 17185420]
- Maass W, Joshi P, Sontag ED. Computational aspects of feedback in neural circuits. *PLoS Comput Biol*. 2007; 3:e165. [PubMed: 17238280]
- Maass W, Natschlagler T, Markram H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput*. 2002; 14:2531–2560. [PubMed: 12433288]
- McNaughton BL, Mizumori SJ, Barnes CA, Leonard BJ, Marquis M, Green EJ. Cortical representation of motion during unrestrained spatial navigation in the rat. *Cereb Cortex*. 1994; 4:27–39. [PubMed: 8180489]
- Mauk MD, Buonomano DV. The neural basis of temporal processing. *Annu Rev Neurosci*. 2004; 27:307–340. [PubMed: 15217335]
- Molgedey L, Schuchhardt J, Schuster HG. Suppressing chaos in neural networks by noise. *Phys Rev Lett*. 1992; 69:3717–3719. [PubMed: 10046895]
- Nadasdy Z, Hirase H, Czurko A, Csicsvari J, Buzsaki G. Replay and time compression of recurring spike sequences in the hippocampus. *J Neurosci*. 1999; 19:9497–9507. [PubMed: 10531452]
- Nitz DA. Tracking route progression in the posterior parietal cortex. *Neuron*. 2006; 49:747–756. [PubMed: 16504949]
- Pastalkova E, Itskov V, Amarasingham A, Buzsaki G. Internally generated cell assembly sequences in the rat hippocampus. *Science*. 2008; 321:1322–1327. [PubMed: 18772431]
- Pearlmutter B. Learning state space trajectories in recurrent neural networks. *Neural Comput*. 1989; 1:263–269.

- Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, Reardon T, Mu Y, Lacefield C, Yang W, Ahrens M, Bruno R, Jessell TM, Peterka DS, Yuste R, Paninski L. Simultaneous denoising, deconvolution, and demixing of Calcium imaging data. *Neuron*. 2016; 89(2):285–299. [PubMed: 26774160]
- Pulvermutter F, Shtyrov Y. Spatiotemporal signatures of large-scale synfire chains for speech processing as revealed by MEG. *Cereb. Cortex*. 2009; 19:79–88. [PubMed: 18460449]
- Rajan, K.; Abbott, LF.; Sompolinsky, H. Inferring Stimulus Selectivity from the Spatial Structure of Neural Network Dynamics. In: Lafferty; Williams; Shawne-Taylor; Zemel; Culotta, editors. *Advances in Neural Information Processing Systems*. Vol. 23. 2011.
- Rajan K, Abbott LF, Sompolinsky H. Stimulus-dependent Suppression of Chaos in Recurrent Neural Networks. *Phys Rev E*. 2010; 82:01193.
- Schwartz AB, Moran DW. Motor cortical activity during drawing movements: population representation during lemniscate tracing. *J Neurophysiol*. 1999; 82:2705–2718. [PubMed: 10561439]
- Seidemann E, Meilijson I, Abeles M, Bergman H, Vaadia E. 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*. 1996; 16:752–768. [PubMed: 8551358]
- Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol*. 2001; 86:1916–1936. [PubMed: 11600651]
- Snyder LH, Batista AP, Andersen RA. Coding of intention in the posterior parietal cortex. *Nature*. 1997; 386:167–170. [PubMed: 9062187]
- Sompolinsky H, Crisanti A, Sommers HJ. Chaos in Random Neural Networks. *Phys Rev Lett*. 1988; 61:259–262. [PubMed: 10039285]
- Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLoS Biol*. 2005; 3(10):e350.
- Sussillo D, Abbott LF. Generating coherent patterns of activity from chaotic neural networks. *Neuron*. 2009; 63:544–557. [PubMed: 19709635]
- Sussillo D. Neural circuits as computational dynamical systems. *Curr Op Neurobiol*. 2014; 25:156–163. [PubMed: 24509098]
- Tang A, Jackson D, Hobbs J, Chen W, Smith JL, Patel H, Prieto A, Petrusca D, Grivich MI, Sher A, Hottowy P, Dabrowski W, Litke AM, Beggs JM. A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. *J Neurosci*. 2008; 28:505–518. [PubMed: 18184793]
- 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. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat Methods*. 2009; 6(12):875–881. [PubMed: 19898485]
- Vogelstein J, Packer A, Machado T, Sippy T, Babadi B, Yuste R, Paninski L. Fast non-negative deconvolution for spike train inference from calcium imaging. *J. Neurophysiol*. 2010; 104(6): 3691–3704. [PubMed: 20554834]
- White B, Abbott LF, Fiser J. Suppression of cortical neural variability is stimulus-and state-dependent. *J. Neurophysiol*. 2012; 108:2383–2392. [PubMed: 22896720]
- Whitlock JR, Sutherland RJ, Witter MP, Moser MB, Moser EI. Navigating from hippocampus to parietal cortex. *Proc Natl Acad Sci USA*. 2008; 105:14755–14762. [PubMed: 18812502]
- Williams RJ, Zipser D. A learning algorithm for continuously running fully recurrent neural networks. *Neural Comput*. 1989; 1:270–280.
- Zhang K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A Theory. *Journal of Neuroscience*. 1996; 16:2112–2126. [PubMed: 8604055]

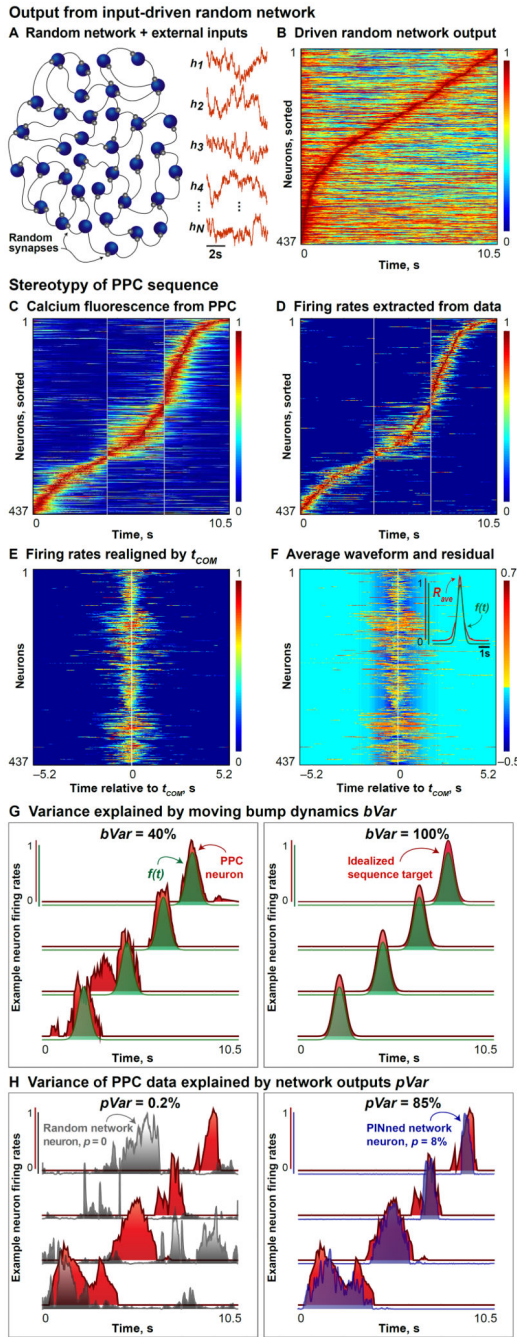


Figure 1. Stereotypy of PPC Sequences and Random Network Output

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 are at least as big as the size of the data under consideration [Harvey, Coen & Tank, 2012], and are 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) are normalized by the maximum per neuron and sorted by their t_{COM} , the activity appears time ordered. However, this sequence contains large amounts of background activity, compared to the PPC data ($bVar = 12\%$, $pVar = 0.2\%$).

C. Calcium fluorescence (i.e., normalized F/F) data collected from 437 trial-averaged PPC neurons during an 10.5s-long 2AFC experiment from both left and right correct-choice outcomes, pooled from 6 mice.

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) are realigned by their time of center-of-mass (abbreviated t_{COM}) and plotted here.

F. **Inset:** A “typical” waveform (R_{ave} in red) obtained by averaging the realigned rates from (C) over neurons, and a Gaussian curve with mean = 0 and variance = 0.3 ($f(t)$, green), that best fits 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 4 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 from 4 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 4 example neurons here. Left panel shows the normalized rates from 4 example PPC neurons (red) picked from (D) and 4 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, $pVar = 0.2\%$. Right panel shows the same PPC neurons as in the left panel in red, along with 4 model neuron outputs from a PINned network with $p = 12\%$ plastic synapses. For this example, $pVar = 85\%$.

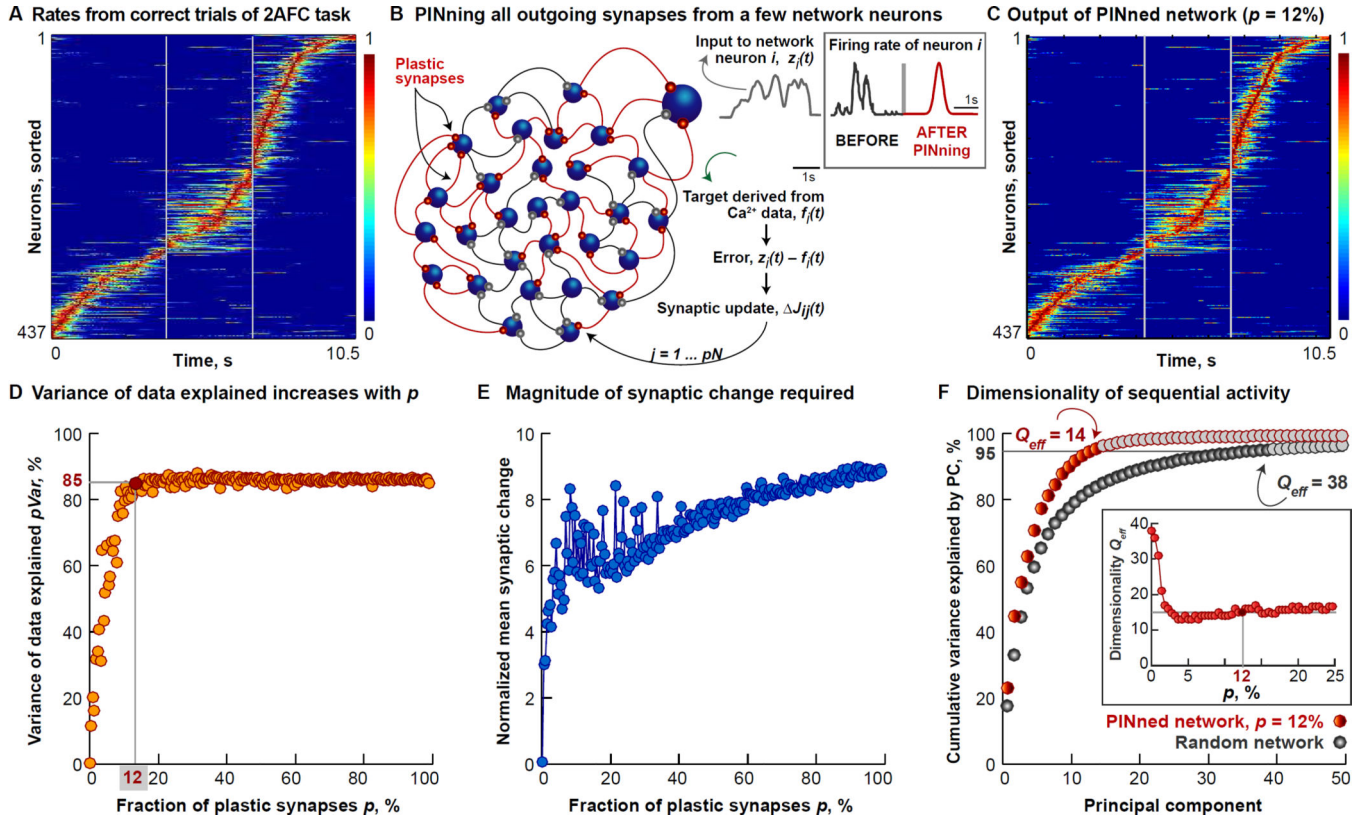


Figure 2. Partial In-Network Training (PINning) Matches PPC-like sequences

A. 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) are 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 P). Here, the target functions for PINning are extracted from the rates shown in (A).

C. Normalized activity from the network with $p = 12\%$ plastic synapses (red circle in (D)). This activity is temporally constrained, has a relatively small amount of extra-sequential activity ($bVar = 40\%$ for (C) and (A)), and shows 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. $pVar$ increases from 0 for random networks with no plasticity (i.e., $p = 0$) to $pVar = 50\%$ for $p = 8\%$ (not shown), and asymptotes at $pVar = 85\%$ as $p = 12\%$ (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 grows 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 change more in small- p networks but the total change across the synaptic matrix is 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 accounting for $>95\%$ for the random untrained network. **Inset** shows Q_{eff} (depicted in red circles) of the manifold of the overall network activity, as p increases.

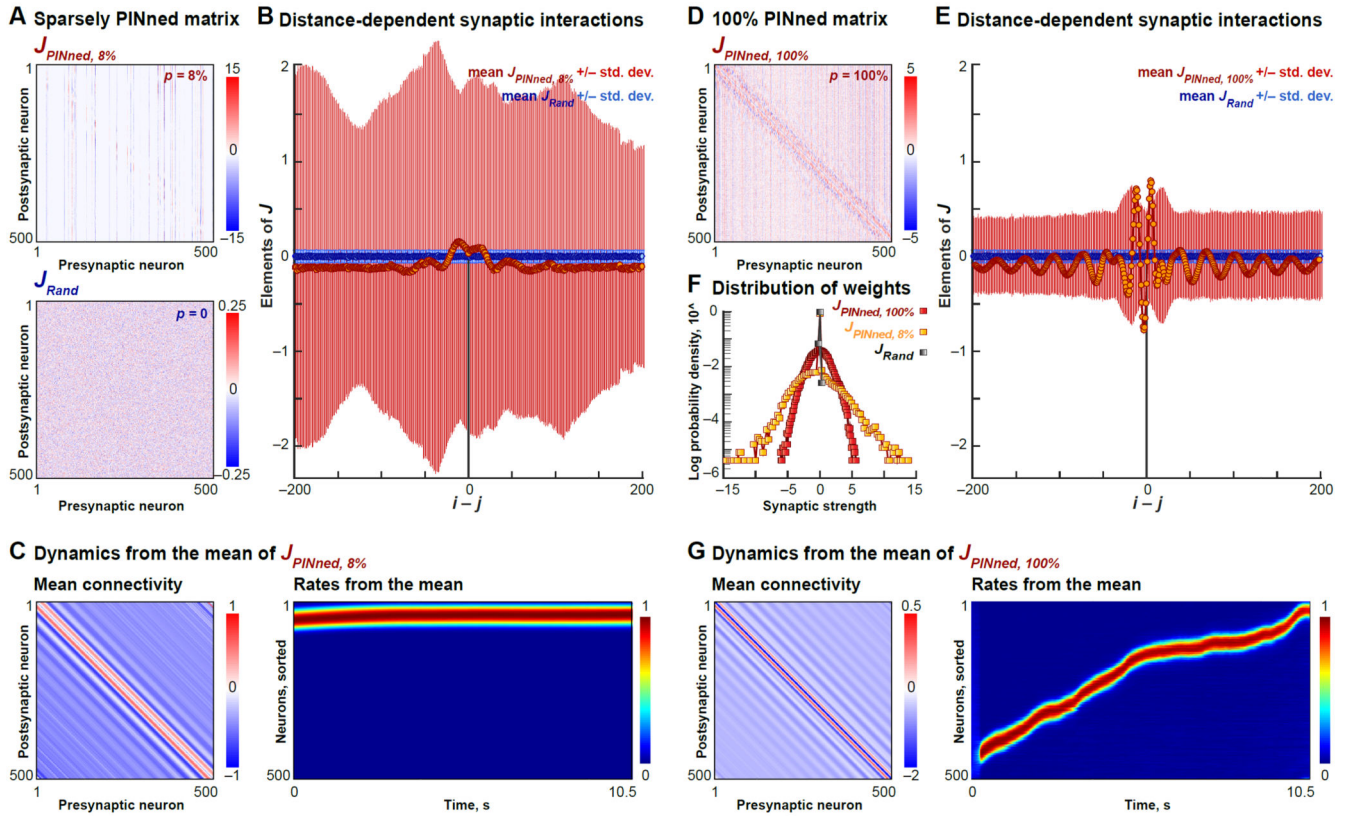


Figure 3. Properties of PINned Connectivity Matrices

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 $J_{PINned, 8\%}$, and that of the randomly initialized network with $p = 0$, denoted by J_{Rand} , are shown here. Colorbars indicate, in pseudocolor, the magnitudes of synaptic strengths after PINning.

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. 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 a distance $i - j$ away from the primary sequential neurons.

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

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

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

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

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

Formation and propagation mechanisms for an idealized sequence

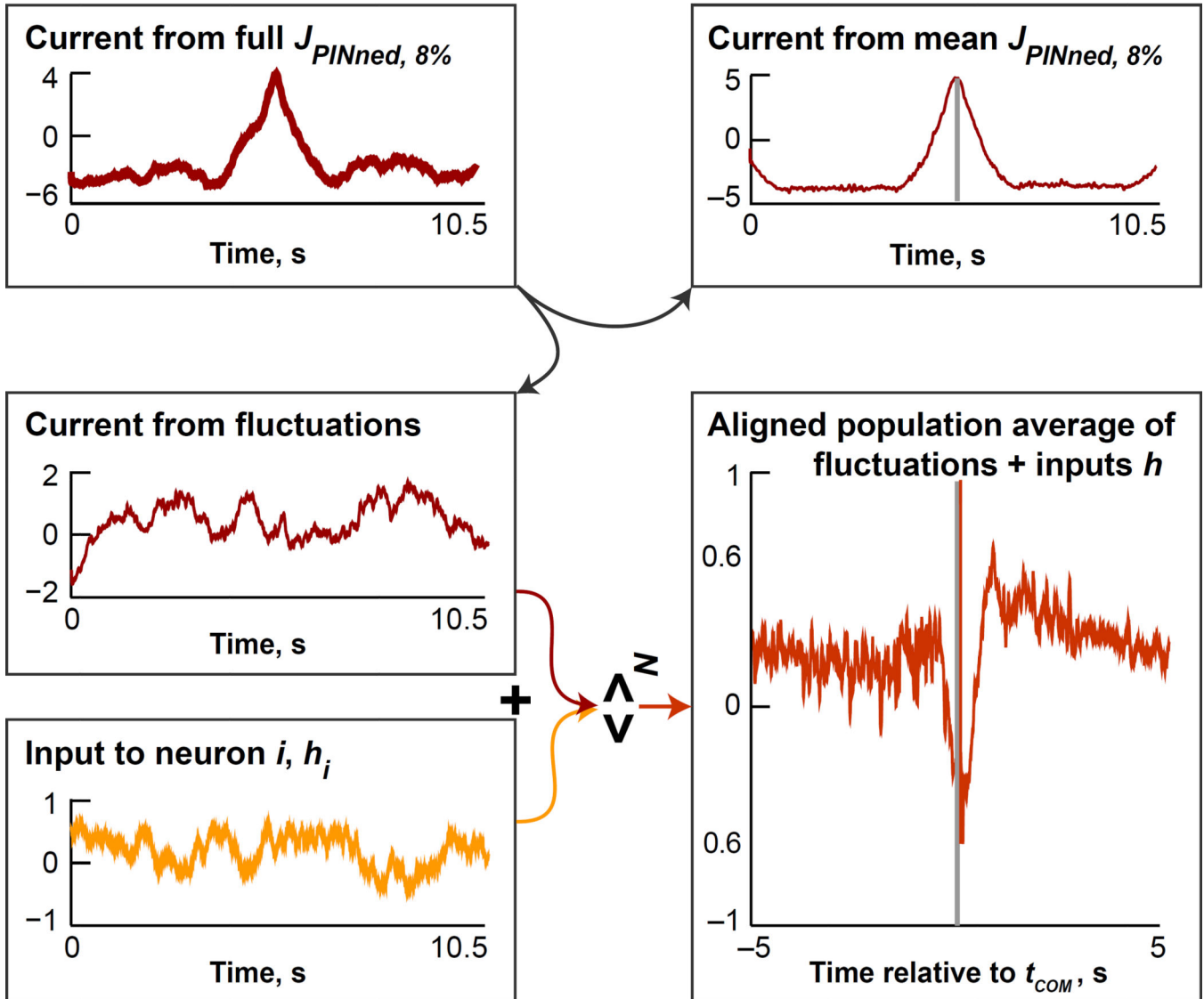


Figure 4. Mechanism for Formation and propagation of sequence

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 $J_{PINned, 8\%}$ cause 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) causes the bump to move. We demonstrate 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 are 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 the bottom panel on the right, and reveals the asymmetry responsible for the movement of the bump across the network.

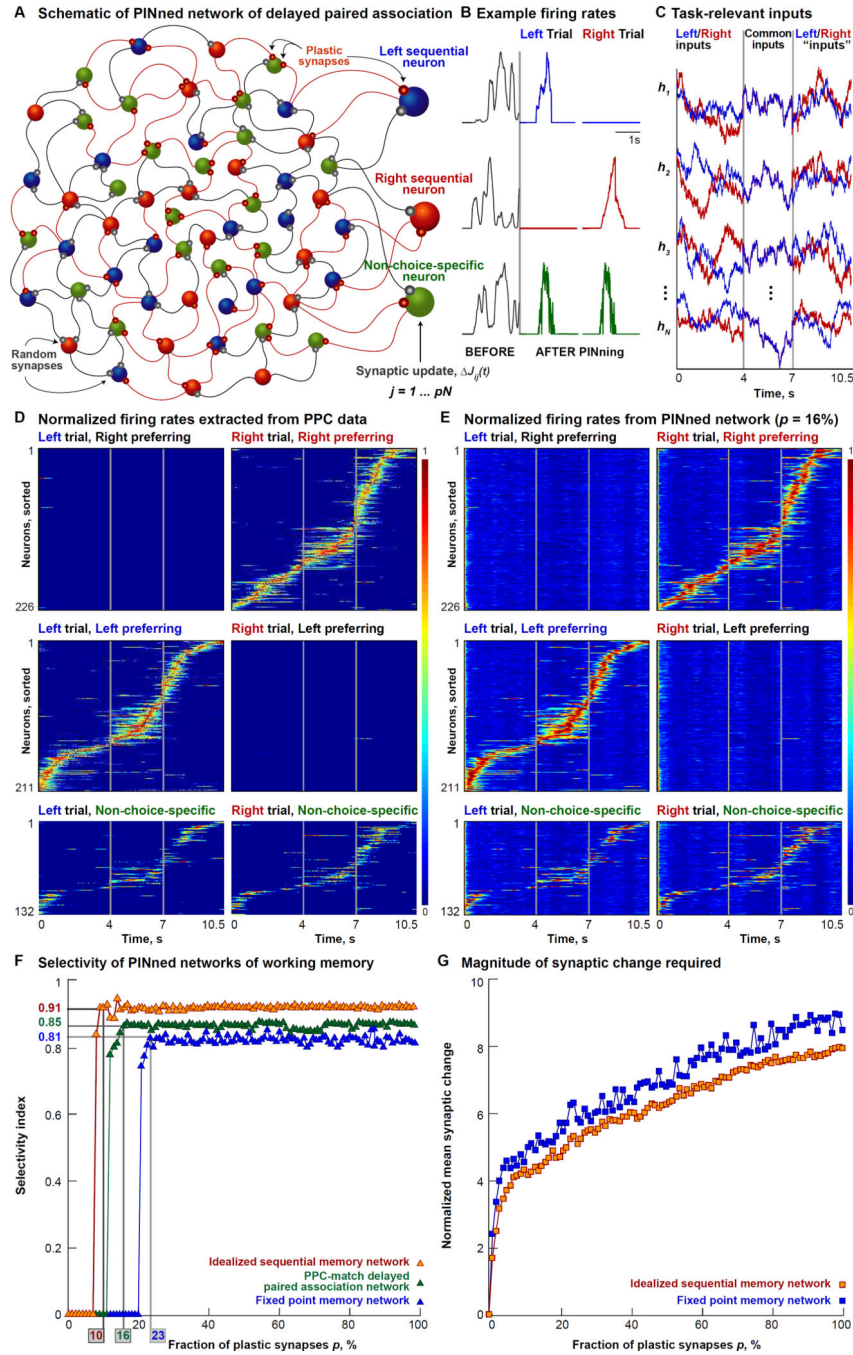


Figure 5. Delayed Paired Association In PINned Networks Of Working Memory

A. Schematic of a sparsely PINned network implementing delayed paired association through a 2AFC task [Harvey, Coen and Tank, 2012]. The targets for PINning are 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 is 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 $i = 1, 2, 3, \dots, N$) are shown here. Each neuron gets a different irregular, spatially delocalized, filtered white-noise input, but receives 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 are extracted by deconvolution 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, Coen and Tank, 2012]. These firing rates are used to extract the target functions for PINning (schematized in (A)). Traces are normalized to the peak of the mean firing rate of each neuron on preferred trials and sorted by the time of center-of-mass (t_{COM}). Vertical gray lines indicate the time points corresponding to the different epochs of the task – the cue period ending at 4s, the delay period ending at 7s, and the turn period concluding at the end of the trial, at 10.5s.

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, showing a match with the data (D). For this network, $pVar = 85\%$.

F. Selectivity index, shown here, is computed as the ratio of the difference and the sum of the mean activities of preferred neurons at the 10s time-point during preferred trials and the mean activities of preferred neurons during opposite trials. Task performance of 3 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 exhibits long-duration population dynamics and memory activity through idealized sequences (orange triangles) has selectivity = 0.91, when $p = 10\%$ of its synapses are plastic; the PPC-like network needs $p = 16\%$ plastic synapses for selectivity = 0.85, and the fixed point network needs $p = 23\%$ of its synapses to be plastic for selectivity = 0.81.

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

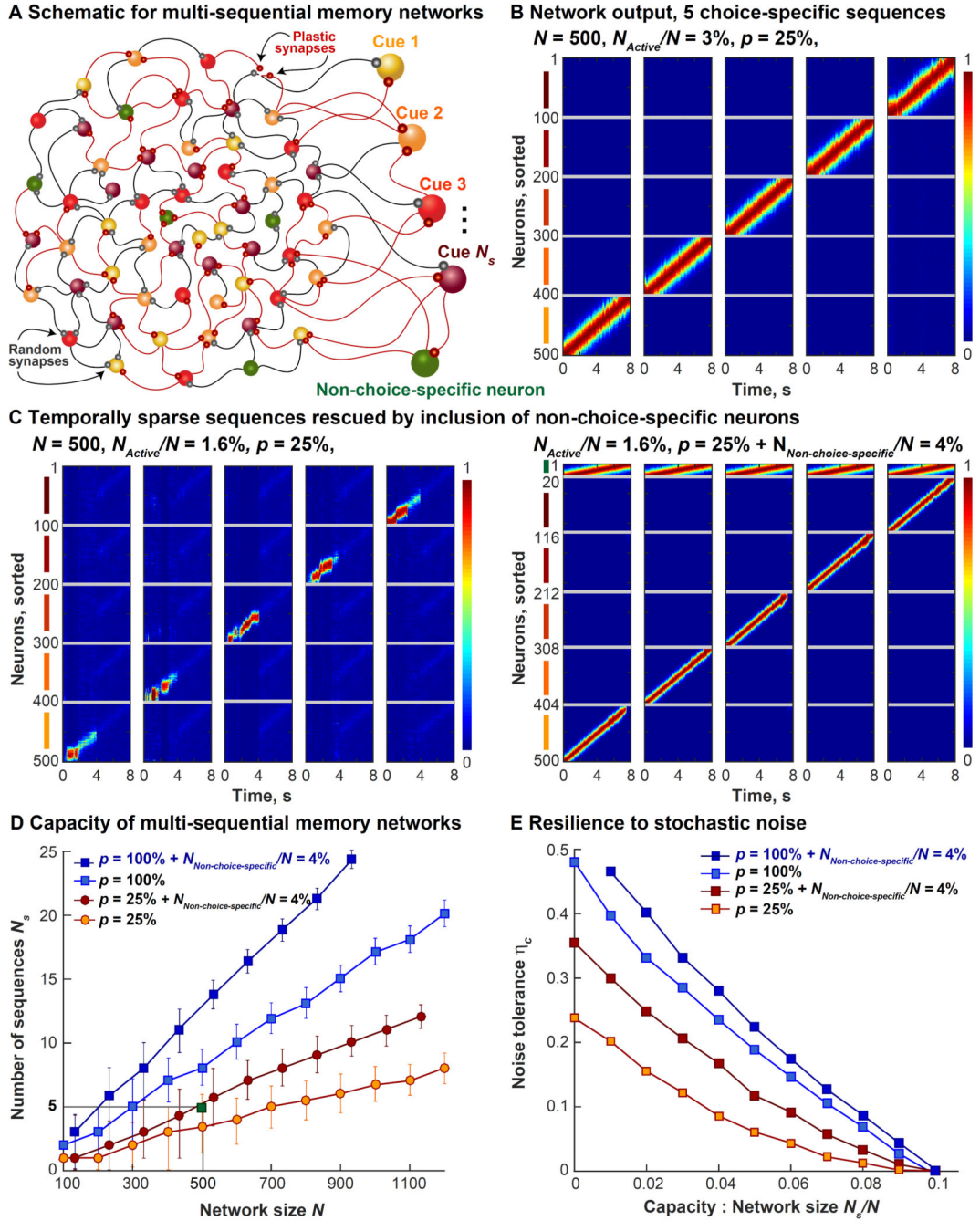


Figure 6. Capacity of Multi-Sequential Memory Networks

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

executed through 5 choice-specific sequences; during the delay period (4–8s), neurons fire in a sequence only on trials of the same type as their cue preference.

C. Left panel shows the same network as (B) failing to perform the task correctly when the widths of the targets was halved ($N_{Active}/N = 1.6\%$), sparsifying the sequences in time. Memory of the cue identity is not maintained during the delay period. Right panel shows the result of including a small number of non-choice-specific neurons ($N_{Non-choice-specific}/N = 4\%$) that fire in the same order in trials of all 5 types. Including non-choice-specific neurons restores 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 $p = 25\% + N_{Non-choice-specific}/N = 4\%$, light blue squares for $p = 100\%$ and dark blue squares for $p = 100\% + N_{Non-choice-specific}/N = 4\%$ neurons. PINned networks containing additional non-choice-specific neurons have temporally sparse target functions with $N_{Active}/N = 1.6\%$, the rest have $N_{Active}/N = 3\%$. Error bars are calculated over 5 different random instantiations each and decrease under the following conditions: as network size is increased, as fraction of plastic synapses p is 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 is 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 , decreases as p is lowered and as capacity increases, although it falls slower with the inclusion of non-choice-specific neurons. Only mean values shown here for clarity.