2011

Cell Assembly Sequences Arising from Spike Threshold Adaptation Keep Track of Time in the Hippocampus

Vladimir Itskov
University of Nebraska-Lincoln, vladimir.itskov@psu.edu

Carina Curto
University of Nebraska - Lincoln, ccurto2@math.unl.edu

Eva Pastalkova
Howard Hughes Medical Institute

György Buzsáki
Rutgers University

Follow this and additional works at: https://digitalcommons.unl.edu/mathfacpub
Behavioral/Systems/Cognitive

Cell Assembly Sequences Arising from Spike Threshold Adaptation Keep Track of Time in the Hippocampus

Vladimir Itskov, Carina Curto, Eva Pastalkova, and György Buzsáki

Hippocampal neurons can display reliable and long-lasting sequences of transient firing patterns, even in the absence of changing external stimuli. We suggest that time-keeping is an important function of these sequences, and propose a network mechanism for their generation. We show that sequences of neuronal assemblies recorded from rat hippocampal CA1 pyramidal cells can reliably predict elapsed time (15–20 s) during wheel running with a precision of 0.5 s. In addition, we demonstrate the generation of multiple reliable, long-lasting sequences in a recurrent network model. These sequences are generated in the presence of noisy, unstructured inputs to the network, mimicking stationary sensory input. Identical initial conditions generate similar sequences, whereas different initial conditions give rise to distinct sequences. The key ingredients responsible for sequence generation in the model are threshold-adaptation and a Mexican-hat-like pattern of connectivity among pyramidal cells. This pattern may arise from recurrent systems such as the hippocampal CA3 region or the entorhinal cortex. We hypothesize that mechanisms that evolved for spatial navigation also support tracking of elapsed time in behaviorally relevant contexts.

Introduction

Tracking time is of fundamental importance in a wide range of brain operations, including sensory perception and motor actions, learning, memory, planning, decision-making and language (Gibbon et al., 1997; Buonomano and Karmarkar, 2002; Ivry and Spencer, 2004; Mauk and Buonomano, 2004; Buhusi and Meck, 2005). Despite the central importance of temporal processing, its underlying neural mechanisms remain unknown. At the systems level, two competing ideas have been put forward: timing is generated by a central mechanism and distributed to various brain regions (Church, 1984), or each subsystem produces its own timing (Mauk and Buonomano, 2004). With regard to timing and duration, a distinction is made between subsecond (perceptual-motor) and suprasecond (cognitively mediated) scales (Michon, 1985; Lewis et al., 2003). At the level of mechanisms, two models are typically distinguished; clocks and ramping time keepers, with neuronal substrates in the cerebellum, basal ganglia, prefrontal, motor, and parietal cortical regions (cf. Mauk and Buonomano, 2004; Buhusi et al., 2005). The hippocampus has also been implicated in timing (Clark and Isaacson, 1965; Thompson and Krupa, 1994; Young and McNaughton, 2000), although the mechanism has remained elusive.

We report here on a novel form of time-tracking mechanism, which is manifested by evolving transiently active cell assemblies and is accurate for periods of tens of seconds. First, we examine the ability of evolving neuronal sequences to predict elapsed time in a memory task. Second, we propose a simple network model with Mexican-hat-type connectivity and adaptation of the membrane potential thresholds for action potential generation that is similar to what has been observed in rodent hippocampus (Henze and Buzsáki, 2001) and in fish (Chacron et al., 2007). The threshold-adaptation model reproduces the key properties of the observed sequences, suggesting that time-keeping in the hippocampus may arise from the same cellular and network mechanisms that support spatial navigation.

Materials and Methods

The experimental data used in this paper were adopted from Pastalkova et al. (2008), where all relevant experimental methods and protocols are described. The animals were all male rats.

Time prediction from experimental data. Two versions of time prediction models were fit from experimental data: a “rate-only” model and a “phase-only” model. These models used firing rates or theta phases of spikes to fit a probability distribution for population spiking activity in 0.5 s time windows. Maximum-likelihood estimation was then used on single trials to predict the most likely time reflected by the population activity at each time bin. In all cases, the models were fit from complementary trials, never including the trial on which time was subsequently inferred. The time estimates corresponding to each time bin were obtained independently, so that only knowledge of the current population activity was needed to estimate elapsed time. See supplemental Text
The error of time estimate is shown in Figure 1D. The reliability of individual trials in the control, home cage condition is displayed in Figure 1E. The reliability of the model predictions within the home cage condition is displayed in Figure 1F. The model dynamics in layer 2 were governed by the same Equations 1 and 2 in the previous layer, with two differences. First, the connectivity matrix $J$ in layer 2 represented an additional ingredient of the model is the activity-dependent adaptation of the “spike” thresholds of individual neurons, represented by the dynamic variables $\theta(t) = (\theta_1(t), \ldots, \theta_N(t))$. The model can thus be described by a system of 2N equations:

$$\tau_x \dot{x}_i = -x_i + \left( \sum_{j=1}^{N} J_{ij} x_j + I_i - h_i \right)$$

(1)

$$\tau_h \dot{h}_i = -h_i + c x_i$$

(2)

where $\tau_x = 30 \text{ ms}$ and $\tau_h = 2 \text{ s}$ are membrane and threshold-adaptation time constants, respectively. $J$ is the matrix of synaptic weights for the recurrent network, and $I$ is a (time-dependent) vector of external inputs. The constant $c = 0.5$ controls the strength of the activity-dependent adaptation, whereas $\tau_h$ determines the time scale with which the thresholds recover in the absence of firing.

The connectivity matrix $J$ is constant in time and can be written as a sum of two components, $J = J^f + J^{inh}$, where $J^f$ has Mexican-hat-type connectivity on a two-dimensional lattice of neurons with periodic (torus) boundary conditions, and $J^{inh}$ is a matrix of heterogeneous weights, sampled randomly and independently from a normal distribution with mean zero. In simulations, we generated three different instances of $J^{inh}$, resulting in three different connectivity matrices $J^f$, $J^{inh}$, and $J$.

In simulations, two types of behavioral conditions were distinguished: “task” and “home cage.” In all cases, initial conditions consisted of the same “bump” of firing rate activity $\ddot{x}(0) = x_0$ in the center of the two-dimensional lattice of neurons, and differed in the adaptation variables $\ddot{h}(0) = h_0$ only. Neurons with adapted thresholds were chosen to lie at the left, bottom, right, or top of the initial bump of activity, resulting in initial conditions $A$ to $D$, respectively. For task trials, the initial conditions were consistent from trial to trial, and the model was driven by temporally and spatially unstructured noise $\ddot{I}(t)$; different instances of noise was thus the only difference between trials of the same initial condition type. In the home cage trials, the initial conditions $A$ to $D$ were randomized across trials, and the model was driven by spatially unstructured noise that had temporal correlations on the order of 125 ms (see supplemental Text, available at www.jneurosci.org as supplemental material, for further information).

Layer 2 simulations. To investigate the ability of a downstream layer to “inherit” the sequence generated by the threshold adaptation model (“layer 1”), we simulated activity in a second layer connected to the first via sparse and random feedforward projections. The dynamics in this layer were governed by the same Equations 1 and 2 in the previous layer, with two differences. First, the connectivity matrix $J$ in layer 2 represented an overall global inhibition (to ensure sparse firing) and had no spatial structure. Second, the input vector $\dot{I}(t)$ had two components: temporally and spatially unstructured noise, and feedforward input derived from the activity in the previous layer via random and sparse connections (10% connection probability).
Reliability measure. A reliability measure was used to compare population vectors from individual trials to the mean. First, we normalized firing rates so that each neuron had the same average firing rate, calculated over all time bins and all trials. Let \( \mathbf{v}(t) \) be the population vector corresponding to the \( t \)th trial at time \( t \). For each fixed time bin \( t_0 \) we computed a “reliability score” \( R_i(t_0) \) for each trial by computing the squared-distance between the population vector \( \mathbf{v}(t_0) \) and the mean vector across trials \( \langle \mathbf{v}(t_0) \rangle \),

\[
R_i(t_0) = \frac{1}{N_t - 1} \sum_{t \neq t_0} \| \mathbf{v}(t_0) - \langle \mathbf{v}(t_0) \rangle \|^2,
\]

where \( N_t \) is the total number of time bins. If a trial is reliable at time \( t_0 \), then \( W_i(t_0) < B_i(t_0) \). The reliability was thus defined to be

\[
R_i(t_0) = \frac{W_i(t_0) + B_i(t_0)}{W_i(t_0) + B_i(t_0)},
\]

with the denominator chosen so that \(-1 < R_i(t_0) < 1\). A reliability of 1 corresponds to the “best case” scenario of \( \mathbf{v}(t_0) = \langle \mathbf{v}(t_0) \rangle \), and -1 corresponds to the “worst case” scenario of \( \mathbf{v}(t_0) = \langle \mathbf{v}(t_0) \rangle \) for all \( t \neq t_0 \), and \( \mathbf{v}(t_0) \neq \langle \mathbf{v}(t_0) \rangle \). If \( \mathbf{v}(t_0) \) is closer to the mean vector \( \langle \mathbf{v}(t_0) \rangle \) than it is to the mean vectors from other time bins \( \langle \mathbf{v}(t_0) \rangle \), then \( R_i(t_0) \) will be positive. For random (and thus completely unreliable) population vectors, \( R_i(t_0) \approx 0 \). The average reliability \( R(t) \) was computed as the trial-average of the time series vectors \( R_i(t) \).

Results

Elapsed time is well estimated by the sequence of cell assembly activation

A reliable pattern of sequential activation of neuronal activity was observed in the CA1 region of hippocampus during the delay period of a memory task (Pastalkova et al., 2008; cf. Gill et al., 2010; Kraus et al., 2010; Macdonald et al., 2010). Three rats were trained to run for \(-20\) s in a running wheel during the delay period before making a choice (left or right) for the next run through a T maze. Action potentials for pyramidal cells were recorded together with local field potentials. The pattern of sequential activation of simultaneously recorded neurons for a given trial type (Fig. 1A,B) was reliable across trials and lasted \(-10\) to \(-20\) s without repeating itself. Therefore, we hypothesized that the population spiking activity of pyramidal neurons in CA1 at any point in time during a trial could be used to infer elapsed time.

To verify this hypothesis, we designed two probabilistic models for inferring elapsed time from instantaneous neural activity: one based on the cells’ firing rates, and the other using phases of spikes with respect to the theta oscillation (see Materials and Methods). Both models were good predictors of elapsed time on single trials (Fig. 1C). The average error of time estimation by the rate model was \(-2\) s in rats 2 and 3 and \(-2\) to \(-3\) s in rat 1 (Fig. 1D). Similar errors were observed using the phase model (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). The accuracy of time estimation from the models increased with the number of cells used in each animal (data not shown).

This observation suggests that by recording from a much larger fraction of hippocampal neurons, the accuracy of time estimation can be improved further. It also suggests that a greater amount of information is available to structures downstream from CA1 than what was available for our method of inference. Since we obtained timing precision on a behaviorally relevant scale, we hypothesize that the brain could use population activity to estimate elapsed time.

To investigate how behavioral relevance might influence timekeeping, we also recorded the activity of CA1 cells during running in a wheel placed adjacent to the home cage of the rat (control condition). The rat could enter the wheel and run at its leisure and was not required to keep track of elapsed time. While the patterns of neuronal activity on individual runs displayed some semblance of sequential activation near the beginning of the wheel run, the overall sequences were not consistent from trial to trial (supplemental Fig. 2, available at www.jneurosci.org as supplemental material); as a result, the time prediction models did not yield any statistically significant prediction of time on the control data (data not shown).

While the reliability \( R(t) \) of sequential activity during the memory task was well above \( 0 \) (the expected value for random, unreliable data; see Materials and Methods) (Fig. 1E), \( R(t) \) was not significantly positive for the control (home cage) data (Fig. 1F). This suggests that the patterns of neuronal activity during wheel runs reflect timing information only in behaviorally relevant contexts.
The experimentally observed sequences have several important features that make them particularly suitable for timekeeping. First, they are internally generated; that is, the sequences are not depend on the fine-tuned symmetry of the Mexican-hat connectivity; this ensures that our results do not depend on the on the distance between neurons arranged on a two-dimensional torus-like grid (Fig. 2A) and follow a Mexican-hat pattern of connectivity reflecting short-range excitation and long range effective inhibition. Jhet is a random matrix that adds heterogeneity to the pattern of connections. Although the synapses represented by J0 are chosen to be weak relative to Jhet (see supplemental Text, available at www.jneurosci.org as supplemental material), they strongly influence network dynamics due to their highly correlated structure. The matrix Jhet disrupts the perfect symmetry of the Mexican-hat connectivity; this ensures that our results do not depend on the fine-tuned symmetry of J0, as this symmetry is unrealistic and may produce misleading results (Zhang, 1996; Seung et al., 2000; Renart et al., 2003).

Mexican-hat connectivity, as in J0, and the associated continuous attractor dynamics have been hypothesized as an underlying network mechanism of spatial working memory, spatial navigation and path integration (Samsonovich and McNaughton, 1997; Tsodyks, 1999; Constantinidis and Wang, 2004; McNaughton et al., 2006; Burak and Fiete, 2009). For a wide range of perturbed Mexican-hat connectivity matrices J = J0 + Jhet, the network activity will quickly converge to a “bump attractor” in the presence of constant input (Seung et al., 2000; Renart et al., 2003). If the input stays approximately constant, the bump will not move. Therefore, the dynamics of Equation 1 alone cannot produce self-generated sequences, since the “bump” only moves in response to significant changes in external inputs.

To overcome this limitation, we added an activity-dependent adaptation of the spike thresholds (Materials and Methods, Eq.

**A possible network mechanism for sequence generation: the threshold adaptation model**

The experimentally observed sequences have several important features that make them particularly suitable for timekeeping. First, they are internally generated; that is, the sequences are not brought about by changing, temporally structured environmental or body-derived inputs. Second, the sequences are reliable from trial to trial, which allows for time inference on single trials. Third, the sequences are context-dependent, with “left” and “right” trials producing different sequences. Fourth, the sequences are long, on the order of 10–20 s, lasting for the entire duration of the required delay period. To explore how the entorhino-hippocampal circuit might be capable of generating these sequences, we considered a minimalistic model of a recurrent network that we call the threshold adaptation model. This model captures all of the above properties.

We began with a firing rate model (Materials and Methods, Eq. 1). Firing rate models provide fairly accurate descriptions of the dynamics of large recurrent networks when exact spike timing is not important. In our model, the recurrent network structure is a superposition of two components: J = J0 + Jhet, where J0 is the “correlated” component and Jhet is a component uncorrelated across neurons. J0 has connection strengths that depend only on the distance between neurons arranged on a two-dimensional torus-like grid (Fig. 2A) and follow a Mexican-hat pattern of connectivity reflecting short-range excitation and long range effective inhibition. Jhet is a random matrix that adds heterogeneity to the pattern of connections. Although the synapses represented by J0 are chosen to be weak relative to Jhet (see supplemental Text, available at www.jneurosci.org as supplemental material), they strongly influence network dynamics due to their highly correlated structure. The matrix Jhet disrupts the perfect symmetry of the Mexican-hat connectivity; this ensures that our results do not depend on the fine-tuned symmetry of J0, as this symmetry is unrealistic and may produce misleading results (Zhang, 1996; Seung et al., 2000; Renart et al., 2003).
Threshold adaptation in hippocampal pyramidal neurons was observed experimentally (Henze and Buzsáki, 2001), and evolves on a relatively slow time scale (~1 s). In our model, threshold adaptation has the important effect of destabilizing the “bump” attractor states of the network. In the fast-time scale dynamics (Eq. 1), the system still evolves to a bump attractor, but as the firing rates of the neurons in the bump of activity increase, so do the corresponding thresholds, and this in turn decreases each neuron’s ability to continue firing. The threshold adaptation thus forces the bump to move to a new location, at which point the process repeats itself, resulting in a continuously moving bump that never stabilizes (Fig. 2A, C; see also supplemental Movie, available at www.jneurosci.org as supplemental material). The moving bump of activity is what produces sequential firing of the neurons.

Cell assembly sequences generated by the threshold-adaptation model are context-dependent, long-lasting, and reliable

To better assess the length and reliability of cell assembly sequences produced by the model, we simulated multiple task trials for each of four initial conditions (A to D). Reliability was quantified using the same reliability measure as in Figure 1, E and F (see Materials and Methods). Sequential activity on the order of 15–20 s can be seen for different initial conditions (Fig. 3A, B), and the sequences are quite different (Fig. 3C). As expected from the reliability of bump trajectories (Fig. 2D), sequential activity on single trials closely resembled that of the average (Fig. 3D, E) and had a high degree of reliability (Fig. 3G, H). Average and single-trial bump trajectories also showed reliability in the task conditions (Fig. 3J, K). Note that the time scale of the adaptation controls the speed of the moving bump. A shorter time scale would result in a faster-moving trajectory, leading to shorter sequences (data not shown). We also simulated a home cage condition (see Materials and Methods) that resulted in large trial-to-trial variability, which bore little resemblance to the average activity across trials (Fig. 3F, I). The lack of sequential structure in the home cage condition resulted from the lack of consistency in initial conditions, and the temporally correlated noise (task trials had uncorrelated noise). Finally, we investigated the reliability of sequences as a function of the strength of the input noise to each neuron. As expected, sequences generated in the presence of fivefold and tenfold increases in input noise were less reliable, and the reliability decreased more rapidly as a function of time (Fig. 3L). Average and single-trial sequences in these higher noise conditions were tightly clustered around the data points for the average.

Figure 4. Sequential activity is inherited by a second layer via sparse and random connections. A, average sequence over 25 trials of simulated data in a second layer with 1000 neurons. Layer 2 has no recurrent excitation and receives sparse, random feedforward inputs from the torus-like layer simulated in Figure 3. B, a single trial, with neuron ordering as in A. C, the reliability of the sequence as a function of time follows a profile similar to what is seen in the first layer (cf. Fig. 3G).
Cell assembly sequences can be inherited by a downstream layer

We have shown that reliable and context-dependent sequences of neuronal activation similar to what we have observed in CA1 may arise from a recurrent network with torus-like architecture and a weakly correlated pattern of Mexican-hat connectivity. However, the architecture of the CA1 region, with its supersparsity recurrent excitation, does not fit with this pattern of connectivity. For this reason, we investigated whether or not reliable sequences generated in one layer can be inherited by a downstream layer. In contrast to the first layer, the second layer we devised had no recurrent excitation and only a global, nonspecific recurrent inhibition. Layer 2 was driven by both the output of the previous layer and noisy, temporally and spatially unstructured inputs. The feedforward connections between the first, torus-like layer and the second layer were random and sparse (see Materials and Methods). Figure 4 shows that despite the lack of structure in layer 2, the sequential activity from the first layer was perfectly inherited by the second layer with a similar reliability profile. The reliability remained unchanged even when the magnitude of the noisy inputs to the second layer was increased 5- or tenfold (see supplemental Fig. 7, available at www.jneurosci.org as supplemental material).

Discussion

Neural correlates of elapsed time on a suprasecond scale have been documented in several cortical regions (Kojima and Goldman-Rakic, 1982; Fuster, 2001; Brody et al., 2003; Jansen and Shahlen, 2005; Lebedev et al., 2008; Mitra et al., 2009). Surprisingly, hippocampal circuits have not been considered as timers, despite the critical role of the hippocampus in timing behavior (Clark and Isaacson, 1965; Young and McNaughton, 2000) and the key importance of temporal context in episodic memory (Tulving, 1972) and navigation (McNaughton et al., 1996).

Our experimental observations and modeling results suggest that in the hippocampus the same cell populations that keep information about past memories and planned travel directions of the animal (Pastalkova et al., 2008) also provide information about elapsed time. Elapsed time was reliably inferred from the population firing rate vector of the recorded neurons at any time point of the memory task. Although the time estimation error from the neuronal population increased over time, it did not increase proportionally to the duration of elapsed time, in contrast to Weber’s law (cf. Staddon, 2005). In our network model (threshold adaptation model), the sequences emerged as a natural byproduct of a network with a perturbed Mexican-hat connectivity pattern and adaptation of the spike thresholds. The threshold adaptation model does not involve learning of sequences, synfire chains, or a “hidden” feedforward network structure (Abeles, 1991; Levy et al., 2005; Ganguli et al., 2008; Liu and Buonomano, 2009; Fiete et al., 2010). In principle, however, any network that allows for self-sustained, sequential activation of neurons can potentially be a substrate for time keeping. The adapting spike threshold mechanism was adopted because of its simplicity and because dependence of the spike threshold on prior spiking activity has been demonstrated experimentally (Henze and Buzsáki, 2001). We emphasize though that other mechanisms such as short-term synaptic depression may play a similar role (Abbott and Regehr, 2004). Although the mechanism for sequence generation in our model relies on connectivity patterns unlikely to be present in CA1, we have shown that sequences generated in one area with this kind of architecture (potentially in the entorhinal cortex or CA3) can be inherited by another area via sparse, random connections.

The model we have described here may be specific to the hippocampal system, where time keeping is needed on the scale of tens of seconds, and is different from the timing mechanisms in sensory and motor systems. Since evolving neuronal assemblies, or sequences, have been observed in other systems (Luczak et al., 2007; Fujisawa et al., 2008; Long and Fee, 2008; Johnson et al., 2010), it is possible that our modeling results apply to them as well. In general, our findings support the view that each neuronal system generates its own timing, providing temporal frames for its operations (Buonomano and Karmarkar, 2002). In summary, we found that a simple network mechanism can generate long-lasting, reliable sequences that may be used for timekeeping in the hippocampus.

References

Supplementary Figure 1. Error of time estimate using phase prediction.
Same as in Figure 1D, using the time prediction model via phase rather than rate.
Supplementary Figure 2. Single trials from the same session of the homecage recordings, all sorted according to the best ordering for the average across trials (top left). Sequential activity is not evident on individual trials.
Supplementary Figure 3. Trajectory dependence on initial conditions. (A) For each of the four initial conditions, A, B, C and D, three trials were chosen at random. The trajectories of the center of the bump for each trial (red, blue and yellow curves), resulting from 10 seconds of simulated data, are reliable for each initial condition but distinct across initial conditions. (B, C) Same as in (A), but using matrices $J^2$ and $J^3$ produced by using different instances of the random heterogeneity matrix $J^{het}$, where $J^i = J^0 + J^{het}$ for $i = 1, 2, 3$. Note that for identical initial conditions, the particular trajectory differs for different synaptic connectivity matrices $J^i$. 
Supplementary Figure 4. Simulation results for the synaptic connectivity matrices $J^1$ and $J^0$.

(J1) Same as Figure 3, but with all 4 initial conditions shown (A, B, C and D) and no homeage condition. The first two columns are identical to those in Figure 3. The heterogeneous synaptic weight matrix used in all cases is $J^1 = J^0 + J_{het}$. Note that the initial conditions C and D produced sequences that are less reliable at later times than those for initial conditions A and B. (J0) Same as (J1), but with homogeneous synaptic connectivity matrix $J^0$. While reliability is very good using the perfectly symmetric torus architecture $J^0$, the sequences are all short-lived, repeating after 5 seconds. This makes them unsuitable for timekeeping as multiple time points during a 20 second trial yield similar activity patterns, introducing ambiguity in the representation of time.
Supplementary Figure 5. Simulation results using matrices $J^2$ and $J^3$. Same as Supp Fig. 4, this time with heterogeneous synaptic connectivity matrices $J^2$ (top) and $J^3$ (bottom). A variety of sequence lengths are produced, depending on the combination of matrix and initial condition.
**Supp Fig 6. Simulation results using greater input noise strength.** Same as the left column of Figure 3 (panels A, D, G, and J, with connectivity matrix $J^1$ and initial condition ‘A’), repeated using 5-fold (left column) and 10-fold (right column) increases in the input noise to each neuron. Sequence is still present, but reliability is lower and drops off more quickly as a function of time.
Supp Fig 7. Sequence reliability is preserved for different levels of input noise to layer 2. Same as Figure 4 (left column), but including simulations with 5-fold and 10-fold increases in input noise to layer 2 (middle and right columns, respectively). Although increasing levels of input noise to the first layer, where the sequence is generated, decreases sequence reliability, this is not the case for increased noise levels to layer 2, where the sequence is inherited.
Supplementary Movie.

A movie of the neural activity as a function of time. **(Top)** The activity pattern on the torus resembles a "bump" at all times. Neurons are organized in a 50x50 grid, similar to Figure 2A. White pixels correspond to neurons with firing rate exactly 0, while dark blue corresponds to very small, but nonzero, firing rates. **(Bottom)** Tracing the center of the bump (red dot) one can see the trajectory of activity (black curve). The trial shown here is from simulations with the matrix J1, using initial condition A.
Supplementary text for “Cell assembly sequences arising from spike threshold adaptation keep track of time in the hippocampus”

V. Itskov, C. Curto, E. Pastalkova, G. Buzsáki

Here we describe the models used in the main text in more detail. We have used two different types of models. The first are time prediction models, which are phenomenological models fit to data and used to infer time from experimentally recorded neural activity. The purpose of these models is to show that the sequences observed in the hippocampus have a possible functional role - i.e., they could be used for the animal to keep track of elapsed time. These models do not reflect a cellular or network mechanism for either generating the sequences or for extracting temporal information from them in the brain, rather they show that reliable temporal information is present in the sequential cell assembly activity on a trial-to-trial basis.

The second type of model is a mechanistic model, showing a simple yet biologically plausible network mechanism for generating long-lasting, context-dependent, temporally reliable sequences in a recurrent network such as hippocampus in the absence of structured inputs. We call it the threshold adaptation model (TAM).

1 The time prediction models

For time prediction only one kind of correct trials was used for each session: ‘right’ or ‘left’ trials were chosen depending on which kind of trial had the animal making the fewest number of mistakes in the alternation task. All putative pyramidal cells with average firing rates in the range $0.1 \, \text{Hz} \leq r \leq 15 \, \text{Hz}$, were used in the time prediction models. The models are distinguished by whether they use firing rate (rate-only model) or theta phase (phase-only model) as the feature of neural activity used to estimate elapsed time.

1. Each individual spike train was smoothed with a Gaussian of width $\sigma = 0.25 \, \text{s}$ and then binned into time bins of size $0.5 \, \text{s}$. This resulted in a time series of rates $r_k(t)$ for each $k$-th cell.

2. For each cell, each time bin was assigned a discrete activity state using one of the following two rules:

a) Rate-only. For each $k$-th cell an average maximal firing rate (across trials), $r_k^{\text{max}}$, was computed. Then, for each time-bin $t$, the cell was assigned one of 3 states $u_k(t)$ based on the relative firing rate $\tilde{r}_k(t) = r_k(t)/r_k^{\text{max}}$ as follows: $u = 1$ if $\tilde{r}_k(t) \leq 0.05$, $u = 2$ if $0.05 < \tilde{r}_k(t) \leq 0.25$, and $u = 3$ if $\tilde{r}_k(t) > 0.25$. The values $u_k(t) = 1, 2$ or $3$ represent ‘low’, ‘medium’, of ‘high’ firing rates, respectively.

b) Phase-only. For each $k$-th cell and each time bin $t$ the average phase $\theta_k(t)$ of spike times with respect to the theta oscillation in that time bin was computed. Note that each time bin contains several periods of theta. For each time bin and each cell one of 4 states $u_k(t)$ was assigned. One state was reserved for having no spikes in that time-bin ($u = 0$), and the other
three states \((u = 1, 2, or 3)\) were computed based on the phase \(\theta_k(t)\) by dividing the unit circle into three equal parts.

3. For each individual neuron, \(k\), the following state-based probabilistic model was used for both the *rate-only* and *phase-only* versions of time prediction. The population model assumed that the states of different neurons are independent from each other, and thus the probability of observing a state vector \(u(t) = (u_1, .., u_N)\) at a time-bin \(t\) is

\[
P_t(u) = \prod_{k=1}^{N} P^k_t(u_k), \quad \text{where} \quad P^k_t(u) = \text{Prob}(u_k(t) = u), \quad \text{with } u = 0, 1, 2, \text{or } 3.
\]

The independence is of course an incorrect assumption. However, the prediction based on this assumption can be no better than that of a model taking correlations into account, and is less susceptible to overfitting.

4. For each trial, we obtained a time prediction for each time bin as follows. The trial was selected as a “test set,” with all the other trials serving the role of “training set.” The probabilistic model \(P_t(u)\) was fit on the training set by computing the frequency of each state for each neuron in each time bin, and using the above product formula to generate from this the probability of observing a given state vector \(u\). The model was then used on the selected “test set” trial with population vectors \(u(t)\) to obtain a time prediction for each time bin via the maximum-likelihood estimate

\[
\hat{t}(t) = \arg \max_{t'} P_{t'}(u(t)),
\]

which selects the time \(\hat{t}(t)\) at which the population pattern \(u(t)\) was most likely to occur. Notice that the estimates at different times are computed independently of each other since they used only instantaneous neuronal activity; therefore the inference method we used does not require any kind of working memory or integration. The time-estimate errors were computed for each time bin as the average error across trials.

We have chosen to use a model with a few discrete states, rather than analog firing rate values, in order to reduce overfitting of the probabilistic model. Overfitting could also be avoided using other techniques such as smoothing in continuous models. We have found that this simple model is sufficient to show that elapsed time can in principle be inferred by brain structures downstream from the CA1 area of hippocampus (on behaviorally relevant time scales), but a different model may have served this purpose equally well.
2 The threshold-adaptation model (TAM)

Model equations

We model network dynamics using a standard firing rate model, with threshold nonlinearity. At any point in time, the vector $x(t) = (x_1(t), \ldots, x_N(t))$ represents a population vector of firing rates for each of $N$ neurons. Another ingredient is the activity-dependent adaptation of the thresholds for each cell, represented by the dynamic variables $h(t) = (h_1(t), \ldots, h_N(t))$. The model equations are thus:

$$
\tau_m \frac{dx_i}{dt} = -x_i + \left( \sum_{j=1}^{N} J_{ij} x_j + I_i - h_i \right)_+, \quad i = 1, \ldots, N \quad (1)
$$

$$
\tau_a \frac{dh_i}{dt} = -h_i + c x_i, \quad i = 1, \ldots, N \quad (2)
$$

where $\tau_m$ and $\tau_a$ are membrane and threshold-adaptation time constants, respectively, $J$ is the matrix of synaptic weights for the recurrent network, and $I = (I_1, \ldots, I_N)$ is a (time-dependent) vector of external inputs to the considered recurrent network. Here the brackets denote the threshold-linear function $[y]_+ = \begin{cases} y & \text{if } y > 0, \\ 0 & \text{if } y \leq 0. \end{cases}$

The constant $c$ controls the strength of the activity-dependent adaptation, whereas $\tau_a$ determines the timescale with which the thresholds recover in the absence of the cell’s firing.

The synaptic matrix $J$: torus architecture with heterogeneity

The neurons in the recurrent network were organized on a two-dimensional sheet, with periodic boundary conditions yielding a torus-like grid of neurons ($N = 50^2 = 2500$ in simulations). Each neuron was indexed by a location $(x_i, y_i)$ on this sheet, with the $\{x_i\}$ and $\{y_i\}$ values evenly partitioning the interval $[0, 2\pi]$ with periodic boundary condition ($0 \equiv 2\pi$.)

The matrix $J = J^0 + J^\text{het}$ is the sum of a matrix $J^0$, with perfect torus-like topography, and a matrix $J^\text{het}$, which represents heterogeneity of the synaptic weights. By the “perfect torus-like topography” we mean that the strength of the (symmetric) connection between neurons $i$ and $j$ is purely a function of their distance on the torus:

$$
J^0_{ij} = \frac{4\pi^2}{N} \varphi \left( \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right),
$$

where $\varphi(d) = j_0 + j_1 \cos(d/\sqrt{2})$ was chosen similar\footnote{One difference is the $\sqrt{2}$, which is added here to compensate for the fact that the longest distance on the torus is $\sqrt{2}\pi$, rather than $\pi$ as it is on the circle.} to the “ring model” [1], however a wide class of Mexican-hat-type connectivity functions $\varphi(d)$ may be chosen to yield bump attractor solutions (c.f. [2]).

The parameters $j_0$ and $j_1$ ($j_0 = -0.6, j_1 = 1$ in simulations) were chosen to to yield recurrent excitation for nearby neurons, and inhibition for far away neurons on the torus, so that the resulting dynamics of equation (1) with spatially and temporally constant inputs and thresholds ($I_i - h_i = \text{const}$) results in “bump attractor” dynamics.

The matrix of heterogeneities is constant in time and is drawn from the normal distribution:

$$
J^\text{het}_{ij} = \frac{\varepsilon}{\sqrt{N}} n_{ij},
$$
where \( n_{ij} \sim N(0,1) \) are independent gaussian-distributed variables with mean 0 and variance 1. The scaling of \( N^{-\frac{1}{2}} \) is chosen so that the spectral radius (\( = \varepsilon \)) of this matrix does not change with finer and finer discretizations of the torus sheet. Note that due to the scaling by \( N \), the individual matrix elements \( J_{ij}^{\text{het}} \) are much larger than \( J_{ij}^{0} \) in the limit of large \( N \); nevertheless, the correlated structure of the matrix \( J^{0} \) determines the “bump attractor” dynamics of equation (1). The heterogeneity matrix \( J^{\text{het}} \) is important not just to avoid fine-tuning of the model, but also for the reliability. By breaking the symmetry in the two-dimensional sheet of bump attractors, \( J^{\text{het}} \) has the effect of “carving out” out an irregular trajectory for the movement of the bump; this makes the sequences last longer without repeating the same neurons.

**Simulations with TAM**

Here we describe in detail all of the particular choices made in simulations (parameters, noise, synaptic weight matrix, etc.). In all, we generated 500 simulated data sets, each consisting of a single 20-second ‘trial’ with 2500 neurons. We used 4 different synaptic-weight matrices \( J^{0}, J^{1}, J^{2}, \text{ and } J^{3} \), where \( J^{0} \) had perfect torus-like symmetry in the connections, and the others had heterogeneities introduced by the addition of different ‘noise’ matrices \( J_{ij}^{\text{het}} \). For each matrix \( J^{l}, l = 0, \ldots, 4 \), we ran 25 trials for each of 5 different trial-types: 4 trial types were ‘task’ trials with different adaptation initial conditions (tasks A, B, C, and D), and the fifth trial-type represented ‘homecage’ trials (hc).

**Parameters used in simulations**

\[
\begin{align*}
N & = 50^2 = 2500 \quad \text{(number of neurons, in a 50 by 50 grid)} \\
T & = 20s \quad \text{(time length for simulations)} \\
\tau_m & = 30ms \quad \text{(membrane time constant)} \\
\tau_a & = 2s \quad \text{(timescale of the adaptation)} \\
c & = 0.5 \quad \text{(adaptation constant, controls ‘ergodicity’ of the trajectory)} \\
J_{ij}^{0} & = \frac{4\pi^2}{N} \left( j_0 + j_1 \cos \sqrt{\frac{1}{2}(x_i-x_j)^2 + \frac{1}{2}(y_i-y_j)^2} \right), \quad \text{and} \\
J_{ij}^{l} & = J_{ij}^{0} + \frac{\varepsilon}{\sqrt{N}} n_{ij}, \quad \text{where} \\
\varepsilon & = 0.5, \quad j_0 = -0.6, \quad j_1 = 1, \quad \text{and } n_{ij} \sim N(0,1). \\
I_{i}(t) & = I_{0} + I_{i}^{\text{noise}}(t) \quad \text{(input noise varies over neurons, time bins, and trials)} \\
I_{0} & = 1 \quad \text{(constant across neurons, time bins, and trials)} \\
I_{i}^{\text{noise}}(t) & \sim N(0,1)
\end{align*}
\]

**What changed between trials and trial-types in the ‘task’ conditions**

For each choice of matrix \( J^{l} \) in the ‘task’ conditions, we simulated trials with initial conditions ‘A’, ‘B’, ‘C’, and ‘D’. The differences in initial conditions are for the adaptation variables only. The initial conditions ‘A’, ‘B’, ‘C’, and ‘D’ have neurons with adapted thresholds to the left, bottom, right and top of the initial activity bump, respectively.

Within a fixed matrix choice \( J^{l} \) and fixed initial condition, the only aspect of the simulation that changed from trial to trial was the white noise \( I_{i}^{\text{noise}}(t) \). Each neuron was driven by a random input,
normally distributed with mean 1 and variance 1. The drive was independent across neurons and from one 1ms-time bin to the next. Each noise instance was re-sampled from the distribution for each trial.

**Simulations for the second layer.**

In order to investigate if the reliable sequences can be “inherited” by a second layer via random projections, we have considered a second layer of $N_2 = 1000$ cells that were receiving downstream projections from the layer with torus-like mexican-hat connectivity having $N = 2500$ cells (as described above). Each cell in the second layer was receiving an input from the first layer as well as a noisy unstructured input. The second layer also had non-specific recurrent inhibition. The dynamics of the second layer are thus described by the following equations:

\[
\tau_m \frac{dy_k}{dt} = -y_k + \left[ \sum_{i=1}^{N} M_{ki} x_i - \sum_{l=1}^{N_2} y_l + \tilde{I}_k - g_k \right]_+, \quad k = 1, \ldots, N_2
\]

\[
\tau_a \frac{dg_k}{dt} = -g_k + cy_k, \quad k = 1, \ldots, N_2.
\]

Here $y_k(t)$ are the firing rates of the cells in the second layer, $x_i(t)$ are the firing rates in the first layer, $\tau_m$ and $\tau_a$ are the membrane time-constant and the timescale of the synaptic adaptation respectively (the same as in the first layer). The synaptic weights $M_{kl}$ of the projection from the first to the second layer were chosen to be nonzero with probability $p = 0.1$, and the nonzero elements were drawn from a uniform distribution on the unit interval (i.e. $0 \leq M_{kl} \leq 1$). The input noise $\tilde{I}_k = \tilde{I}_k(t)$ in the feedforward synapses was drawn from a Gaussian distribution with zero mean and the same variance as the input noise $I_{\text{noise}}(t)$ in the first layer. The second layer was assumed to have the same dynamics of threshold adaptation (the variables $g_k(t)$) as in the first layer.

**The ‘homecage’ condition**

There are two differences between our ‘homecage’ condition simulations and the above simulations for the ‘task’ conditions: (1) On each trial, we randomly selected one of the initial conditions ‘A’, ‘B’, ‘C’, or ‘D’, to reflect the fact that – unlike in the task conditions – the specific initial condition on adaptation variables in the homecage will vary unpredictably from trial to trial; (2) Instead of having noise inputs $I(t)$ that are independently sampled on each 1ms time step of the simulation, the noise input to each neuron is constant over a duration of 125ms (approx. one theta-cycle). This reflects the fact that in inattentive states, cortical activity often shows higher temporal correlations than in attentive states [3]. These two changes destroyed the sequences produced by the model.

**References**

