1

Intrinsic dynamics of randomly clustered networks generate place fields and preplay of novel environments

Jordan Breffle[1], Hannah Germaine[1], Justin D. Shin[1,2,3], Shantanu P. Jadhav[1,2,3], Paul Miller[1,2,4]

[1]Neuroscience Program, [2]Volen National Center for Complex Systems, [3]Department of Psychology, and [4]Department of Biology, Brandeis University, 415 South St., Waltham, MA 02454

Corresponding Authors: shantanu@brandeis.edu (S.P.J.) pmiller@brandeis.edu (P.M.)

ORCiD:
Jordan Breffle https://orcid.org/0000-0001-5793-4427
Hannah Germaine https://orcid.org/0000-0002-7624-2431
Justin D. Shin https://orcid.org/0000-0002-7959-7772
Shantanu P. Jadhav https://orcid.org/0000-0001-5821-0551
Paul Miller https://orcid.org/0000-0002-9280-000X

Abstract

During both sleep and awake immobility, hippocampal place cells reactivate time-compressed versions of sequences representing recently experienced trajectories in a phenomenon known as replay. Intriguingly, spontaneous sequences can also correspond to forthcoming trajectories in novel environments experienced later, in a phenomenon known as preplay. Here, we present a model showing that sequences of spikes correlated with the place fields underlying spatial trajectories in both previously experienced and future novel environments can arise spontaneously in neural circuits with random, clustered connectivity rather than pre-configured spatial maps. Moreover, the realistic place fields themselves arise in the circuit from minimal, landmark-based inputs. We find that preplay quality depends on the network's balance of cluster isolation and overlap, with optimal preplay occurring in small-world regimes of high clustering yet short path lengths. We validate the results of our model by applying the same place field and preplay analyses to previously published rat hippocampal place cell data. Our results show that clustered recurrent connectivity can generate spontaneous preplay and immediate replay of novel environments. These findings support a framework whereby novel sensory experiences become associated with preexisting "pluripotent" internal neural activity patterns.

Impact Statement

Neural circuits with small-world connectivity spontaneously emit sequences of spikes that are correlated with any of the distinct sequences of realistic place fields produced by location-modulated, monotonically varying input.

2

Contributions:

3

Introduction

54 The hippocampus plays a critical role in spatial and episodic memory in mammals (Morris
55 et al., 1982; Squire et al., 2004). Place cells in the hippocampus exhibit spatial tuning, firing
56 selectively in specific locations of a spatial environment (Moser et al., 2008; O'Keefe and
57 Nadel, 1978). During sleep and quiet wakefulness, place cells show a time-compressed
58 reactivation of spike sequences corresponding to recent experiences (Wilson and
59 McNaughton, 1994; Foster and Wilson, 2006), known as replay. These replay events are
60 thought to be important for memory consolidation, often referred to as memory replay
61 (Carr et al., 2011).

62 The CA3 region of the hippocampus is a highly recurrently connected region that is the
63 primary site of replay generation in the hippocampus. Input from CA3 supports replay in
64 CA1 (Csicsvari et al., 2000; Yamamoto and Tonegawa, 2017; Nakashiba et al., 2008;
65 Nakashiba et al., 2009), and peri-ripple spiking in CA3 precedes that of CA1 (Nitzan et al.,
66 2022). The recurrent connections support intrinsically generated bursts of activity that
67 propagate through the network.

68 Most replay models rely on a recurrent network structure in which a map of the
69 environment is encoded in the recurrent connections of CA3 cells, such that cells with
70 nearby place fields are more strongly connected. Some models assume this structure is pre-
71 existing (Haga and Fukai, 2018; Pang and Fairhall, 2019), and some show how it could
72 develop over time through synaptic plasticity (Theodoni et al., 2018; Jahnke et al., 2015).
73 Related to replay models based on place-field distance-dependent connectivity is the
74 broader class of synfire-chain-like models. In these models, neurons (or clusters of
75 neurons) are connected in a 1-dimensional feed-forward manner (Diesmann et al., 1999;
76 Chenkov et al., 2017). The classic idea of a synfire-chain has been extended to included
77 recurrent connections, such as by Chenkov et al., 2017, however such models still rely on
78 an underlying 1-dimensional sequence of activity propagation.

79 A problem with these models is that in novel environments place cells remap immediately
80 in a seemingly random fashion (Leutgeb et al., 2005; Muller and Kubie, 1987). The CA3
81 region, in particular, undergoes pronounced remapping (Leutgeb et al., 2004; Leutgeb et al.,
82 2005; Alme et al., 2014). A random remapping of place fields in such models that rely on
83 environment-specific recurrent connectivity between place cells would lead to recurrent
84 connections that are random with respect to the novel environment, and thus would not
85 support replay of the novel environment.

86 Rather, these models require a pre-existing structure of recurrent connections to be
87 created for each environment. A proposed solution to account for remapping in
88 hippocampal models is to assume the existence of multiple independent and uncorrelated
89 spatial maps stored within the connections between cells. In this framework, the maximum
90 number of maps is reached when the noise induced via connections needed for alternative
91 maps becomes too great for a faithful rendering of the current map (Samsonovich and
92 McNaughton, 1997; Battaglia and Treves, 1998; Azizi et al., 2013). However, experiments
93 have found that hippocampal representations remain uncorrelated, with no signs of

4

94 representation re-use, after testing as many as 11 different environments in rats (Alme et
95 al., 2014).

96 Rather than re-using a previously stored map, another possibility is that a novel map for a
97 novel environment is generated *de novo* through experience-dependent plasticity while in
98 the environment. Given the timescales of synaptic and structural plasticity, one might
99 expect that significant experience within each environment is needed to produce each new
100 map. However, replay can occur after just 1-2 laps on novel tracks (Foster and Wilson,
101 2006; Berners-Lee et al., 2022), which means that the synaptic connections that allow the
102 generation of the replayed sequences must already be present. Consistent with this
103 expectation, it has been found that decoded sequences during sleep show significant
104 correlations when decoded by place fields from future, novel environments. This
105 phenomenon is known as preplay and has been observed in both rodents (Dragoi and
106 Tonegawa, 2011; Dragoi and Tonegawa, 2013; Grosmark and Buzsaki, 2016; Liu et al.,
107 2018) and humans (Vaz et al., 2023).

108 The existence of both preplay and immediate replay in novel environments suggests that
109 the preexisting recurrent connections in the hippocampus that generate replay are
110 somehow correlated with the pattern of future place fields that arise in novel
111 environments. To reconcile these experimental results, we propose a model of intrinsic
112 sequence generation based on randomly clustered recurrent connectivity, wherein place
113 cells are connected within multiple overlapping clusters that are random with respect to
114 any future, novel environment. Such clustering is a common motif across the brain,
115 including the CA3 region of the hippocampus (Guzman et al., 2016) as well as cortex (Song
116 et al., 2005; Perin et al., 2011), naturally arises from a combination of Hebbian and
117 homeostatic plasticity in recurrent networks (Bourjaily and Miller, 2011; Litwin-Kumar
118 and Doiron, 2014; Lynn et al., 2022), and spontaneously develops in networks of cultured
119 hippocampal neurons (Antonello et al., 2022).

120 As an animal gains experience in an environment, the pattern of recurrent connections of
121 CA3 would be shaped by Hebbian plasticity (Debanne et al., 1998; Mishra et al., 2016).
122 Relative to CA1, which has little recurrent connectivity, CA3 has been found to have both
123 more stable spatial tuning and a stronger functional assembly organization, consistent with
124 the hypothesis that spatial coding in CA3 is influenced by its recurrent connections
125 (Sheintuch et al., 2023). Gaining experience in different environments would then be
126 expected to lead to individual place cells participating in multiple formed clusters. Such
127 overlapping clustered connectivity may be a general feature of any hippocampal and
128 cortical region that has typical Hebbian plasticity rules. Sadovsky and MacLean, 2014,
129 found such structure in the spontaneous activity of excitatory neurons in primary visual
130 cortex, where cells formed overlapping but distinct functional clusters. Further, such
131 preexisting clusters may help explain the correlations that have been found in otherwise
132 seemingly random remapping (Kinsky et al., 2018; Whittington et al., 2020) and support
133 the rapid hippocampal representations of novel environments that are initially generic and
134 become refined with experience (Liu et al., 2021). Such clustered connectivity likely
135 underlies the functional assemblies that have been observed in hippocampus, wherein
136 groups of recorded cells have correlated activity that can be identified through
137 independent component analysis (Peyrache et al., 2010; Farooq et al., 2019).

138   Since our model relies on its random recurrent connections for propagation of activity
139   through the network during spontaneous activity, we also sought to assess the extent to
140   which the internal activity within the network can generate place cells with firing rate
141   peaks at a location where they do not receive a peak in their external input. While the total
142   input to the network is constant as a function of position, each cell only receives a peak in
143   its spatially linearly varying feedforward input at one end of the track. Our reasoning is that
144   landmarks in the environment, such as boundaries or corners, provide location-specific
145   visual input to an animal, but locations between such features are primarily indicated by
146   their distance from them, which in our model is represented by reduction in the landmark-
147   specific input. One can therefore equate our model's inputs as corresponding to boundary
148   cells (Savelli et al., 2008; Solstad et al., 2008; Bush et al., 2014), and the place fields
149   between boundaries are generated by random internal structure within the network.
150   Further, variations in spatial input forms do not affect the consistency and robustness of
151   the model.

152   In our implementation of this model, we find that spontaneous sequences of spikes
153   generated by a randomly clustered network can be decoded as spatial trajectories without
154   relying on pre-configured, environment-specific maps. Because the network contains
155   neither a preexisting map of the environment nor experience-dependent plasticity, we
156   refer to the spike-sequences it generates as preplay. However, the model can also be
157   thought of as a preexisting network in which immediate replay in a novel environment can
158   be expressed and then reinforced through experience-dependent plasticity. We find that
159   preplay in this model occurs most strongly when the network parameters are tuned to
160   generate networks that have a small-world structure (Watts and Strogatz, 1998;
161   Humphries et al., 2006; Humphries et al., 2008). Our results support the idea that preplay
162   and immediate replay could be a natural consequence of the preexisting recurrent
163   structure of the hippocampus.

6

Results

The model
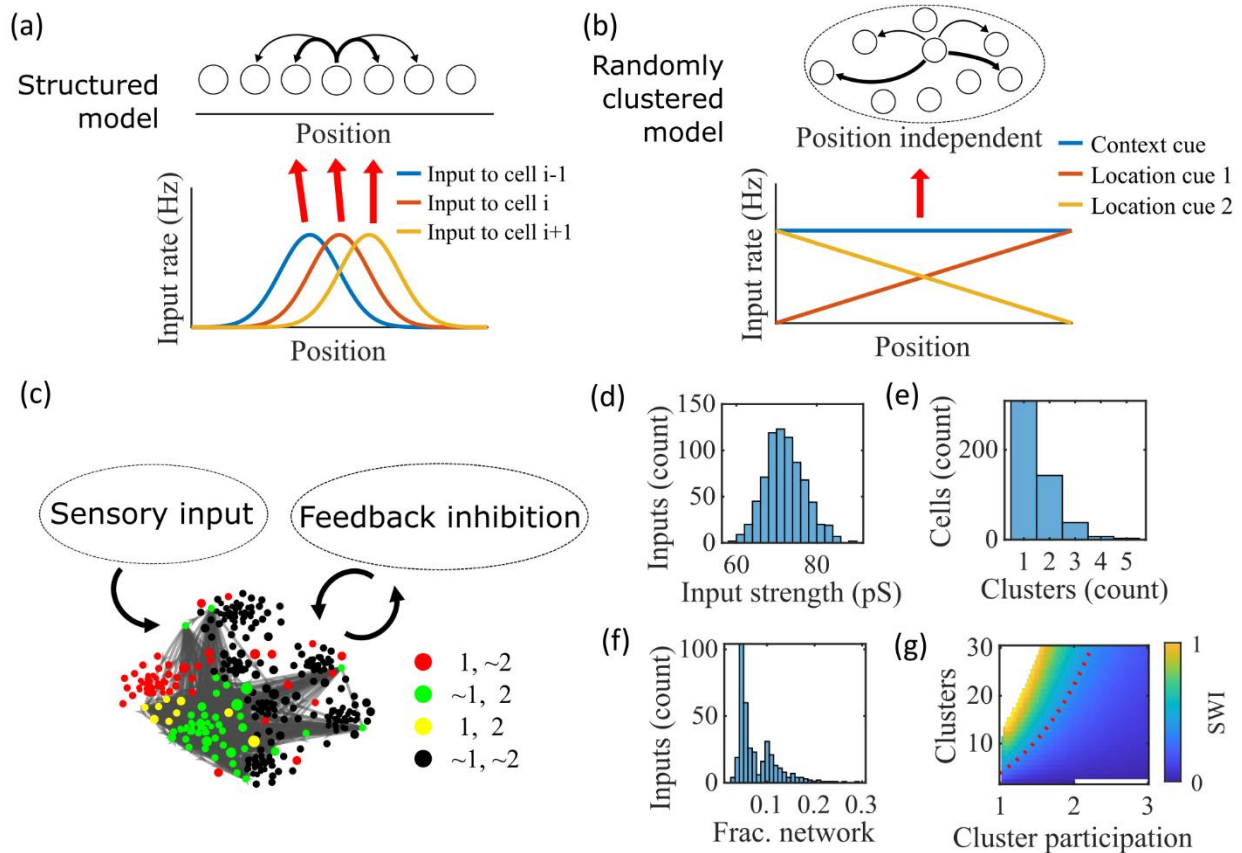
Figure 1: Illustration of the randomly clustered model
(a) Schematic diagram of prior replay models that rely on preexisting environment-specific structure, wherein each cell receives uniquely tuned Gaussian-shaped feed-forward inputs to define the place fields, and cells with nearby place fields are recurrently connected. Pairs of cells with closest place fields are connected most strongly (thicker arrows). **(b)** Schematic diagram of our model, where neurons are randomly placed into clusters and all neurons receive the same spatial and contextual information but with random, cluster-dependent input strengths. **(c)** Example representation of the network (8 clusters, mean cluster participation per cell of 1.5). Excitatory cells (each symbol) are recurrently connected with each other and with inhibitory cells ("Feedback inhibition", individual inhibitory cells not shown) and receive feed forward input ("Sensory input"). Symbol color indicates neurons' membership in clusters 1 and 2, with ~ meaning not in the cluster. Symbol size scales with the number of clusters a neuron is in. Lines show connections between neurons that are in cluster 2. Symbol positions are plotted based on a t-distributed stochastic neighbor embedding (t-SNE) of the connection matrix, which reveals the randomly overlapping clusters. **(d-f)** Histograms based on the network in (c) of: **(d)** the distribution of input strengths; **(e)** the number of clusters that each neuron is a member of; and **(f)** the fraction of the excitatory cells to which each excitatory cell connects. **(g)** The

185   Small-World Index (SWI) of the excitatory connections varies with the number of clusters
186   and the mean number of clusters of which each neuron is a member ("cluster
187   participation"). The median value of the SWI from 10 networks at each parameter point is
188   plotted. The red dashed line shows a contour line where SWI = 0.4. Regions in white are
189   not possible due to either cluster participation exceeding the number of clusters (lower
190   right) or cells not being able to connect to enough other cells to reach the target global
191   connectivity p_c (upper left).

192   _____

193   We propose a model of preplay and immediate replay based on randomly clustered
194   recurrent connections (Figure 1). In prior models of preplay and replay, a preexisting map
195   of the environment is typically assumed to be contained within the recurrent connections
196   of CA3 cells, such that cells with nearby place fields are more strongly connected (Figure
197   1a). While this type of model successfully produces replay (Haga and Fukai, 2018; Pang and
198   Fairhall, 2019), such a map would only be expected to exist in a familiar environment, after
199   experience-dependent synaptic plasticity has had time to shape the network (Theodoni et
200   al., 2018). It remains unclear how, in the absence of such a preexisting map of the
201   environment, the hippocampus can generate both preplay and immediate replay of a novel
202   environment.

203   Our proposed alternative model is based on a randomly clustered recurrent network with
204   random feed-forward inputs (Figure 1b). In our model, all excitatory neurons are randomly
205   assigned to overlapping clusters that constrain the recurrent connectivity, and they all
206   receive the same linear spatial and contextual input cues which are scaled by randomly
207   drawn, cluster-dependent connection weights (see Methods). This bias causes cells that
208   share cluster memberships to have more similar place fields during the simulated run
209   period, but, crucially, this bias is not present during sleep simulations so that there is no
210   environment-specific information present when the network generates preplay.

211   An example network with 8 clusters and cluster participation of 1.5 (the mean number of
212   clusters to which an excitatory neuron belongs) is depicted in Figure 1c. Excitatory neurons
213   are recurrently connected to each other and to inhibitory neurons. Inhibitory cells have
214   cluster-independent connectivity, such that all E-to-I and I-to-E connections exist with a
215   probability of 0.25. Feed-forward inputs are independent Poisson spikes with random
216   connection strength for each neuron (Figure 1d). Excitatory cells are randomly,
217   independently assigned membership to each of the clusters in the network. All neurons are
218   first assigned to one cluster, and then randomly assigned additional clusters to reach the
219   target cluster participation (Figure 1e). Given the number of clusters and the cluster
220   participation, the within-cluster connection probability is calculated such that the global
221   connection probability matches the parameter $p_c = 0.08$ (Figure 1f). The left peak in the
222   distribution shown in Figure 1f is from cells in a single cluster and the right peak is from
223   cells in two clusters, with the long tail corresponding to cells in more than two clusters.

224   For a given p_c, excitatory connectivity is parameterized by the number of clusters in the
225   network and the mean cluster participation. The small-world index (SWI; Neal, 2015; Neal,
226   2017) systematically varies across this 2-D parameterization (Figure 1g). A high SWI

8

227 indicates a network with both clustered connectivity and short path lengths (Watts and
228 Strogatz, 1998). A ring lattice network (Figure 1—figure supplement 1a) exhibits high
229 clustering but long path lengths between nodes on opposite sides of the ring. In contrast, a
230 randomly connected network (Figure 1—figure supplement 1c) has short path lengths but
231 lacks local clustered structure. A network with small world structure, such as a Watts-
232 Strogatz network (Watts and Strogatz, 1998) or our randomly clustered model (Figure 1—
233 figure supplement 1b), combines both clustered connectivity and short path lengths. In our
234 clustered networks, for a fixed connection probability, SWI increases with more clusters
235 and lower cluster participation, so long as cluster participation is greater than one to
236 ensure sparse overlap of (and hence connections between) clusters. Networks in the top
237 left corner of Figure 1g are not possible, since in that region all within-cluster connections
238 are not sufficient to match the target global connectivity probability, p_c. Networks in the
239 bottom right are not possible because otherwise mean cluster participation would exceed
240 the number of clusters. The dashed red line shows an example contour line where $SWI =$
241 0.4.

Example activity

243
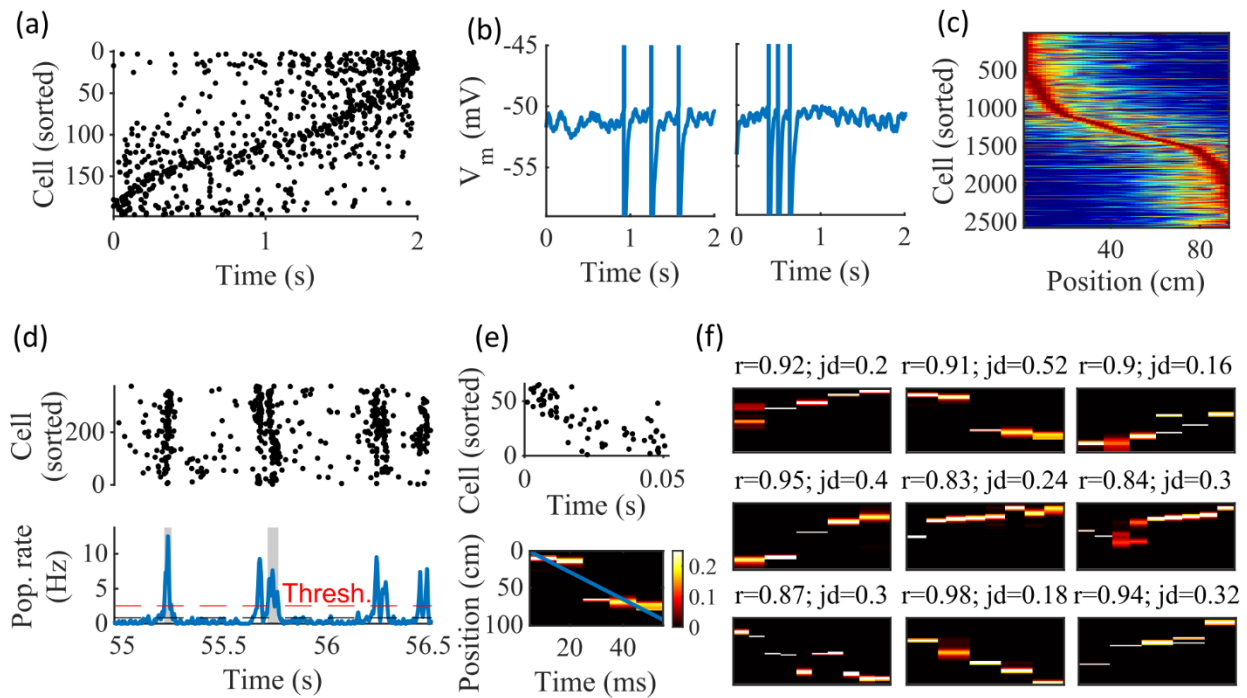
244 **Figure 2: Spatially correlated reactivations in networks without environment-**
245 **specific connectivity or plasticity**
246 **(a-f)** Example activity from the fiducial parameter set (15 clusters, mean cluster
247 participation of 1.25). **(a)** Example raster plot from one place-field trial. Cells sorted by trial
248 peak. **(b)** Example membrane traces from two of the cells in (a). **(c)** Place fields from 10
249 different networks generated from the same parameter set, sorted by peak location and
250 normalized by peak rate. **(d)** Example raster plot (top) and population firing rate (bottom;

251 blue line) showing preplay in a simulation of sleep. Horizontal dashed black line is the
252 mean population rate across the simulation. Horizontal dashed red line is the threshold for
253 detecting a population-burst event (PBE). PBEs that exceeded the threshold for at least 50
254 ms and had at least 5 participating cells were included in the preplay decoding analysis.
255 Grey bars highlight detected events. **(e)** Example preplay event (Top, raster plot. Bottom,
256 Bayesian decoding of position). Event corresponds to the center event in (d). Raster
257 includes only participating cells. The blue line shows the weighted correlation of decoded
258 position across time. **(f)** Nine example decoded events from the same networks in (c). The
259 width of each time bin is 10 ms. The height spans the track length. Same color scale as in
260 (e). r is each event's absolute weighted correlation. jd is the maximum normalized jump in
261 peak position probability between adjacent time bins. The same event in (e) is shown with
262 its corresponding statistics in the center of the top row. Preplay statistics calculated as in
263 Farooq et al., 2019.

---

265 Our randomly clustered model produces both place fields and preplay with no
266 environment-specific plasticity or preexisting map of the environment (Figure 2). Example
267 place cell activity shows spatial specificity during linear track traversal (Figure 2a-c).
268 Although the spatial tuning is noisy, this is consistent with the experimental finding that
269 the place fields that are immediately expressed in a novel environment require experience
270 in the environment to stabilize and improve decoding accuracy (Tang and Jadhav, 2022;
271 Shin et al., 2019; Hwaun and Colgin, 2019). Raster plots of network spiking activity (Figure
272 2a) and example cell membrane potential traces (Figure 2b) demonstrate selective firing in
273 specific track locations. Place fields from multiple networks generated from the same
274 parameters, but with different input and recurrent connections, show spatial tuning across
275 the track (Figure 2c).

276 To test the ability of the model to produce preplay, we simulated sleep sessions in the same
277 networks. Sleep sessions were simulated in a similar manner to the running sessions but
278 with no location cue inputs active and a different, unique set of context cue inputs active to
279 represent the sleep context. The strength of the context cue inputs to the excitatory and
280 inhibitory cells were scaled in order to generate an appropriate level of network activity, to
281 account for the absence of excitatory drive from the location inputs (see Methods). During
282 simulated sleep, sparse, stochastic spiking spontaneously generates sufficient excitement
283 within the recurrent network to produce population burst events resembling preplay
284 (Figure 2d-f). Example raster and population rate plots demonstrate spontaneous transient
285 increases in spiking that exceed 1 standard deviation above the mean population rate
286 denoting population burst events (PBEs; Figure 2d). We considered PBEs that lasted at
287 least 50 ms and contained at least 5 participating cells candidates for Bayesian decoding
288 (Shin et al., 2019). Bayesian decoding of an example PBE using the simulated place fields
289 reveals a spatial trajectory (Figure 2e). We use the same two statistics as Farooq et al.
290 (2019) to quantify the quality of the decoded trajectory: the absolute weighted correlation
291 (r) and the maximum jump distance (jd; Figure 2f). The absolute weighted correlation of a
292 decoded event is the absolute value of the linear Pearson's correlation of space-time
293 weighted by the event's derived posteriors. Since sequences can correspond to either
294 direction along the track, the sign of the correlation simply indicates direction while the

10

295 absolute value indicates the quality of preplay. The maximum jump distance of a decoded
296 event is the maximum jump in the location of peak probability of decoded position across
297 any two adjacent 10-ms time bins of the event's derived posteriors. A high-quality event
298 will have a high absolute weighted correlation and a low maximum jump distance.

299 Together, these results demonstrate that the model can reproduce key dynamics of
300 hippocampal place cells, including spatial tuning and preplay, without relying on
301 environment-specific recurrent connections.

302 **Place Fields**

303
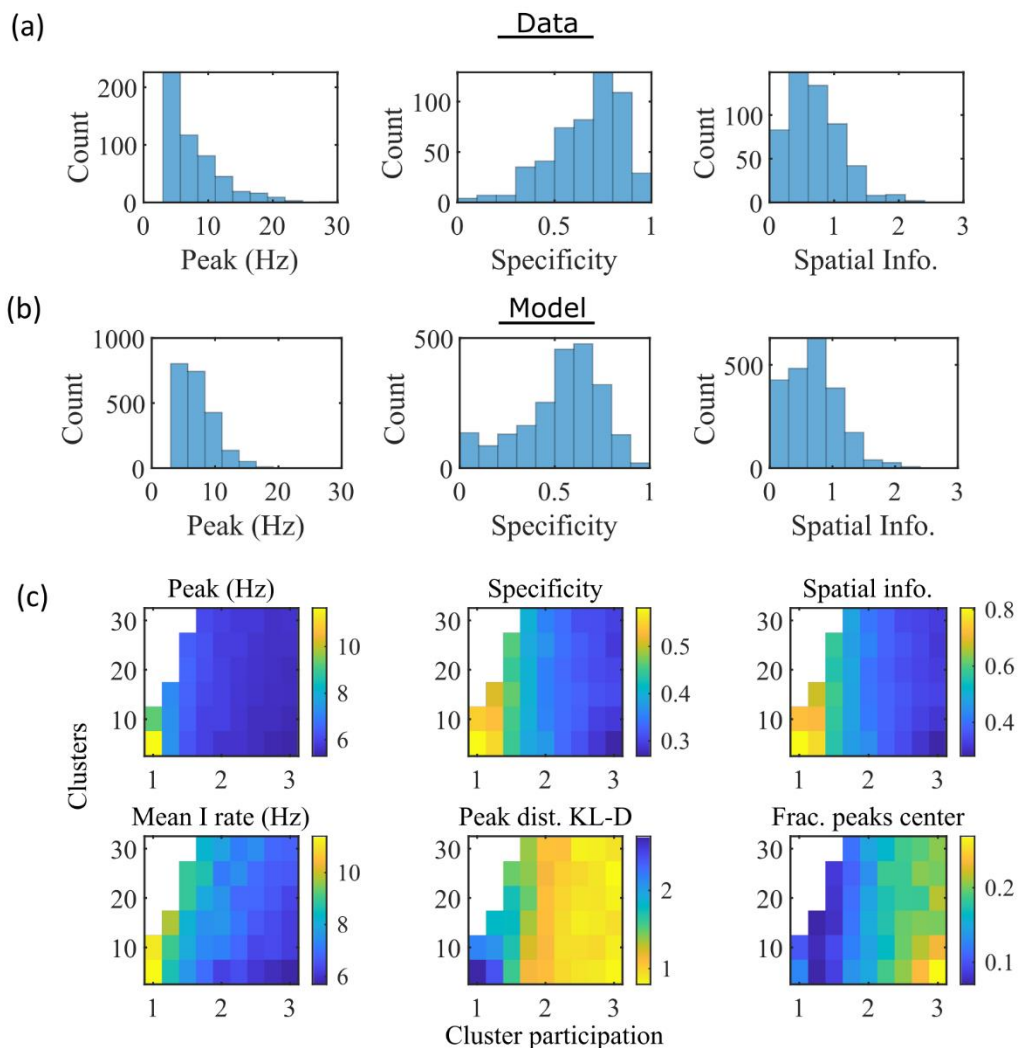
304 **Figure 3: The model produces place fields with similar properties to hippocampal**
305 **place fields**
306 **(a)** Place field statistics for hippocampal place fields recorded in rats upon their first
307 exposure to a W-track (Shin et al., 2019). Left, place-field peak rate (Hz). Center, place-field
308 specificity (fraction of track). Right, place-field spatial information (bits/spike). **(b)** Same as
309 (a) but for place fields from a set of 10 simulated networks at one parameter point (15
310 clusters and mean cluster participation of 1.25). **(c)** Network parameter dependence of

11

311 place-field statistics. For each parameter point, the color indicates the mean over all place
312 fields from all networks. Top row: mean statistics corresponding to the same measures of
313 place fields used in panels (a, b). Bottom left: mean firing rate of the inhibitory cells. Bottom
314 center: the KL-divergence of the distribution of place-field peaks relative to a uniform
315 spatial distribution. Bottom right: fraction of place-field peaks peaked in the central third of
316 the track.

317 ───────────────────────────────────

318 To compare the place fields generated by the model to those from hippocampal place cells
319 of rats, we calculated several place-field statistics for both simulated and experimentally
320 recorded place fields (Figure 3). Because our model assumes no previous environment-
321 specific plasticity, we analyzed data from place cells in rats on their first exposure to a W-
322 track (Shin et al., 2019). Equivalent statistics of place-field peak rate, sparsity, and spatial
323 information are shown for experimental data (Figure 3a) and simulations (Figure 3b). We
324 found that the model produces qualitatively similar (but not quantitatively identical)
325 distributions for the fiducial parameter set.

326 These place-field properties depend on the network parameters (Figure 3c). With fewer
327 clusters and lower cluster overlap (lower cluster participation), place fields have higher
328 peak rates, sparsity, and spatial information (Figure 3c, top row and bottom left). However,
329 lower overlap reduces the uniformity of place-field locations, measured by KL-divergence
330 (Figure 3c bottom middle) and the fraction of place fields in the central third of the track
331 (Figure 3c bottom right).

332 To verify that our simulated place cells were more strongly coding for spatial location than
333 for elapsed time, we performed simulations with additional track traversals at different
334 speeds and compared the resulting place fields and time fields in the same cells. We find
335 that there is significantly greater place information than time information (Figure 3—
336 figure supplement 1)
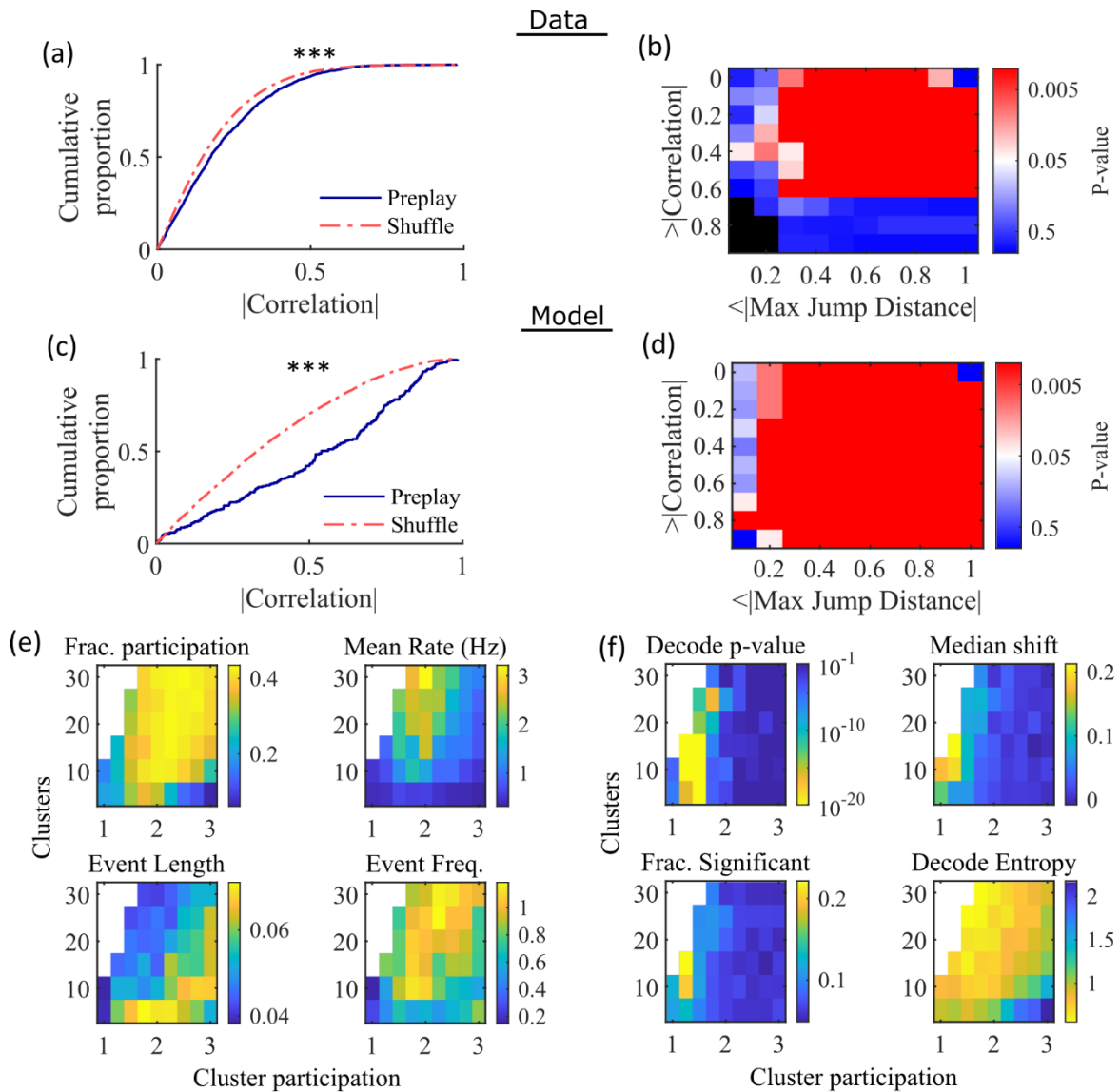
Preplay

Figure 4: Preplay depends on modest cluster overlap
(a,c) The cumulative distribution function (CDF) of the absolute weighted correlations for actual events (blue line) versus shuffled events (red dashed line) of experimental data from Shin at al., 2019 (a; KS-test, p=2×10⁻¹², KS-statistic=0.078) and simulated data (c; KS-test, p=3×10⁻¹⁶, KS-statistic=0.29) reveal results similar to those in Figure 1h of Farooq et al., 2019. *** p<0.001. **(b,d)** P-value grids (p-value indicated logarithmically by color) showing that the actual decoded events are higher quality sequences than shuffles across a wide range of quality thresholds for both experimental data from Shin et al., 2019 (b) and simulated data (d). For each point on the grid the fraction of events that exceed the absolute weighted correlation threshold (y-axis) and don't exceed the maximum jump distance (x-axis) is calculated, and the significance of this fraction is determined by comparison against a distribution of corresponding fractions from shuffled events. Black

351   squares indicate criteria that were not met by any events (either shuffled or actual). The
352   panel is equivalent to Figure 1e of Farooq et al., 2019. **(e)** Network parameter dependence
353   of several statistics quantifying the population-burst events. Top left, fraction of excitatory
354   cells firing per event. Top right, mean excitatory cell firing rate (Hz). Bottom left, mean
355   event duration (s). Bottom right, mean event frequency (Hz). Each point is the mean of data
356   combined across all population-burst events of all networks at each parameter point. Data
357   from the same simulations as Figure 3. **(f)** Network parameter dependence of several
358   statistics quantifying the Bayesian decoding. Top left, p-value of the absolute weighted
359   correlations (from a KS-test as calculated in (c)). Top right, the shift in the median absolute
360   weighted correlation of actual events relative to shuffle events. Bottom left, the fraction of
361   events with significant absolute weighted correlations relative to the distribution of
362   absolute weighted correlations from time bin shuffles of the event. Bottom right, the mean
363   entropy of the position probability of all time bins in decoded trajectories.

364   ───

365   Having found that the model produces realistic place-field representations with neither
366   place-field like inputs nor environment-specific spatial representation in the internal
367   network connectivity (Figure 3), we next examined whether the same networks could
368   generate spontaneous preplay of novel environments. To test this, for the same set of
369   networks characterized by place-field properties in Figure 3, we simulated sleep activity by
370   removing any location-dependent input cues and analyzed the resulting spike patterns for
371   significant sequential structure resembling preplay trajectories (Figure 4). We find
372   significant preplay in both our reference experimental data set (Shin et al., 2019; Figure 4a,
373   b; see Figure 4—figure supplement 1 for example events) and our model (Figure 4c, d)
374   when analyzed by the same methods as Farooq et al., 2019, wherein the significance of
375   preplay is determined relative to time-bin shuffled events (see Methods). The distribution
376   of absolute weighted correlations of actual events was significantly greater than the
377   distribution of absolute weighted correlations of shuffled events for both the experimental
378   data (Figure 4a, KS-test, $p=2\times10^{-12}$, KS-statistic=0.078) and the simulated data (Figure 4c,
379   KS-test, $p=3\times10^{-16}$, KS-statistic=0.29). Additionally, we found that this result is robust to
380   random subsampling of cells in our simulated data (Figure 4—figure supplement 2). Our
381   analyses of the hippocampal data produce similar results when analyzing each trajectory
382   independently (Figure 4—figure supplement 3).

383   For each event, we also calculated the maximum spatial jump of the peak probability of
384   decoded position between any two adjacent time bins as a measure of the continuity of the
385   decoded trajectory. The absolute weighted correlation (high is better) and maximum jump
386   (low is better) were then two different measures of the quality of a decoded trajectory. We
387   performed a bootstrap test that took both of these measures into account by setting
388   thresholds for a minimum absolute weighted correlation and a maximum jump distance
389   and then calculating the fraction of events meeting both criteria of quality. The significance
390   of the fraction of events meeting both criteria was then determined by comparing it against
391   a distribution of such fractions generated by sets of the time-bin shuffled events. We
392   systematically varied both thresholds and found that the actual events are of significantly
393   higher quality than chance for a wide range of thresholds in both the hippocampal (Figure
394   4b) and simulated (Figure 4d) data. The upper right corner of these grids cannot be

14

395   significant since 100% of all possible events would be included in any shuffle or actual set.
396   Points in the left-most column are not all significant because the strictness of the maximum
397   jump distance means that very few events in either the actual or shuffled data sets meet the
398   criterion, and therefore the analysis is underpowered. This pattern is similar to that seen in
399   Farooq et al., 2019 (as shown in their Figure 1e).

400   Both PBEs and preplay are significantly affected by the two network parameters (Figure 4e,
401   f). The number of clusters and the extent of cluster overlap (indicated via mean cluster
402   participation) affects PBE participation (Figure 4e, top left), firing rates (Figure 4e, top
403   right), event durations (Figure 4e, bottom left), and event frequency (Figure 4e, bottom
404   right). We find that significant preplay occurs only at moderate cluster overlap (Figure 4f,
405   top left), where we also find the greatest increase from chance in the linearity of decoded
406   trajectories (Figure 4f, top right). The fraction of events that are individually significant
407   (determined by comparing the absolute weighted correlation of each decoded event
408   against the set of absolute weighted correlations of its own shuffles) is similarly highest for
409   modest cluster overlap (Figure 4f, bottom left). The mean entropy of position probability of
410   each time bin of decoded trajectories is also highest for modest cluster overlap (Figure 4f,
411   bottom right), meaning that high cluster overlap leads to more diffuse, less precise spatial
412   decoding.

413   To test the robustness of our results to variations in input types, we simulated alternative
414   forms of spatially modulated feedforward inputs. We found that with no parameter tuning
415   or further modifications to the network, the model generates robust preplay with
416   variations on the spatial inputs, including inputs of three linearly varying cues (Figure 4—
417   figure supplement 4a) and two stepped cues (Figure 4—figure supplement 4b-c). The
418   network is impaired in its ability to produce preplay with binary step location cues (Figure
419   4—figure supplement 4d), when there is no cluster bias (Figure 4—figure supplement 4e),
420   and at greater values of cluster participation (Figure 4—figure supplement 4f)

421   **Preplay is due to successive activations of individual clusters**

422

15
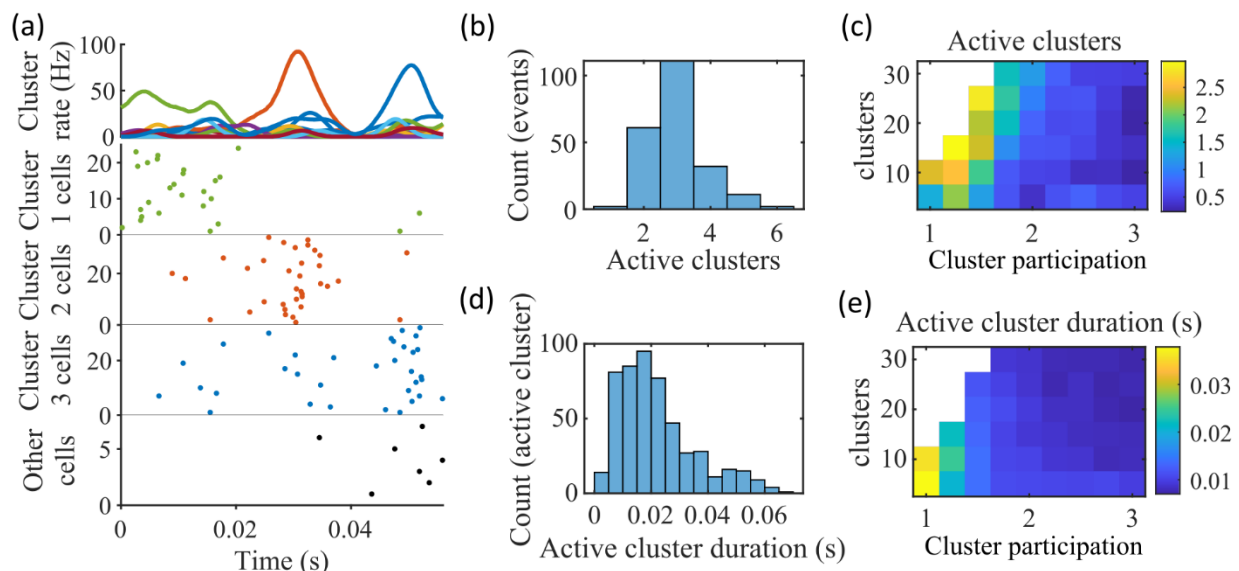
Figure 5: Coherent spiking within clusters supports preplay

(a) Example event. Top, spike rates averaged across neurons of individual clusters: Each firing rate curve is the smoothed mean firing rate across the population of cells belonging to each cluster. We defined clusters as "active" if at any point their rates exceed twice that of any other cluster. Three clusters meet the criterion of being active (green, then red, then blue). Bottom, raster plots: Cells belonging to each of the active clusters are plotted separately in the respective colors. Cells in multiple clusters contribute to multiple population curves, and cells in multiple active clusters appear in multiple rows of the raster plot. Cells that participate but are not in any active clusters are labeled "Other cells" and plotted in black. Only active cells are plotted. **(b)** For the fiducial parameter set (15 clusters, mean cluster participation of 1.25), the distribution over events of the number of active clusters per event. **(c)** The mean number of active clusters per event as a function of the network parameters. Same data as that used for the parameter grids in earlier figures. **(d)** For the fiducial parameter set (15 clusters, mean cluster participation of 1.25), the distribution of durations of active clusters for all active cluster periods across all events. The active duration was defined as the duration for which an active cluster remained the most-active cluster. **(e)** The mean active cluster duration as a function of the network parameters.


Figure 4f indicates that PBEs are best decoded as preplay when cluster participation is only slightly above one, indicating a small, but non-zero, degree of cluster overlap. We hypothesized that this can be explained as balancing two counteracting requirements: 1) Sufficient cluster overlap is necessary for a transient increase in activity in one cluster to induce activity in another cluster, so as to extend any initiated trajectory; and 2) Sufficient cluster isolation is necessary so that, early in a transient, spikes from an excited cluster preferentially add excitement to the same cluster. A network with too much cluster overlap will fail to coherently excite individual clusters—rendering decoded positions to be spread randomly throughout the track—while a network with too little cluster overlap will fail to excite secondary clusters—rendering decoded positions to remain relatively localized.

We find that the dependence of preplay on cluster overlap can indeed be explained by the manner in which clusters participate in PBEs (Figure 5). An example PBE (Figure 5a) shows transient recruitment of distinct clusters, with only one cluster prominently active at a time. We define a cluster as 'active' if its firing rate exceeds twice the rate of any other cluster. We calculated the number of active clusters per event (Figure 5b) and the duration of each active cluster period (Figure 5d). We find that these statistics vary systematically with the network parameters (Figure 5c, e), in a manner consistent with the dependence of preplay on cluster overlap (Figure 4f). When there is modest overlap of an intermediate number of clusters, events involve sequential activation of multiple clusters that are each active sufficiently long to correspond to at least one of the time bins used for decoding (10 ms). Figures 4 and 5 together indicate that high-quality preplay arises via a succession of individually active clusters. Such succession requires a moderate degree of cluster overlap, but this must be combined with sufficient cluster isolation to promote independent activation of just one cell assembly for the duration of each time-bin used for decoding.

16

466　The results of Figure 5 suggest that cluster-wise activation may be crucial to preplay. One
467　possibility is that the random overlap of clusters in the network spontaneously produces
468　biases in sequences of cluster activation which can be mapped onto any given environment.
469　To test this, we looked at the pattern of cluster activations within events. We found that
470　sequences of three active clusters were not more likely to match the track sequence than
471　chance (Figure 5—figure supplement 1a). This suggests that preplay is not dependent on a
472　particular biased pattern in the sequence of cluster activation. We then asked if the number
473　of clusters that were active influenced preplay quality. We split the preplay events by the
474　number of clusters that were active during each event and found that the median preplay
475　shift relative to shuffled events with the same number of active clusters decreased with the
476　number of active clusters (Spearman's rank correlation, p=0.0019, ρ=-0.13; Figure 5—
477　figure supplement 1b).

478　**Cluster identity is sufficient for preplay**

479
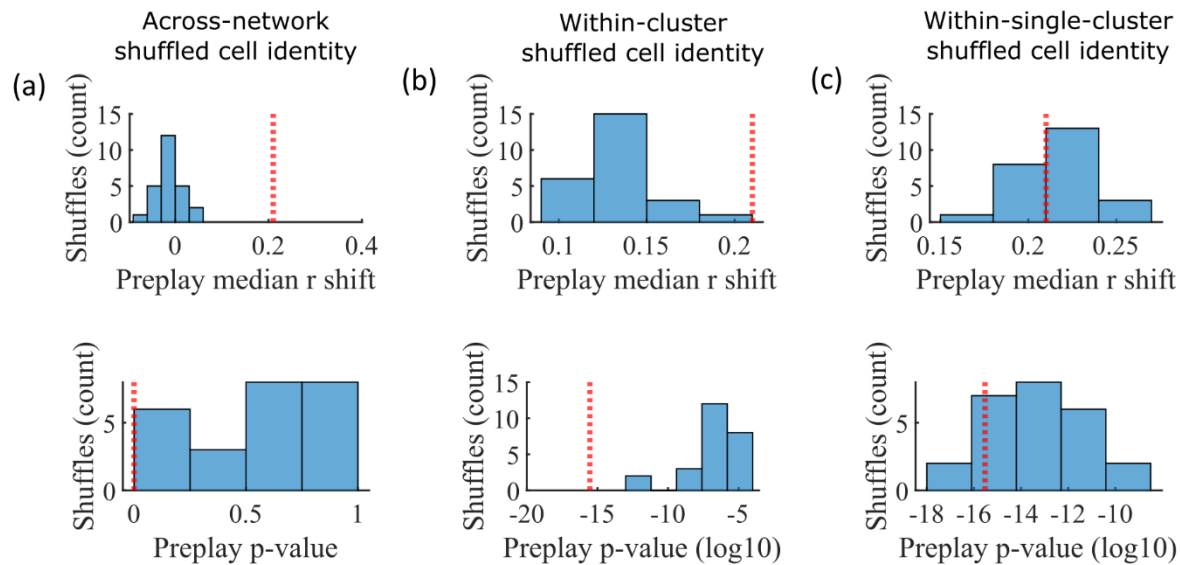
480　**Figure 6: Preplay is abolished when events are decoded with shuffled cell identities**
481　**but is preserved if cell identities are shuffled only within clusters.**
482　We decoded the population burst events from the fiducial parameter set simulations after
483　randomly shuffling cell identities in three different manners (a-c, 25 replicates for each
484　condition) and compared the resulting preplay statistics to the unshuffled result (red line).
485　**(a)** Randomly shuffling cell identities results in median preplay correlation shifts near zero
486　(top, 100th percentile of shuffles), with p-values distributed approximately uniformly
487　(bottom, 0th percentile of shuffles). **(b)** Randomly shuffling cell identities within clusters
488　reduces the magnitude of the median preplay correlation shifts (top, 100th percentile of
489　shuffles) but preserves the statistical significance of preplay (bottom, 0th percentile of
490　shuffles). **(c)** Randomly shuffling cell identities within clusters for only cells that belong to
491　a single cluster results in median preplay correlation shifts that are similar to the
492　unshuffled result (top, 36th percentile of shuffles) and are all statistically significant

17

493   (bottom, 12th percentile of shuffles).

494   ――

495   The pattern of preplay significance across the parameter grid in Figure 4f shows that
496   preplay only occurs with modest cluster overlap, and the results of Figure 5 show that this
497   corresponds to the parameter region that supports transient, isolated cluster-activation.
498   This raises the question of whether cluster-identity is sufficient to explain preplay. To test
499   this, we took the sleep simulation population burst events from the fiducial parameter set
500   and performed decoding after shuffling cell identity in three different ways. We found that
501   when the identity of all cells within a network are randomly permuted the resulting median
502   preplay correlation shift is centered about zero (t-test 95% confidence interval, -0.2018 to
503   0.0012) and preplay is not significant (distribution of p-values is consistent with a uniform
504   distribution over 0 to 1, chi-square goodness-of-fit test p=0.4436, chi-square statistic=2.68;
505   Figure 6a). However, performing decoding after randomly shuffling cell identity between
506   cells that share membership in a cluster does result in statistically significant preplay for all
507   shuffle replicates, although the magnitude of the median correlation shift is reduced for all
508   shuffle replicates (Figure 6b). The shuffle in Figure 6b does not fully preserve cell's cluster
509   identity because a cell that is in multiple clusters may be shuffled with a cell in either a
510   single cluster or with a cell in multiple clusters that are not identical. Performing decoding
511   after doing within-cluster shuffling of only cells that are in a single cluster results in
512   preplay statistics that are not statistically different from the unshuffled statistics (t-test
513   relative to median shift of un-shuffled decoding, p=0.1724, 95% confidence interval of -
514   0.0028 to 0.0150 relative to the reference value; Figure 6c). Together these results
515   demonstrate that cluster-identity is sufficient to produce preplay.

18

Mean relative spike rank correlates with place field location
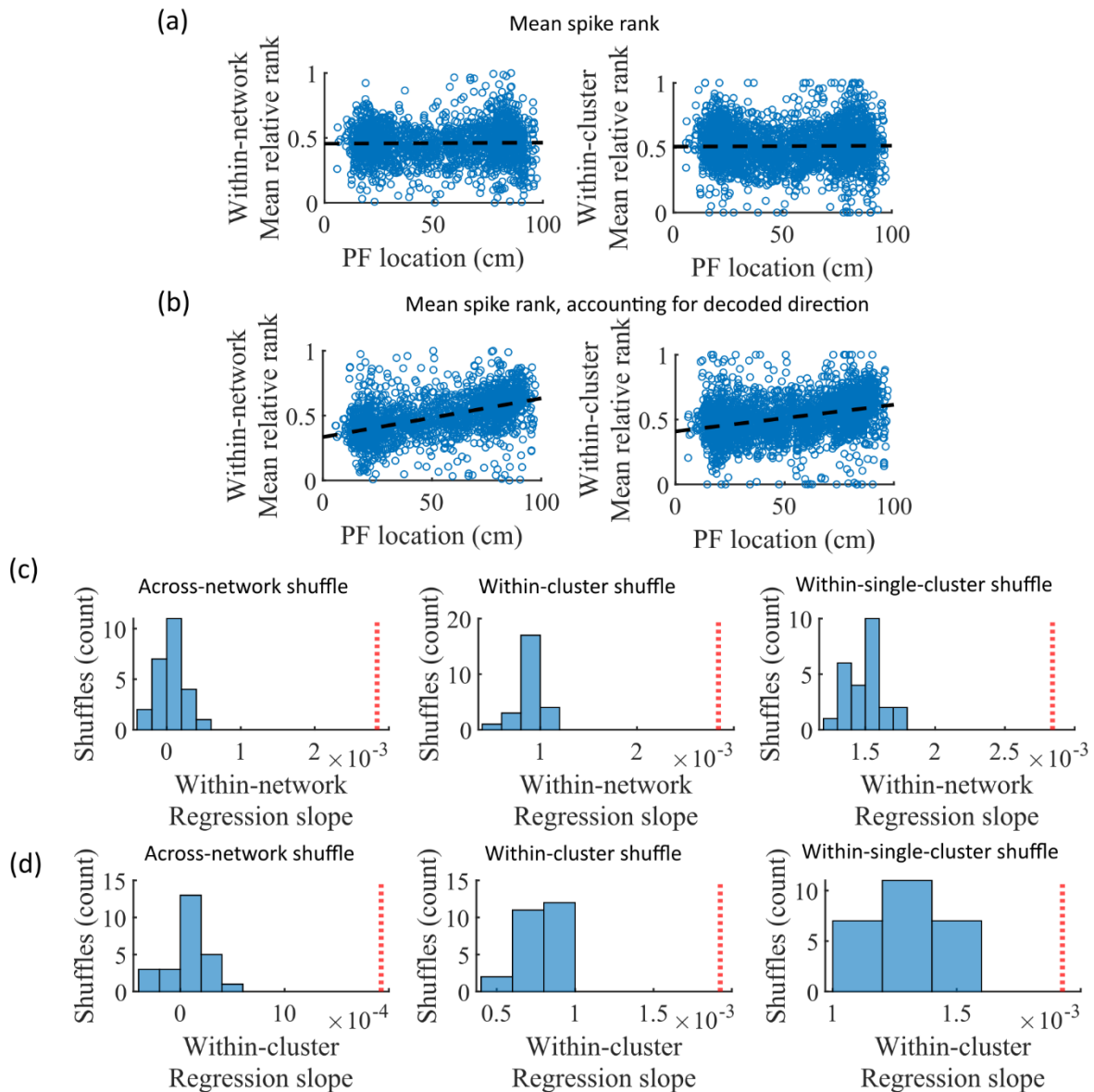
Figure 7: Place cells' mean event rank are correlated with their place field location when accounting for decode direction.
(a) Mean within-event relative spike rank of all place cells as a function of the location of their mean place field density on the track for networks at the fiducial parameter set. Left, mean relative rank with respect to all cells in each network. Right, mean relative rank with respect to only cells that share cluster membership. **(b)** Same as (a), but after accounting for the direction of each events' decoded trajectory. If the decoded slope for a given event was negative, then the order of spiking in that event was reversed. **(c-d)** Comparison of the regression slopes from (b) to the distribution of slopes that results from applying the same analysis after shuffling cell identities as in Figure 6. **(c)** The within-network regression

19

528   slope is significant relative to all three methods of shuffling cell identity. **(d)** Same as (c),
529   but for the within-cluster regression slope.

530   ————————————————————————————————————

531   While cluster-identity is sufficient to produce preplay (Figure 6b), the shuffle of Figure 6c is
532   incomplete in that cells belonging to more than one cluster are not shuffled. Together, these
533   two shuffles leave room for the possibility that individual cell-identity may contribute to
534   the production of preplay. It might be the case that some cells fire earlier than others, both
535   on the track and within events. To test the contribution of individual cells to preplay, we
536   calculated for all cells in all networks of the fiducial parameter point their mean relative
537   spike rank and tested if this is correlated with the location of their mean place field density
538   on the track (Figure 7). We find that there is no relationship between a cell's mean relative
539   within-event spike rank and its mean place field density on the track (Figure 7a). This is the
540   case when the relative rank is calculated over the entire network (Figure 7, "Within-
541   network") and when the relative rank is calculated only with respect to cells with the same
542   cluster membership (Figure 7, "Within-cluster"). However, because preplay events can
543   proceed in either track direction, averaging over all events would average out the sequence
544   order of these two opposite directions. We performed the same correlation but after
545   reversing the spike order for events with a negative slope in the decoded trajectory (Figure
546   7b). To test the significance of this correlation, we performed a bootstrap significance test
547   by comparing the slope of the linear regression to the slope that results when performing
548   the same analysis after shuffling cell identities in the same manner as in Figure 6. We found
549   that the linear regression slope is greater than expected relative to all three shuffling
550   methods for both the within-network mean relative rank correlation (Figure 6c) and the
551   within-cluster mean relative rank correlation (Figure 6d).

Small-world index correlates with preplay
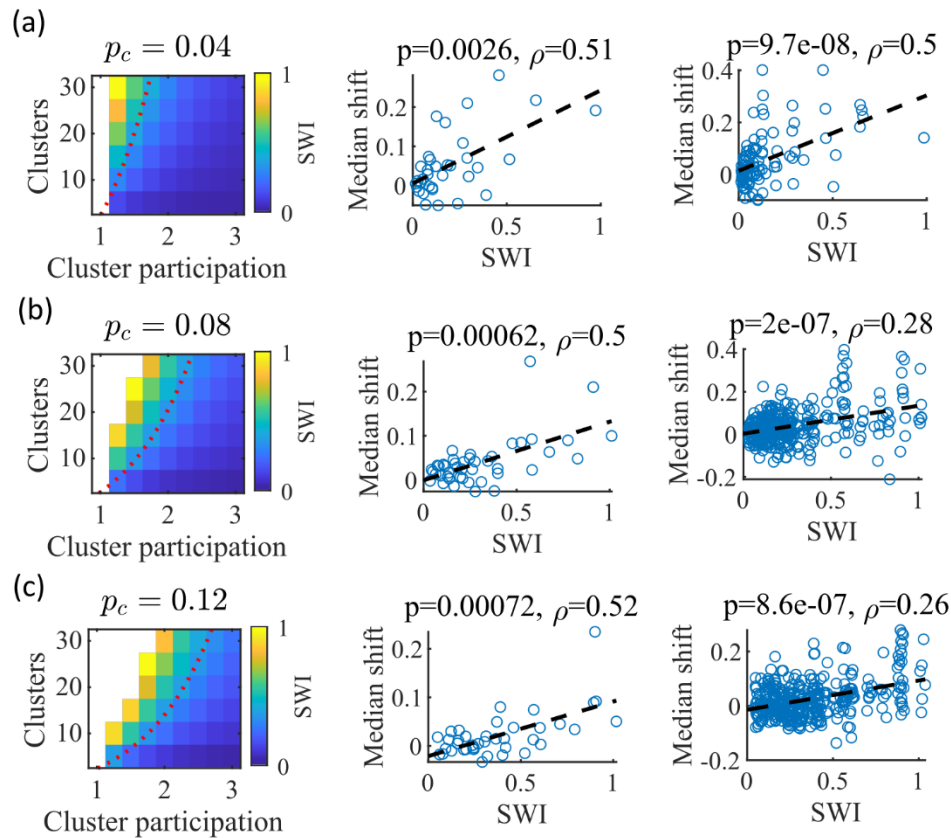
Figure 8: The Small-World Index of networks correlates with preplay quality
(a-c) Left column, the Small-World Index (SWI; plotted as color) is affected by the global E-to-E connection probability, p_c. Red dotted line indicates a contour line of SWI $= 0.4$. This boundary shifts downward as p_c increases. Center column, across parameter points in the network parameter grid, SWI correlates with an increase in the median absolute weighted correlation of decoded trajectories relative to shuffles (e.g. this corresponds in Figure 4c to the rightward shift of the CDF of measured absolute weighted correlations relative to the shuffle events). Each point is produced by analysis of all events across 10 networks from one parameter point in the grid on the left. Right column, same as the center column but each point is data from each of the 10 individual networks per parameter set. P-value and correlation, ρ, are calculated from Spearman's rank-order correlation test. Dashed line is the least-squares fit. **(a)** Data from a parameter grid where the E-to-E connection probability was decreased by 50% and the E-to-E connection strength was doubled from their fiducial values used in prior figures. **(b)** Data from the same parameter grid as Figures 3-5. **(c)** Data from a parameter grid where the E-to-E connection probability was increased by 50% and the E-to-E connection strength scaled by two-thirds from their fiducial values.

We noticed that that the highest quality of decoded trajectories (Figure 4f) seemed to arise in networks with the highest small-world index (SWI; Figure 1g). In order to test this, we

573   simulated different sets of networks with both increased and decreased global E-to-E
574   connection probability, p_c. Changing p_c, in addition to varying the number of clusters and
575   the mean cluster participation, impacted the SWI of the networks (Figure 8, left column).

576   We hypothesized that independent of p_c, a higher SWI would correlate with improved
577   preplay quality. To test this, we simulated networks across a range of parameters for three
578   p_c values: a decrease of p_c by 50% to 0.04, the fiducial value of 0.08, and an increase by
579   50% to 0.12 (Figure 8a-c, respectively). For the decreased and increased p_c cases, the E-to-
580   E connection strength was respectively doubled or reduced to 2/3 of the fiducial strength
581   to keep total E-to-E input constant. For each parameter combination, we quantified preplay
582   quality as the rightward shift in median absolute weighted correlation of decoded preplay
583   events versus shuffled events (as in Figure 4f, top right). We then asked if there was a
584   correlation between that quantification of preplay quality and SWI.

585   Across all three p_c values, SWI significantly correlated with improved preplay both across
586   parameter sets (Figure 8, center column) and across individual networks (Figure 8, right
587   column). These results support our prediction that higher small-world characteristics
588   correspond to higher-quality preplay dynamics regardless of average connectivity.

Preplay significantly decodes to linear trajectories in arbitrary environments
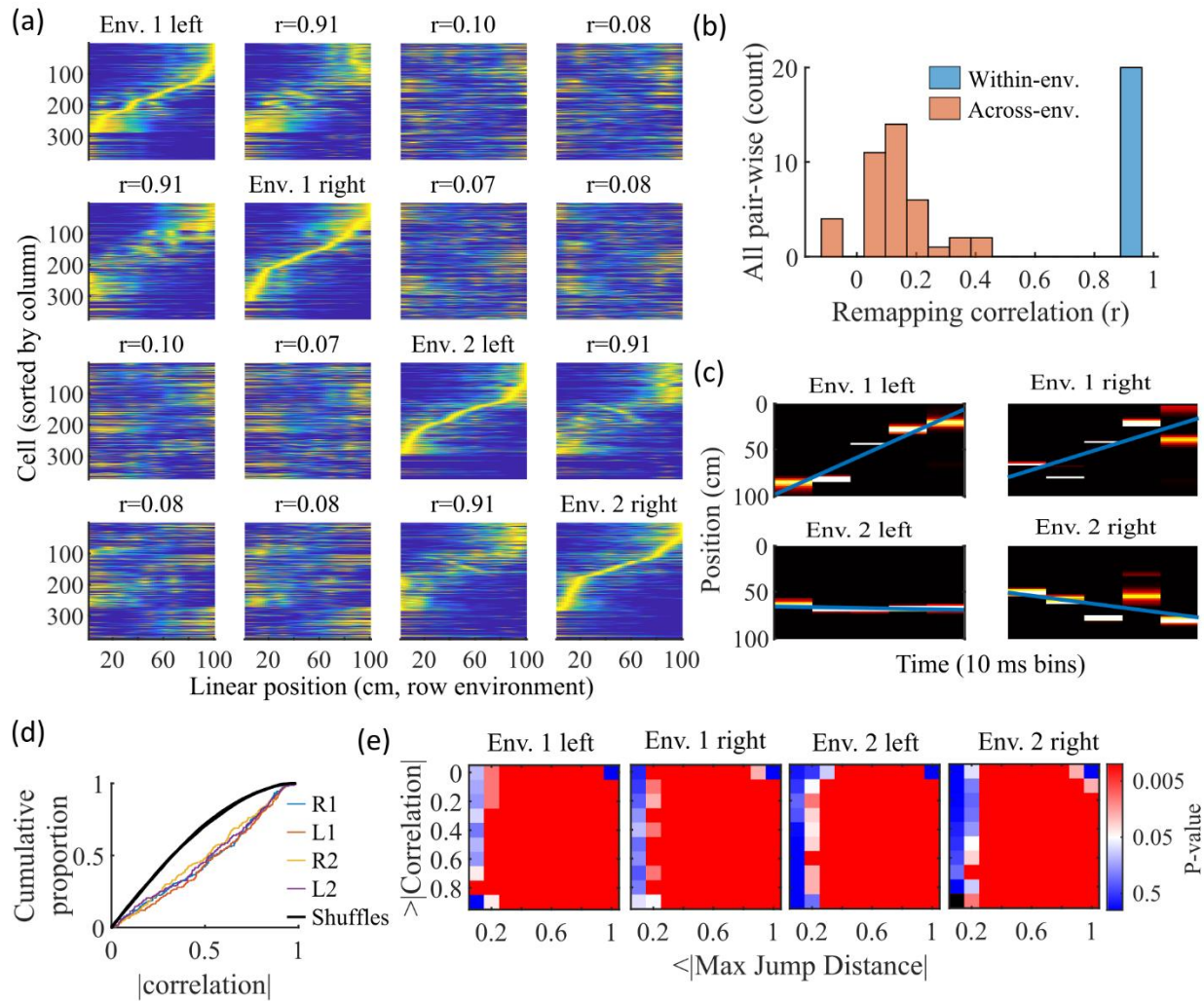
Figure 9: Trajectories decoded from population-burst events are significantly correlated with linear trajectories in arbitrary environments
(a) Place fields from a single network with simulated runs in both directions of travel on a linear track in two different environments. Each column of panels is the set of place fields for the trajectory labeled on the diagonal. Each row of panels has cells sorted by the order of place-field peaks for the trajectory labeled on the diagonal. The r values are the correlations between the corresponding remapped trajectory with its comparison on the diagonal. Note that correlations mirrored across the diagonal are equal because they correspond only to a change in the labels of the dimensions of the population rate vectors, which does not affect the vector correlation. **(b)** Distribution of the place-field map correlations across trajectories from both directions of travel on a linear track in two environments for 10 networks. Blue is the distribution of correlations for all left vs right place-field maps from the same environment. Red is the correlations from all pair-wise comparisons of trajectories from different environments. **(c)** An example event with a statistically significant trajectory when decoded with place fields from Env. 1 left (absolute

606 correlation at the 99th percentile of time-bin shuffles) but not when decoded with place
607 fields of the other trajectories (78th, 45th, and 63rd percentiles for Env. 1 right, Env. 2 left,
608 and Env. 2 right, respectively). **(d)** An entire set of PBEs shows similar levels of absolute
609 weighted correlations when decoded with different sets of place fields. In color are CDFs of
610 absolute weighted correlations of decoded trajectories with leftward and rightward linear
611 trajectories in each of the two environments (R1 and L1 are the rightward and leftward
612 trajectories of environment one. R2 and L2 are the rightward and leftward trajectories of
613 environment two). In black (all overlapping) are the corresponding absolute weighted
614 correlations with each of the 4 trajectories arising from decoding of shuffled events. **(e)**
615 The significance of linearity of decoded trajectories indicated by p-value in color (as in
616 Figure 4b) from decoding the same PBEs with the four different environment place fields.
617 Black squares indicate criteria that were not met by any events (either shuffled or actual).
618 Env. 1 left is the same as that shown in Figure 4d.

619 _____

620 Information about each environment enters the network via the feed-forward input
621 connection strengths, which contain cluster-dependent biases. A new environment is
622 simulated by re-ordering those input biases. We first wished to test that a new
623 environment simulated in such a manner produced a distinct set of place fields. We
624 therefore simulated place maps for leftward and rightward trajectories on linear tracks in
625 two distinct environments (Figure 9a). The two maps with different directions of motion
626 showed very high correlations when in the same environment (Figure 9b, blue) while the
627 comparisons of trajectories across environments show very low correlations (Figure 9b,
628 red). Cells that share membership in a cluster will have some amount of correlation in their
629 remapping due to the cluster-dependent cue bias, which is consistent with experimental
630 results (Hampson et al., 1996; Pavlides et al., 2019), but the combinatorial nature of cluster
631 membership renders the overall place field map correlations low (Figure 9b). We also
632 performed simulations with extra laps of running and calculated the correlations between
633 paired sets of place fields produced by random, independent splits of trials of the same
634 trajectory. The distribution of these correlations was similar to the distribution of within-
635 environment correlations (comparing opposite trajectories with the same spatial input),
636 showing no significant *de novo* place-field directionality. This is consistent with
637 hippocampal data in which place-field directionality is initially low in novel environments
638 and increases with experience (Frank et al., 2004; Navratilova et al., 2012; Shin et al.,
639 2019).

640 Because we simulated preplay without any location-specific inputs, we expected that the
641 set of spiking events that significantly decode to linear trajectories in one environment
642 (Figure 4) should decode with a similar fidelity in another environment. Therefore, we
643 decoded each PBE four times, once with the place fields of each trajectory (Figure 9c-e).
644 Since the place field map correlations are high for trajectories on the same track and near
645 zero for trajectories on different tracks, any individual event would be expected to have
646 similar decoded trajectories when decoding based on the place fields from different
647 trajectories in the same environment and dissimilar decoded trajectories when decoding
648 based on place fields from different environments. A given event with a strong decoded
649 trajectory based on the place fields of one environment would then be expected to have a

24

650 weaker decoded trajectory when decoded with place fields from an alternative
651 environment (Figure 9c). The distributions of absolute weighted correlations arising from
652 decoding of PBEs according to each of the four sets of place fields was consistent across
653 environments (Figure 9d, colored lines) and all were significantly rightward shifted
654 (indicating greater absolute weighted correlation) when compared to those absolute
655 weighted correlations arising from the corresponding shuffled events (Figure 9d,
656 overlapping black lines). If we consider both absolute weighted correlation and jump-
657 distance thresholds as in Figure 4d, we find that the matrices of p-values are consistent
658 across environments (Figure 9e). In summary, without environment-specific or place-field
659 dependent pre-assigned internal wiring, the model produces population-burst events,
660 which, as an ensemble, show significant preplay with respect to any selected environment.

Discussion

661

662 Our work shows that spontaneous population bursts of spikes that can be decoded as
663 spatial trajectories can arise in networks with clustered random connectivity without pre-
664 configured maps representing the environment. In our proposed model, excitatory neurons
665 were randomly clustered with varied overlap and received feed-forward inputs with
666 random strengths that decayed monotonically from the boundaries of a track (Figure 1).
667 Even though the model neural circuit lacked place-field like input and lacked environment-
668 specific internal wiring, the network exhibited both realistic place fields (Figures 2,3) and
669 spontaneous preplay of novel, future environments (Figures 2,4).

670 We validated our modeling results by applying the same analyses to a previously collected
671 experimental data set (Shin et al., 2019). Indeed, we replicated the general finding of
672 hippocampal preplay found previously in Farooq et al., 2019, although the p-value matrix
673 for our experimental data (Figure 4b) is significant across a smaller range of threshold
674 values than found in their prior work. This is likely due to differences in statistical power.
675 The pre-experience sleep sessions of Shin et al., 2019 were not longer than half an hour for
676 each animal, while the pre-experience sleep sessions of Farooq et al., 2019 lasted 2-4 hours.
677 However, finding statistically significant hippocampal preplay in an experiment not
678 designed for studying preplay shows that the general result is robust to a number of
679 methodological choices, including shorter recording sessions, use of a W-track rather than
680 linear track, and variations in candidate event detection criterion.

681 Although our model is a model of the recurrently connected CA3 region and the data set we
682 analyze (Shin et al., 2019) comes from CA1 cells, the qualitative comparisons we make here
683 are nevertheless useful. Despite some statistically significant quantitative differences, the
684 general properties of place fields that we consider are qualitatively similar across CA1 and
685 CA3 (Sheintuch et al., 2023; Harvey et al., 2020), and CA3 and CA1 generally reactivate in a
686 coordinated manner (O'Neil et al., 2008; Karlsson and Frank, 2009).

687 The model parameters that controlled the clustering of the recurrent connections strongly
688 influenced preplay and place-field quality. Moderate overlap of clusters balanced the
689 competing needs for both a) sufficiently isolated clusters to enable cluster-wise activation
690 and b) sufficiently overlapping clusters to enable propagation of activity across clusters
691 (Figure 5). In our clustered network structure, such a balance in cluster overlap produces
692 networks with small-world characteristics (Watts and Strogatz, 1998) as quantified by a
693 small-world index (SWI; Neal, 2015; Neal, 2017). Networks with a high SWI, indicating high
694 clustering (if two neurons are connected to the same third neuron, they are more likely
695 than chance to be connected to each other) yet short paths (the mean number of
696 connections needed to traverse from one neuron to any other), showed optimal preplay
697 dynamics (Figure 8). The same networks could flexibly represent distinct remapped
698 environments (Leutgeb et al., 2004; Leutgeb et al., 2005; Alme et al., 2014) solely through
699 differences in scaling of feed-forward spatially linear input (Figure 9).

700 Across many species, small-world properties can be found at both the local neuronal
701 network scale and the gross scale of the network of brain regions. At the neuronal
702 connection scale, small-world properties have been reported in a number of networks,

703   such as the C. elegans connectome (Watts and Strogatz, 1998; Humphries et al., 2008), the
704   brainstem reticular formation (Humphries et al., 2006), mouse visual cortex (Sadovsky and
705   MacLean, 2014), cultured rat hippocampal neurons (Antonello et al., 2022), mouse
706   prefrontal cortex (Luongo et al., 2016), and connectivity within the entorhinal-
707   hippocampal region in rats (She et al., 2016). At the level of connected brain regions, small-
708   world properties have been reported across the network of brain regions activated by fear
709   memories in mice (Vetere et al., 2016), in the hippocampal-amygdala network in humans
710   (Zhang et al., 2022), and across the entire human brain (Liao et al., 2010).

711   Our results suggest that the preexisting hippocampal dynamics supporting preplay may
712   reflect general properties arising from randomly clustered connectivity, where the
713   randomness is with respect to any future, novel experience. The model predicts that
714   preplay quality will depend on the network's balance of cluster isolation and overlap, as
715   quantified by small-world properties. Synaptic plasticity in the recurrent connections of
716   CA3 may primarily serve to reinforce and stabilize intrinsic dynamics, which could be
717   established through a combination of developmental programming (Perin et al., 2011;
718   Druckmann et al., 2014; Huszar et al., 2022) and past experiences (Bourjaily and Miller,
719   2011), rather than creating spatial maps *de novo*. The particular neural activity associated
720   with a given experience would then selectively reinforce the relevant intrinsic dynamics,
721   while leaving the rest of the network dynamics unchanged.

722   Our model provides a general framework for understanding the origin of pre-configured
723   hippocampal dynamics. Hebbian plasticity on independent, previously experienced place
724   maps would produce effectively random clustered connectivity. The spontaneous dynamics
725   of such networks would influence expression of place fields in future, novel environments.
726   Together with intrinsic sequence generation, this could enable preplay and immediate
727   replay generated by the preexisting recurrent connections.

728   Future modeling work should explore how experience-dependent plasticity may leverage
729   and reinforce the dynamics initially expressed through preexisting clustered recurrent
730   connections to produce higher-quality place fields and decoded trajectories during replay
731   (Shin et al., 2019; Farooq et al., 2019). Plasticity may strengthen connectivity along
732   frequently reactivated spatiotemporal patterns. Clarifying interactions between intrinsic
733   dynamics and experience-dependent plasticity will provide key insights into hippocampal
734   neural activity. Additionally, the *in vivo* microcircuitry of CA3 is complex and includes
735   aspects such as nonlinear dendritic computations and a variety of inhibitory cell types
736   (Rebola et al., 2017). This microcircuitry is crucial for explaining certain aspects of
737   hippocampal function, such as ripple and gamma oscillogenesis (Ramirez-Villegas et al.,
738   2017), but here we have focused on a minimal model that is sufficient to produce place cell
739   spiking activity that is consistent with experimentally measured place field and preplay
740   statistics.

27

Methods

To investigate what network properties could support preplay, we simulated recurrently connected networks of spiking neurons and analyzed their dynamics using standard hippocampal place cell analyses.

Neuron model

We simulate networks of Leaky Integrate-and-Fire (LIF) neurons, which have leak conductance, g_L, excitatory synaptic conductance, g_E, inhibitory synaptic conductance, g_I, spike-rate adaptation (SRA) conductance, g_{SRA}, and external feed-forward input synaptic conductance, g_{ext}. The membrane potential, V, follows the dynamics

$$\tau_m \frac{dV}{dt} = -g_L(V - E_L) - g_E(V - E_E) - g_I(V - E_I) - g_{SRA}(V - E_{SRA}) - g_{ext}(V - E_E)$$

where τ_m is the membrane time constant, E_L is the leak reversal potential, E_E is the excitatory synapse reversal potential, E_I is the inhibitory synapse reversal potential, E_{SRA} is the SRA reversal potential, and E_{ext} is the external input reversal potential. When the membrane potential reaches the threshold V_{th}, a spike is emitted and the membrane potential is reset to V_{reset}.

The changes in SRA conductance and all synaptic conductances follow

$$\tau_i \frac{dg_i}{dt} = -g_i$$

to produce exponential decay between spikes for any conductance i. A step increase in conductance occurs at the time of each spike by an amount corresponding to the connection strength for each synapse (W_{E-E} for E-to-E connections, W_{E-I} for E-to-I connections, and W_{I-E} for I-to-E connections), or by δ_{SRA} for g_{SRA}. Initial feed-forward input conductances were set to values approximating their steady-state values by randomly selecting values from a Gaussian with a mean of $W_{in}r_G\tau_E$ and a standard deviation of $\sqrt{W_{in}^2 r_G \tau_E}$. Initial values of the recurrent conductances and the SRA conductance were set to zero.

Parameter	Value	Description
τ_m	40 ms	Membrane time constant
C_m	0.4 nF	Membrane capacitance
d_t	0.1 ms	Simulation time step
g_L	10 nS	Leak conductance
E_L	−70 mV	Leak reversal potential
E_E	0 mV	Excitatory synaptic reversal potential
E_I	−70 mV	Inhibitory synaptic reversal potential
E_{SRA}	−80 mV	SRA reversal potential

28

V_{th}	-50 mV	Spike threshold
V_{reset}	-70 mV	Reset potential
τ_E	10 ms	Excitatory time constant
τ_I	3 ms	Inhibitory time constant
τ_{SRA}	30 ms	Spike-rate adaptation time constant
δ_{SRA}	3 pS	Spike-rate adaptation strength

766

Network structure

768 We simulated networks of $n = 500$ neurons, of which 75% were excitatory. Excitatory
769 neurons were randomly, independently assigned membership to each of n_c clusters in the
770 network. First, each neuron was randomly assigned membership to one of the clusters.
771 Then, each cluster was assigned a number—$n_E(\mu_c - 1)/n_c$ rounded to the nearest
772 integer—of additional randomly selected neurons such that each cluster had identical
773 numbers of neurons, $n_{E,clust} = n_E(\mu_c/n_c)$, and mean cluster participation, μ_c, reached its
774 goal value.

775 E-to-E recurrent connections were randomly assigned on a cluster-wise basis, where only
776 neurons that shared membership in a cluster could be connected. The within-cluster
777 connection probability was configured such that the network exhibited a desired global E-
778 to-E connection probability p_c. Given the total number of possible connections between
779 excitatory neurons is $C_{tot} = n_E(n_E - 1)$ and the total number of possible connections
780 between excitatory neurons within all clusters is $C_{clust} = n_{E,clust}(n_{E,clust} - 1)n_c$, we
781 calculated the within-cluster connection probability as $p_c(C_{tot}/C_{clust})$. That is, given the
782 absence of connections between clusters (clusters were coupled by the overlap of cells) the
783 within-cluster connection probability was greater than p_c so as to generate the desired
784 total number of connections equal to $p_c C_{tot}$.

785 All E-to-I and I-to-E connections were independent of cluster membership and existed with
786 a probability p_{c_I}. There were no I-to-I connections. p_c, n_c, and μ_c were varied for some
787 simulations. Except where specified otherwise, all parameters took the fiducial value
788 shown in the table below.

789 The network visualization in Figure 1c was plotted based on the first 2 dimensions of a t-
790 distributed stochastic neighbor embedding of the connectivity between excitatory cells
791 using the MATLAB function *tsne*. The feature vector for each excitatory cell was the binary
792 vector indicating the presence of both input and output connections.

Parameter	Fiducial Value	Description
n	500	Number of neurons
n_E	375	Number of excitatory neurons

29

n_c or "clusters"	15	Number of clusters
μ_c or "cluster participation"	1.25	Mean cluster membership per neuron
p_c	0.08	E-to-E connection probability
p_{c_I}	0.25	E-to-I and I-to-E connection probability
$W_{E\text{-}E}$	220 pS	E-to-E synaptic conductance step increase
$W_{E\text{-}I}$	400 pS	E-to-I synaptic conductance step increase
$W_{I\text{-}E}$	400 pS	I-to-E synaptic conductance step increase

Network inputs

794 All excitatory neurons in the network received three different feed-forward inputs (Figure
795 1b). Two inputs were spatially modulated, with rates that peaked at either end of the track
796 and linearly varied across the track to reach zero at the opposite end. One input was a
797 context cue that was position independent. All excitatory cells received unique Poisson
798 spike trains from each of the three inputs at their position-dependent rates. Inhibitory cells
799 received only the context input.

800 The connection strength of each feed-forward input to each neuron was determined by an
801 independent and a cluster-specific factor.

802 First, strengths were randomly drawn from a log-normal distribution $e^{\mu+\sigma\mathcal{N}}$, where $\mathcal{N}$ is a

803 zero-mean, unit variance Normal distribution, $\mu = ln\left(\frac{W_{in}^2}{\sqrt{\sigma_{in}+W_{in}^2}}\right)$ and $\sigma = \sqrt{ln\left(\frac{\sigma_{in}}{W_{in}^2+1}\right)}$ for

804 mean strength W_{in} and standard deviation σ_{in} for the location cues, with σ_{in} replaced by
805 $\sigma_{context}$ for the context cue. Each environment and the sleep session had unique context cue
806 input weights. For model simplicity, the mean input strength W_{in} for all inputs was kept the
807 same for both E and I cells in both the awake and sleep conditions, but the strength of the
808 resulting context input was then scaled by some factor f_x for each of the 4 cases to
809 accommodate for the presence, or lack thereof, of the additional current input from the
810 location cues. These scaling factors were set at a level that generated appropriate levels of
811 population activity. During simulation of linear track traversal, the context cue to excitatory
812 cells was scaled down by $f_{E\text{-awake}}$ to compensate for the added excitatory drive of the
813 location cue inputs, and the context cue input to I cells was not changed ($f_{I\text{-awake}} = 1$).
814 During sleep simulation, the context cue input to E cells was not scaled ($f_{E\text{-awake}} = 1$) but
815 the context cue input to I cells was scaled down by $f_{I\text{-sleep}}$.

816 Second, to incorporate cluster-dependent correlations in place fields, a small ($\leq 4\%$)
817 location cue bias was added to the randomly drawn feed-forward weights based on each
818 neuron's cluster membership. For each environment, the clusters were randomly shuffled
819 and assigned a normalized rank bias value, such that the first cluster had a bias of -1
820 (corresponding to a rightward cue preference) and the last cluster had a bias of +1
821 (leftward cue preference). A neuron's individual bias was calculated as the mean bias of all

30

822   clusters it belonged to, multiplied by the scaling factor σ_{bias}. The left cue weight for each
823   neuron was then scaled by 1 plus its bias, and the right cue weight was scaled by 1 minus
824   its bias. In this way, the feed-forward input tuning was biased based on the mean rank of a
825   neuron's cluster affiliations for each environment. The addition of this bias produced
826   correlations in cells' spatial tunings based on cluster membership, but, importantly, this
827   bias was not present during the sleep simulations, and it did not lead to high correlations of
828   place-field maps between environments (Figure 9b).

Parameter	Value	Description
r_G	5000 Hz	Peak Poisson input rate
W_{in}	72 pS	Mean strength of the input synapses
σ_{in}	5 pS	Standard deviation of the location cue input synapses
$\sigma_{context}$	1.25 pS	Standard deviation of the context cue input synapses
σ_{bias}	0.04	Location bias scale
$f_{\text{E-awake}}$	0.1	E-cell context cue input scaling during awake simulation
$f_{\text{E-sleep}}$	1	E-cell context cue input scaling during sleep simulation
$f_{\text{I-awake}}$	1	I-cell context cue input scaling during awake simulation
$f_{\text{I-sleep}}$	0.75	I-cell context cue input scaling during sleep simulation

Simulation

830   For a given parameter set, we generated 10 random networks. We simulated each network
831   for one sleep session of 120 s and for five 2-s long traversals of each of the two linear
832   trajectories on each track. For the parameter grids in Figures 3 and 4 we simulated 20
833   networks with 300 s long sleep sessions in order to get more precise empirical estimates of
834   the simulation statistics. For analysis comparing place-field reliability, we simulated 10
835   traversals of each trajectory.

836   To compare coding for place vs time, we performed repeated simulations for the same
837   networks at the fiducial parameter point with 1.0x and 2.0x of the original track traversal
838   speed. We then combined all trials for both speed conditions to calculate both place fields
839   and time fields for each cell from the same linear track traversal simulations. The place
840   fields were calculated as described below (average firing rate within each of the fifty 2-cm
841   long spatial bins across the track) and the time fields were similarly calculated but for fifty
842   40-ms time bins across the initial two seconds of all track traversals.

843

Place field analysis

Place-field rate maps

846   We followed the methods of Shin et al., 2019 to generate place fields from the spike trains.
847   We calculated for each excitatory cell its trial-averaged occupancy-discounted firing rate in

31

848   each 2 cm spatial bin of the 1 m long linear track. Note that the occupancy-discounting term
849   is uniform across bins, so it has no impact in our model, because we simulated uniform
850   movement speed. We then smoothed this with a Gaussian kernel with a 4 cm standard
851   deviation. For statistics quantifying place-field properties and for Bayesian decoding, we
852   considered only excitatory cells with place-field peaks exceeding 3 Hz as in Shin et al.,
853   2019.

854

855   **Place-field specificity**

856   Place-field specificity was defined as 1 minus the fraction of the spatial bins in which the
857   place field's rate exceeded 25% of its maximum rate (Shin et al., 2019).

858

859   **Place-field spatial information**

860   The spatial information of each cells' place field was calculated as

861
$$\text{Spatial Information} = \sum_i p_i \left(\frac{r_i}{\bar{r}}\right) log_2 \left(\frac{r_i}{\bar{r}}\right)$$

862   where p_i is the probability of being in spatial bin i, r_i is the place field's rate in spatial bin i,
863   and $\bar{r}$ is the mean rate of the place field (Sheintuch et al., 2023). Given the division of the
864   track into 50 spatial bins, spatial information could vary between 0 for equal firing in all
865   bins and $log_2(50) \cong 5.6$ for firing in only a single bin. Spatial information of 1 is equivalent,
866   for example, to equal firing in exactly one half of the bins and no firing elsewhere.

867

868   **Distribution of peaks**

869   We used two measures to quantify the extent to which place-field peaks were uniformly
870   distributed across the track. In our first measure, we calculated the Kullback-Leibler
871   divergence of the distribution of peaks from a uniform distribution, as

872
$$D_{KL} = -\sum_i p_i^{\text{data}} log_2 \left(\frac{p_i^{\text{uniform}}}{p_i^{\text{data}}}\right)$$

873   where p_i^{data} is the fraction of cells with peak firing rates in the ith spatial bin and $p_i^{uniform}$
874   is 1/50, *i. e.,* the fraction expected from a uniform distribution (Sheintuch et al., 2023).
875   Similarly, the range for spatial information, D_{KL} is bounded between zero for a perfectly
876   uniform distribution of peaks and $log_2(50) \cong 5.6$ if all peaks were in a single bin. D_{KL} of 1
877   is equivalent, for example, to all peaks being uniformly spread over one half of the bins in
878   the track.

879   For our second measure, we calculated the fraction of place cells whose peak firing rate
880   was in the central third of the track. Since inputs providing spatial information only peaked

881 at the boundaries of the track, the central third was ubiquitously the most depleted of high
882 firing rates.

883

Place-field map correlations

885 To compare the similarity of place fields across different trajectories, we calculated the
886 correlation between the place-field rate maps of each pair of trajectories. For each spatial
887 bin, we calculated the Pearson correlation coefficient between the vector of the population
888 place-field rates of the two trajectories. We then averaged the correlation coefficients
889 across all spatial bins to get the correlation between the two trajectories.

890

PBE detection

892 We detected candidate preplay events in the simulated data by identifying population-
893 burst events (PBEs). During the simulated sleep period, we calculated the mean rate of the
894 population of excitatory cells, which defines the population rate, smoothed with a Gaussian
895 kernel (15 ms standard deviation). We then detected PBEs as periods of time when the
896 population rate exceeded 1 standard deviation above the mean population rate for at least
897 30 ms. We also required the peak population rate to exceed 0.5 Hz (corresponding to 5-6
898 spikes per 30ms among excitatory cells) in order for the rate fluctuation to qualify as a PBE.
899 We then combined PBEs into a single event if their start and end times were separated by
900 less than 10 ms.

901

Sharp-wave ripple detection

903 Because of the reduced number of recorded cells relative to the simulated data, we
904 detected candidate events in the Shin et al., 2019 data with a method that incorporated the
905 ripple band oscillation power in the local field potential (LFP) in addition to the population
906 spiking activity. We first calculated the smoothed firing rate for each excitatory neuron by
907 convolving its spikes with a Gaussian kernel (100 ms standard deviation) and capping at 1
908 to prevent bursting dominance. We then computed the z-scored population firing rate from
909 the capped, smoothed single-neuron rates. Additionally, we calculated the z-scored, ripple-
910 filtered envelope of the tetrode-averaged LFP. We then summed these two z-scores and
911 detected peaks that exceeded 6 for at least 10 ms and exceeded the neighboring regions by
912 at least 6 (*MinPeakHeight*, *MinPeakWidth*, and *MinPeakProminence* of the MATLAB function
913 *findpeaks*, respectively). Candidate events were defined as periods around detected peaks,
914 spanning from when the z-score sum first dipped below 0 for at least 5 ms before the peak
915 to after the peak when it again dipped below 0 for at least 5 ms. We additionally required
916 that the animal be immobile during the event.

917

33

Bayesian decoding

We performed Bayesian decoding of candidate preplay events following the methods of Shin et al., 2019. We performed decoding on all candidate events that had at least 5 active cells and exceeded at least 50 ms in duration. Spikes in the event were binned into 10 ms time bins. We decoded using the place fields for each trajectory independently. The description provided below is for the decoding using the place fields of one particular trajectory.

For each time bin of each event, we calculated the location on the track represented by the neural spikes based on the place fields of the active cells using a memoryless Bayesian decoder

$$P(x|s) = \frac{P(s|x)P(x)}{P(s)}$$

where $P(x|s)$ is the probability of the animal being in spatial bin x given the set of spikes s that occurred in the time bin, $P(s|x)$ is the probability of the spikes s given the animal is in spatial bin x (as given by the place fields), $P(x)$ is the prior probability of the animal being in spatial bin x, and $P(s)$ is the probability of the spikes s.

We assumed a uniform prior probability of position, $P(x)$. We assumed that the N cells firing during the event acted as independent Poisson processes in order to calculate

$$P(s|x) = \prod_i^N \frac{(\tau r_i(x))^{s_i} e^{-\tau r_i(x)}}{s_i!}$$

where τ is the time bin window duration (10 ms), $r_i(x)$ is the place-field rate of cell i in spatial bin x and s_i is the number of spikes from cell i in the time bin.

This allows us to calculate the posterior probability of position for each time bin as

$$P(x|s) = C \left(\prod_i^N r_i(x)^{s_i} \right) e^{-\tau \sum_i^N r_i(x)}$$

where C is a normalization constant, which accounts for the position-independent term, $P(s)$.

Bayesian decoding statistical analyses

We analyzed the significance of preplay using the methods of Farooq et al., 2019 (see also Silva et al., 2015). We computed two measures of the sequence quality of each decoded event: the event's absolute weighted correlation and its jump distance. The absolute weighted correlation is the absolute weighted Pearson's correlation of decoded position across the event's time bins. For each decoded event, we calculate the weighted correlation between space and time with MATLAB's *fitlm* function using the decoded probability in

950 each space-time bin (10 ms by 2 cm) as the weight for the corresponding location in the
951 correlation. The absolute value of the weighted correlation is used in order to account for
952 both forward and reverse preplay. The jump distance is the maximum of the distance
953 between the positions of peak probability for any two adjacent 10-ms time bins in the
954 event, quantified as fraction of the track length.

955 For each event, we generated 100 shuffled events by randomly permuting the order of the
956 10-ms time bins. We then calculated the weighted correlation and jump distance for each
957 shuffled event in the same manner as for the actual events. For each simulated parameter
958 set, we combined all events from the 10 simulated networks.

959 Following the methods of Farooq et al., 2019, we calculated the statistical significance of
960 the population of preplay events using two different methods. First, we used the
961 Kolmogorov-Smirnov (KS) test to compare the distributions of absolute weighted
962 correlations obtained from the actual events and the shuffled events (Figure 4a, c).

963 Second, we used a bootstrap test to compare the fraction of high-quality events—defined as
964 having both high absolute weighted correlations and low maximum jump distance—
965 relative to shuffles (Figure 4b,d). To perform the bootstrap test, we created a grid of
966 thresholds for minimum absolute weighted correlation and maximum jump distance, and
967 for each combination of thresholds we calculated the fraction of actual events that
968 exceeded the minimum absolute weighted correlation threshold and did not exceed the
969 maximum jump distance threshold. Then, we generated 100 data sets of shuffled events by
970 randomly permuting the order of the 10-ms time bins for each actual event and calculated
971 the fraction of events meeting the same pairs of thresholds for each shuffled data set. The
972 p-value of the fraction of high-quality events was then calculated as the fraction of shuffled
973 data sets with a higher fraction of high-quality events.

974 To test the significance of each event's absolute weighted correlation individually, we
975 calculated the event's p-value as the fraction of the event's own shuffles that had a higher
976 absolute weighted correlation than the un-shuffled event (Figure 4f, bottom left).

977 The spatial entropy H of a decoded event was calculated as the mean over its time bins of
978 the entropy of the decoded position probability in each time bin, using the equation

$$H = - \sum_i p_i \ log_2(p_i)$$

979

980 for each time bin, where p_i is the decoded position probability for spatial bin i.

Cell identity shuffled decoding

982 We performed Bayesian decoding on the fiducial parameter set after shuffling cell
983 identities in three different manners (Figures 6 and 7). To shuffle cells in a cluster-
984 independent manner ("Across-network shuffle"), we randomly shuffled the identity of cells
985 during the sleep simulations. To shuffle cells within clusters ("Within-cluster shuffle"), we
986 randomly shuffled cell identity only between cells that shared membership in at least one
987 cluster. To shuffle cells within only single clusters ("Within-single-cluster shuffle"), we

35

988   shuffled cells in the same manner as the within-cluster shuffle but excluded any cells from
989   the shuffle that were in multiple clusters.

990   To test for a correlation between spike rank during sleep PBEs and the order of place fields
991   on the track (Figure 7), we calculated for each excitatory cell in each network of the fiducial
992   parameter set its mean relative spike rank and correlated that with the location of its mean
993   place field density on the track (Figure 7a). To account for event directionality, we
994   calculated the mean relative rank after inverting the rank within events that had a
995   negatively sloped decoded trajectory (Figure 7b). We calculated mean relative rank for
996   each cell relative to all cells in the network ("Within-network mean relative rank") and
997   relative to only cells that shared cluster membership with the cell ("Within-cluster mean
998   relative rank"). We then compared the slope of the linear regression between mean relative
999   rank and place field location against the slope that results when applying the same analysis
1000   to each of the three methods of cell identify shuffles for both the within-network regression
1001   (Figure 7c) and the within-cluster regression (Figure 7d).

1002

Small-world index

1004   The small-world index (SWI) was calculated following the method of Neal, 2015 (see also
1005   Neal, 2017). It was defined as

1006
$$\text{SWI} = \frac{(L - L_l)}{(L_r - L_l)} \times \frac{(C - C_r)}{(C_l - C_r)}$$

1007   where L is the mean path distance and C is the clustering coefficient of the network. We
1008   calculate L as the mean over all ordered pairs of excitatory cells of the shortest directed
1009   path length from the first to the second cell. We calculate C as the ratio of the number of all
1010   triplets of excitatory cells that are connected in either direction over the number of all
1011   triplets that could form, following the methods of Fagiolo, 2007 for directed graphs. L_l and
1012 C_l are the expected values for a one-dimensional ring lattice network with the same size
1013   and connection probability (in which connections are local such that there are no
1014   connections between cells with a greater separation on the ring than that of any pairs
1015   without a connection). And L_r and C_r are the expected values for a random network of the
1016   same size and connection probability. A network with a high SWI index is therefore a
1017   network with both a high clustering coefficient, similar to a ring lattice network, and small
1018   mean path length, similar to a random network.

1019   For directed graphs of size n, average degree k, and global connection probability p

1020 $C_r = p$ (Fagiolo, 2007),

1021 $L_r = \frac{ln(n) - \gamma}{ln(k)} + 0.5$ (Fronczak et al., 2004),

1022 $C_l = \frac{3(k-2)}{4(k-1)}$ (Neal, 2015)

1023 $L_l = \frac{n}{2k} + 0.5$ (Neal, 2015; Fronczak et al., 2004)

36

1024   where γ is the Euler-Mascheroni constant.

1025

Active cluster analysis

1027   To quantify cluster activation (figure 5), we calculated the population rate for each cluster
1028   individually as the mean firing rate of all excitatory cells belonging to the cluster smoothed
1029   with a Gaussian kernel (15 ms standard deviation). A cluster was defined as 'active' if at
1030   any point its population rate exceeded twice that of any other cluster during a PBE. The
1031   active clusters' duration of activation was defined as the duration for which it was the most
1032   active cluster.

1033   To test whether the sequence of activation in events with three active clusters matched the
1034   sequence of place fields on the track, we performed a bootstrap significance test (Figure
1035   5—figure supplement 1). For all events from the fiducial parameter set that had three
1036   active clusters, we calculated the fraction in which the sequence of the active clusters
1037   matched the sequence of the clusters' left vs right bias on the track in either direction. We
1038   then compared this fraction to the distribution expected from randomly sampling
1039   sequences of three clusters without replacement.

1040   To determine if there was a relationship between the number of active clusters within an
1041   event and it's preplay quality we performed a Spearman's rank correlation between the
1042   number of active clusters and the normalized absolute weighted correlation across all
1043   events at the fiducial parameter set. The absolute weighted correlations were z-scored
1044   based on the absolute weighted correlations of the time-bin shuffled events that had the
1045   same number of active clusters.

Experimental data

1047   Electrophysiological data was reanalyzed from the hippocampal CA1 recordings first
1048   published in Shin et al., 2019. All place-field data (Figure 3a) came from the six rats' first
1049   experience on the W-track spatial alternation task. All preplay data (Figure 4a,b) came
1050   from the six rats' first sleep-box session, which lasted 20-30 minutes and occurred
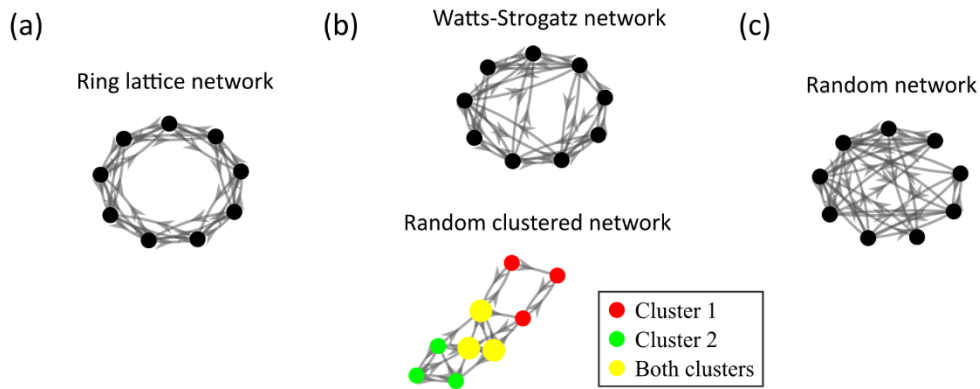1051   immediately before their first experience on the W-track.

1052

Code

1054   Simulations and analysis were performed in MATLAB with custom code. Code available at
1055   https://github.com/primon23/Preplay_paper.

37

Supplemental figures

Figure 1—figure supplement 1: Comparison of the randomly clustered network and the canonical Watts-Strogatz small-world network
(a) A small ring-lattice network. **(b)** Example small-world networks. Top, a Watts-Strogatz network with re-wiring parameter $\beta = 0.2$. Bottom, a randomly clustered network with two clusters and a cluster participation of 1.25. **(c)** Example randomly connected network.
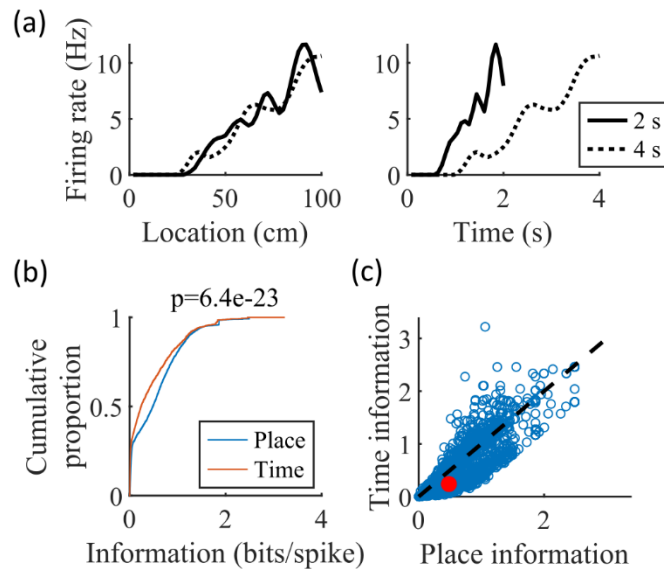

Figure 3—figure supplement 1: The simulated cells have greater place information than time information.
(a) Place fields (left) and time fields (right) for an example cell calculated from simulated trajectories that took 2 seconds (solid line) or 4 seconds (dotted line) to traverse the track.
(b) CDFs of the information content of the place fields ("Place") and time fields ("Time") of all cells. The spatial information is significantly greater than the temporal information (KS-test, p=6.4e-23). **(c)** Scatter plot of the data in (b), with the median values marked in red.

38

1072

1073
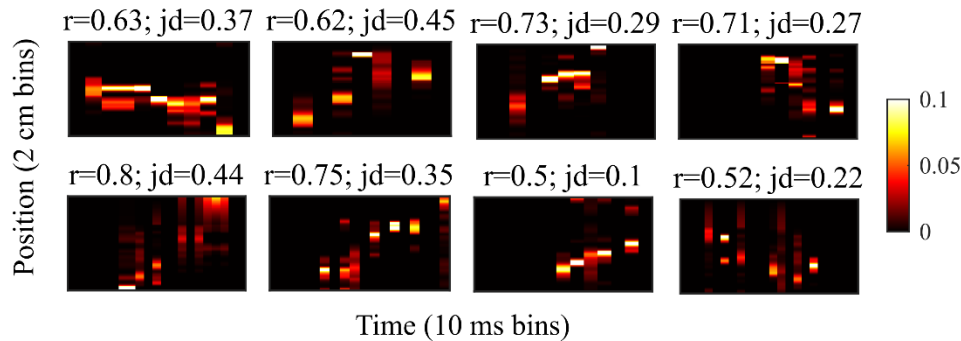
Figure 4—figure supplement 1: Example preplay events from the Shin et al., 2019 data

Example preplay events. Same as Figure 2f but for events from the hipopcampal data from Shin et al., 2019. The height of each plot spans the length of the trajectory used for decoding, divided into 2 cm spatial bins. The width of each plot spans the duration of the detected event, divided into 10 ms time bins. Probability is show in color.

1080

1081
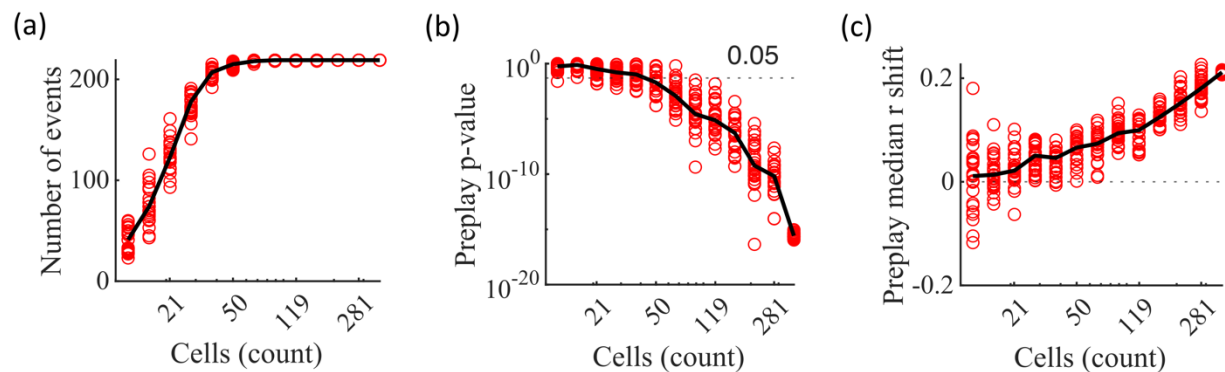
Figure 4—figure supplement 2: Significant preplay can typically be identified with as few as 50 cells

(a-c) Results from performing the same Bayesian decoding on the same simulated population burst events (PBEs) in Figure 4c but using only random subsets of the excitatory cells for performing the decoding analysis. Each circle is the result of an analysis performed on one random subset of the cells. 25 random subsets were analyzed for each analyzed cell count. The subset sizes are logarithmically spaced. Black lines show the median value. The variability at N=375 is due to the variation in the randomness of the time-bin shuffles. **(a)** Number of events meeting the inclusion criterion for decoding analysis. **b)** P-value of the KS-test comparing actual vs shuffled event absolute weighted correlations. A majority of the random subsets of 50 cells (17 out of 25) produce preplay p-values below 0.05. **(c)** Shift in the median absolute weighted correlation of actual events relative to shuffled events.

39
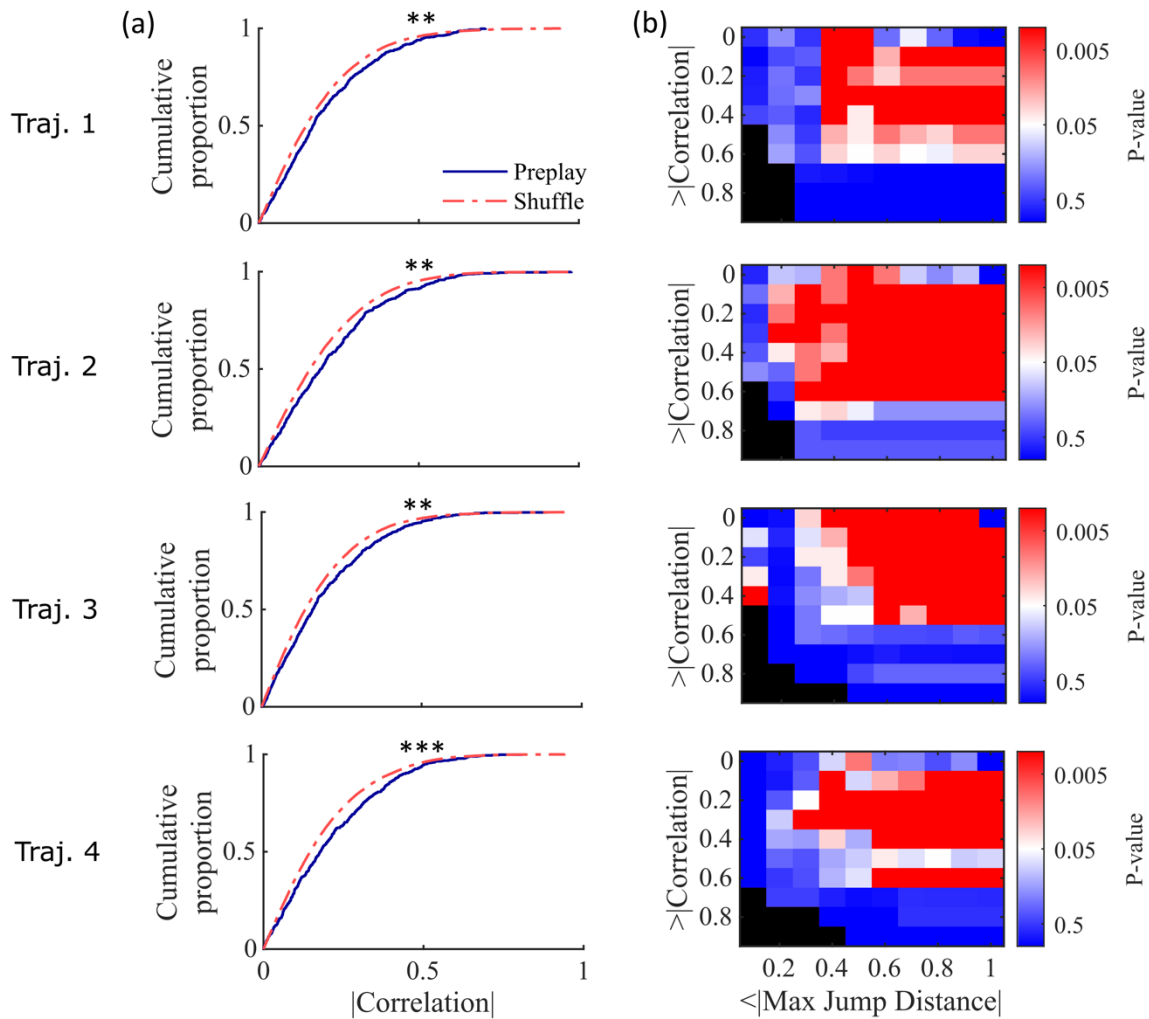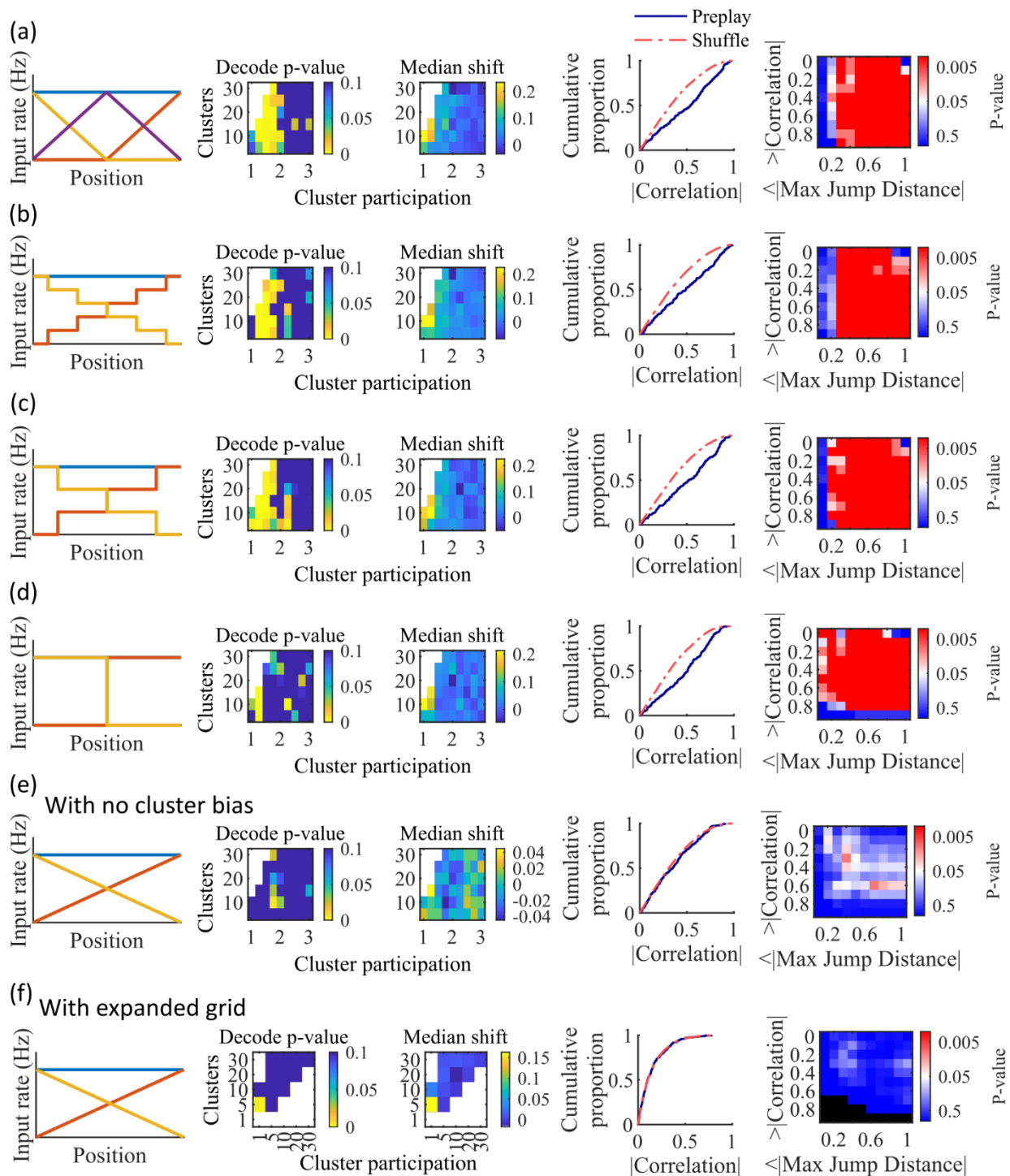

Figure 4—figure supplement 3: Preplay statistics by trajectory for Shin et al., 2019 data.
(a) Same as Figure 4a but separated by results from decoding by each of the 4 trajectories of the W-track individually (trajectory 1, center arm to right arm; trajectory 2, right arm to center arm; trajectory 3, center arm to left arm; trajectory 4, left arm to center arm). KS-test for each trajectory: trajectory 1, p=0.0030; trajectory 2, p=0.0028; trajectory 3, p=0.0027; trajectory 4, p=5.461×10^{-5}. ** p<0.01, *** p<0.001. **b)** Same as Figure 4b but separated by results from decoding by each of the 4 trajectories individually.

40

1107

Figure 4—figure supplement 4: Additional simulations support the consistency and robustness of the model to variations in spatial input forms.
Each row corresponds to a different parameter grid simulation, with statistics calculated as in the corresponding panel from Figure 4. **(a)** Preplay statistics are similar to the main

1112   simulation results when a third linearly varying spatial cue is included in the inputs to the
1113   network (CDF KS-test, p=3.9e-13, KS-statistic=0.26). **(b)** Preplay statistics are similar to the
1114   main simulation results when a stepped input is used (CDF KS-test, p=2.5e-08, KS-
1115   statistic=0.20). The stepped input is less spatially informative since stretches of adjacent
1116   locations on the track have identical spatial input. **(c)** Same as (b), but with three step
1117   increments (CDF KS-test, p=6.2e-13, KS-statistic=0.26). **(d)** Same as (c), but with a single
1118   step increment (CDF KS-test, p=4.9e-13, KS-statistic=0.26). With this input the fiducial
1119   parameter set still shows significant preplay (right two columns), but most of the
1120   parameter grid loses significant preplay. **(e)** When the bias in cluster spatial input location
1121   is removed preplay is no longer significant (CDF KS-test, p=0.34, KS-statistic=0.063). **(f)** A
1122   parameter grid that shows greater values of cluster participation do not have significant
1123   preplay. Values along the diagonal where clusters equals cluster participation are
1124   equivalent to a random cluster-less network. Example parameter point is at clusters=5 and
1125   cluster participation=5 (CDF KS-test, p=0.99, KS-statistic=0.02).

1126   _____

1127

1128   _____
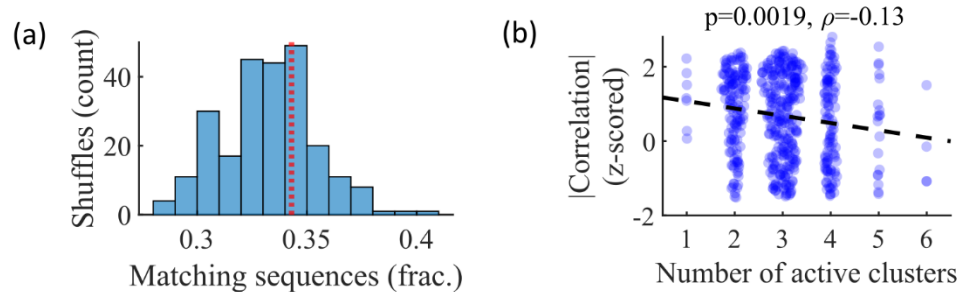
1129   **Figure 5—figure supplement 1: Relationship between cluster activation and preplay.**
1130   **(a)** Out of all events from the fiducial parameter set simulations where 3 unique clusters
1131   were active, the fraction of those events with sequences that match the order of cluster
1132   biases on the track (red line) is consistent with the values expected by randomly sampling
1133   clusters (blue). **(b)** Z-scored absolute weighted preplay correlation is negatively correlated
1134   with the number of active clusters (Spearman's rank correlation).

1135   _____

42

References

Alme CB, Miao C, Jezek K, Treves A, Moser EI, Moser M-B. 2014. Place cells in the hippocampus: Eleven maps for eleven rooms. Proceedings of the National Academy of Sciences. 111(52):18428–18435. doi:10.1073/pnas.1421056111.

Antonello PC, Varley TF, Beggs J, Porcionatto M, Sporns O, Faber J. 2022. Self-organization of in vitro neuronal assemblies drives to complex network topology. eLife. 11:e74921. doi:10.7554/eLife.74921.

Azizi AH, Wiskott L, Cheng S. 2013. A computational model for preplay in the hippocampus. Front Comput Neurosci. 7. doi:10.3389/fncom.2013.00161.

Battaglia FP, Treves A. 1998. Attractor neural networks storing multiple space representations: A model for hippocampal place fields. Phys Rev E. 58(6):7738–7753. doi:10.1103/PhysRevE.58.7738.

Berners-Lee A, Feng T, Silva D, Wu X, Ambrose ER, Pfeiffer BE, Foster DJ. 2022 Apr. Hippocampal replays appear after a single experience and incorporate greater detail with more experience. Neuron.:S089662732200246X. doi:10.1016/j.neuron.2022.03.010.

Bourjaily MA, Miller P. 2011. Excitatory, Inhibitory, and Structural Plasticity Produce Correlated Connectivity in Random Networks Trained to Solve Paired-Stimulus Tasks. Front Comput Neurosci. 5. doi:10.3389/fncom.2011.00037.

Bush D, Barry C, Burgess N. 2014. What do grid cells contribute to place cell firing? Trends in Neurosciences. 37(3):136–145. doi:10.1016/j.tins.2013.12.003.

Carr MF, Jadhav SP, Frank LM. 2011. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. Nature Neuroscience. 14(2):147–153. doi:10.1038/nn.2732.

Chenkov N, Sprekeler H, Kempter R. 2017. Memory replay in balanced recurrent networks. Gutkin BS, editor. PLOS Computational Biology. 13(1):e1005359. doi:10.1371/journal.pcbi.1005359.

Csicsvari J, Hirase H, Mamiya A, Buzsáki G. 2000. Ensemble Patterns of Hippocampal CA3-CA1 Neurons during Sharp Wave–Associated Population Events. Neuron. 28(2):585–594. doi:10.1016/S0896-6273(00)00135-5.

Debanne D, Gähwiler BH, Thompson SM. 1998. Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures. The Journal of Physiology. 507(1):237–247. doi:10.1111/j.1469-7793.1998.237bu.x.

Diesmann M, Gewaltig M-O, Aertsen A. 1999. Stable propagation of synchronous spiking in cortical neural networks. Nature. 402(6761):529–533. doi:10.1038/990101.

Dragoi G, Tonegawa S. 2011. Preplay of future place cell sequences by hippocampal cellular assemblies. Nature. 469(7330):397–401. doi:10.1038/nature09633.

Dragoi G, Tonegawa S. 2013. Distinct preplay of multiple novel spatial experiences in the rat. Proceedings of the National Academy of Sciences. 110(22):9100–9105. doi:10.1073/pnas.1306031110.

Farooq, Usman, Jeremie Sibille, Kefei Liu, and George Dragoi. 2019. Strengthened Temporal Coordination within Pre-Existing Sequential Cell Assemblies Supports Trajectory Replay. Neuron 103, no. 4: 719-733.e7. doi:10.1016/j.neuron.2019.05.040.

1179 Fagiolo G. 2007. Clustering in complex directed networks. Phys Rev E. 76(2):026107.
1180   doi:10.1103/PhysRevE.76.026107.
1181 Foster DJ, Wilson MA. 2006. Reverse replay of behavioural sequences in hippocampal place
1182   cells during the awake state. Nature. 440(7084):680–683. doi:10.1038/nature04587.
1183 Frank LM, Stanley GB, Brown EN. 2004. Hippocampal Plasticity across Multiple Days of
1184   Exposure to Novel Environments. J Neurosci. 24(35):7681–7689.
1185   doi:10.1523/JNEUROSCI.1958-04.2004.
1186 Fronczak A, Fronczak P, Holyst JA. 2004. Average path length in uncorrelated random
1187   networks with hidden variables. Phys Rev E. 70(5):056110.
1188   doi:10.1103/PhysRevE.70.056110.
1189 Grosmark AD, Buzsaki G. 2016. Diversity in neural firing dynamics supports both rigid and
1190   learned hippocampal sequences. Science. 351(6280):1440–1443.
1191   doi:10.1126/science.aad1935.
1192 Guzman SJ, Schlogl A, Frotscher M, Jonas P. 2016. Synaptic mechanisms of pattern
1193   completion in the hippocampal CA3 network. Science. 353(6304):1117–1123.
1194   doi:10.1126/science.aaf1836.
1195 Haga T, Fukai T. 2018. Recurrent network model for learning goal-directed sequences
1196   through reverse replay. eLife. 7:e34171. doi:10.7554/eLife.34171. Humphries MD,
1197   Gurney K, Prescott TJ. 2006. The brainstem reticular formation is a small-world, not
1198   scale-free, network. Proc R Soc B. 273(1585):503–511. doi:10.1098/rspb.2005.3354.
1199 Harvey RE, Berkowitz LE, Savage DD, Hamilton DA, Clark BJ. 2020. Altered Hippocampal
1200   Place Cell Representation and Theta Rhythmicity following Moderate Prenatal Alcohol
1201   Exposure. Current Biology. 30(18):3556-3569.e5. doi:10.1016/j.cub.2020.06.077.
1202 Humphries MD, Gurney K. 2008. Network 'Small-World-Ness': A Quantitative Method for
1203   Determining Canonical Network Equivalence. Sporns O, editor. PLoS ONE.
1204   3(4):e0002051. doi:10.1371/journal.pone.0002051.
1205 Hwaun E, Colgin LL. 2019. CA3 place cells that represent a novel waking experience are
1206   preferentially reactivated during sharp wave-ripples in subsequent sleep.
1207   Hippocampus. 29(10):921–938. doi:10.1002/hipo.23090.
1208 Jahnke S, Timme M, Memmesheimer R-M. 2015. A Unified Dynamic Model for Learning,
1209   Replay, and Sharp-Wave/Ripples. Journal of Neuroscience. 35(49):16236–16258.
1210   doi:10.1523/JNEUROSCI.3977-14.2015.
1211 Karlsson MP, Frank LM. 2009. Awake replay of remote experiences in the hippocampus.
1212   Nature Neuroscience. 12(7):913–918. doi:10.1038/nn.2344.
1213 Kinsky NR, Sullivan DW, Mau W, Hasselmo ME, Eichenbaum HB. 2018. Hippocampal Place
1214   Fields Maintain a Coherent and Flexible Map across Long Timescales. Current Biology.
1215   28(22):3578-3588.e6. doi:10.1016/j.cub.2018.09.037.
1216 Leutgeb S, Leutgeb JK, Treves A, Moser M-B, Moser EI. 2004. Distinct Ensemble Codes in
1217   Hippocampal Areas CA3 and CA1. Science. 305(5688):1295–1298.
1218   doi:10.1126/science.1100265.
1219 Leutgeb S, Leutgeb JK, Barnes CA, Moser EI, McNaughton BL, Moser M-B. 2005.
1220   Independent Codes for Spatial and Episodic Memory in Hippocampal Neuronal
1221   Ensembles. Science. 309(5734):619–623. doi:10.1126/science.1114037.

44

Liao W, Ding J, Marinazzo D, Xu Q, Wang Z, Yuan C, Zhang Z, Lu G, Chen H. 2011. Small-world directed networks in the human brain: Multivariate Granger causality analysis of resting-state fMRI. NeuroImage. 54(4):2683–2694. doi:10.1016/j.neuroimage.2010.11.007.

Litwin-Kumar A, Doiron B. 2014. Formation and maintenance of neuronal assemblies through synaptic plasticity. Nat Commun. 5(1):5319. doi:10.1038/ncomms6319.

Liu K, Sibille J, Dragoi G. 2019. Preconfigured patterns are the primary driver of offline multi-neuronal sequence replay. Hippocampus. 29(3):275–283. doi:10.1002/hipo.23034.

Liu K, Sibille J, Dragoi G. 2021 Sep. Orientation selectivity enhances context generalization and generative predictive coding in the hippocampus. Neuron.:S0896627321006103. doi:10.1016/j.neuron.2021.08.013.

Luongo FJ, Zimmerman CA, Horn ME, Sohal VS. 2016. Correlations between prefrontal neurons form a small-world network that optimizes the generation of multineuron sequences of activity. Journal of Neurophysiology. 115(5):2359–2375. doi:10.1152/jn.01043.2015.

Lynn CW, Holmes CM, Palmer SE. 2022. Heavy–tailed neuronal connectivity arises from Hebbian self–organization. Neuroscience. [accessed 2022 Oct 27]. http://biorxiv.org/lookup/doi/10.1101/2022.05.30.494086.

Mishra RK, Kim S, Guzman SJ, Jonas P. 2016. Symmetric spike timing-dependent plasticity at CA3–CA3 synapses optimizes storage and recall in autoassociative networks. Nat Commun. 7(1):11552. doi:10.1038/ncomms11552.

Morris RGM, Garrud P, Rawlins JNP, O'Keefe J. 1982. Place navigation impaired in rats with hippocampal lesions. Nature. 297(5868):681–683. doi:10.1038/297681a0.

Moser EI, Kropff E, Moser M-B. 2008. Place Cells, Grid Cells, and the Brain's Spatial Representation System. Annu Rev Neurosci. 31(1):69–89. doi:10.1146/annurev.neuro.31.061307.090723.

Muller R, Kubie J, Ranck J. 1987. Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. J Neurosci. 7(7):1935–1950. doi:10.1523/JNEUROSCI.07-07-01935.1987.

Nakashiba T, Young JZ, McHugh TJ, Buhl DL, Tonegawa S. 2008. Transgenic Inhibition of Synaptic Transmission Reveals Role of CA3 Output in Hippocampal Learning. Science. 319(5867):1260–1264. doi:10.1126/science.1151120.

Nakashiba T, Buhl DL, McHugh TJ, Tonegawa S. 2009. Hippocampal CA3 Output Is Crucial for Ripple-Associated Reactivation and Consolidation of Memory. Neuron. 62(6):781–787. doi:10.1016/j.neuron.2009.05.013.

Navratilova Z, Hoang LT, Schwindel CD, Tatsuno M, McNaughton BL. 2012. Experience-dependent firing rate remapping generates directional selectivity in hippocampal place cells. Front Neural Circuits. 6. doi:10.3389/fncir.2012.00006.

Neal Z. 2015. Making Big Communities Small: Using Network Science to Understand the Ecological and Behavioral Requirements for Community Social Capital. American Journal of Community Psychology. 55(3–4):369–380. doi:10.1007/s10464-015-9720-4.

Neal ZP. 2017. How small is it? Comparing indices of small worldliness. Net Sci. 5(1):30–44. doi:10.1017/nws.2017.5.

1266 Nitzan N, Swanson R, Schmitz D, Buzsáki G. 2022. Brain-wide interactions during
1267   hippocampal sharp wave ripples. Proc Natl Acad Sci USA. 119(20):e2200931119.
1268   doi:10.1073/pnas.2200931119.

1269 O'Keefe J, Nadel L. 1978. The hippocampus as a cognitive map. Oxford : New York:
1270   Clarendon Press ; Oxford University Press.

1271 O'Neill J, Senior TJ, Allen K, Huxter JR, Csicsvari J. 2008. Reactivation of experience-
1272   dependent cell assembly patterns in the hippocampus. Nature Neuroscience.
1273   11(2):209–215. doi:10.1038/nn2037.

1274 Pang R, Fairhall AL. 2019. Fast and flexible sequence induction in spiking neural networks
1275   via rapid excitability changes. eLife. 8:e44324. doi:10.7554/eLife.44324.

1276 Perin R, Berger TK, Markram H. 2011. A synaptic organizing principle for cortical neuronal
1277   groups. Proc Natl Acad Sci USA. 108(13):5419–5424. doi:10.1073/pnas.1016051108.

1278 Peyrache A, Benchenane K, Khamassi M, Wiener SI, Battaglia FP. 2010. Principal
1279   component analysis of ensemble recordings reveals cell assemblies at high temporal
1280   resolution. J Comput Neurosci. 29(1–2):309–325. doi:10.1007/s10827-009-0154-6.

1281 Ramirez-Villegas JF, Willeke KF, Logothetis NK, Besserve M. 2018. Dissecting the Synapse-
1282   and Frequency-Dependent Network Mechanisms of In Vivo Hippocampal Sharp Wave-
1283   Ripples. Neuron. 100(5):1224-1240.e13. doi:10.1016/j.neuron.2018.09.041.

1284 Rebola N, Carta M, Mulle C. 2017. Operation and plasticity of hippocampal CA3 circuits:
1285   implications for memory encoding. Nat Rev Neurosci. 18(4):208–220.
1286   doi:10.1038/nrn.2017.10.

1287 Sadovsky A, MacLean J. 2014. Mouse Visual Neocortex Supports Multiple Stereotyped
1288   Patterns of Microcircuit Activity. Journal of Neuroscience. 34(23):7769–7777.
1289   doi:10.1523/JNEUROSCI.0169-14.2014.

1290 Samsonovich A, McNaughton BL. 1997. Path Integration and Cognitive Mapping in a
1291   Continuous Attractor Neural Network Model. J Neurosci. 17(15):5900–5920.
1292   doi:10.1523/JNEUROSCI.17-15-05900.1997.

1293 Savelli F, Yoganarasimha D, Knierim JJ. 2008. Influence of boundary removal on the spatial
1294   representations of the medial entorhinal cortex. Hippocampus. 18(12):1270–1282.
1295   doi:10.1002/hipo.20511.

1296 She Q, Chen G, Chan RHM. 2016. Evaluating the Small-World-Ness of a Sampled Network:
1297   Functional Connectivity of Entorhinal-Hippocampal Circuitry. Sci Rep. 6(1):21468.
1298   doi:10.1038/srep21468.

1299 Sheintuch L, Geva N, Deitch D, Rubin A, Ziv Y. 2023. Organization of hippocampal CA3 into
1300   correlated cell assemblies supports a stable spatial code. Cell Reports. 42(2):112119.
1301   doi:10.1016/j.celrep.2023.112119.

1302 Shin JD, Tang W, Jadhav SP. 2019. Dynamics of Awake Hippocampal-Prefrontal Replay for
1303   Spatial Learning and Memory-Guided Decision Making. Neuron. 104(6):1110-1125.e7.
1304   doi:10.1016/j.neuron.2019.09.012.

1305 Silva D, Feng T, Foster DJ. 2015. Trajectory events across hippocampal place cells require
1306   previous experience. Nat Neurosci. 18(12):1772–1779. doi:10.1038/nn.4151.

1307 Solstad T, Boccara CN, Kropff E, Moser M-B, Moser EI. 2008. Representation of Geometric
1308   Borders in the Entorhinal Cortex. Science. 322(5909):1865–1868.
1309   doi:10.1126/science.1166466.

46

1310 Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. 2005. Highly Nonrandom Features of
1311   Synaptic Connectivity in Local Cortical Circuits. Friston KJ, editor. PLoS Biology.
1312   3(3):e68. doi:10.1371/journal.pbio.0030068.
1313 Squire LR, Stark CEL, Clark RE. 2004. The medial temporal lobe. Annual Review of
1314   Neuroscience. 27(1):279–306. doi:10.1146/annurev.neuro.27.070203.144130.
1315 Tang, W., & Jadhav, S. P. (2022). Multiple-Timescale Representations of Space: Linking
1316   Memory to Navigation. Annual Review of Neuroscience, 45(1), 1–21.
1317   https://doi.org/10.1146/annurev-neuro-111020-084824
1318 Theodoni P, Rovira B, Wang Y, Roxin A. 2018. Theta-modulation drives the emergence of
1319   connectivity patterns underlying replay in a network model of place cells. eLife. 7.
1320   doi:10.7554/eLife.37388.
1321 Vaz AP, Wittig JH, Inati SK, Zaghloul KA. 2023. Backbone spiking sequence as a basis for
1322   preplay, replay, and default states in human cortex. Nat Commun. 14(1):4723.
1323   doi:10.1038/s41467-023-40440-5.
1324 Vetere G, Kenney JW, Tran LM, Xia F, Steadman PE, Parkinson J, Josselyn SA, Frankland PW.
1325   2017. Chemogenetic Interrogation of a Brain-wide Fear Memory Network in Mice.
1326   Neuron. 94(2):363-374.e4. doi:10.1016/j.neuron.2017.03.037.
1327 Watts DJ, Strogatz SH. 1998. Collective dynamics of 'small-world' networks. Nature.
1328   393(6684):440–442. doi:10.1038/30918.
1329 Whittington JCR, Muller TH, Mark S, Chen G, Barry C, Burgess N, Behrens TEJ. 2020. The
1330   Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through
1331   Generalization in the Hippocampal Formation. Cell. 183(5):1249-1263.e23.
1332   doi:10.1016/j.cell.2020.10.024.
1333 Wilson MA, McNaughton BL. 1994. Reactivation of Hippocampal Ensemble Memories
1334   During Sleep. Science. 265(5172):676–679. doi:10.1126/science.8036517.
1335 Yamamoto J, Tonegawa S. 2017. Direct Medial Entorhinal Cortex Input to Hippocampal CA1
1336   Is Crucial for Extended Quiet Awake Replay. Neuron. 96(1):217-227.e4.
1337   doi:10.1016/j.neuron.2017.09.017.
1338 Zhang L, Hu X, Hu Y, Tang M, Qiu H, Zhu Z, Gao Y, Li H, Kuang W, Ji W. 2022. Structural
1339   covariance network of the hippocampus–amygdala complex in medication-naïve
1340   patients with first-episode major depressive disorder. Psychoradiology. 2(4):190–198.
1341   doi:10.1093/psyrad/kkac023.