Hippocampal Reactivation of Random Trajectories Resembling Brownian Diffusion

Highlights

- Hippocampal replay can represent Brownian diffusion-like random trajectories

- Reactivated trajectories cover positions over wide ranges of spatiotemporal scales

- Replay event statistics are incompatible with actual behavioral trajectories

- Expression dynamics of replayed assemblies was linked to specific oscillatory bands

Authors

Federico Stella, Peter Baracskay, Joseph O'Neill, Jozsef Csicsvari

Correspondence

federico.stella@ist.ac.at (F.S.), jozsef.csicsvari@ist.ac.at (J.C.)

In Brief

Stella et al. examine the dynamical properties of reactivated spatial trajectories in the hippocampus following non-stereotypical exploration and find that reactivated trajectories are governed by a Brownian diffusion process that occur at varying lengths and timescales, without directly reflecting behavior.

Article

Hippocampal Reactivation of Random Trajectories Resembling Brownian Diffusion

Federico Stella,[1,*] Peter Baracskay,[1,2] Joseph O'Neill,[1,3] and Jozsef Csicsvari[1,4,*]

[1]Institute of Science and Technology Austria (IST Austria) Am Campus 1, 3400 Klosterneuburg, Austria
[2]Present address: Department of Biochemistry, Institute of Biology, Eötvös Loránd University, Budapest 1117, Hungary
[3]Present address: School of Psychology, Cardiff University, Park Place, Cardiff CF10 3AT, UK
[4]Lead Contact
*Correspondence: federico.stella@ist.ac.at (F.S.), jozsef.csicsvari@ist.ac.at (J.C.)
https://doi.org/10.1016/j.neuron.2019.01.052

SUMMARY

Hippocampal activity patterns representing movement trajectories are reactivated in immobility and sleep periods, a process associated with memory recall, consolidation, and decision making. It is thought that only fixed, behaviorally relevant patterns can be reactivated, which are stored across hippocampal synaptic connections. To test whether some generalized rules govern reactivation, we examined trajectory reactivation following non-stereotypical exploration of familiar open-field environments. We found that random trajectories of varying lengths and timescales were reactivated, resembling that of Brownian motion of particles. The animals' behavioral trajectory did not follow Brownian diffusion demonstrating that the exact behavioral experience is not reactivated. Therefore, hippocampal circuits are able to generate random trajectories of any recently active map by following diffusion dynamics. This ability of hippocampal circuits to generate representations of all behavioral outcome combinations, experienced or not, may underlie a wide variety of hippocampal-dependent cognitive functions such as learning, generalization, and planning.

INTRODUCTION

The hippocampus plays a crucial role in navigation and episodic memory by forming a cognitive map of space, encoded by the spatially selective activity of place cells (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978). At any given moment, the combined activity of place cells provides a signal that encodes the instantaneous location of the animal, while the sequential firing of these cells can represent entire movement paths through space (O'Keefe and Recce, 1993; Skaggs et al., 1996). Such neural representations of behavior likely underpin the spatial component for memories of places and events, that, when reinstated, can be used for navigation in previously explored environments. Consistent with this notion, firing sequences that encode behavioral trajectories during exploration can be subsequently replayed

by the network. This takes the form of place cell representations of movement paths that can last several seconds or more, compressed into brief bursts of activity, lasting tens to hundreds of milliseconds (Lee and Wilson, 2002). This replay occurs in a variety of behavioral and network states, which may reflect different underlying mechanisms and cognitive functions (Foster and Wilson, 2006; O'Neill et al., 2010). The replay of temporally compressed firing sequences of neurons was first described during transient network synchronization epochs associated with sharp-wave ripples (SWRs) in sleep (Lee and Wilson, 2002). However, subsequently, it has been shown that similar phenomena can also occur during exploration simultaneously with theta oscillations (Feng et al., 2015; Gupta et al., 2012; Zheng et al., 2016). Moreover, not only have 200 Hz ripple-band oscillations been seen during compressed trajectory replay, but simultaneously occurring slower gamma-beta (20–80 Hz) band oscillatory components have been seen as well. The power of the slower oscillatory components of SWRs predicted the fidelity of replay with different oscillatory cycles separating trajectory segments with a discontinuous trajectory jump (Carr et al., 2012; Pfeiffer and Foster, 2015; Yamamoto and Tonegawa, 2017).

While replay has been associated with a number of mnemonic processes, including consolidation in sleep, goal-directed navigation, and decision making at maze choice points (Ólafsdóttir et al., 2018), the mechanisms underlying its generation remain controversial. Competing lines of evidence suggest that sequence firing in replay is generated from either experience-dependent mechanisms or, alternatively, reflects hardwired predetermined assemblies (Dragoi and Tonegawa, 2011; Grosmark and Buzsáki, 2016; Shen and McNaughton, 1996; Silva et al., 2015). On the one hand, both the stability of novel spatial maps (Kentros et al., 1998) and the accurate replay of novel linear tracks during sleep (Silva et al., 2015) require NMDA receptors. This suggests that the associations between spatial cell assemblies during behavior are formed through synaptic plasticity. Consistent with this idea, sequences expressed on novel linear tracks during theta oscillations only emerge after the first lap (Feng et al., 2015), indicating that they are generated through activity-dependent mechanisms. However, novel spatial maps and their reactivated trajectories may in part use hardwired circuit connections, which are then refined with the animals' experience (Dragoi and Tonegawa, 2011; Shen and McNaughton, 1996). Indeed, it has been shown that firing patterns resembling that of novel paths can also be seen in sleep prior to the animal experiencing that path,

termed preplay (Dragoi and Tonegawa, 2011; Grosmark and Buz-sáki, 2016; Ólafsdóttir et al., 2015). However, it is still unclear to what degree preplay patterns are refined as a result of exploration to provide a more accurate replay in sleep after the experience (Silva et al., 2015). Nevertheless, several reports describe replay that is not simply a product of experience, particularly replay seen during exploration (Gupta et al., 2010; Wu and Foster, 2014).

During exploration, the reactivation seen in transient SWR-associated network bursts tends to occur when animals stop briefly (Davidson et al., 2009; Foster and Wilson, 2006). These trajectories tend to originate from the animals' current position and point toward future goal locations, even in open two-dimensional (2D) environments where a fixed goal is to be reached (Pfeiffer and Foster, 2013). However, during replay in more complex maze environments, maze segment combinations that the animal never took before may reactivate together in the form of a novel combined trajectory (Gupta et al., 2010). These observations may challenge the theory that reactivated trajectories need to be first experienced, in order to be stored in the hippocampal circuits. However, unlike replay in sleep, it cannot be ruled out that upstream spatial sensory inputs, channeled through the entorhinal cortex, may still be involved in driving reactivation events observed during active waking periods (Yamamoto and Tonegawa, 2017). Indeed, this replay of inexperienced path combinations through the integration of maze segments may reflect the imagination of novel situations or map refinement (Gupta et al., 2010). These observations do not provide a clear answer as to whether, in the absence of sensory inputs, reactivated patterns represent the exact spatial experience of the animal or, alternatively, are generated by a more generalized process that does not require the direct storage of patterns. We set out to examine the mechanisms of replay by testing for its occurrence in conditions that should not support sequence formation through experience-dependent mechanisms. The behavior used in replay studies typically utilizes narrow linear tracks or mazes, in which the animal performs repeated stereotyped trajectories. Instead, we utilized an open-field enclosure in which the animal performed a random pellet-chasing task and then detected replay during a subsequent rest period, recorded in a separate sleep box. We found that reactivated trajectories followed rules of random movement governed by Brownian diffusion that did not directly reflect the past behavioral trajectories of the animal. However, our results were also not consistent with a model of replay solely based on a limited number of preexisting cell assemblies, since the number of replay trajectories was only constrained by the size of the environment and the diffusion dynamics. Instead, the data indicate that hippocampal circuits have the built-in ability to generate sequences that link assemblies together across the entire cognitive map, on all temporal and spatial scales. Such a mechanism likely provides a framework on which past experiences can be reactivated in sleep, or possible future paths can be compared during navigation.

RESULTS

Reactivation of 2D Trajectories after Random Foraging

We examined reactivation following a random exploration of a large (1.2 m diameter) circular arena during quiet immobility pe-

riods in the absence of sensory drive of the explored environment. So far, only one study has been able to see the reactivation of 2D trajectories (Pfeiffer and Foster, 2013), in that case during awake, exploration-associated SWRs. Therefore, the replay shown in that study may have been driven by an upstream sensory drive (Yamamoto and Tonegawa, 2017), while the animal was performing a goal-seeking task. The encoding of these reactivated trajectories, however, required the simultaneous recordings from many neurons. Therefore, we reasoned that the accurate reconstruction of 2D trajectories in big environments requires the simultaneous recording from large assemblies. Accordingly, in four rats we recorded CA1 multiple-unit activities in 128 channels (32 tetrodes), yielding simultaneous recording of 383, 243, 206, and 100 putative pyramidal and 31, 18, 30, and 24 putative interneuron units during each animal's recording session (Figures S1 and S2). Trajectory reactivation tends to occur during SWR oscillatory patterns (Buzsáki, 1989; Lee and Wilson, 2002; Figure 1A). We detected these SWR periods during the quiet immobility period after exploration in the absence of theta oscillatory periods (see STAR Methods). We subdivided each SWR into prediction time windows and used a Bayesian method to estimate the location encoded by putative pyramidal neurons (Davidson et al., 2009; Zhang et al., 1998; Figures 1A–1D). Only SWRs that met a quality threshold for the encoded probability (see STAR Methods) and contained at least five encoded trajectory points were included in the subsequent analysis to ensure that only events with reliably encoded places were considered. Altogether n = 558, 693, 544, and 238 SWR events passed the quality threshold for fixed spike number encoding. Each prediction window contained the same number of spikes (n = 15) to ensure uniform coding accuracy. However, similar results were seen using fixed duration windows (8 ms; see STAR Methods) in which case n = 519, 624, 79, and 65 SWRs passed the quality threshold. Reactivation speed was measured by the jump distance between neighboring predicted locations (Figures 1E and S3). Although the reactivation speed remained relatively similar within a SWR, different SWRs yielded wide ranges of mean reactivation speeds, which were lognormally distributed (see Figure 2A inset; all p > 0.3, Pearson chi-square test).

To verify that the encoded SWR trajectories represented place assemblies of the previous environment and were not from chance coding of temporally organized spike trains, we shuffled the cluster identities of the active cells or performed a random 2D rotation of place fields for the events already passing the selection criteria. Both randomized events yielded distributions with higher mean reactivation speeds (all p < 10e−277, Kolmogorov-Smirnov [KS] test) and significantly weaker prediction confidence (all p < 10e−200, KS test; Figures 2A–2C, S4A, and S4B). Next, the role of spike timing in encoding was tested by randomizing them within each SWR. As above, spike time-jittered events exhibited higher mean speeds (all p < 10e−46, KS test) and weaker encoding confidence (all p < 10e−45, KS test). Therefore, the original SWR trajectories were encoded from temporally organized spike trains representing the previous environment. Using larger time-step intervals to compute the reactivation speeds yielded relatively similar mean speed values for the spike-jittered data. However, the mean speed of the original data increased with time-step interval length, making
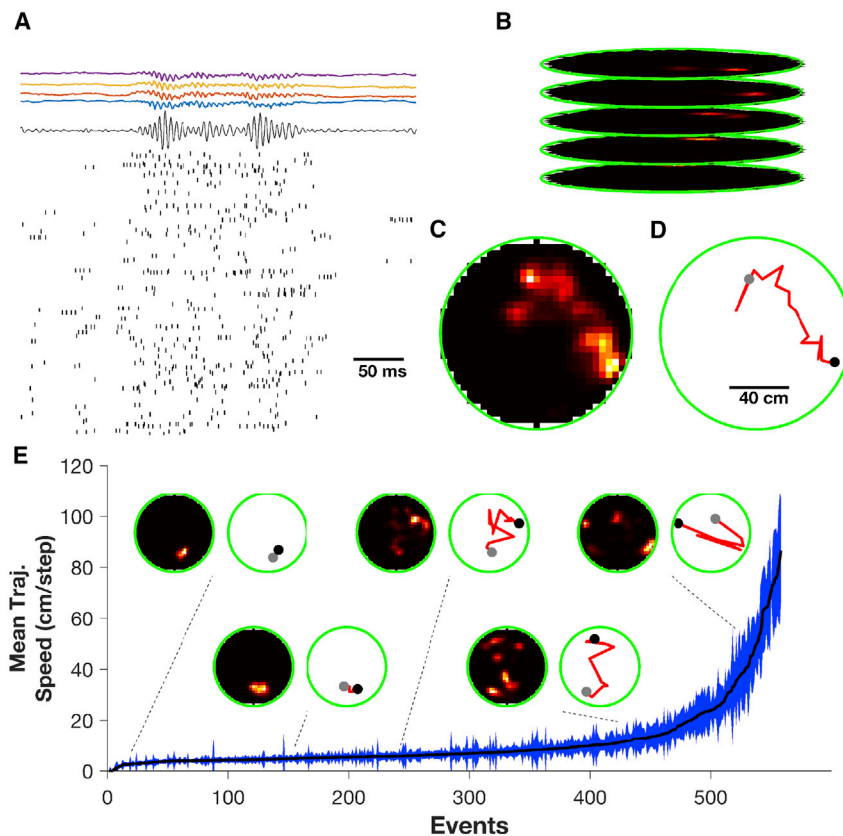
Figure 1. Reactivation of Random 2D Trajectories during SWRs

(A) Neuronal population activity during a SWR. Top four traces show low-pass filtered (<2.5 kHz) local field potential traces from different tetrodes along with one ripple-band (150–250 Hz) filtered trace. Below, the raster plot marks the spike timing of n = 185 putative pyramidal neurons that were active during the SWR.

(B and C) Stacked Bayesian probability maps (B) from five representative prediction windows from the SWR in (A) and the corresponding cumulative map (C) representing the sum of all probability maps. Hotter colors correspond to higher probability values.

(D) Trajectory reconstructed from probability maps of the SWR in (A).

(E) Sorted mean (±SD) reactivation speed of SWR events (n = 558) from one animal. The cumulative map and the encoded trajectory of five events are highlighted representing different reactivation speeds. As above, the heatmap shows the posterior probability density. Gray and black dots on the plotted trajectories indicate the start and end points, respectively. Examples of encoded trajectories from additional animals are shown in Figure S3.

See also Figures S1 and S2.

it eventually larger than the shuffled values for the spike-jittered data (Figure S5A). We then examined whether the places expressed during SWRs were organized. First, we shuffled the order of encoded places within a trajectory, which yielded shifted speed distributions (all p < 10e−57, KS test), indicating that each encoded location influenced the next. Second, we replaced each reactivation speed by randomly drawn speeds from other SWRs (Figures 2D and S4C). These speed-randomized events exhibited a higher coefficient of variation in their jump speed than the unshuffled data (all p < 10e−31, KS test), indicating that each replay expressed a consistent speed throughout the SWR. To test how the use of a quality threshold influenced our results, we performed the same analysis using different quality thresholds. Higher threshold values lead to a larger separation between original and shuffled data distributions for reactivation speed (Figures S5B and S5C), but the shuffled distributions were still significantly different from the original distribution even when no threshold was used (i.e., threshold = 0; all p < 0.01, KS test). Significant differences in the distributions were seen as well when we examined events that fell below a quality threshold (all p < 0.04, KS test; Figure S5D). Given that with the reduction of the quality threshold the original and shuffled distributions got more similar, it is possible that some of the low encoding quality events expressed places or trajectories of another environment, perhaps even "preplayed" assemblies of the environment the animal explored next. The animal explored a novel environment following the rest session we tested (see STAR Methods). The

proportion of SWR events that passed our quality threshold for encoding the subsequent novel environment was low (<0.005) in the previous rest session and was not significantly different from the proportion of events detected after place fields rotation or spike time jittering generated probability maps (Figure S5E; all p > 0.1, Binomial test). These data question whether preplay of trajectories of the subsequently explored novel environment can be seen in our data. Nevertheless, we compared the speed distributions of the original and shuffled data at different threshold levels to check whether any preplay can be detected in our data. At any of the quality thresholds we tested [0:0.12], the speed distributions were significantly different for the place field rotation generated trajectories (all p < 10e−4, KS). However, at any threshold levels the speed distributions of the original trajectories were not significantly different from those of the spike time jittering trajectories (Figure S5F; all p > 0.15, KS test). The speed distribution differences with the place field rotation generated trajectories suggest that in the rest session before the exploration of a novel environment a small subset of the recorded units maintain consistent correlations with that in the subsequent novel environment exploration. Therefore, a possible explanation for these findings is that places expressed in some events at least partially represented patches of the subsequent novel environment. But, considering that spike time-jittered speed distributions were not different from those of the original data, these preplayed trajectories did not exhibit a consistent order in the expression of places to form organized trajectories in which one location will influence the next one. Similar to SWRs representing the familiar environment, a novel environment encoding SWRs expressed consistent trajectories in the subsequent
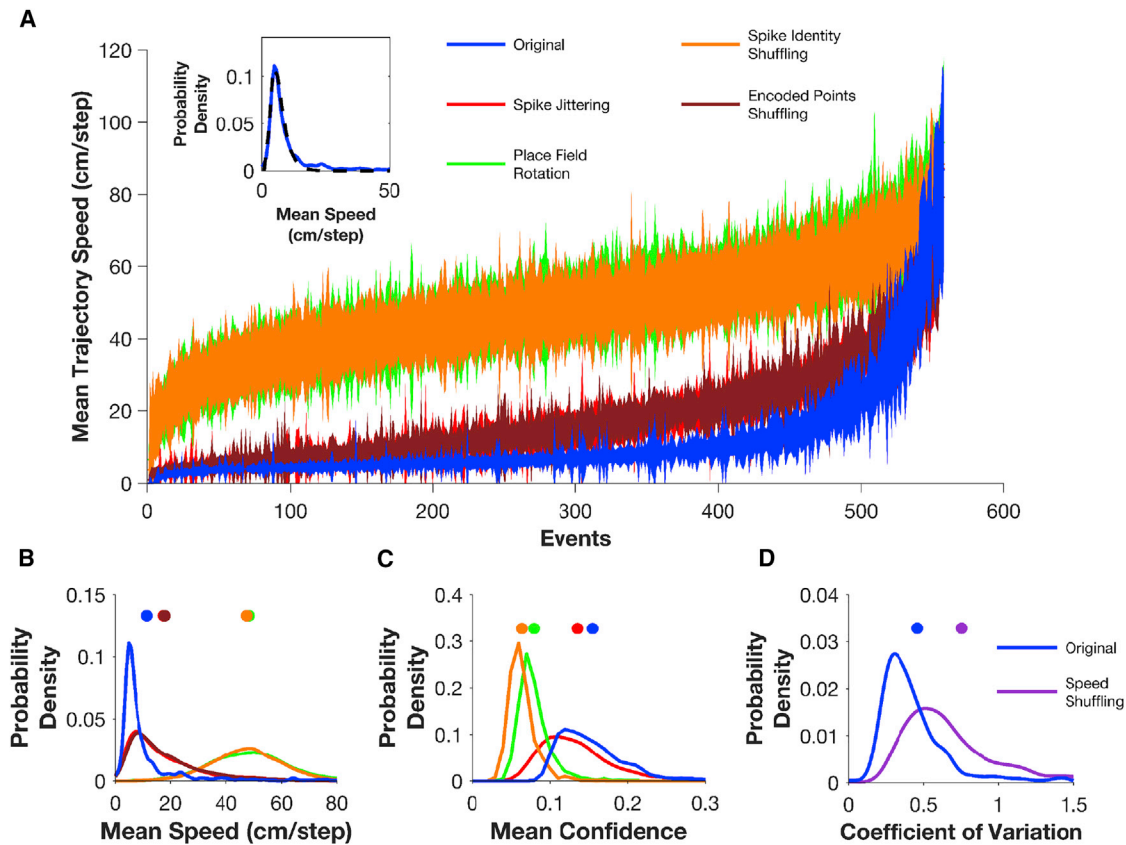
Figure 2. Comparison of Original and Shuffled Trajectories

(A) Sorted mean (±SD) reactivation speed of SWRs and shuffled counterparts for an example session. An equal number of original and randomized events are displayed. Inset: lognormal distribution (dashed line) fits the distribution of mean reactivation speeds for the original events (p = 0.7, Pearson chi-square test).
(B and C) Distribution of the mean reactivation speed (B) and encoding confidence (C) for the original (n = 558) and shuffled (n = 279,000) events.
(D) Distribution of coefficient of variation of the reactivation speed for original events and those in which step sizes representing reactivation speeds were randomly assigned from other events. Filled circles display the median of different distributions. Results for the additional animals are shown in Figure S4. See also Figure S5.

post-novel rest session (Figure S5C). Hence, the apparent lack of trajectory preplay was not due to an impairment of replaying these trajectories after the novel environment exploration.

Reactivation Followed Brownian Diffusion Dynamics
Given that each replay location depended on the previous locations and that place changes maintained a defined speed range, we then asked whether these trajectories can be approximated by a random walk process akin to random particle movements during diffusion (Rudnick and Gaspari, 2004). In physics, diffusion processes are described by a power law between mean distance and time, which represents an exponential relationship characterized by the value of the exponent (Metzler et al., 2012; Vlahos et al., 2008). In Brownian trajectories, the exponent is equal to 0.5. To check the relationship between time and distance, first SWR trajectories were subdivided into homogeneous groups according to their mean reactivation speed. For each reactivation speed group, within each SWR, we took all possible pairs of time windows and measured the distance between predicted places and the time interval between them. On a log-log

scale, they showed a linear relationship (Figure 3A). The power law exponents (α), calculated by the slope of the fits, consistently fell near 0.5 (all p > 0.2, Wilcoxon signed-rank test; Figures 3A, 5A, and 5D) with the slope estimates with the smallest confidence intervals yielding slopes closest to 0.5 (r = 0.64 p < 10e−4; Figures 5B and 5E). Shuffling the places within a trajectory resulted in almost flat lines with slopes significantly smaller than the original ones (all p < 0.0003, Wilcoxon one-tailed signed-rank test). Moreover, at any time interval, the distances obtained from the original data were fitted better by a Rayleigh distribution than the shuffled data, as expected for a 2D Brownian diffusion process (Rudnick and Gaspari, 2004) (all p < 0.0002, Mann-Whitney test; Figures 3B, 3C, 5C, and 5F). These distance distributions also showed a significant overlap (all p > 0.1, Wilcoxon signed-rank test) with those produced by a simulated Brownian process with the same speed and length statistics (Figures 3B, 3C, 5C, and 5F). Crucially, the same analysis applied to the trajectories of the animal and to trajectories reconstructed using neural activity expressed during active behavior yielded different results, with both exponents α
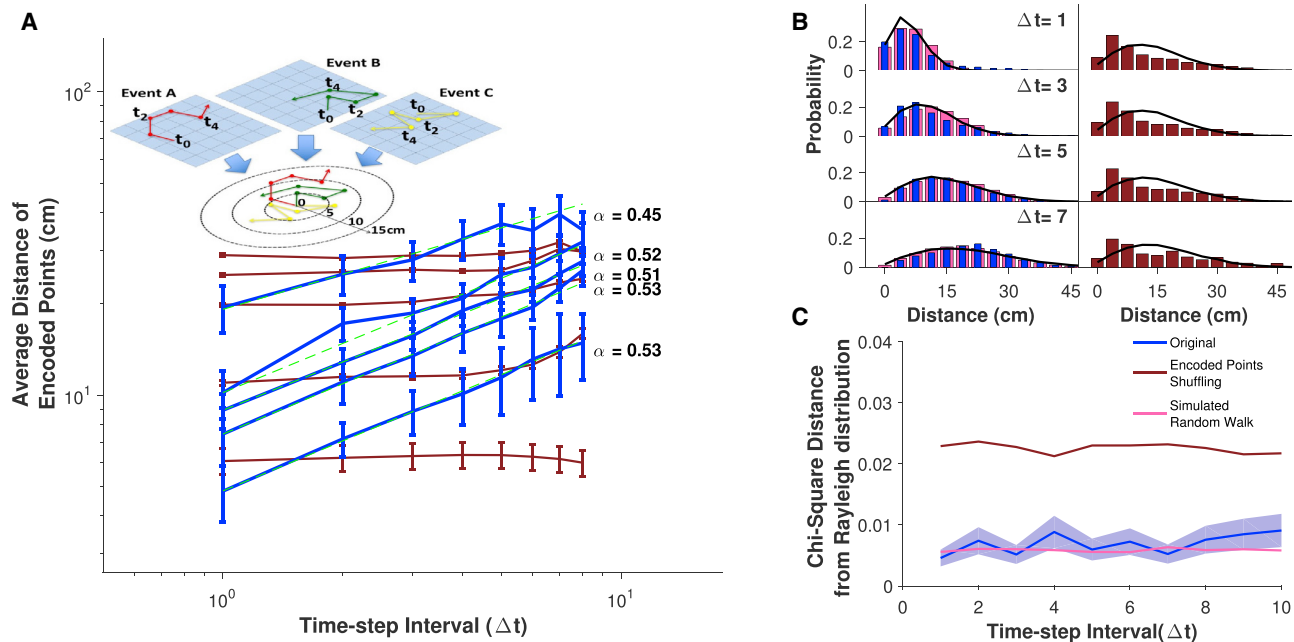
Figure 3. Reactivated Trajectories Resemble Brownian Diffusion

(A) Log-log plot of the time interval (i.e., encoded window distance) and the distance of reactivated positions for different mean reactivation speed groups (n = 59, 287, 87, 33, and 32, respectively, for each). Mean (±SEM) distance is shown for the original events (blue) and those in which the position order was shuffled (red) along with linear regression fit for original events (dashed line) and slopes values (α). Scheme on the top explains how the time-distance relationship was measured on different trajectories. The distances were measured from the reference point t_0 to the next trajectory points ($t_2, t_3 ...$). Four trajectories are displayed on the sides that were all aligned together to originate from the same location. The circles around common starting location illustrate the distance measurement to the next trajectory points relative to t_0.

(B) Distribution of trajectory distances and the best-fit Rayleigh distribution for different time intervals for original (n = 1,915, 1,084, 562, and 250) and shuffled events (n = 584,331, 392,171, 223,826, and 107,002). Left column: original data; right column: encoded points shuffling. Pink overlaid bars show the distribution obtained from a simulated Brownian motion with equal diffusion constant (n equal to original events, all p > 0.1, Wilcoxon signed-rank test).

(C) Average chi-square distance for different time-step intervals between the distribution of trajectory distances and the Rayleigh best fit (blue, with ±SD shade). Red and pink lines: the same measure for shuffled events and simulated Brownian motion, respectively. Results for additional animals are summarized in Figures 5A–5C.

See also Figure S7.

(all p < 10e−4, Wilcoxon signed-rank test) and Rayleigh fit (all p < 0.02, Mann-Whitney test) being incompatible with a Brownian diffusion (Figures 4 and 5G–5I). Also, comparison with a simulated Brownian motion indicated a significant discrepancy in the distance distribution for Δt > 2 (all p < 0.001, Wilcoxon signed-rank test) (Figures 4B and 4C).

We showed above that reactivated trajectories, once started from a point were random; however, their starting position and the direction of the initial trajectory may express bias. Therefore, first we evaluated to what extent the reactivation of starting points was randomly arranged over the environment. By computing the entropy of their spatial density distribution (Figure 6A), we found that, although not completely uniform, their degree of randomness was close to the maximum entropy (>90%; Figure 6B) and similar to that of the rat's behavior (either while running or during the entire exploratory session). However, the distribution of the trajectory starting locations was not related to that of behavioral occupation time during behavioral periods of running (>5 cm/s, running all p > 0.2) or periods including the entire exploration session (all p > 0.08; Figure 6C). Second, we examined the distribution of the direction of the initial reacti-

vated trajectory (i.e., the first step from the starting point to the second encoded point). We computed this distribution for events starting from different zones of the environment (distinguishing the central zone blocks from those close the environmental boundary; Figures 6D and 6E). For each of these groups of events, we compared the distribution of reactivation starting directions to a uniform distribution (Figure 6F). After correcting for multiple comparisons, we found that the first steps were not significantly directionally tuned for events originating in the central blocks (all p > 0.1, Rayleigh-test, Benjamini-Hochberg correction), while we found some degree of tuning in the initial propagation of events originating in blocks in the vicinity of boundaries (all p < 0.05, Rayleigh-test, Benjamini-Hochberg correction).

Network Oscillations Predict Reactivation Dynamics

Co-modulation of different network oscillations may highlight differences in circuit processes during reactivation. Past work showed that beta-gamma band (20–30 Hz) oscillatory power correlated with large jumps in reactivated trajectories (Pfeiffer and Foster, 2015). Therefore, we examined whether the presence
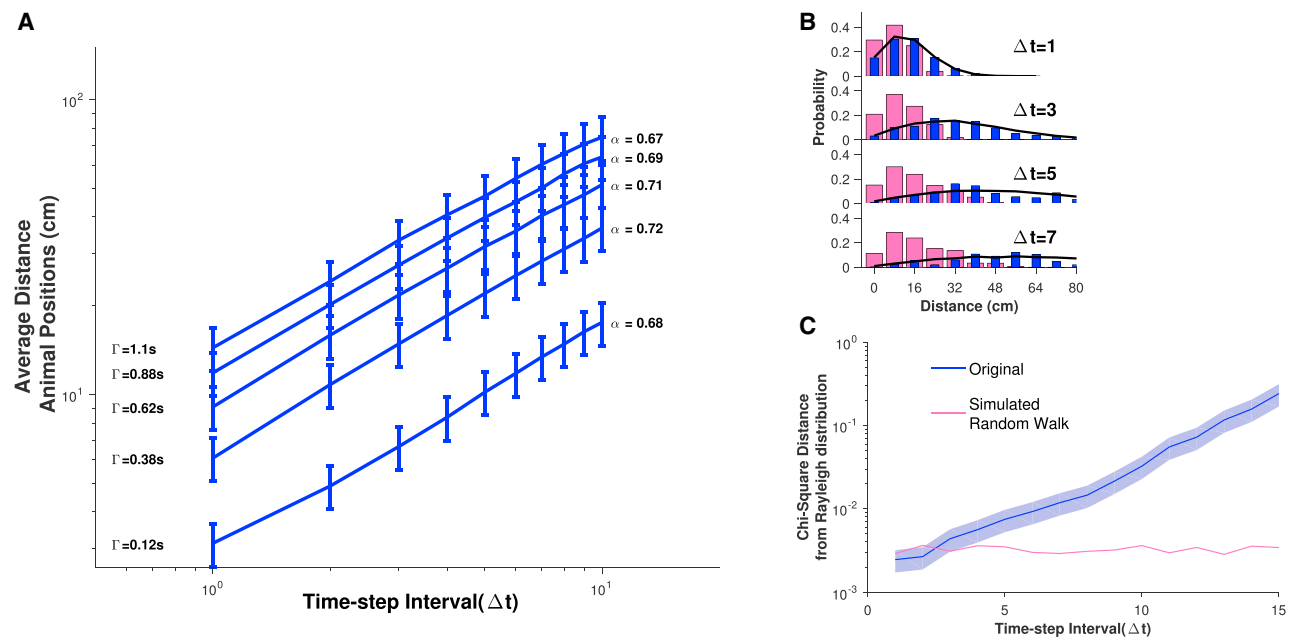
Figure 4. Animal Trajectories during Exploration Are Different from Brownian Diffusion

(A) Log-log plot of the time interval and the mean distance (±SEM) traveled by the animal for different temporal resolutions (Γ on the left, representing the time between trajectory points for Δt = 1). Linear regression fits (dashed line) and slopes values (α) are also shown.

(B) The example demonstrates the comparison of Rayleigh best fit (black line) and the actual distribution (blue bars). Time interval Γ = 1.1 s is shown with representative time-step intervals. Pink overlaid bars show the distribution obtained from a simulated Brownian motion with equal diffusion constant (n equal to original events) (for Δt > 1 p < 0.001, Wilcoxon signed-rank test).

(C) Average (shaded region: ±SD) chi-square distance between the distribution of trajectory distances and the best-fit Rayleigh distribution for different time-step intervals. Pink line: same measure for simulated Brownian motion. Note that the chi-square distance diverges for larger time intervals. Results for additional animals are summarized in Figures 5G–5I.

See also Figure S7.

of an underlying network oscillation predicted such jumps. Oscillations may show a relationship with additional assembly measures because assembly differences are not solely related to the encoded distance (Leutgeb et al., 2005). We assessed population activity difference (PAD) by a modified cosine distance between spike-count vectors as measured across prediction time windows (see STAR Methods). We chose this measure because it quantifies the orthogonality of assemblies and reactivated assemblies have been suggested to be orthogonal (Malvache et al., 2016) (Figures 7A and S6). Although the PAD (i.e., Figure 7A, Δt = 1) was increased with the distance of the encoded positions across neighboring windows, it saturated beyond a certain distance (Figure 7A, inset). The curves with Δt > 1 window distance were similar, but these progressively up-shifted with Δt (Figure 7A). Accordingly, PAD correlated with window distance Δt (all p < 0.0001), and this correlation was independent of the encoded distance (all p < 0.0002, partial correlation). This partial correlation analysis was performed by taking all prediction window pairs in each SWR and calculating window distance, PAD, and encoded distance and tested whether window distances can be predicted by PAD even when the correlations between PAD and encoded distance were taken into account. In addition, we examined the increase of PAD with window distance in cases when the same place was encoded over a series of consecutive prediction windows (see encoded distance

0 at Figure 7A). Because PAD correlated with elapsed reactivation time (i.e., window distance) independently from reactivated positions, reactivation time and trajectory positions may be coded separately during SWRs. Interestingly, the substantial orthogonality of the two ensemble measures was confirmed by looking at the distance-time relationship (as in Figure 3) but this time using a PAD metric. To do so, the PAD distances measured across different encoding windows were nonlinearly embedded in a two-dimensional space for each of the speed groups (Figure S7A). Two major differences appeared in this case with respect to the embedded PAD distance measure: first, the speed of the PAD increase was similar in different reactivation speed groups, and, second, rather than following a power law, an exponential function fitted the curves best. Both findings point to major differences in the mechanisms regulating the evolution of the two ensemble measures during SWRs.

Finally, we assessed whether SWR events associated with different average reactivation speed (i.e., average encoding distance at Δt = 1) or average PAD (Δt = 1) are related to other aspects of network activity. Putative interneuron firing rate correlated with the average PAD (all p < 10e−7) but not with average reactivation speed (all p > 0.07) during SWRs (Figures 7C, 7D, and S6) indicating that putative interneuron activity was stronger when the orthogonality of place cell population patterns was higher during SWRs, independent of the average reactivation
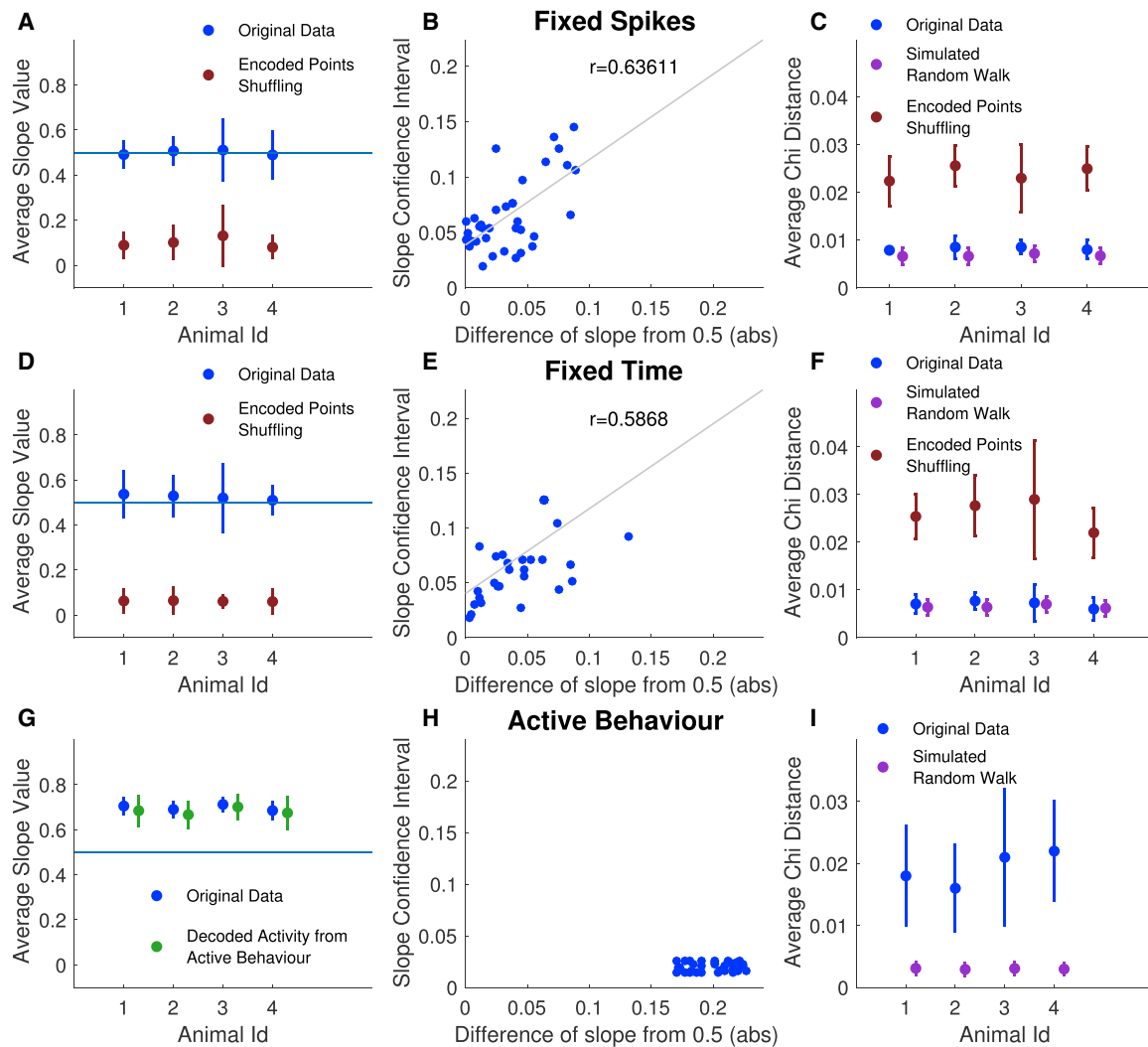
Figure 5. Testing for Brownian Diffusion Properties and Comparison of Different Conditions

(A) Average regression slope obtained (see Figure 3A) from original data (blue) and encoded positions shuffling (magenta) for each animal's session using fixed-spike windows. The theoretical prediction of a slope equal to 0.5 is shown for reference (blue line).

(B) Results for individual regression lines. For each regression, we show the relationship between the discrepancy of the slope from 0.5 (absolute value) and the size of the confidence interval on the estimated slope value for the original data using fixed-spikes windows. For every session, the estimated slopes were gathered by using different reactivation speed group subdivisions starting from 3 up to the maximum number of subdivision in which at least a minimum number of SWRs events belong to each group (see STAR Methods). The correlation line between the two quantities is also shown (r = 0.64, p < 10−4, n = 39).

(C) Average chi-square distance between the distribution of reconstructed trajectory distances and the Rayleigh best fit (maximum likelihood). Results are shown for every session from different animals using fixed-spikes windows. Results for the original data (blue) and simulated Brownian motion (purple) were not significantly different (all p > 0.4, Wilcoxon signed-rank test), but original data chi-square distances were significantly different from those resulted from encoded positions shuffling (red, all p < 0.005, Wilcoxon signed-rank test).

(D–F) Same as above but using fixed-time windows. Average regression slopes (D), results for individual regression lines (E), and average chi-square distance between the distribution of reconstructed trajectory distances and the Rayleigh best-fit (F) are shown. Regression result for (E): r = 0.59, p < 10−4, n = 36. Results for the original data and simulated Brownian motion were not significantly different (all p > 0.35, Wilcoxon signed-rank test), but original data chi-square distances were significantly different from those resulting from encoded positions shuffling (red, all p < 0.01, Wilcoxon signed-rank test).

(G–I) Same as above but showing results from the analysis of the trajectories of the animals. Average regression slopes (G), results for individual regression lines (H), and average chi-square distance between the distribution of reconstructed trajectory distances and the Rayleigh best-fit (I) are shown. Slopes in (G) obtained from active (5 cm/s) behavior (blue) and from reconstructed positions from neural activity (purple) were similar (all p > 0.4, Wilcoxon test). Regression result for (H): r = 0.14, p = 0.4, n = 48. Results from original trajectories and those from simulated Brownian motion were different (all p < 0.001, Wilcoxon signed-rank test). All error bars ± SD.
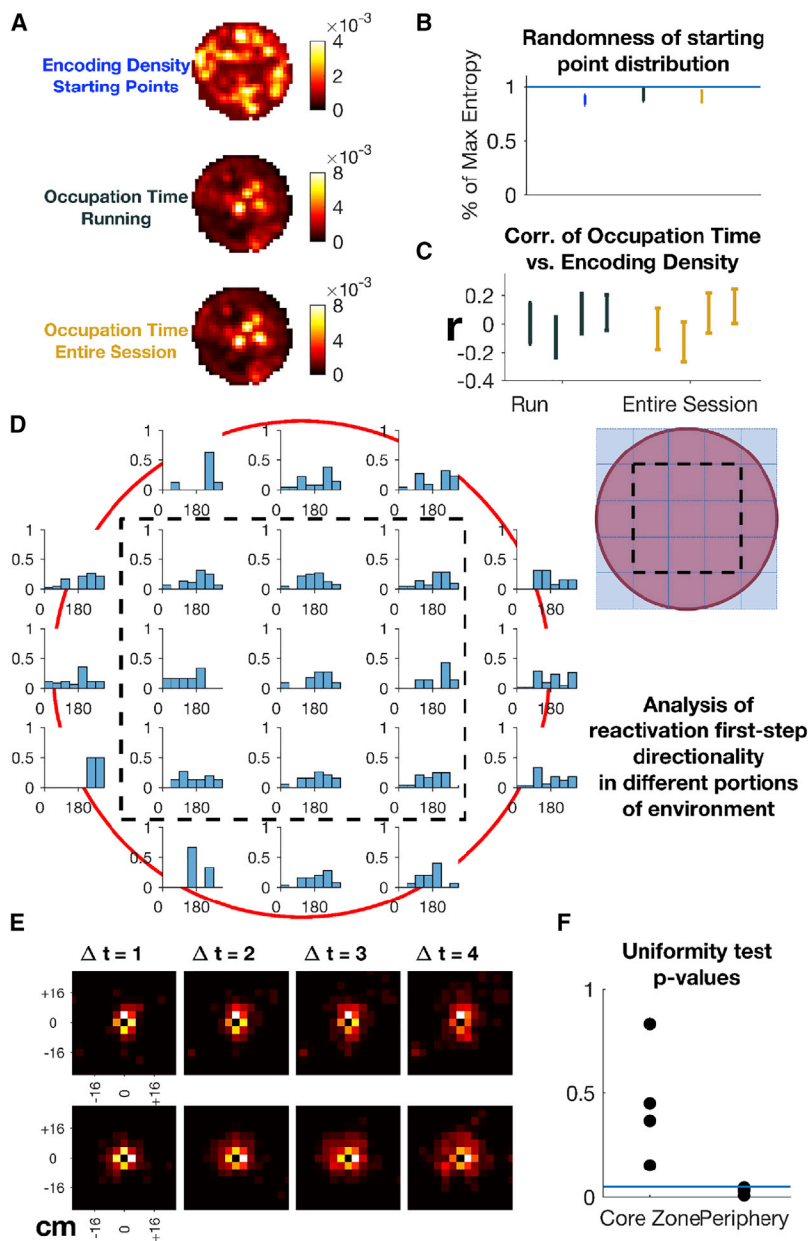
See also Figure S7.

Figure 6. The Start Location and Propagation Directionality of Reactivated Trajectories Are Random

(A) Examples of the spatial distributions representing the density of the encoded starting positions during SWRs, the animal's occupation time for periods of active behavior (speed > 5 cm/s) and over the entire exploration session.

(B) Entropy estimate of the degree of randomness of spatial distributions for trajectory starting position density (blue), active behavior (gray), and behavior in the entire session (ochre). Entropy values were normalized to the maximum entropy of the uniform distribution over the same number of spatial bins (horizontal line). Data for grouped animals are shown.

(C) Spatial correlation of reactivation starting point density with running occupation time (gray) and entire session occupation time (ochre) were not significant in any of the four animals (run: all p > 0.2, entire session: all p > 0.09).

(D) Example of the distribution of the direction of the first encoded trajectory steps. Distributions were calculated separately for different sub-regions of the environment depending on where the trajectory was originated from (see inset on the top right for an illustration of the division of the environment in 25 sub-regions). The red circle illustrates the boundaries of the environment, while the dashed rectangle illustrates the separation of histogram zones near the center or the boundaries.

(E) The displacement distribution of the reactivation trajectories measured relative to the trajectory origin. The first four trajectory point distributions (columns) are shown from two different zones of trajectory origin (rows) as in (D). Intensity levels represent probability values with hot colors indicating higher relative probabilities than colder colors.

(F) The session-by-session lowest p value (each point corresponds to separate session per animal) for testing uniformity of the directionality of reactivation first-steps (Rayleigh-test, Benjamini-Hochberg correction for multiple comparisons). p values were evaluated separately for portions of the environment either laying at its center (Core zone, left) or along its border ones (Periphery, right; all p < 0.03) (see STAR Methods). Lines mark the 0.05 significance level.

oscillations (Figures 7E and S6). Therefore, both ripple-band power and enhanced putative interneuron firing rate predicted reactivation events involving larger changes in assembly pattern orthogonality. In contrast, fast gamma oscillations predicted reactivation speed within an event.

DISCUSSION

We showed that, following non-stereotypical exploration of open environments, SWR events reactivated random, 2D trajectories of the previous environment that were governed by a Brownian diffusion process. The pattern of the replayed trajectories was not a direct representation of behavior since the animals' trajectory did not exhibit Brownian statistics. This suggests that these reactivated 2D trajectories were not a replay of directly learned and stored patterns. Moreover, the flexible expression of our patterns, covering wide ranges of temporal and spatial scales,
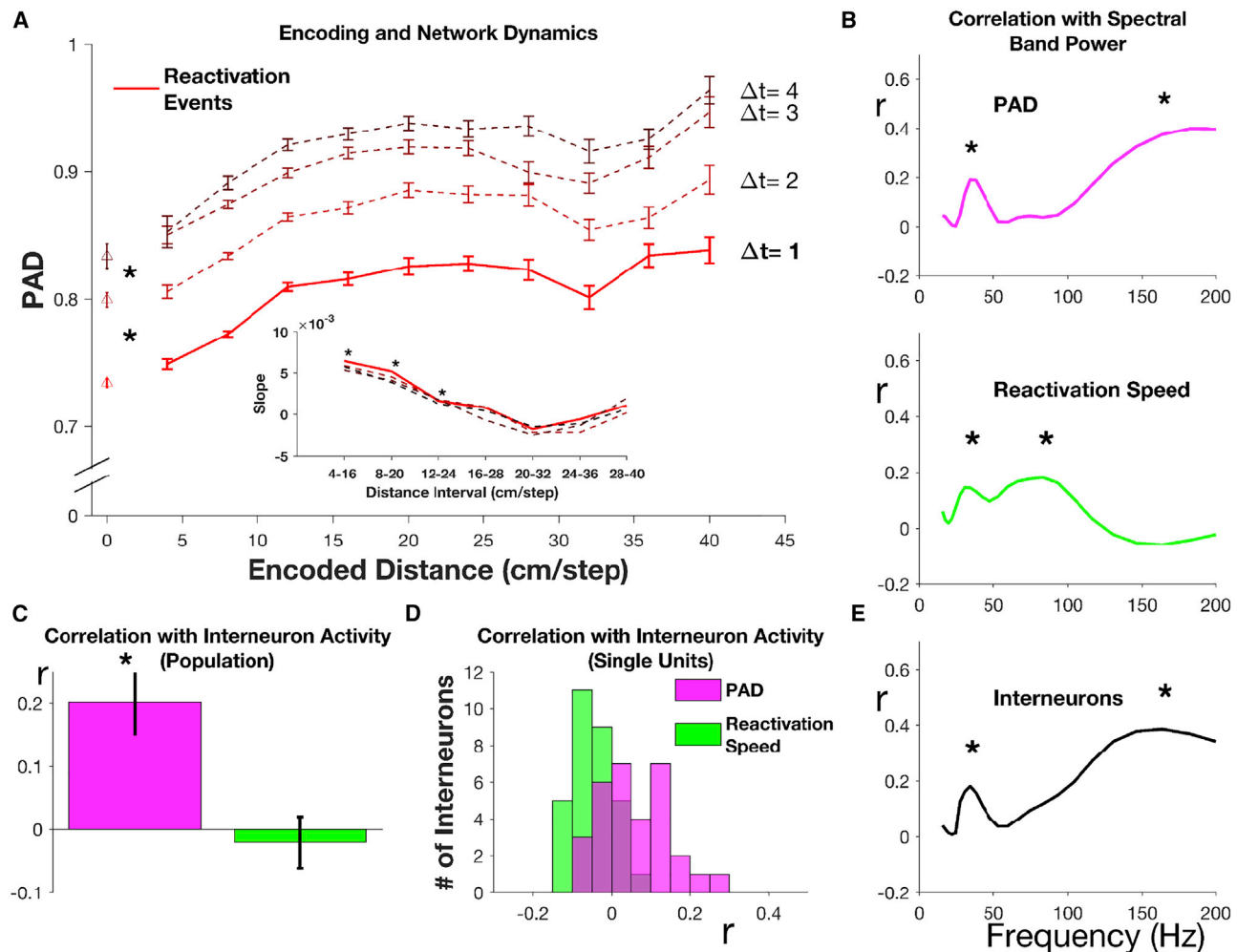
speed. Similar to putative interneuron activity, we also found that the PAD but not reactivation speed correlated with the number of active putative pyramidal cells in the SWR and with their firing rate (all p < 10e−4) during the reactivation event (Figure S7B). Both the average PAD and reactivation speed measured within each SWR showed a distinct relationship with local field potential oscillatory power. While the average PAD showed a strong correlation with the ripple band (150–200 Hz, all p < 10e−12), reactivation speed was correlated with fast gamma power (≈80 Hz, all p < 10e−5; Figures 7B and S6). In addition, both the reactivation speed (all p < 0.0004) and PAD (all p < 0.005) were correlated with a slow ≈30 Hz-peaked gamma band power. Putative interneuron SWR firing rate correlated with both slow gamma (all p < 10e−8) and ripples (all p < 10e−13) but not with fast gamma

Figure 7. Network Oscillations Associated with Different Reactivation Events

(A) The relationship between the encoded distance of trajectory positions and PAD during SWR for an example session is shown. Solid line: neighboring trajectory positions (i.e., time interval Δt = 1); dashed lines represent Δt = 2,3,4. Triangles: average PAD for Δt = 1,2,3 calculated from continuous sequences encoding the same spatial bin (Δt = 1 to Δt = 2 all p < 10e−6; Δt = 2 to Δt = 3 all p < 0.015, Wilcoxon signed-rank test). Inset: linear regression slopes obtained from subsets of data of increasing encoded distance interval (*p < 0.01; rest: p > 0.15).

(B) Correlation of the average PAD (top) and average reactivation speed (bottom) with the different bands of the SWR-associated power spectrum (single session). The correlation was calculated across different SWR events (n = 558) (PAD ≈ 30 Hz p < 0.003, PAD 150–200 Hz p < 10e−13; reactivation speed ≈ 30 Hz p < 10e−4, reactivation speed ≈ 80 Hz p = 10e−6).

(C) Correlation of summed putative interneuron firing rate with the average PAD or with reactivation speed for a single session (±SEM). For each SWR, the average PAD and reactivation speed were measured for Δt = 1 (PAD, p < 10e−6, reactivation speed p = 0.44).

(D) Same as (C) but correlations were calculated for single putative interneurons, and the distribution of correlations coefficients is shown.

(E) Correlation of summed putative interneuron rate with different components of the SWR-associated power spectrum for a single session (≈30 Hz p < 10e−9, 150–200 Hz p < 10e−14).

Results for the additional animals are shown in Figure S6.

makes it unlikely that they reflected a limited repertoire of pre-wired assemblies. It is also unlikely that sensory experience has driven these patterns because the animal rested and slept in a different sleep box without any visual contact with the explored environment. Instead, our data suggest that the hippocampus is capable of expressing random trajectories of any stored map, although preexisting connection topology of hippocampal principal cells may influence expression dynamics (Guzman et al., 2016).

We used a random foraging behavioral paradigm in order to disentangle neural representations from stereotyped behavioral motifs. Under such circumstances, one might expect that SWR firing encodes individual places or path segments from prior exploration. Alternatively, when the animal does not experience similar, stereotyped trajectories multiple times, replay might be disordered, where the set of assemblies active during a given SWR are unrelated to each other, and their expression of a location at one instance will not influence the next one. However, the

assembly firing observed here represented a random walk through the environment, where each encoded place within the SWR was separated by a similar distance (i.e., had a consistent reactivation speed). Such trajectories can be analogous to that of gaseous particle trajectories in which the temperature of the media influences the trajectory speed, and, each time the particle collides with another, it will change its movement to a random direction. The Brownian reactivated trajectories we observed can be modeled as the sequential recruitment of place cell assemblies in which activation of one assembly triggers the expression of the next (Monasson and Rosay, 2015; Romani and Tsodyks, 2015) in any spatial direction, but assembly transitions of a SWR allow similar distances only. This ability enables the hippocampus to reactivate memory traces covering diverse spatiotemporal scales to meet mnemonic demand flexibly. Such activity could represent the default mode of reactivated output during SWRs when unshaped by experience or sensory inputs. This would form a part of the generalized network processes that endow the hippocampus with the flexibility necessary to provide the support for more demanding cognitive tasks.

The capacity of the hippocampus to generate unconstrained propagation of activity through neural representations of space could underpin a number of potential cognitive functions of SWRs, including prospective novel route planning (van de Ven et al., 2016) and replaying alternative routes at spatial decision points (Johnson and Redish, 2007; Papale et al., 2016). It could also serve to facilitate the assimilation of new maps into existing related representations or integrate previously unexplored path segment combinations within a familiar environment (Gupta et al., 2010). However, the replay of memory episodes could arise from a mechanism that biases the spontaneous propagation of map trajectories toward salient or stereotyped behavioral episodes. Stereotypical trajectories may be stored within hippocampal circuits through a process that makes the reactivation of behaviorally relevant events more probable than other random trajectories. A prior replay of stereotyped trajectories during SWRs and theta oscillation in waking exploration (Davidson et al., 2009; Foster and Wilson, 2006; Gupta et al., 2010; Zheng et al., 2016) may reinforce this biasing process, as well as activity-dependent strengthening between adjacent spatial assemblies during exploration (Ekstrom et al., 2001; Mehta et al., 2000). A central feature of episodic experience is that it can cover various time and spatial scales. Replay events with different reactivation speeds can construct representations on multiple scales, providing the degree of signal compression to convey complex neural representations efficiently.

Not only was the expression of ongoing assembly patterns controlled by diffusion dynamics and the reactivation speed range of the event, but the orthogonality of the expressed assemblies was also regulated during the course of a SWR event. That is, assemblies expressed further apart within a SWR were more orthogonal (i.e., had larger PADs) than those in the vicinity (Figure 7A). Such an increase in the PAD (i.e., increase in orthogonality) is shown by the positive correlation between the reactivation time (i.e., encoding window distance during a SWR) and PAD, even when the PAD versus encoded distance correlation was taken into account. Such a relationship can provide the means to encode reactivation time. Our data show that different

assemblies can represent the same reactivated place, with varying degrees of orthogonality. In this way, orthogonality between assemblies can independently represent the relative time elapsed between each encoded location and can even signal the duration of staying at a similar place. Over longer periods, such as days or weeks, a subpopulation of assemblies encoding places may be altered suggesting that the change of spatial assemblies can code for time (Mankin et al., 2012; Ziv et al., 2013). Moreover, sequential activation of place cells can signal the time elapsed during a task (MacDonald et al., 2011). However, the increase of orthogonality with reactivation time may additionally relate to the functional limitation of hippocampal circuit computations. Here, multiple orthogonal assemblies that represent overlapping locations would also allow similar places represented throughout the SWR, without the same pyramidal cells firing together for long durations.

Considering that assembly orthogonality (i.e., PAD) and reactivation speed are correlated with different bands of network oscillations, multiple mechanisms may work in parallel during reactivation. In particular, the PAD was correlated with ripple-band power. The putative interneuron rate was also correlated with the PAD, suggesting that interneurons play a role in assembly turnover during replay. Although CA2/CA3a region activity triggers sharp waves and SWR emission (Csicsvari et al., 2000; Oliva et al., 2016), the ripple oscillation itself is generated by local CA1 pyramidal cell-interneuron interactions (Gan et al., 2017; Schlingloff et al., 2014; Stark et al., 2015). Therefore, it seems possible that changes in assembly orthogonality are modulated partially within the CA1 area through the control of local inhibitory interneurons. The local emergence of CA1 sequences is supported by the observation that the CA1 network is capable of autonomously generating sequences (Stark et al., 2015). When ripples were artificially generated with optogenetic stimulation of CA1 pyramidal cells, this gave rise to sequences that were similar to that observed during spontaneous SWR emission (Stark et al., 2015). Moreover, the consistency between waking sequential coding and SWR activity, artificial or spontaneous, was dependent on interneuron firing.

While the power of ripple oscillations was associated with PADs and was not related to reactivation speed, fast gamma power was related only to reactivation speed and not to PADs. Interestingly, waking gamma oscillation power is correlated with the actual speed of the animal (Ahmed and Mehta, 2012). Fast gamma oscillations have been associated with interactions between CA1 and neighboring regions such as the dentate gyrus and the entorhinal cortex (Colgin et al., 2009; Sullivan et al., 2011). This suggests that reactivation dynamics regulated by reactivation speed is correlated with interregional interactions and not with local activity modulation mechanisms like inhibitory activity or ripple power. Therefore, some of the regulatory mechanisms determining reactivated trajectories might originate from nonlocal mechanisms such as the medial entorhinal cortex inputs, including those that encode speed (Yamamoto and Tonegawa, 2017; Ye et al., 2018). At the same time, however, slow gamma band oscillations show a complementary behavior, relating to both reactivation speed and PADs. In agreement with this finding large jumps in reactivation places have been suggested to occur in different slow gamma packets (Pfeiffer and Foster, 2015), and even

CellPress

theta sequences represent more stretched trajectories when slow gamma oscillations were nested on theta oscillations (Zheng et al., 2016). In our case, larger-amplitude, slow gamma events predicted larger jump sizes, but they also predicted larger PAD distance. Our results point to a complex system of information routing at work during the unfolding of SWR events. Understanding the mechanism of information routing and the underlying system interactions will require a more detailed description of the activity state not only of CA1, but also of neighboring areas (Ólafsdóttir et al., 2016; O'Neill et al., 2017).

The ability of the hippocampal network to spontaneously reactivate random trajectories of the previously active map provides not only a scaffold on which to build declarative memory traces, but also a means to generate potential future behavioral choices. Attractor networks are able to optimize cost functions to solve complex problems (Hopfield and Tank, 1985). Such random replay may reflect such an optimization ability of the hippocampal attractor network representing space (Samsonovich and McNaughton, 1997). The ability of hippocampal circuits to replay random trajectories of a recently experienced environment may provide better behavioral solutions and help behavioral adaptation through initiating optimization processes during sleep.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHODS DETAILS
  - Animals and Surgery
  - Data Acquisition and Behavior
  - Spike Sorting
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Rate Map Generation
  - Place Cell Classification
  - SWR Detection
  - Time-Binning of the Data
  - Reconstruction of Position (Bayesian Decoding)
  - Analysis of Reactivation Events
  - Shuffling Procedures
  - Reactivated Trajectory Classification
  - Novel Environment Analysis
  - Brownian Diffusion Analysis
  - Behavioral Data
  - Measures of Spatial Distribution Randomness
  - Measures of Reactivation Variability
  - PAD Embedding Analysis
  - Interneuron Activity
  - Network Oscillations
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and can be found with this article online at https://doi.org/10.1016/j.neuron.2019.01.052.

REFERENCES

Ahmed, O.J., and Mehta, M.R. (2012). Running speed alters the frequency of hippocampal gamma oscillations. J. Neurosci. *32*, 7373–7383.

Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience *31*, 551–570.

Carr, M.F., Karlsson, M.P., and Frank, L.M. (2012). Transient slow gamma synchrony underlies hippocampal memory replay. Neuron *75*, 700–713.

Colgin, L.L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., Moser, M.B., and Moser, E.I. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. Nature *462*, 353–357.

Csicsvari, J., Hirase, H., Czurkó, A., Mamiya, A., and Buzsáki, G. (1999a). Fast network oscillations in the hippocampal CA1 region of the behaving rat. J. Neurosci. *19*, RC20.

Csicsvari, J., Hirase, H., Czurkó, A., Mamiya, A., and Buzsáki, G. (1999b). Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving Rat. J. Neurosci. *19*, 274–287.

Csicsvari, J., Hirase, H., Mamiya, A., and Buzsáki, G. (2000). Ensemble patterns of hippocampal CA3-CA1 neurons during sharp wave-associated population events. Neuron *28*, 585–594.

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

Dragoi, G., and Tonegawa, S. (2011). Preplay of future place cell sequences by hippocampal cellular assemblies. Nature *469*, 397–401.

Dunn, B., Mørreaunet, M., and Roudi, Y. (2015). Correlations and functional connections in a population of grid cells. PLoS Comput. Biol. *11*, e1004052.

Ekstrom, A.D., Meltzer, J., McNaughton, B.L., and Barnes, C.A. (2001). NMDA receptor antagonism blocks experience-dependent expansion of hippocampal "place fields". Neuron *31*, 631–638.

Feng, T., Silva, D., and Foster, D.J. (2015). Dissociation between the experience-dependent development of hippocampal theta sequences and single-trial phase precession. J. Neurosci. *35*, 4890–4902.

Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440*, 680–683.

Gan, J., Weng, S.-M., Pernía-Andrade, A.J., Csicsvari, J., and Jonas, P. (2017). Phase-locked inhibition, but not excitation, underlies hippocampal ripple oscillations in awake mice in vivo. Neuron *93*, 308–314.

Grosmark, A.D., and Buzsáki, G. (2016). Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. Science *351*, 1440–1443.

Gupta, A.S., van der Meer, M.A.A., Touretzky, D.S., and Redish, A.D. (2010). Hippocampal replay is not a simple function of experience. Neuron 65, 695–705.

Gupta, A.S., van der Meer, M.A.A., Touretzky, D.S., and Redish, A.D. (2012). Segmentation of spatial experience by hippocampal θ sequences. Nat. Neurosci. 15, 1032–1039.

Guzman, S.J., Schlögl, A., Frotscher, M., and Jonas, P. (2016). Synaptic mechanisms of pattern completion in the hippocampal CA3 network. Science 353, 1117–1123.

Harris, K.D., Henze, D.A., Csicsvari, J., Hirase, H., and Buzsáki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J. Neurophysiol. 84, 401–414.

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

Johnson, A., and Redish, A.D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. J. Neurosci. 27, 12176–12189.

Kentros, C., Hargreaves, E., Hawkins, R.D., Kandel, E.R., Shapiro, M., and Muller, R.V. (1998). Abolition of long-term stability of new hippocampal place cell maps by NMDA receptor blockade. Science 280, 2121–2126.

Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36, 1183–1194.

Leutgeb, S., Leutgeb, J.K., Barnes, C.A., Moser, E.I., McNaughton, B.L., and Moser, M.B. (2005). Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. Science 309, 619–623.

MacDonald, C.J., Lepage, K.Q., Eden, U.T., and Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. Neuron 71, 737–749.

Malvache, A., Reichinnek, S., Villette, V., Haimerl, C., and Cossart, R. (2016). Awake hippocampal reactivations project onto orthogonal neuronal assemblies. Science 353, 1280–1283.

Mankin, E.A., Sparks, F.T., Slayyeh, B., Sutherland, R.J., Leutgeb, S., and Leutgeb, J.K. (2012). Neuronal code for extended time in the hippocampus. Proc. Natl. Acad. Sci. USA 109, 19462–19467.

Mehta, M.R., Quirk, M.C., and Wilson, M.A. (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. Neuron 25, 707–715.

Metzler, R., Chechkin, A.V., and Klafter, J. (2012). Levy statistics and anomalous transport: levy flights and subdiffusion. Computational Complexity (Springer), pp. 1724–1745.

Monasson, R., and Rosay, S. (2015). Transitions between spatial attractors in place-cell models. Phys. Rev. Lett. 115, 098101.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. 34, 171–175.

O'Keefe, J., and Nadel, L. (1978). Hippocampus as a Cognitive Map (Clarindon).

O'Keefe, J., and Recce, M.L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus 3, 317–330.

O'Neill, J., Senior, T., and Csicsvari, J. (2006). Place-selective firing of CA1 pyramidal cells during sharp wave/ripple network patterns in exploratory behavior. Neuron 49, 143–155.

O'Neill, J., Pleydell-Bouverie, B., Dupret, D., and Csicsvari, J. (2010). Play it again: reactivation of waking experience and memory. Trends Neurosci. 33, 220–229.

O'Neill, J., Boccara, C.N., Stella, F., Schoenenberger, P., and Csicsvari, J. (2017). Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. Science 355, 184–188.

Ólafsdóttir, H.F., Barry, C., Saleem, A.B., Hassabis, D., and Spiers, H.J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. eLife 4, e06063.

Ólafsdóttir, H.F., Carpenter, F., and Barry, C. (2016). Coordinated grid and place cell replay during rest. Nat. Neurosci. 19, 792–794.

Ólafsdóttir, H.F., Bush, D., and Barry, C. (2018). The role of hippocampal replay in memory and planning. Curr. Biol. 28, R37–R50.

Oliva, A., Fernández-Ruiz, A., Buzsáki, G., and Berényi, A. (2016). Role of hippocampal CA2 region in triggering sharp-wave ripples. Neuron 91, 1342–1355.

Papale, A.E., Zielinski, M.C., Frank, L.M., Jadhav, S.P., and Redish, A.D. (2016). Interplay between hippocampal sharp-wave-ripple events and vicarious trial and error behaviors in decision making. Neuron 92, 975–982.

Pfeiffer, B.E., and Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. Nature 497, 74–79.

Pfeiffer, B.E., and Foster, D.J. (2015). PLACE CELLS. Autoassociative dynamics in the generation of sequences of hippocampal place cells. Science 349, 180–183.

Romani, S., and Tsodyks, M. (2015). Short-term plasticity based network model of place cells dynamics. Hippocampus 25, 94–105.

Rudnick, J., and Gaspari, G. (2004). Elements of the Random Walk: An Introduction for Advanced Students and Researchers (Cambridge University Press).

Sammon, J.W. (1969). A nonlinear mapping for data structure analysis. IEEE Trans. Comput. 18, 401–409.

Samsonovich, A., and McNaughton, B.L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. J. Neurosci. 17, 5900–5920.

Schlingloff, D., Káli, S., Freund, T.F., Hájos, N., and Gulyás, A.I. (2014). Mechanisms of sharp wave initiation and ripple generation. J. Neurosci. 34, 11385–11398.

Shen, B., and McNaughton, B.L. (1996). Modeling the spontaneous reactivation of experience-specific hippocampal cell assembles during sleep. Hippocampus 6, 685–692.

Silva, D., Feng, T., and Foster, D.J. (2015). Trajectory events across hippocampal place cells require previous experience. Nat. Neurosci. 18, 1772–1779.

Skaggs, W.E., McNaughton, B.L., Wilson, M.A., and Barnes, C.A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus 6, 149–172.

Stark, E., Roux, L., Eichler, R., and Buzsáki, G. (2015). Local generation of multineuronal spike sequences in the hippocampal CA1 region. Proc. Natl. Acad. Sci. USA 112, 10521–10526.

Sullivan, D., Csicsvari, J., Mizuseki, K., Montgomery, S., Diba, K., and Buzsáki, G. (2011). Relationships between hippocampal sharp waves, ripples, and fast gamma oscillation: Influence of dentate and entorhinal cortical activity. J. Neurosci. 31, 8605–8616.

Tenenbaum, J.B., de Silva, V., and Langford, J.C. (2000). A global geometric framework for nonlinear dimensionality reduction. Science 290, 2319–2323.

van de Ven, G.M., Trouche, S., McNamara, C.G., Allen, K., and Dupret, D. (2016). Hippocampal offline reactivation of recently formed cell assembly patterns during sharp wave-ripples. Neuron 92, 968–974.

Vlahos, L., Isliker, H., Kominis, Y., and Hizanidis, K. (2008). Normal and anomalous diffusion: a tutorial. arXiv, arXiv:08050419 https://arxiv.org/abs/0805.0419.

Wu, X., and Foster, D.J. (2014). Hippocampal replay captures the unique topological structure of a novel environment. J. Neurosci. 34, 6459–6469.

Yamamoto, J., and Tonegawa, S. (2017). Direct medial entorhinal cortex input to hippocampal ca1 is crucial for extended quiet awake replay. Neuron 96, 217–227.

Ye, J., Witter, M.P., Moser, M.-B., and Moser, E.I. (2018). Entorhinal fast-spiking speed cells project to the hippocampus. Proc. Natl. Acad. Sci. USA 115, E1627–E1636.

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting neuronal population activity by reconstruction: Unified framework with application to hippocampal place cells. J. Neurophysiol. 79, 1017–1044.

Zheng, C., Bieri, K.W., Hsiao, Y.-T., and Colgin, L.L. (2016). Spatial sequence coding differs during slow and fast gamma rhythms in the hippocampus. Neuron 89, 398–408.

Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nat. Neurosci. 16, 264–266.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Long Evans Rats	Janvier, France	RjOrl:LE
Software and Algorithms		
KlustaKwik	Harris et al., 2000	https://github.com/klusta-team/klustakwik/
Python	https://www.python.org	Python
MATLAB	https://de.mathworks.com/	MATLAB
Other		
12um tungsten wires	California Fine Wire	M294520
Headstage amplifier	Axona, St. Albans, UK	http://www.axona.com/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and datasets should be directed to and will be fulfilled by the Lead Contact, Jozsef Csicsvari (jozsef.csicsvari@ist.ac.at).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All procedures involving experimental animals were carried out in accordance with Austrian animal law (Austrian federal Law for experiments with live animals) under a project license approved by the Austrian Federal Science Ministry. A total of four adult male Long-Evans rats (Janvier, St-Isle, France) were used for the experiments. Rats were housed individually in standard rodent cages (56 × 40 × 26 cm) in a temperature and humidity controlled animal room. All rats were maintained on a 12 hr light/dark cycle and all testing performed during the light phase. Food and water were available *ad libitum* prior to the recording procedures and body weight at the time of surgery was 300-375 g.

METHODS DETAILS

Animals and Surgery

Animals were implanted with microdrives housing 32 (2x16) independently movable tetrodes targeting the dorsal CA1 region of the hippocampus bilaterally. Each tetrode was fabricated out of four 10 um tungsten wires (H-Formvar insulation with Butyral bond coat California Fine Wire Company, Grover Beach, CA) that were twisted and then heated to bind them into a single bundle. The tips of the tetrodes were then gold-plated to reduce the impedance to 200-400 kΩ. During surgery, the animal was under deep anesthesia using isoflurane (0.5%–3% MAC), oxygen (1-2l/min), and an initial injection of buprenorphine (0.1mg/kg). Two rectangular craniotomies were drilled at relative to bregma (centered at AP = −3.2; ML = ±1.6), the dura mater removed and the electrode bundles implanted into the superficial layers of the neocortex, after which both the exposed cortex and the electrode shanks were sealed with paraffin wax. Five to six anchoring screws were fixed on to the skull and two ground screws (M1.4) were positioned above the cerebellum. After removal of the *dura*, the tetrodes were initially implanted at a depth of 1-1.5 mm relative to the brain surface. Finally, the microdrive was anchored to the skull and screws with dental cement (R*efobacin Bone Cement R*, Biomet, IN, USA). Two hours before the end of the surgery the animal was given the analgesic Metacam (5mg/kg). After a one-week recovery period, tetrodes were gradually moved into the dorsal CA1 cell layer (*stratum pyramidale*). After completion of the experiments, the rats were deeply anesthetized and perfused through the heart with 0.9% saline solution followed by a 4% buffered formalin phosphate solution for the histological verification of the electrode tracks.

Data Acquisition and Behavior

The animals were housed individually in a separate room under a 12h light/12h dark cycle. Following the postoperative recovery period, rats were reduced to and maintained at 85% of their age-matched preoperative weight. Water was available *ad libitum*. Each animal was handled and familiarized with the recording room and with the general procedures of data acquisition. Four to five days before the start of recording, animals were familiarized at least 30 min with a circular open-field environment (diameter = 120 cm). On each recording day (1 or 2 recording days were performed per animal), the animal underwent a behavioral protocol in the following order: exploration of the familiar circular open-field environment (40 mins), sleep/rest in rest box (diameter = 26cm, 50 mins). Directly after this rest session the animals also explored a novel environment for an additional 40 min and rested

after for 50 mins. The novel environment recordings were performed in the same recording room but in an enclosure of a different geometric shape but similar size (e.g., a square environment of 100cm width). The wall of both the familiar and novel environment enclosures was 30cm in height, which limited the ability of the animal to access distal room cues. In addition, in two animals a 50 mins sleep/rest session was performed before the familiar exploration. During open-field exploration sessions, food pellets (MLab rodent tablet 12mg, TestDiet) were scattered on the floor to encourage foraging and therefore good coverage of the environment.

A headstage with 128 channels (4 × 32 channels, Axona Ltd, St. Albans, UK) was used to preamplify the extracellular electric signals from the tetrodes. Wide-band (0.4 Hz–5 kHz) recordings were taken and the amplified local field potential and multiple-unit activity were continuously digitized at 24 kHz using a 128-channel data acquisition system (Axona Ltd St. Albans, UK). A small array of three light-emitting diode clusters mounted on the preamplifier headstage was used to track the location of the animal via an overhead video camera. The animal's location was constantly monitored throughout the daily experiment. The data were analyzed offline. Only one recording day with the best unit yield was included in the subsequent spike clustering and data analysis procedures.

Spike Sorting

The spike detection and sorting procedures were performed as previously described (O'Neill et al., 2006). Action potentials were extracted by first computing power in the 800-9000 Hz range within a sliding window (12.8 ms). Action potentials with a power of > 5 SD from the baseline mean were selected and spike features were then extracted by using principal components analyses. The detected action potentials were segregated into putative multiple single units by using automatic clustering software (http://klustakwik. sourceforge.net/). These clusters were manually refined by a graphical cluster cutting program. Only units with clear refractory periods in their autocorrelation and well-defined cluster boundaries were used for further analysis. We further confirmed the quality of cluster separation by calculating the Mahalanobis distance between each pair of clusters (Harris et al., 2000). Periods of waking spatial exploration, immobility, and sleep were clustered together and the stability of the isolated clusters was examined by visual inspection of the extracted features of the clusters over time. Putative pyramidal cells and putative interneurons in the CA1 region were discriminated by their autocorrelations, firing rate, and waveforms, as previously described (Csicsvari et al., 1999a). In this way, we were able to identify the activity of 933 CA1 putative pyramidal units and 103 putative interneurons in 4 recording sessions performed in four different animals.

QUANTIFICATION AND STATISTICAL ANALYSIS

Rate Map Generation

As described in previous work (Dunn et al., 2015) we used a maximum entropy model inference paradigm to reconstruct the two-dimensional spatial distribution of each cell's firing probability. As a statistical model, we considered the maximum entropy model known as kinetic Ising model. We first separated running periods from periods of quiescence by applying a 5 cm/s speed filter. The activity of the cells was binned in 12.8 ms bins, and a binary variable $S_i(t)$ was assigned to each neuron for each bin. $S_i(t)$ has $+1/-1$ values, denoting the presence/absence of spikes emitted by neuron i within time bin t. Letting the state of each neuron at time t depend on the state of the population in the previous time step $t-1$, the maximum entropy distribution over the state $S_i(t)$ of neuron i at time t is

$$p(S_i(t)) = \frac{\exp[S_i(t)H(t-1)]}{2\cosh[H(t-1)]} \tag{1}$$

where H(t) would be identified as the time-varying covariate having the role of the external field in statistical physics. Equation 1 defines a GLM (Generalized Linear Model), where, in each time bin, mostly only one or zero spikes per bin are observed and the interaction kernel extends one time step in the past.

To find what values of H(t) are the most likely to generate the observed data given Equation 1, we maximized the log-likelihood function

$$L[S,H] = \sum_{ij}[S_i(t)H_i(t-1) - \log(2\cosh H_i(t-1))] \tag{2}$$

with respect to H(t).

The log-likelihood measures how well the model explains the statistics in the observed data. In our analysis, we have used the natural logarithm. Since the external field, H(t), can explain the variations in the firing rate as the rat navigates in space, it becomes important to model it appropriately.

Here we assumed that the spatial input arises as the sum of two-dimensional Gaussian basis functions centered on an evenly spaced M × M square lattice covering the recording environment. The spatial field of cell i at time t is then

$$H_i(t) = \sum_{jk}\alpha_{ijk}\exp\left[-\left((x(t) - x_{jk})^2 + (y(t) - y_{jk})^2\right)\Big/r^2\right] + h_i \tag{3}$$

where h_i is a unit-specific spatially and temporally constant baseline, and (x_{jk}, y_{jk}) and r are the vertices of the regular lattice and the widths of the basis functions, respectively. An accurate representation of the cell activity in space can be found by inferring the parameters α_{ijk} of the linear combination of the Gaussian basis function. We first optimized the values of M (the number of Gaussian basis functions in the lattice) and r (their widths). We maximized the likelihood over a range of values of M (from 12 to 20) and r (from 6 to 30 cm) and chose the values of the parameters that gave the highest Akaike-adjusted likelihood value. In all of the models, the parameters (α's and h_i), were found by maximizing the likelihood function given in (1.2) by gradient ascent. When comparing the models, we first Akaike-corrected the log-likelihood. The Akaike information criterion (AIC) is a measure to compensate for overfitting by models with more parameters, where the preferred model is that with the minimum AIC value, defined as

$$AIC = 2\ln\left(L\left[S, H, (M, r)_{ML}\right]\right) + 2k \tag{4}$$

where L is the likelihood at the maximum likelihood (ML) estimates of the parameters (α's and h_i) for a given value of M and r. k is the number of parameters (here r does not affect this number as it is a scale factor for Gaussian basis functions, while larger values of M result in more parameters α included in the model). The procedure was performed over all available sessions at once so that the resulting optimal parameters (M = 15 and r = 20 cm) were applied to all of them.

Firing rate maps can then be expressed for an arbitrary choice of spatial bin size as

$$H_i(t) = \sum_{jk} \tilde{\alpha}_{ijk} \exp\left[-\left((x(t) - x_{jk})^2 + (y(t) - y_{jk})^2\right)\Big/r^2\right] + \tilde{h}_i \tag{5}$$

where $\tilde{\alpha}$ and $\tilde{h}$ are the optimized parameters and (x,y) is any position in the environment. In the following, we consider a partition of the environment in bins of 4cm.

Place Cell Classification

To classify putative pyramidal cells as place cells, we applied a criterion based on sparsity. For each cell, we computed the quantity $a = \sum_i (r_i o_i)^2 / \sum_i r_i^2 o_i$ where r_i is the cell firing rate in bin i and o_i is the normalized amount of time spent by the animal in the same spatial bin. The sum runs over all spatial bins. To establish a significance threshold, we compared the sparsity score for each cell with the distribution obtained from shuffling the spike times. Namely, we produced 10000 surrogates by wrapping the spike times list of a random time comprised between [20 s END-20 s]. Cells whose sparsity score exceeded the 95th percentile of the surrogate distribution were label as place cells.

SWR Detection

SWR detection was performed in the following way. We first excluded rapid eye movement periods (REM). REM periods were detected based on the theta/delta ratio as described in Csicsvari et al. (1999b). To identify periods of theta activity, the theta/delta power ratio was measured in 1600 ms segments (800 ms steps in between measurement windows), using Thomson's multi-taper method.

In the remaining periods of sleep, we computed a wavelet transform (Morlet wavelet) of the local field potential. We used wavelet scales spanning a 5Hz – 500Hz range. We then calculated the power (root mean square) for each electrode in the 150Hz – 250Hz range, z-scored it and took the maximum across the set of electrodes identified as being in CA1. Ripple events were then classified by taking peaks in the score passing 5 SD and extending the event window until the score dropped below 2 SD.

The power spectrum for each identified SWR event was then obtained by taking the average power over electrodes for different frequency bands covering the 5Hz – 500Hz range.

Time-Binning of the Data

Our subdivision of temporal spike trains was performed according to two criteria. We either used (1) a traditional subdivision of time in windows of equal duration, or we used (2) windows of adaptive size to constraint the number of spikes (from any putative pyramidal cell in the population) falling in each of them to be constant. Results are shown for windows of size 15 spikes and 8ms respectively.

Reconstruction of Position (Bayesian Decoding)

We used Bayesian place prediction (Zhang et al., 1998) first to establish whether our population of CA1 units would provide sufficient spatial information to reconstruct position during exploration.

$$P(x|n) = P(n|x)P(x)/P(n). \tag{6}$$

P($\mathbf{x}$) represents the probability that the animal is at a given location. It was set to a uniform distribution not to bias our analysis by any place preference of the animal. P($\mathbf{n}|\mathbf{x}$) represents the conditional probability that a given spike count occurs at a location. This was estimated using the firing probability obtained from the place-rate maps, and adjusted to the assuming that the number of spikes follows a Poisson distribution. P($\mathbf{n}$), the normalizing constant, was used to ensure that P($\mathbf{x}|\mathbf{n}$) summed up to 1. The location with the maximum probability was selected as the reconstructed position. Error measurements represented the absolute distance between the middle of the reconstructed bin to the real position of the animal. We compared the quality of our reconstruction across

different window-sizes. The mean error obtained from decoding using fixed-time windows or fixed-spike windows was found to be comparable and to have a similar dependency on the size of the window used.

Analysis of Reactivation Events

SWR detection produced candidate events from which to detect reactivation. To reconstruct the content of the SWR-associated activity, we used a Bayesian reconstruction procedure equivalent to the one described above and the rate maps obtained from the environment exploration session. For each SWR-detected event, we generated a series of population vectors by subdividing the period into sub-windows. The sub-widows did not overlap with each other and covered the entire SWR duration. Depending on the window-type in use, windows had either a fixed number of spikes (n = 15) or the duration of each window was fixed (8ms). Events containing less than five windows were removed from further analysis.

Shuffling Procedures

If not otherwise stated, shuffling procedures are always applied to the sub-set of SWR reactivation events that were selected after the application of the quality criteria.

Spike Jittering

Within each reactivation event, the decoding procedure was performed after each spike's occurrence time was randomly and independently reassigned by drawing it from a uniform distribution within the event window. The new occurrence times were chosen with a resolution of 0.05ms and the samples drawing was performed with replacement.

Spike Identity Reassignment

Decoding of position is performed after each spike was randomly and independently reassigned to another CA1 putative pyramidal cell in the simultaneously recorded population while keeping the spike occurrence times constant.

Place Field Rotation

To apply a randomizing procedure equivalent to the one-dimensional place field rotation (Grosmark and Buzsáki, 2016), we proceeded in the following way. For each unit, we projected (Lambert azimuthal equal-area projection) the associated rate map to the bottom-half of a sphere. The top half of the sphere was completed with a copy of the bottom-half, symmetric with respect to the center of the sphere. We then applied a random rotation Φ to this spherical map by randomly selecting 3 Euler angles. We used a *zyz* convention so that $\Phi = \theta_z \times \theta_y \times \theta_z \in [0, 2\pi] \times [0, \pi] \times [0, 2\pi]$. After the rotation we transformed back the bottom half of the rotated spherical map by projecting it onto a plane, to finally obtain our new firing rate map for the environment. The procedure provided us with a new set of firing rate maps. In turn, these maps were used to perform position decoding.

Encoded Positions Shuffling

For each detected reactivation event we generated a new decoded trajectory by using the same set of encoded positions but rearranging the order of occurrence. The shuffling procedure is equivalent to generate random paths through a constant set of given locations in space. These paths visit all locations within the decoded trajectory and each location once.

Reactivation Speed Shuffling

This shuffling procedure works on the statistics of the average reactivation speed of different SWRs across the dataset. We first generated a reactivation speed pool by gathering the values of the distance between subsequent encoded locations across all reconstructed events in the session. We then reassigned to each event its original number of reactivation speeds by redrawing them (without replacement) from this pool and the new average reactivation speed was calculated for the shuffled event.

Simulated Brownian Motion

To directly compare the propagation dynamics of our reconstructed spatial trajectories with the expected results obtainable from a diffusive random walk of the same duration, we produced artificial sequences of points using the same number of events in our dataset and their respective lengths. For each event selected from our analysis we took its duration in time-steps d and its average step length v. Then starting from an arbitrary location $d-1$ points were generated sequentially by iteratively picking random steps of length L in the direction Θ, with L drawn from a normal distribution $N(v, v/4)$ and Θ drawn from a uniform distribution $[0 \ 2\pi]$.

Reactivated Trajectory Classification

To ensure that we were considering sleeping activity with space-related content, we applied an encoding quality criterion to select only those SWR reactivation events in which place encoding was reliable. First, we computed the average likelihood of the encoded position (the maximum of the posterior probability) over the prediction windows comprised in the SWR period. To do so, we normalized the Bayesian posterior probability so that its sum over all spatial bins at any time was equal to 1. Then, for every prediction window, we took the maximum value of the normalized probability. If the average maximum value over the windows of a SWR event exceeded a threshold of 0.1, the SWR event was admitted for subsequent analysis. Furthermore, within the admitted events, we considered only windows whose associated likelihood value individually passed the same threshold.

When analyzing the effects on the reactivation statistics of different quality thresholds, the same procedure was applied while varying the value of the threshold in the range between 0 and 0.12. After the thresholding, either the events passing the criterion or those failing it were selected for further analysis. Selection criteria were generally applied to original data only, with the exception of the data in Figure S5E where selection criteria were applied separately to original data and all types of shuffled surrogate data.

We then classified each SWR event that passed the encoding quality criterion by computing the average step size occurring between positions encoded from pairs of neighboring, non-overlapping windows. If the number of available pairs had fallen below three, we removed the SWR event from the analysis.

To compare original results with the shuffling ones obtained from spike jittering, identity shuffling, and place field rotation, we considered the shuffled versions of the original SWR quality-filtered events and computed the average step-size without applying an encoding quality threshold on the shuffled results.

As described above, decoding confidence for each SWR event was obtained as the average over time of the normalized likelihood. To compare the consistency of the trajectory traveling speed within each SWR reactivation event, we computed its coefficient of variation (CV). By calculating the mean μ_i and the standard deviation σ_i of all step-sizes within an event we could then obtain the ratio $CV_i = \sigma_i / \mu_i$ that defines the coefficient of variation associated with event i.

Novel Environment Analysis

In our experimental paradigm, the exploratory session of the novel environment was both directly preceded and directly followed by a quiet rest periods. The two sleep periods took place in the same rest box and had similar duration (see above). Therefore we analyzed reactivation events coming from these two periods separately.

Brownian Diffusion Analysis

We tested to what extent the trajectory reactivation dynamics could be approximated by a diffusion process described by a power law equation of the type:

$$\left\langle d(t)^2 \right\rangle^{1/2} = G t^\alpha \tag{7}$$

where d represents the distance between two encoded points and t is the interval between the two associated decoding windows. (Note that in the case of fixed-time windows t corresponds to an actual time interval, while in the case of fixed-spike windows this interval should be measured as the number of spikes intervening between the two windows.). G is a constant tuning the average speed of the trajectory propagation. To be able to compare the behavior of events sharing a similar average propagation speed, we separated the set of selected reactivation events by subdividing them into sub-groups of homogeneous average speed. Moreover, since we wished to avoid the boundary effects given by the finite dimensions of the environment used for the recording, we limited the speed range of our dataset to those reactivation events having a mean step size below 20cm/step. This threshold was set to ensure that reactivation trajectories could be approximately considered as events propagating freely without boundary limitations for a consistent number of steps. Thus, the speed groups were defined by chunking the events according to their average speed in evenly spaced intervals between 0 and 20cm/s. For each session, the number of speed groups was varied from a minimum of 3 to the maximum number of subdivisions that granted at least 20 events for each of them.

For each speed group, we then built a time-distance relationship. We considered all window pairs that occurred at a distance equal to an increasing multiple of the window size, therefore, only considering population vectors originating from non-overlapping groups of spikes. For each window-distance (i.e., time) group we generated all the squared distances traveled by the set of reconstructed trajectories in that segment. We then computed the mean $\widetilde{d_t^2}(t)$, estimating mean squared traveled distance by trajectories in each speed group (s) after a time interval t. From Equation (7) it follows that the logarithm of the average traveled distance has a linear dependency on the logarithm of the time interval elapsed, where G is the y axis intercept and α is the slope of the line. Therefore, to extract the parameters of the power law fit to our data, we perform a linear regression over the $(\log(\widetilde{d_i^2}(t)), \log(t))$ pairs of points.

The case of $\alpha = 1/2$ is of special interest since it corresponds to the mathematical description of the so-called Brownian diffusive process. We thus tested the consistency of the regression slopes obtained from our data with the Brownian diffusion hypothesis. First, we looked at the distribution of the values obtained from individual sessions and for the different speed groups (see above) (Figures 5A and 5D). For each session, this distribution did not significantly differ from a normal distribution centered in 0.5 (all p > 0.6 t test for fixed-spikes windows all p > 0.2 t test for fixed-time windows).

Then we considered each regression slope individually. For each slope, we compared the difference of the associated α value from 0.5 to the size of the fit 95% confidence interval (Figures 5B and 5E). A positive relationship between the two quantities was found (r = 0.64 p < 10e-4 for fixed-spikes windows, r = 0.59 p < 10e-3 for fixed-time windows).

To further test whether the data follow Brownian diffusion dynamics, we examined whether to $\widetilde{d_i}(t)$ follow a Rayleigh distribution, which is the theoretical limit of diffusion processes of this type in two dimensions. Therefore, for each speed group and time-interval, we fitted a Rayleigh distribution. The goodness of fit was then computed as a Chi-square distance by subdividing distance traveled into 12 bins of equal size. Chi-square scores were then averaged for each session with each possible time interval and speed group (Figures 5C and 5F) and compared to encoded positions shuffling results.

Behavioral Data

To compare the statistical properties of the reactivated ensembles with those expected from the animal's behavior, we applied the same power law analysis to the paths followed by the animal during exploration. To do so, we took trajectory segments continuously above a speed threshold (5 cm/s) and we sampled the animal's position at regular intervals with different temporal resolutions (time

intervals between 0.12 s and 1.1 s). For each temporal resolution, we then performed an equivalent analysis to the one described above, by treating trajectory segments as separate reactivation events (Figures 4 and 5). α values were found to center around 0.7 and to be significantly different from a normal distribution centered in 0.5 (t test p < 10e-4) (Figures 5G and 5H). Chi-square distance from the Rayleigh distribution best-fit was significantly higher than that observed from reactivation events (Figure 5I) (t test p < 10e-3).

We then used the same set of periods in which the animal was running to perform the power law analysis on the locations encoded by neural activity recorded during active behavior. We proceeded similarly to the case of sleep reactivation data, by again treating trajectory segments as separate reactivation events and using as spatial locations the positions obtained by the Bayesian decoding of the associated neural activity. The only difference with the sleep case was the size of the time windows used for binning the spike data. To compensate for the change in the temporal scale between sleep reactivation and running-associated activity we increased the size of both spike-fixed and time-fixed windows taking it to 100 spikes and 200ms respectively. Consistently with the results of the animal behavioral trajectories and as expected from a neural activity mostly encoding the actual animal position during navigation, the values of α were found to be again close to 0.7 (Figure 5G). This result suggests that the obtained values of the power law exponents are not a product of the analysis method we used but that they instead vary reflecting changes in the cognitive state.

Measures of Spatial Distribution Randomness

To quantify the degree of uniformity of spatial distributions, either those for encoded point density or behavioral occupation time, we used a measure of entropy. Spatial density distributions were normalized to one and the associated Entropy was calculated as $H_{Max} = -\sum_i P_i \log(P_i)$ where the sum runs over all the N_b bins of the map and P_i is the value of the distribution in the bin i. We also computed the theoretical maximum of the Entropy given the number of bins as $H_{Max} = -\log(1/N_b) = \log(N_b)$ obtained in the case of a perfectly uniform distribution. The entropy score reported in the figures is then defined as H_{Max}/H_{Max} and it runs from 0 to 1 (where 1 correspond to complete randomness).

We then measured the angular distribution of the direction taken during the first step of a set of reactivation events originating in different portions of the environment. To do so, we defined a grid of 5x5 macro square bins tiling the entire environment. We then divided our entire sample of selected reactivation events according to the macro-bin containing the position encoded by the spikes the first decoding window. The central set of 3x3 macro-bins were considered to be core bins as they were not in contact with the environment border and therefore presented no geometrical constrain to the directional evolution of the encoded trajectory. The remaining frame of 16 macro-bins was instead labeled as peripheral bins as they overlapped with the environmental border.

Measures of Reactivation Variability

To investigate the dynamics of neuronal activity within each SWR event and their relationship to network oscillations we computed two scores for each SWR event: encoded distance score, measured as the spatial distance between pairs of encoded locations, and population activity distance (PAD), calculated by the cosine distance between spike-count vectors (where any value > 0 was replaced by a 1) from decoding window pairs. Given A and B (where A and B are any two vectors of 1's and 0's). the cosine similarity is computed using a dot product $S_{AB} = \cos(\theta) = \sum_i A_i B_i / \sum_i A_i^2 \sum_i B_i^2$. The PAD is obtained by $d_{AB} = 1 - S_{AB}$, so that identical vectors have a score of 0, uncorrelated vectors a score of 1 and completely anticorrelated vectors a score of 2.

Both scores were separately computed for increasing interval between decoding windows (Δt from 1 to 4). For each interval size, we examined the relationship between encoded distance and PAD by computing the average PAD associated with average encoding distance in different ranges.

We further computed the overall correlation between the two measures over all pairs of windows, increasing Δt-s and calculated the correlation between the average PAD and the Δt interval size.

PAD Embedding Analysis

We analyzed the time-propagation properties of an ensemble based on the PAD, as opposed to the one obtained from encoded positions. The PAD was obtained from cosine distance and similarly to the Bayesian decoding allows us to compute the relative distances between pairs of time-windows. To be able to compare the propagation dynamics of this ensemble with that of the encoded locations, we run a non-linear dimensionality reduction algorithm to embed the PAD values in a 2-dimensional space with Euclidean metric. We transformed the original distance matrix of all PAD pairs using the "Sammon" mapping (Sammon, 1969), a variant of non-classical multi-dimensional scaling (MDS), but we also ran the same analysis with two different non-classical MDS criteria (stress and squared stress) and additionally with the ISOMAP algorithm (Tenenbaum et al., 2000), obtaining equivalent results. The embedded distance matrix was then used to compute the time-distance relationships shown in Figure S7A.

To further test the relationship between PAD and encoded position distance, we ran the dimensionality reduction procedure separately on the different reactivation speed-subgroups that were defined for the original analysis. No difference was found between the obtained time-distance curves from different groups (all p > 0.05 Wilcoxon test), confirming the observed independence between these two measures of neural variability.

We quantified the properties of the curves obtained from embedding by using either a power law fit (of the kind matching the encoded position data) or an exponential fit $f = a(1 - e^{-1(b/a)})$ where a corresponds to the saturation y-value of the curve and b is its initial slope, close to t = 0. Exponential fits were found to be largely better in describing PAD data as compared to linear fits, considering any goodness of fit measure.

Interneuron Activity

For each putative interneuron, we computed its firing rate expressed during every selected SWR event. From this set of vectors, we also calculated a population activity measure for each SWR event, by averaging over all putative interneuron rates for that event. We then correlated both the individual rate vectors and the cumulative one to PAD and speed scores associated with the SWR events.

Network Oscillations

The power component of each SWR was assessed in the range of frequency intervals within the 5-500Hz band using the Morlet wavelet analysis (see above). This analysis provided us with a reconstruction of the SWR signal in terms of a set of 60 wavelet components associated with different frequency ranges. For each frequency range, we computed the correlation between the oscillatory power provided by the corresponding wavelet component and (i) average encoded distance, (ii) average PAD, (iii) putative interneuron population activity (i.e., an average of activity of all putative interneurons).

DATA AND SOFTWARE AVAILABILITY

Data used in this study will be made available upon request by contacting the lead contact, Jozsef Csicsvari (jozsef.csicsvari@ist.ac.at).

Supplemental Information

Hippocampal Reactivation of Random

Trajectories Resembling Brownian Diffusion

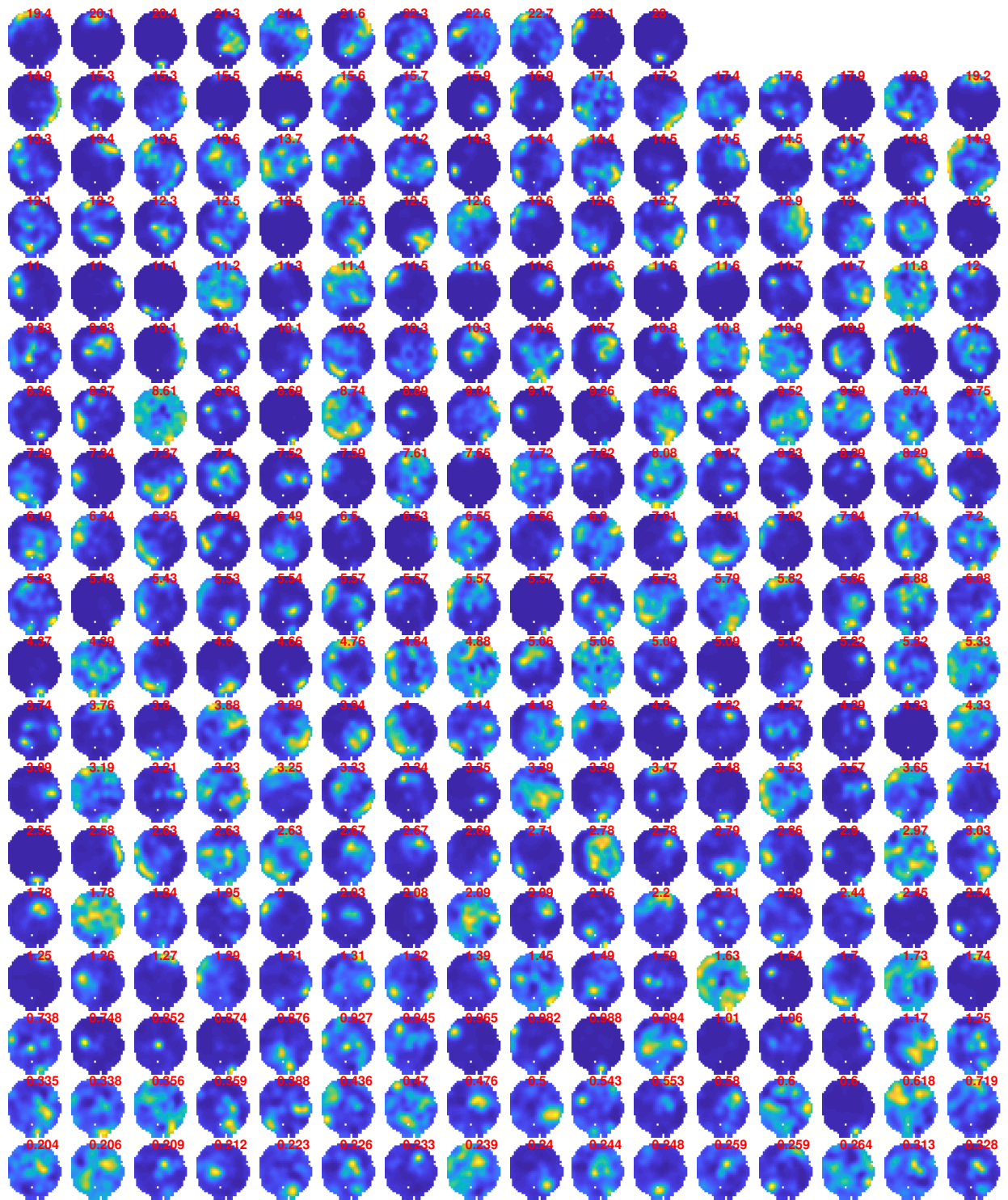Federico Stella, Peter Baracskay, Joseph O'Neill, and Jozsef Csicsvari

Figure S1. Set of simultaneously recorded CA1 place cells, Related to Figure 1

Reconstructed firing rate maps for a set of simultaneously recorded active place cells (n=299) from an example session. Cells were classified as place cells according to a spatial sparsity criterion (see Star Methods). Only cells with minimum average firing rate of 0.1Hz were displayed. Numbers represent the peak rate of the cells (cells were sorted by their peak firing rate).
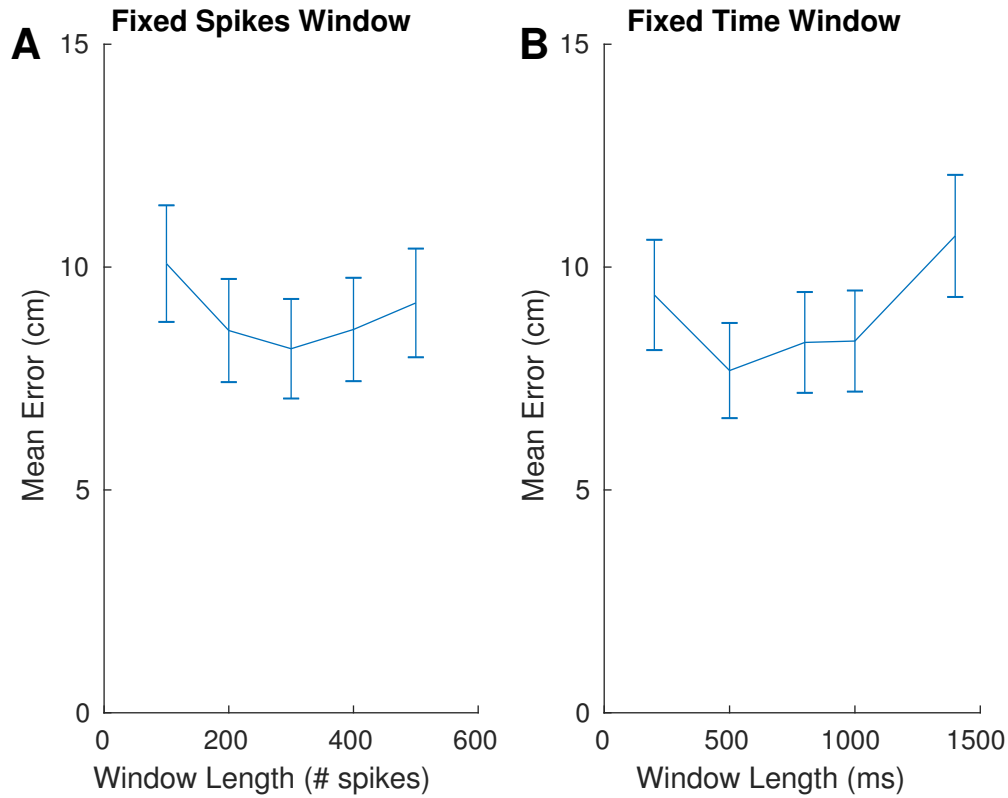
Figure S2. Validation of Bayesian decoding accuracy during exploration, Related to Figure 1

Using previously established firing rate maps, we used a Bayesian approach to reconstruct the position of the animal from instantaneous firing patterns during active exploration. Results are displayed from the representative session with rate maps displayed in (Supplementary Fig S1).
A) Mean reconstruction error for fixed-spike windows of different size.
B) Mean reconstruction error for fixed-time windows of different size.
Error bars: ±SEM.
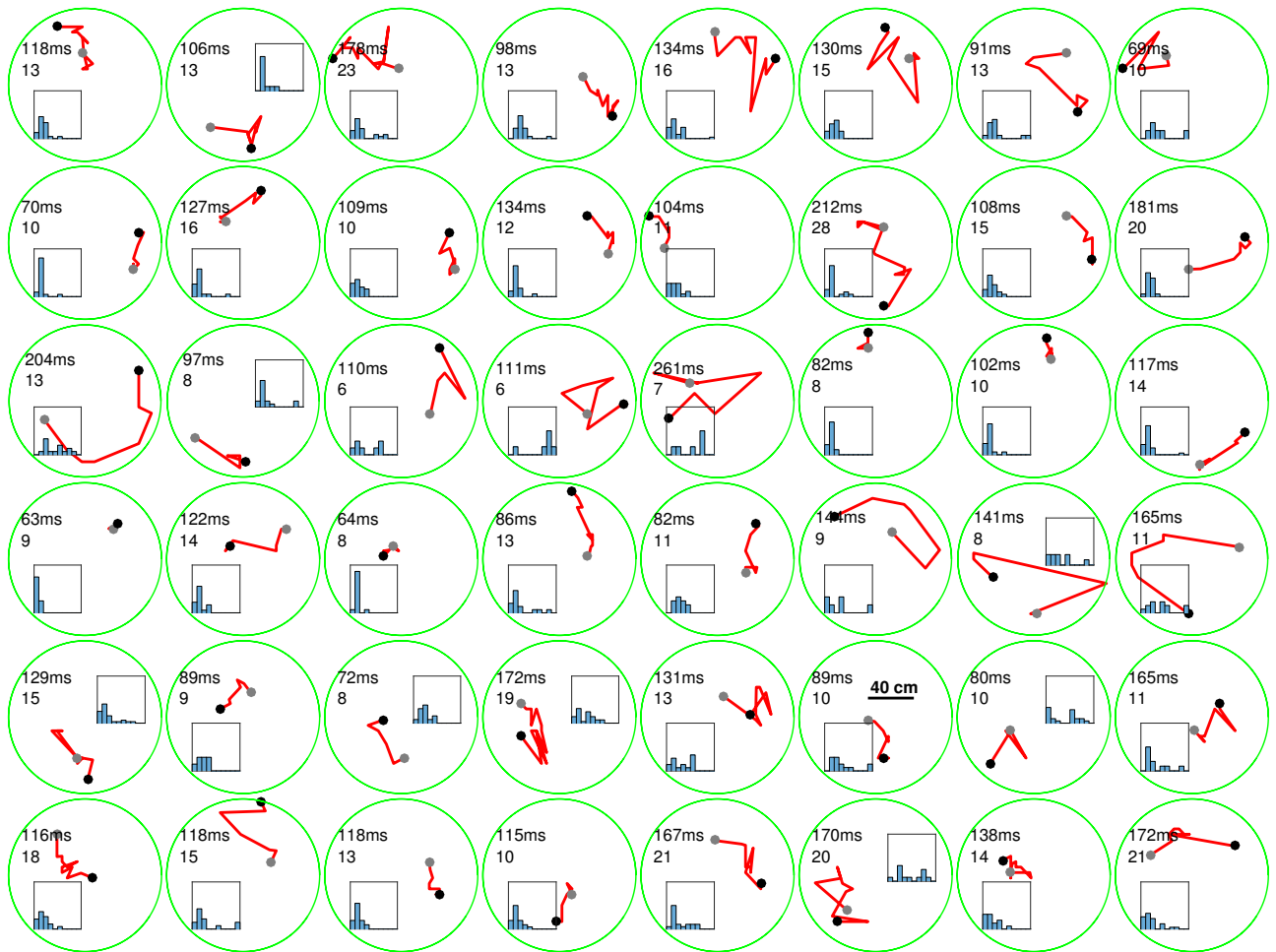
Figure S3. Additional examples of reconstructed trajectories, Related to Figure 1

Examples of trajectories reconstructed from probability maps of SWR. Each red line joins the most likely position obtained from the Bayesian decoding of non-overlapping fixed-spike (n=15) prediction windows. Grey dot: starting position; black dot: last position. For each reconstructed trajectory, we also report the duration of the replay events (top number), the number of trajectory point steps (middle number) and the distribution of step sizes (histogram). Histograms range from probability 0 to 1 on the y axis and from 0 cm/step to 40 cm/step on the x-axis with bin sizes of 4 cm/step.
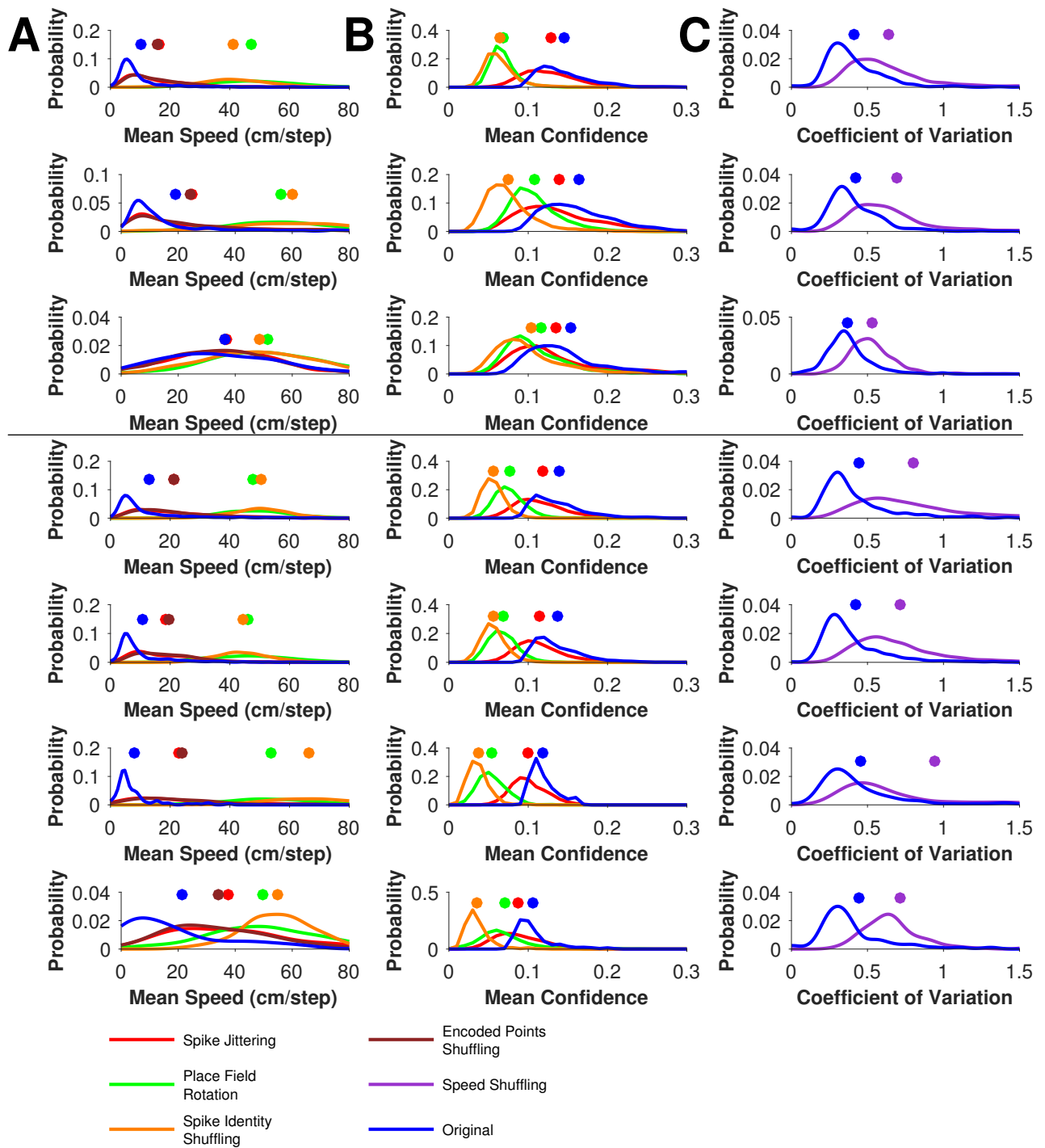
Figure S4. Comparison of shuffled and original trajectories for additional animals using the using fixed-spike decoding and for fixed-time decoding, Related to Figure 2

Each row corresponds to a separate session obtained from different animals. Rows from 1 to 3: further sessions from the remaining three animals using fixed-spike windows for the predictions. Rows from 4 to 7: Results using fixed-time window prediction displaying results from all four animals.
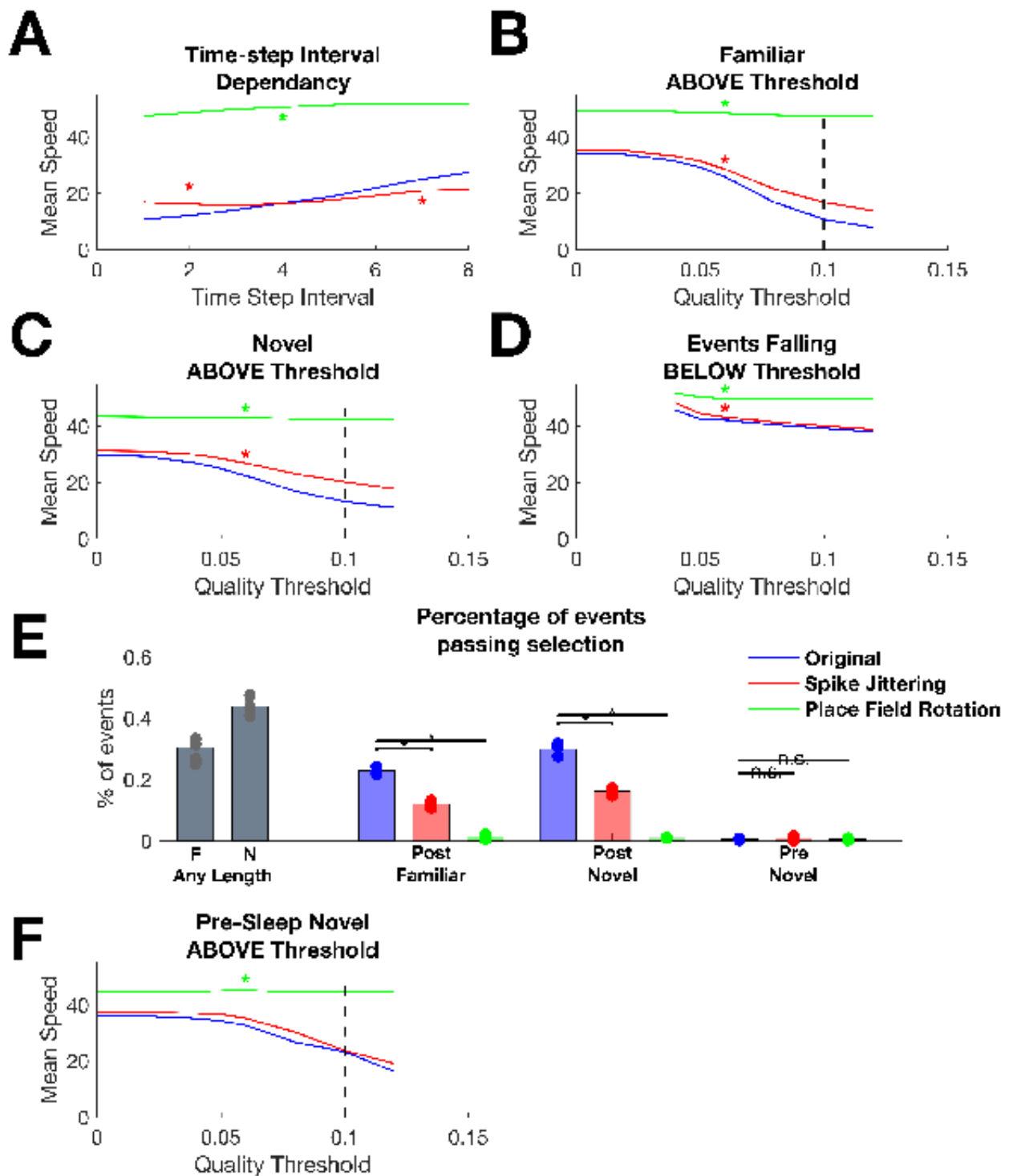
Figure S5. Distribution of reactivation event properties for different quality thresholds, Related to Figure 2

(A) Session-averaged mean of the average reactivation speed distributions when they were measured across different prediction step intervals. Spike jittered data (red) were significantly different from the original (blue) distribution in all sessions for ranges $\Delta t < 4$ $p < 10e\text{-}4$ and $\Delta t > 5$, $p < 10e\text{-}3$ KS, while the place field rotation data distributions (green) were significantly different from the original data's distribution at all tested time steps (all $p < 10e\text{-}5$ KS);
(B) Session-averaged mean of the average reactivation speed distribution for events falling above different quality thresholds. Events from rest sessions following familiar environment exploration are shown.

Decoding quality computed using the familiar environment map. Shuffled distributions were significantly different from the original data. Red: spike jittered data (all sessions and thresholds, p<10e-4 KS); Green: place field rotation data (all p<10e-6 KS); Dashed line: Quality threshold of 0.1, used in reactivation analysis.

(C) Same as (B) for events from rest sessions following novel environment exploration. Decoding quality computed using the novel environment map. Shuffled distributions were significantly different from original data: spike jittered data (all p<0.01 KS), place field rotation data (all p<10e-6 KS).

(D) Same as (B) for events falling below different quality thresholds. Events from rest sessions following familiar environment exploration. Decoding quality computed using the familiar environment map. Shuffled distributions were significantly different form that of the original data: spike jittered data (all p<0.04 KS); place field rotation data (all p<10e-5 KS).

(E) Proportion of events passing our selection criteria of 0.1 used in the reactivation analysis. Grey bars on the left represent the percentage of events passing quality threshold without restriction on minimal length of the number of trajectory point steps. Familiar (F) novel (N) post-rest data are shown. Blue: percentage of original events passing length and quality criteria for different sleep rest sessions; Red: same for spike jittered data; Green: same for place field rotation data. (Color code here and in the other panels as in Main Figure 2). (*: p<10e-3 and n.s. p>0.6, Binomial test).

(F) Same as (B) but SWR events from the rest session preceding the novel environment exploration were tested using the novel environment map. Therefore, the preplay of novel environment trajectories were tested here. Spike jittered data distributions were not significantly different from the original ones (all, p>0.1 KS) whereas place field rotation data distributions were significantly different (all p<10e-4 KS).
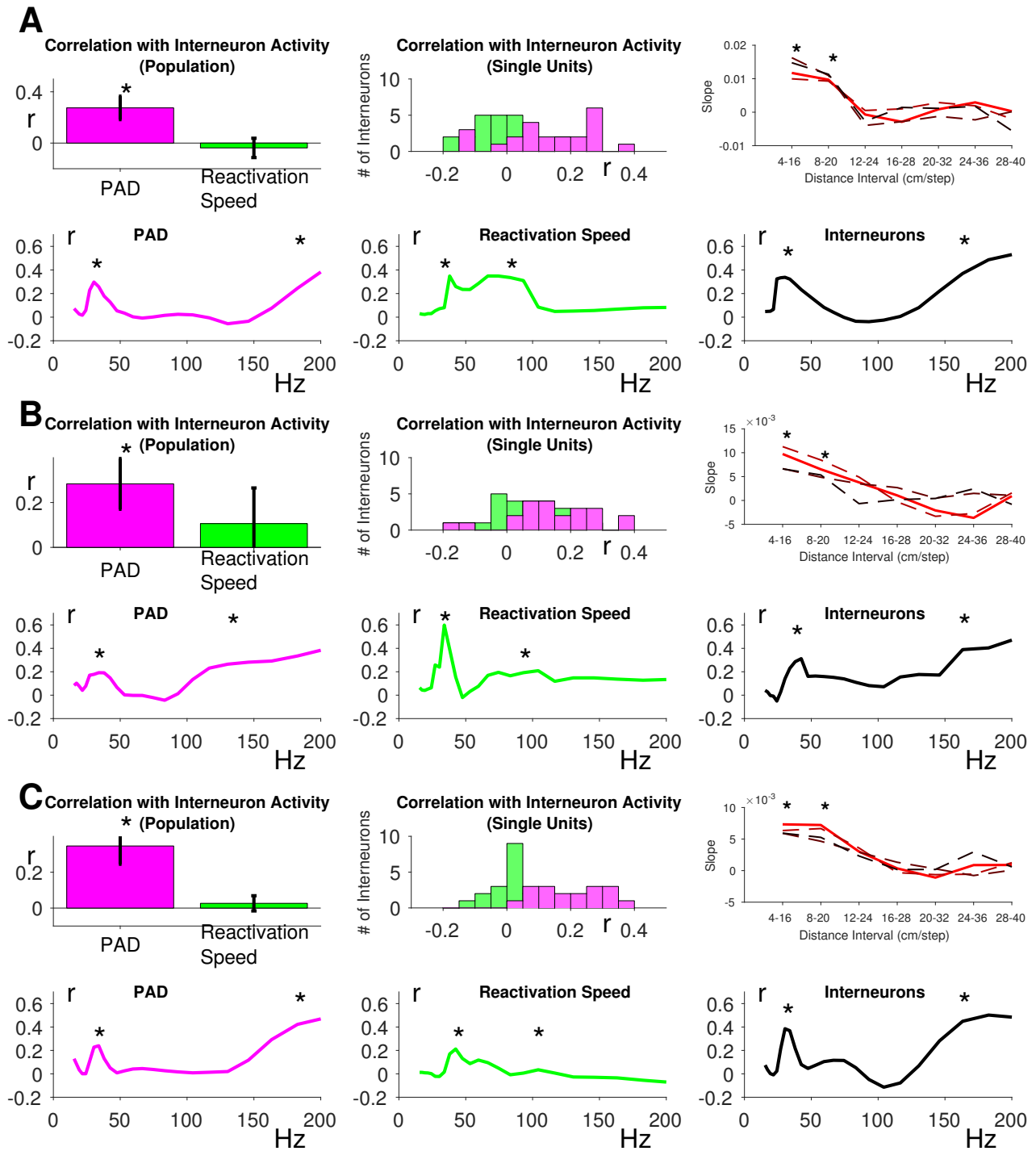
Figure S6. Network oscillations and putative interneuron activity associated with reactivation events, Related to Figure 7

Results of Figure 6 analysis for further recording sessions of different animals. Each session shown, labeled from (A) to (C), corresponds to a different animal. Results are reported similarly to Figure 6.

Top rows left: Correlation of summed putative interneuron rate and the average PAD (all $p<10e-7$) or reactivation speed (all $p>0.07$). For each SWR, putative interneuron rate and PAD and reactivation speed was measured in $\Delta t=1$.

Top Rows center: Same as above but correlations were calculated for single putative interneurons and distribution of correlations is shown.

Top Rows right: linear regression slopes of Decoded Distance vs. PAD data, obtained from subsets of data of encoded distance intervals. Different lines correspond to different time intervals. (*: $p<0.01$, rest: $p>0.25$)

Bottom rows: Correlation of average PAD (magenta) ($\approx$30Hz all p<0.005; 150-200Hz, all p<10e-12), average reactivation speed (green) ($\approx$30Hz all p<0.0004; $\approx$80Hz, all p<10e-5) and summed putative interneuron rate (black) ($\approx$30Hz all p<10e-8; 150-200Hz, all p<10e-13) with the different bands of the associated SWR power spectrum. Correlation was calculated across different SWR events.
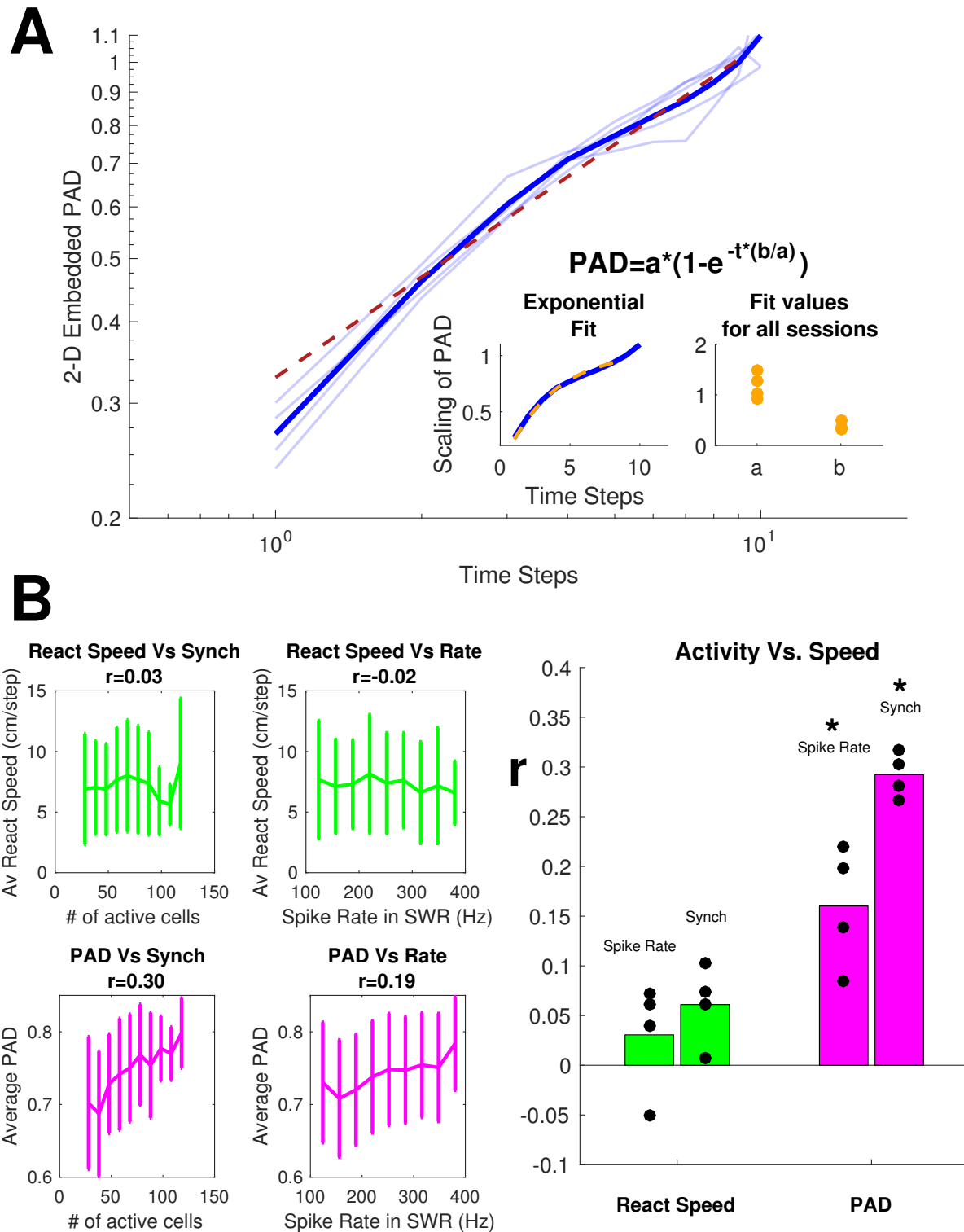
Figure S7. PAD Embedding and relation to activity sparsity, Related to Figures 3, 4 and 7

(A) Log-log plot of the time interval and the 2-D non-linear PAD embedding. Light blue: different mean reactivation speed groups (n=59, 287, 87, 33, 32 respectively for each); blue: mean of the speed groups. Maroon dashed curve: linear fit. (summed square of residuals=2.25, p=0.85 Pearson Chi-Square). Inset: exponential fit for the across-groups mean of the PAD embedding (left, orange dashed, summed square of residuals=0.24, p=0.99 Pearson Chi-Square formula in the figure) and parameters obtained from the exponential fit for all sessions (right).

(B) Relationship of mean reactivation speed (green) and mean PAD (magenta) with measures of activity sparsity: number of putative pyramidal cells active and their combined firing rate during SWR reactivation events. Left: results for one example session (Probability of correlation significance: top row p=0.48 and p=0.44; bottom row p=10e-11 and p=10e-5), all error bars ±SD. Right: summary of results for all sessions. Bars: average correlation coefficient; dots: single session correlation coefficients. (Reactivation Speed p>0.2 PAD p<10e-4)