



Flexible modulation of sequence generation in the entorhinal–hippocampal system

Daniel C. McNamee^{1,2,3}  , Kimberly L. Stachenfeld⁴ , Matthew M. Botvinick^{4,5}  and Samuel J. Gershman^{3,6,7} 

Exploration, consolidation and planning depend on the generation of sequential state representations. However, these algorithms require disparate forms of sampling dynamics for optimal performance. We theorize how the brain should adapt internally generated sequences for particular cognitive functions and propose a neural mechanism by which this may be accomplished within the entorhinal–hippocampal circuit. Specifically, we demonstrate that the systematic modulation along the medial entorhinal cortex dorsoventral axis of grid population input into the hippocampus facilitates a flexible generative process that can interpolate between qualitatively distinct regimes of sequential hippocampal reactivations. By relating the emergent hippocampal activity patterns drawn from our model to empirical data, we explain and reconcile a diversity of recently observed, but apparently unrelated, phenomena such as generative cycling, diffusive hippocampal reactivations and jumping trajectory events.

The entorhinal–hippocampal circuit (EHC) is thought to contribute to a diverse range of cognitive functions^{1,2}. An important motif within this circuit is the sequential nonlocal reactivation of hippocampal place codes^{3,4}. The archetypal instantiation of this functionality is ‘replay’, which refers to a temporally compressed representation of a previously experienced trajectory embedded within hippocampal sharp-wave ripples (SWRs)⁵. Initially observed during sleep, replay is thought to subservise long-term memory consolidation in neocortical networks⁶. More recently, sequential nonlocal hippocampal reactivations have been observed that do not fit with this classical definition of replay⁷. While rodents quietly rest, ensemble place cell activity appears to random walk through a cognitive map of a familiar environment instead of veridically replaying a rodent’s physical traversals⁸. In the awake immobile state, SWR-related trajectory events encode new goal-directed routes⁹. During active movement, an alternative form of sequential nonlocal reactivation may also occur. Theta sequences, which typically phase precess through local positions, may sweep ahead to remote locations along potential paths available to the rodent^{10–12}.

Spanning these diverse forms of sequential hippocampal representation, we consider a unified algorithmic theme conceptualizing the hippocampus as a sequence generator¹³, contributing sample trajectories drawn from cognitive maps to computations being executed downstream in the cortex. Establishing neocortical memory traces via synaptic plasticity may be characterized as a learning process extracting information from experiential replay during sleep¹⁴. Hypothetical environment trajectories, encoded in SWRs or theta sequences, may be thought of as samples of possible future behaviors that are input into a planning algorithm for the purposes of optimizing exploration and prospective decision-making^{15,16}. We suggest that this generalized perspective imposes a substantial computational obligation on the EHC as a generative sampling system since the performance of algorithms

can vary substantially depending on the statistical and dynamical structure of the input samples¹⁷. This motivates our computational hypothesis that hippocampal sequence generation is systematically modulated to optimize the resulting sampling regime for the current cognitive objective. Since previous computational EHC models have tended to focus on relatively specific applications such as localization^{18–20} or vector-based navigation²¹, the necessity to modulate sequence generation between cognitive algorithms is not addressed. Therefore, the broader computational viewpoint taken here raises unique theoretical questions, such as what alternative modes of hippocampal sequence generation are to be expected? By what neural mechanism can such distinctive dynamics be systematically regulated? And how can this be efficiently achieved for large relational spaces (for example, graphs) that may be nonspatial in nature²²?

We develop an algorithmic framework and associated neural mechanism by which distinct dynamical modes of sequence generation may be parsimoniously realized in a pathway between grid cells in the medial entorhinal cortex (MEC) and place cells in the CA1 subregion of the hippocampus². The critical technical innovation is the characterization of grid cells as encoding infinitesimal generators of hippocampal sequence generation. A generator is a mathematical object that specifies how a system evolves in continuous time^{23,24}. Conceptualizing hippocampal sequence generation as a dynamical system operating over a cognitive map, generators then determine the probabilities with which a given position will be reactivated at any time. We propose that MEC grid cells encode generators in a decomposed format. This enables a simple neural mechanism to flexibly interpolate hippocampal sequence generation between qualitatively and quantitatively distinctive regimes from random walks with Lévy jumps to generative cycling (Fig. 1a–e). In a phenomenological linear network model of hippocampal sequence generation, we demonstrate the systematic modulation of grid cells arrayed dorsoventrally as a function of spatial

¹Wellcome Centre for Human Neuroimaging, University College London, London, UK. ²Max Planck UCL Centre for Computational Psychiatry, London, UK. ³Department of Psychology, Harvard University, Cambridge, MA, USA. ⁴Google DeepMind, London, UK. ⁵Gatsby Computational Neuroscience Unit, University College London, London, UK. ⁶Center for Brain Science, Harvard University, Cambridge, MA, USA. ⁷Center for Brains, Minds and Machines, MIT, Cambridge, MA, USA. ✉e-mail: daniel.c.mcnamee@gmail.com

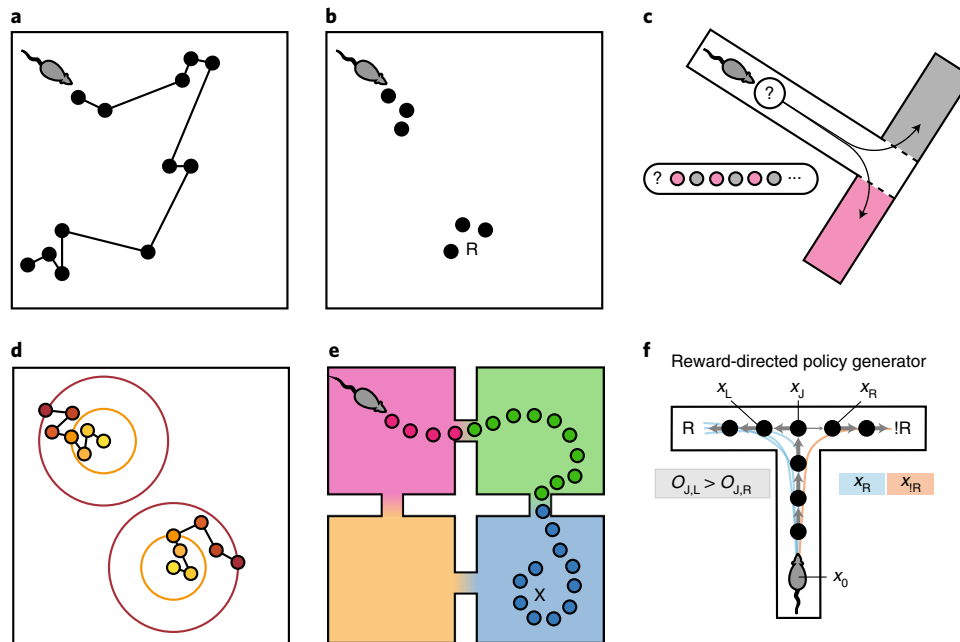


Fig. 1 | In different neurophysiological, behavioral and cognitive states, distinct modes of sequence generation are active in the entorhinal-hippocampal system. Drawing on the empirical and computational literature, we depict several modes of sequence generation supported by our model in a rodent experiment example. Each circle corresponds to a hippocampal place representation, and connecting lines indicate a place activation sequence. **a**, The exploration patterns of animals³⁵ and humans⁴⁸, as well as hippocampal place trajectories³⁶, interleave large ‘jumps’ between environment positions with local steps consistent with search efficiency optimization. This distinctive regime of sequence generation is referred to as superdiffusive. **b**, Motivationally salient locations are over-represented in hippocampal activity, as reflected in place activations at the reward location R ³¹. **c**, Generative cycling refers to the alternating representation of future possibilities in the hippocampal code that occurs during prospective decision-making¹². Here, the rodent evaluates the candidate behavioral trajectories of turning right (pink) or left (gray) at the junction. **d**, Diffusive random-walk trajectories are observed in hippocampal reactivations during rest⁸. Colors indicate the successive time steps at which each position is expected to be sampled along each trajectory (from yellow to red). Diffusion indicates that, on average, each sampled position will be spatially displaced from the previous one by a similar distance. **e**, In hierarchical reinforcement learning and human decision-making⁴⁹, environments may be processed across multiple spatiotemporal scales, for example, using room abstractions in this four-room world. **f**, In this T-maze task, where a rodent seeks reward R to the left (avoiding the right turn leading to no reward, $!R$), the behavioral policy of the rodent is encoded in the generator O , which biases sequence generation to the left x_L at the junction x_J . Sequence generation of hypothetical behavioral trajectories x_R is then biased toward the reward.

scale in the MEC layer and the consequential effects on place cell activity. In simulating this model, we show that our theory reconciles a diversity of empirical observations in sequential hippocampal reactivations.

Results

Spectral modulation in an entorhinal-hippocampal network model of sequence generation. How might sequences of positions in an environment or, more generally, states of an internal world model be simulated within the brain? This question can be posed formally within a generative probabilistic framework as how to sample state sequences $\mathbf{x} = (x_0, x_1, \dots, x_t)$ from a probability distribution $p(\mathbf{x})$ defined over state-space \mathcal{X} . A typical example considered in this study is the sequence distribution $p(\mathbf{x})$ based on the hypothetical decision-making policy of a rodent in an experimental task (Fig. 1f) as this will allow us to relate sequences generated by our model to the sequential nonlocal reactivations encoded in hippocampal place cells. We consider the state-space \mathcal{X} to be discrete as this allows us to study both continuous spatial domains (via discretization) and inherently discrete spaces (for example, graphs) in a common formalism based on matrix-vector products. When studying discretized continuous state-spaces, we interpolate discrete probability distributions where appropriate. Analogous techniques in purely continuous domains replace matrix-vector products with integral transforms²⁴. Therefore, our model may in principle be

applied across a wide range of cognitive maps, relational spaces, mental models or intuitive physics models^{22,25}.

An internal simulation is initialized based on a distribution $p(x_0)$ over states at an initial time $t=0$. We compactly denote this distribution over initial states as $\rho_0 = p(x_0)$. How can this initial distribution (for example, a rodent’s initial position in an experiment) be combined with dynamics information (the rodent’s behavioral policy) to compute the state distribution $\rho_t = p(x_t)$ at an arbitrary timepoint t in the future (where the rodent will be)? To answer this question, we need to understand how the state distribution ρ evolves in time. This is characterized by its time derivative $\dot{\rho} = \frac{d\rho}{dt}$. Assuming that the dynamics depend only on the current state and that they do not change over time, the evolution of a state distribution ρ is determined by a master equation

$$\tau \dot{\rho} = \rho O \quad (1)$$

where O is a matrix known as an infinitesimal generator^{23,26}. The generator O encodes stochastic transitions between states at short timescales (see Supplementary Math Note for details). For example, in a T maze, an entry in O could encode the local bias for a left turn at a critical junction leading to reward acquisition (Fig. 1f). The tempo parameter τ modulates the speed of the simulated evolution. The master equation (equation (1)) indicates that a dynamical system is economically encoded in an initial state distribution ρ_0 and a

generator O since the distribution ρ_t of possible states of the system at any time t can be retrieved from this information.

The master equation (1) has an analytic solution:

$$\rho_{\Delta t} = \rho_0 e^{\tau^{-1} \Delta t O}. \quad (2)$$

Given an initial state distribution ρ_0 , the propagator $e^{\tau^{-1} \Delta t O}$ is a time-dependent matrix that evolves ρ_0 through the time interval Δt to $\rho_{\Delta t}$ by propagating the initial state probability mass ρ_0 across the state-space under the dynamics specified by O . Fixing $\Delta t = 1$, the propagator $P_\tau = e^{\tau^{-1} O}$ can be applied iteratively to generate state distributions on successive time steps as $\rho_{t+1} = \rho_t e^{\tau^{-1} O}$. State sequences characterizing the simulated evolution of the system can therefore be generated by recursively applying this time-step propagator P_τ and sampling

$$x_{t+1} \approx \mathbf{1}_{x_t} P_\tau \quad (3)$$

where $\mathbf{1}_x$ is a one-hot vector indicating that state x is active with a probability of one. This will result in state sequences \mathbf{x} that accurately reflect the generative distribution of sequences $p(\mathbf{x})$ defined by the generator O and initialization ρ_0 (Supplementary Math Note).

Directly computing the propagator $P_\tau = \sum_{n=0}^{\infty} (n!)^{-1} (\tau^{-1} O)^n$ is challenging since it requires an infinite sum of matrix powers. In the Supplementary Math Note, we show that P_τ can be computed efficiently using a generator eigendecomposition $O = G \Lambda W$ as

$$P_\tau = G e^{\tau^{-1} \Lambda} W. \quad (4)$$

Furthermore, the facility to freely modulate the tempo τ of propagation is a highly desirable property. This would enable coarse hierarchical sequence generation to be run expeditiously rather than at the true rate of evolution of the external world or, if time allows, an internal simulation could be slowed down for a more fine-grained analysis^{27–29}. We show that, given this representation (equation (4)), the tempo τ can be efficiently manipulated via a computational mechanism that we refer to as spectral modulation. Since Λ is the diagonal matrix of O eigenvalues, its exponentiation is trivially accomplished by exponentiating the eigenvalues separately along the diagonal $\left[e^{\tau^{-1} \Lambda} \right]_{kk} = e^{\tau^{-1} \lambda_k}$. Multiplication by the G matrix projects a state distribution ρ_t on to the generator eigenvectors $\phi_k = [G]_{\cdot k}$, which we refer to as the spectral components of the propagator. Note that we use this term as a broad reference for generator eigenvectors or components of alternative generator decompositions, potentially constrained by other considerations such as nonnegativity, which facilitate spectral modulation (Supplementary Math Note). In this spectral representation, time shifts simply correspond to rescaling according to the power spectrum

$$s_\tau(\lambda) = e^{\tau^{-1} \lambda}. \quad (5)$$

Each spectral component ϕ_k is scaled by $s_\tau(\lambda_k)$ based on its eigenvalue λ_k . Finally, W projects the spectral representation of the future state distribution ρ_{t+1} back onto the state-space \mathcal{X} . This spectral format factorizes time and position within an environment such that a linear readout can generate a propagator for any timescale.

We apply the spectral propagator (equation (4)) recursively to generate state sequences according to

$$x_{t+1} \approx \mathbf{1}_{x_t} G S W \quad (6)$$

where $S = e^{\tau^{-1} \Lambda}$ is the power spectrum matrix. This generator-based model of sequence generation can be minimally realized in a linear feedback network model of EHC (Fig. 2a) in which the

activity profiles of the network units are qualitatively consistent with those of grid cells in the MEC and place cells in the hippocampus. Specifically, we equate the spectral components ϕ_k (columns of G) with grid cells topographically organized by spatial scale along the dorsoventral axis of MEC (Fig. 2b)³⁰. The linear readout W from the MEC layer to the hippocampal layer embeds the the future state distribution ρ_{t+1} in a predictive place code²⁸. See the Supplementary Math Note for further considerations regarding biological plausibility and connections to other EHC models.

The critical mechanism of spectral modulation, which sets the tempo of propagation, is the systematic regulation of MEC grid cell output as a function of spatial scale according to the power spectrum s . At the implementation level, we hypothesize that spectral modulation may be accomplished via gain control or grid rescaling according to top-down cortical input. An example of the latter would be if a small-scale grid firing map is enlarged to a medium scale to increase MEC power at medium scales and reduce MEC power at small scales. Notably, grid modules appear to have the capacity to rescale independently³⁰, do so as a function of experience and presumed cognitive processing³¹ and cause a consistent rescaling of place fields in the hippocampus when the grid scale is perturbed³². More generally, several empirical results support the critical contribution of entorhinal input toward the coherent temporal organization of hippocampal activity^{33,34}. Indeed, beyond tempo control set by τ , we study several other parametric and nonparametric classes of power spectra with highly distinctive causal effects on hippocampal sequence generation. Each of these EHC settings will be motivated as an optimized operational mode for a particular cognitive process.

Foraging in an open environment. Consider a rodent foraging in a large open environment. Without cues indicating where food may be located, its exploration process must search each location. How might it generate the next environment position to inspect? Approaches to this problem range in terms of computational complexity from serial visitations minimizing repetitions (imposing increasingly burdensome memory and planning costs) to random sampling (requiring no memory). Although generators could be adaptively designed to implement sophisticated forms of uncertainty-driven exploration (for example, based on Gaussian processes), here we focus on maximizing the efficiency of random sampling in the low-complexity limit. Assuming a random-walk generator O^{rw} (suggesting the rodent has no information as to where food may be located), we study the effect of spectrally modulating tempo in generating the next position to visit. If the rodent repeatedly samples target states at large spatial scales ($\tau \rightarrow 0$) then it will repeatedly traverse the environment expending too much energy. In contrast, a small-scale search pattern ($\tau \gg 0$) will lead to the rodent oversampling within a limited area and taking too long to fully explore the environment. Defining exploration efficiency as the fraction of the environment visited per cumulative distance traversed, neither tempo regime delivers a satisfactory return.

This ubiquitous conundrum has been extensively studied in the foraging literature, leading to the Lévy flight foraging hypothesis³⁵. Theoretical analysis and simulations have shown that exploration efficiency is maximized by interleaving jumps (that is, sampling successive positions separated by a large distance) with local search patterns. Effectively, this strikes a balance between local searches ($\tau \gg 0$) and global reorientations to new positions in an environment ($\tau \rightarrow 0$). It is proposed that these distinctive search dynamics are naturally selected for in animals across a wide variety of ecological niches due to their universally advantageous properties and draws its name from the Lévy distribution, which characterizes the distribution of possible next positions in a Euclidean space³⁵. This is a heavy-tailed distribution and therefore in addition to a high probability of sampling a nearby position, it has a small probability of generating a large jump to a more distal region.

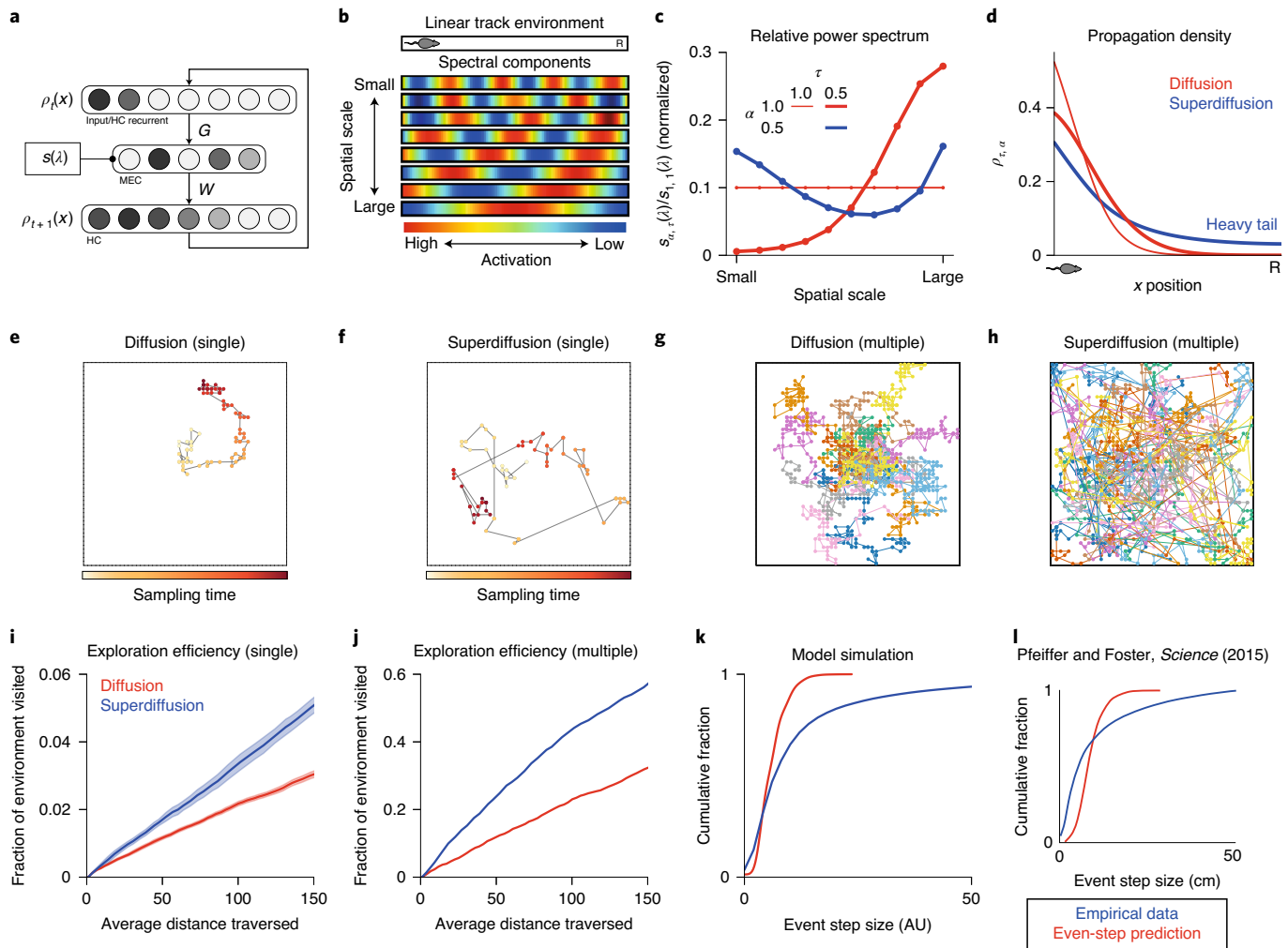


Fig. 2 | Spectral modulation of grid cell activity alters the statistical structure of hippocampal sequence generation. **a**, Abstract EHC model. ρ_t and ρ_{t+1} are the propagation distributions over environment position at time steps t and $t+1$, respectively. The weight matrices G and W are identified with the generator decomposition $O = G\Lambda W$. $s(\lambda)$ is the power spectrum expressed as a function of the generator eigenvalue λ associated with each unit (that is, spectral component) in the MEC layer. HC, hippocampus. **b**, Firing maps of units in the MEC layer that exhibit periodic tuning across a range of spatial scales reminiscent of grid cells. Spectral components are equivalently ordered by their eigenvector numbers k and generator eigenvalue magnitudes $|\lambda_k|$ from large to small scale. **c**, Characteristic changes in the power spectrum associated with the diffusive and superdiffusive regimes are highlighted by the corresponding power ratios relative to a baseline. τ -modulated diffusion ($\alpha=1$, $\tau=0.5$, thick red line) results from a downweighting (or upweighting) of small-scale (or large-scale) spectral components while α -modulated superdiffusion ($\alpha=0.5$, $\tau=1$, thick blue line) upweights both small- and large-scale components but suppresses them at medium scales. **d**, Diffusive ($\alpha=1$, red lines) versus superdiffusive ($\alpha=0.5$, blue line) propagation densities $\rho_{\tau,\alpha} = 1_{x_0} G S_{\tau,\alpha} W$ in a linear track environment where x_0 corresponds to the location of the rodent. By modulating the tempo $\tau=0.5$, diffusive propagation can sample over larger distances (thick red line) compared to the baseline diffusion ($\tau=1$). However, only superdiffusive propagation facilitates an extraordinary jump to the reward (R). Note that these continuous propagation densities have been interpolated from the discrete propagation vectors output by our model. Further technical details are presented in Extended Data Fig. 5. **e, f**, Single diffusive (**e**) and superdiffusive (**f**) sequences in an open-box environment. The color code from yellow to red reflects the sampling iteration (from the initial sample to the final). Sequences are initialized at the center of the environment. **g, h**, Multiple diffusive (**g**) and superdiffusive (**h**) sequences, each distinguished by color. **i**, Mean and standard error of exploration efficiency for individual trajectory simulations ($n=20$). **j**, Parallelized exploration efficiency across all the trajectories in **g** and **h**. **k**, Cumulative histogram of step sizes for diffusive trajectories (red) and superdiffusive trajectories (blue). An event refers to the successive representation of two different locations. Step size is measured by Euclidean distance. Therefore, the x axis reflects the Euclidean distance between successively sampled locations during sequence generation. **l**, Statistical analyses of hippocampal recordings in Pfeiffer and Foster, *Science* (2015)³⁶ indicated that hippocampal trajectory events were superdiffusive since the distribution of step sizes between state activations (blue) was heavy tailed. This qualitatively matched the simulated cumulative histogram for superdiffusions in **k** in contrast to the prediction (red) based on simulated sequences composed of states separated by equal step sizes. In our model, the latter corresponds to diffusive sequence generation with τ varied to match the velocities exhibited by the recorded trajectory events.

We show that analogous heavy-tailed propagation distributions can be accessed in our model via an alternative form of spectral modulation, thus providing a mechanistic account regarding how such sampling may be accomplished in the EHC. Based on mathematical considerations (Supplementary Math Note), we introduce

the stability parameter α , which determines the entorhinal power spectrum $s_{\tau,\alpha}$ according to

$$s_{\tau,\alpha}(\lambda) = e^{-\tau^{-1}|\lambda|^\alpha}. \quad (7)$$

For $\alpha=1$, the power spectrum is unchanged, $s_{c,1}=s_c$ (compare equation (5), Fig. 2c). This results in random walks that correspond to diffusions in continuous spaces; therefore, we refer to the $\alpha=1$ regime as diffusive (Supplementary Math Note). Setting $\alpha < 1$, the linear readout from the MEC layer in our circuit model reflects a propagator with probability mass smoothly redistributed to remote positions from the nearby positions (Fig. 2d). Sequence generation with $\alpha < 1$ is referred to as superdiffusive (Methods). Although both superdiffusive and diffusive sequence generation are ultimately truncated by the limited extent of an environment, they are differentially sensitive to the possible range of spatiotemporal scales. Stability modulation results in a scale-invariant sampling process since sampling can occur at any spatiotemporal scale simultaneously on each iteration. In contrast, τ specifies a limited range of spatiotemporal scales and therefore tempo modulation can never lead to scale invariance. This suggests that α and τ modulation have fundamentally distinct effects on the statistical structure of sequence generation (Extended Data Fig. 1).

We compared diffusive ($\alpha=1$, Fig. 2e) and superdiffusive ($\alpha=0.5$, Fig. 2f) sampling in an open-box environment typically used in rodent foraging experiments^{8,36} postulating that EHC generates candidate exploratory trajectories that the rodent subsequently pursues physically. Whereas diffusive behavioral trajectories failed to fully explore all areas in the simulated window of time (Fig. 2g), superdiffusive trajectories visited positions in an approximately homogeneous distribution across the entire arena (Fig. 2h), thus highlighting the flexibility with which superdiffusions adapt to an arbitrary environment scale. With respect to the standard measure of exploration efficiency, superdiffusions explored more positions in the open environment as a function of distance traversed than diffusions (Fig. 2i,j). This is also the case for structured state-spaces, such as compartmentalized environments with obstacles (Extended Data Fig. 2). Consistent with our theoretical arguments and simulations, superdiffusive sequential activation of hippocampal place cells (Fig. 2k,l) and superdiffusive rodent behavioral trajectories have been observed^{8,36} in foraging experiments where rodents are required to explore environments with essentially random distributions of food locations.

Goal-directed trajectory events with heterogeneous jumps. We relate our model operating in the superdiffusive regime to heterogeneous sequences of hippocampal representations exhibiting jump transitions to motivationally salient locations^{9,11}. In the first experiment we simulated, SWR-related place cell activity in the CA1 subregion of the dorsal hippocampus was recorded from rats engaged in alternating random foraging and spatial memory tasks⁹. In the spatial memory task, a single reward location in the open square arena was repeatedly baited and thus the rat could remember this predictable home location and engage in goal-directed navigation to acquire the reward. The application of neural decoding analyses revealed the rapid sequential encoding of positions across the environment while the rats were task engaged but immobile (Fig. 3a–d). This study provided evidence that the hippocampus encodes new goal-directed paths to memorized locations that were over-represented in the generated trajectories ('away events', Fig. 3b,d). We simulated the generation of trajectory events in this experiment with our model to demonstrate the computational mechanism by which a memorized location may be stored in a generator representation and remotely activated exclusively in the superdiffusive regime. We subtly manipulated a random-walk generator O leveraging a distinguishing feature of the generator-propagator formalism, namely the ability to independently modify the spatial and temporal statistics of sequence generation in a spatially localized manner (Supplementary Math Note). Specifically, we controlled the remote activation of the rewarded home location by scaling the generator transition rates at home states according to $\nu^{-1}O_h$, where h indexes home states $x_h \in \mathcal{X}$ and ν is a

scalar specifying their motivational value (Methods). Initializing the rodent's position away from the home location and activating superdiffusive sequence generation in our EHC model (Fig. 3e–h) results in trajectory events reflecting random walks with biased jumps to the rewarded location (Fig. 3f,h). Using the same generator in the diffusive regime (Fig. 3i–l) does not over-represent the home location (Fig. 3j,l). Comparing propagators between sequence generation regimes explains this remote activation as the localized increase in sampling probability exclusively at the home location in the superdiffusive regime (Fig. 3m–p).

In another study of multi-goal foraging in a circular track environment¹¹, it was observed that theta sequences exhibited a strong dependency on the currently targeted goal. As the rodent initialized their behavioral trajectory, theta sequences exhibited a nonlocal, nondiffusive pattern of activation. Place cells near the goal destination were frequently active along with place cells near the rat's actual position within individual cycles. We constructed a generator O , which encoded a goal-directed policy of clockwise movement around this circular track, with turn-offs to the three goal locations located at the intervals G1, G2 and G3. Assuming that the rodent is located at the start position, we generated sequences in the diffusive ($\alpha=1$, Fig. 4a) and superdiffusive ($\alpha=0.5$, Fig. 4b,c) regimes. We observed that the latter had a strong tendency to generate jumps to goal locations interleaved with local roll-outs as observed empirically (Fig. 4d,e). This effect emerges from the fact that remote goal locations are over-represented in superdiffusive propagation (Fig. 4f). Taking the distance around the track to the furthest encoded location as the look-ahead distance, the distributions of look aheads scaled with the distance to the target goal as recorded in theta sequences (Fig. 4g). After placing the rodent at locations along the circular track before each of the goal turn-offs, and without changing any parameters, the sequences exhibited short-range look aheads over similar distances (Fig. 4h).

These effects emerge from the specific combination of superdiffusive sequence generation and a goal-sensitive generator. In the linear track (Fig. 2b–d) and open-box (Fig. 2e–h) simulations where random-walk generators were used, it was shown that the superdiffusive regime engenders spatially extended jumps between states. Due to the metric correspondence between space and time (set by a specific velocity), this can be equivalently stated as superdiffusions generating large jumps through time¹³: that is, superdiffusive sequence generation interleaves state transitions over short timescales (resulting in small spatial steps) with state transitions over long timescales (resulting in large jumps). Since a goal-directed policy results in goal locations being over-represented in the stationary distribution of internal simulations, the probability of sampling such states in a superdiffusion is relatively high. Therefore, as observed in Fig. 4f, superdiffusive propagation specifically over-weights the goal locations (regardless of the current position of the rodent, Extended Data Fig. 3).

Generative cycling emerges from minimally autocorrelated sampling. A critical element of many planning algorithms is prospectively evaluating possible future trajectories. To accomplish this, model-based simulations are often used such as in Monte Carlo Tree Search³⁷. These methods rely on sampling sequences of states in a state-space, retrieving rewards associated to those states and computing Monte Carlo estimates of choice-relevant objectives. For example, an agent can produce an estimate \hat{r} of the average reward expected to be accrued based on sampled states (x_1, \dots, x_N) and a reward function $r(x)$ as

$$\mathbb{E}[r] \approx \hat{r} = \frac{1}{N} \sum_{i=1}^N r(x_i) \quad (8)$$

and then input such estimates into an action selection algorithm. Sequentially sampling states from a propagator, retrieving the

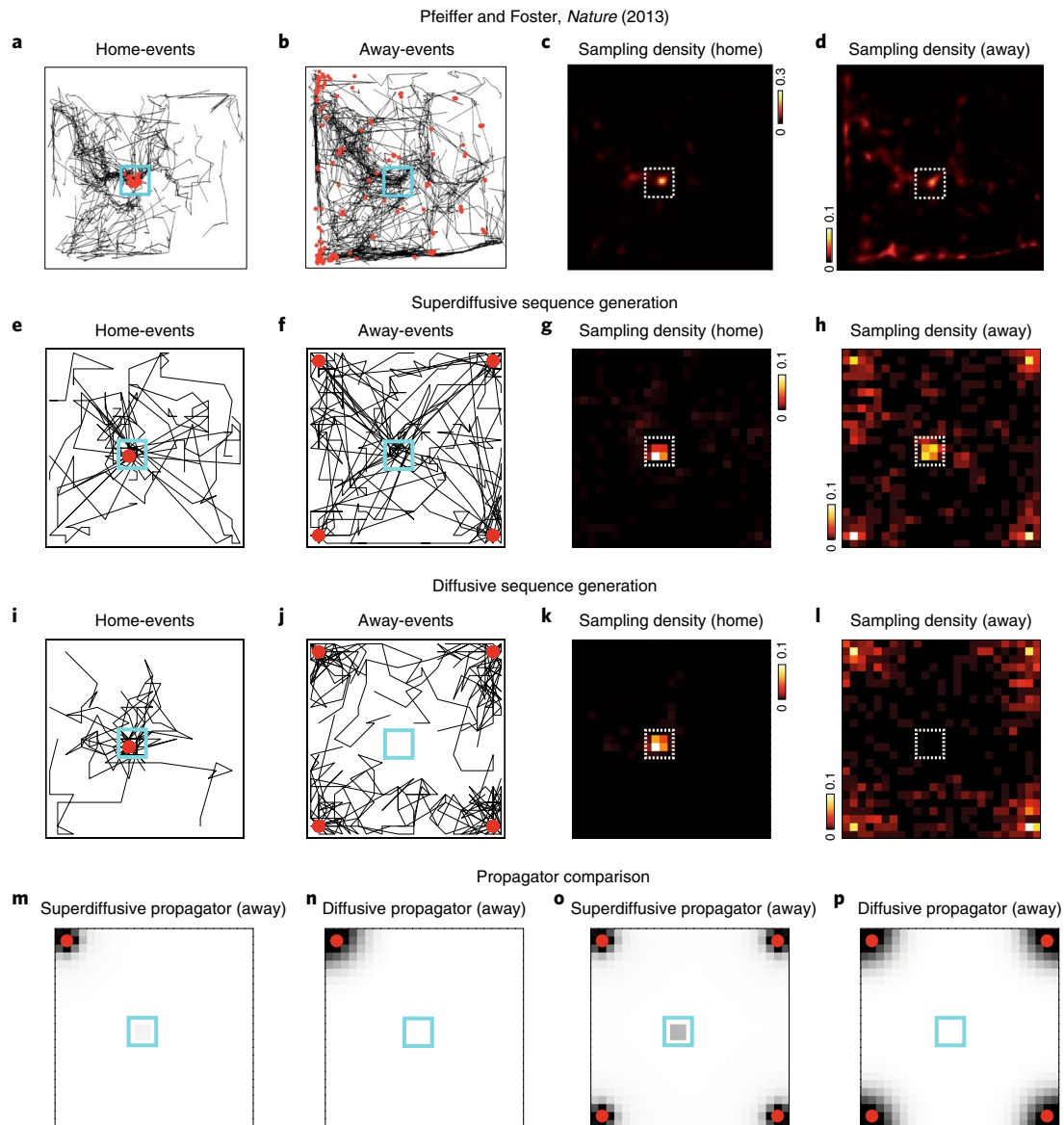


Fig. 3 | Simulating the over-representation of remote, motivationally salient, locations. **a–d**, (first row) Empirical data from Pfeiffer and Foster, *Nature* (2013)⁹ (reproduced from ref. ⁹, Nature Publishing Group). **e–h**, (second row) Superdiffusion simulations. **i–l**, (third row) Diffusion simulations. **m–p**, (fourth row) Propagators for away events in superdiffusive and diffusive regimes. These are plotted for single location initializations (**m,n**) and for multiple location initializations (**o,p**). **a,e,i**, (first column) Empirical and simulated trajectory events (black lines) while the rodent was located (red circles) at the home location (blue square). **b,f,j**, (second column) Empirical and simulated trajectory events while the rodent was located away from home. **c,g,k**, (third column) Estimated sampling density for home events (home location indicated by dotted white square). **d,h,l**, (fourth column) Estimated sampling density for away events. This set of panels contains the key comparison. Note that only superdiffusive away events (**h**) but not diffusive away events (**l**) remotely show over-representation in the home location consistent with the data (**d**). In our generator model, this is explained by the remote propagation probabilities that are exclusively observed in the superdiffusive regime (**m,o**). Diffusive propagators with sufficiently low tempos (thereby sampling over large spatial scales) can jump from away locations to the home location. The critical distinction is that superdiffusive propagation does not require delocalization to a large spatial scale to jump to the home location. Superdiffusions can uncover motivationally salient locations stored locally within the generator regardless of spatial distance.

associated rewards and estimating the expected average reward via equation (8) forms a Markov chain Monte Carlo (MCMC) algorithm in the service of planning (Supplementary Math Note).

The quality of an MCMC estimator (equation (8)) is quantified by its sample variance $\mathbb{V}_{\text{dyn}}[\hat{r}]$, which reports how variable the estimate will be across different sample sequences¹⁷. A major source of sample variance in MCMC, which also afflicts the generator model in both the diffusive and superdiffusive regimes, is generative autocorrelations. Technically, the sample variance is proportional to the

integrated autocorrelation time Δt_{ac} (equation 68, Supplementary Math Note):

$$\Delta t_{\text{ac}} = \sum_{t=0}^{\infty} C_X(t) \quad (9)$$

where $C_X(t)$ is the autocorrelation function of the state variable X at step t . Intuitively, Δt_{ac} is the average number of iterations that a

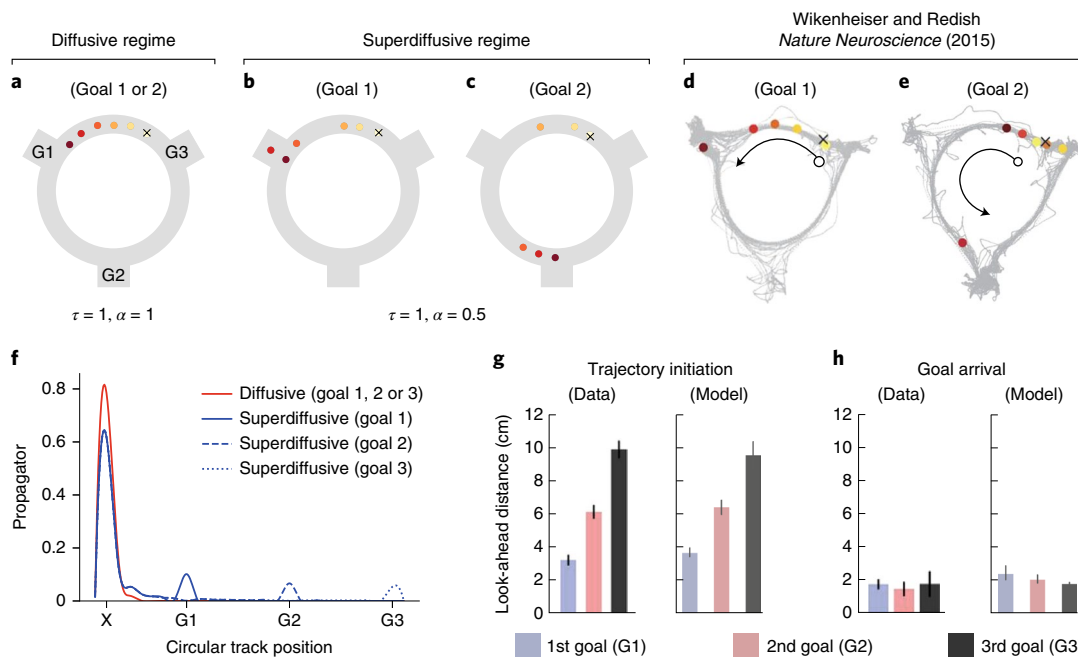


Fig. 4 | Theta sequences generated during a goal-directed foraging experiment exhibited nonlocal transitions in place coding. Within individual theta cycles, sequences of place representations in the hippocampus emanated from the current position of the rodent (X) and proceeded anticlockwise around the track with characteristic jumps to one of three goal locations where food was available (labeled G1, G2, G3) (from Wikenheiser and Redish, *Nature Neuroscience* (2015)¹¹). Each circle corresponds to a decoded location, and the color indicates the temporal order within the sequence (from yellow to red). **a**, Diffusive sequence generation ($\alpha=1$) typically proceeds with localized sequential place activations regardless of the target goal location. **b, c**, Shifting to the superdiffusive regime ($\alpha=0.5$), sequence generation activated local place representations as well as remote goal locations (goal 1 in **b** and goal 2 in **c**) but not intervening locations. **d, e**, Empirically observed patterns of place activation in theta sequences. Superdiffusive goal-directed sequence generation matches the key qualitative features with goal-directed sequences representing local positions near the animal, as well as jumps to goal locations that were over-represented within theta sequences. Panels **d** and **e** reproduced from ref. ¹¹, Nature Publishing Group. **f**, Diffusive propagator and the superdiffusive propagators from the initial X location are plotted. The jumps to goals are explained by the unique bumps in the superdiffusive propagators at the remote goal locations. **g**, The same stability parameter ($\alpha=0.5$) generated nonlocal, goal-jumping trajectories (model) with varying look-ahead distances, as observed in hippocampal theta sequences (data, ref. ¹¹). Bar heights equal the mean across theta sequences (data) or simulated sequences (model) and error bars reflect the standard error ($n=50$ simulated trajectories). **h**, Near the goal locations, look-ahead distances were the same across goals, as superdiffusive sequence generation is attracted to goals and thus automatically alters the look-ahead distance.

sampling algorithm requires to generate a single independent sample. Although samples may be generated on each iteration $t=1,2,3$, we can expect new independent samples to be generated on iterations $t = \Delta t_{ac}, 2\Delta t_{ac}, 3\Delta t_{ac}$. Standard practice in MCMC estimation, such as with the Metropolis–Hastings algorithm, is to simply discard autocorrelated samples (see Supplementary Math Note for further details). This brings into focus a sharp trade-off between sampling time and estimation accuracy, which necessarily burdens any cognitive function dependent on Monte Carlo estimation¹⁵.

We show that a fundamentally different solution to the autocorrelation problem of MCMC estimation is available for generator-based sampling. The need to dispose of correlated samples may be obviated by directly optimizing the power spectrum to minimize autocorrelations in the emitted sequences. The resulting sequences are then composed of approximately independent samples, thus facilitating rapid online simulation, estimation and responsive action loops. Technically, we show that the integrated autocorrelation time Δt_{ac} of the state variable X can be expressed analytically in terms of the power spectrum $s(k)$ (where k indexes the spectral components), and the constraints necessary to ensure that the resulting propagator is valid are linear (Supplementary Math Note). Therefore, the minimally autocorrelated power spectrum $s_{mac}(k)$ can be identified using standard optimization routines. We hypothesized that minimally autocorrelated sequence generation should result in a sampling process that methodically shifts between the most salient points of divergence under

dynamical evolution of the system. As a counterpoint, consider the genesis of autocorrelations in diffusive sequences (Fig. 2e,g). In diffusions, nearby states tend to be closely associated within the sampling dynamics due to overlapping propagation distributions. Therefore, there tend to be relatively likely paths back to previously visited states leading to a large integrated autocorrelation time Δt_{ac} . A minimally autocorrelated sequence generation can avoid this pitfall by restructuring its propagation dynamics to successively sample states that do not admit likely paths between them.

We studied this computational hypothesis in the context of a spatial alternation task where a rodent was required to make a binary decision at a junction leading to a reward or not¹². Alternating representations of hypothetical future trajectories were identified at several levels of neural organization in the dorsal hippocampus while the rats approached the critical junction (but not after the junction turn). In particular, place cells encoding the left and right arms fired in an alternating fashion within the theta band. Assuming that the rodent would initiate planning as it approached the junction where it is required to make a decision, we assumed that hippocampal sequence generation would shift to a regime of minimally autocorrelated sampling (Fig. 5a). The power spectrum s_{mac} , which minimized the integrated autocorrelation time (equation (9)) bore a dissimilar profile to the parametrically modulated power spectra $s_{a,\tau}$ (Fig. 5b). The most notable distinction was the emergence of counterweighted spectral components across spatial scales. The heaviest negative weighting applied to the spectral component encoding a

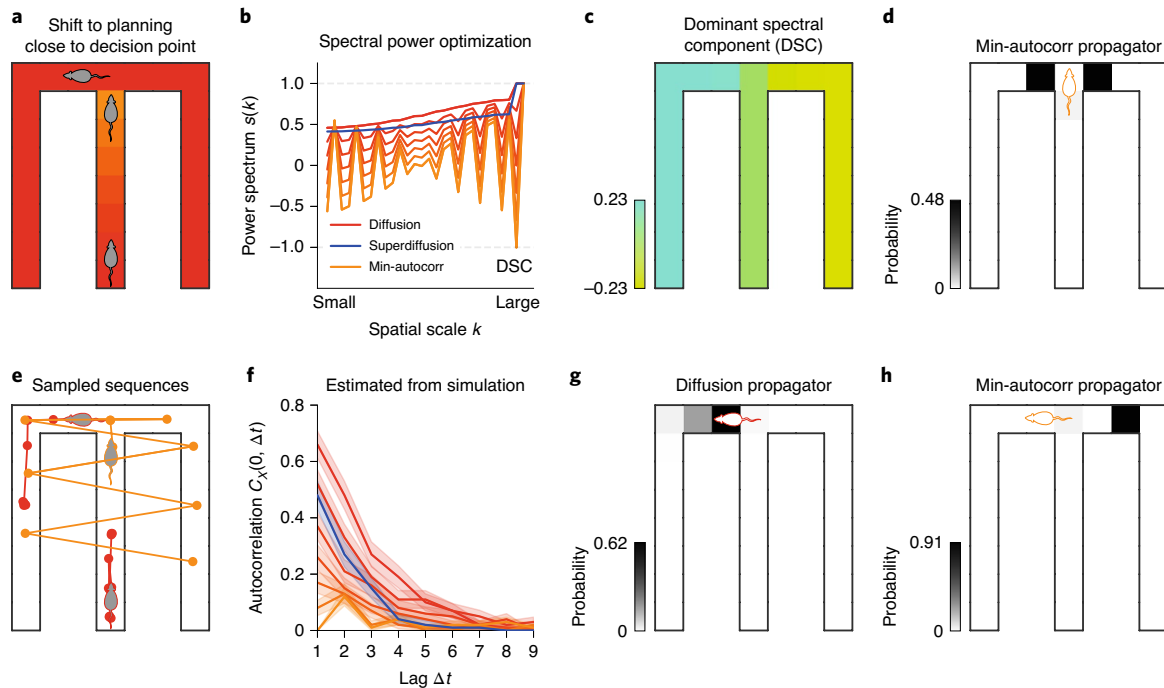


Fig. 5 | Generative cycling emerges from minimally autocorrelated sequence generation. **a**, Spatial alternation task environment from Kay et al.¹². We assumed that the well-trained rodent had a decision-making policy composed of moving to the junction, making a binary decision and then running to the food port at the end of an arm. On approaching the junction where the critical choice is made, we modeled hippocampal sequence generation as shifting from a localized diffusive regime (red) to minimally autocorrelating (min-autocorr) sampling (orange). **b**, The minimal-autocorrelation (orange), diffusive (red) and superdiffusive (blue) power spectra are plotted. Intermediate colors between red and orange reflect intermediate spectra between purely diffusive and minimally autocorrelating. Under autocorrelation minimization, we observed that for each positively weighted spectral component, s_{mac} negatively weighted another spectral component of a similar scale. In the EHC, this corresponds to the up- and down-regulation of MEC input into the hippocampus in alternation across grid scales. In the generator model, negative spectral modulation drives a generative repulsion depending on the structure of the associated spectral component. The DSC is that which undergoes the largest spectral modulation from diffusion to minimal autocorrelation. **c**, The DSC reflects a hierarchical decomposition of the state-space into the three arms of the maze. **d**, At the junction, the minimal-autocorrelation propagator splits its propagation density evenly between the two arms in a similar fashion to diffusive propagation. **e**, While diffusive sequences sweep ahead of the animal in the maze (initialized at the start position and just after a left turn), minimal-autocorrelation sequences cycle back and forth between the two arms. **f**, Estimated autocorrelation $C_x(0, \Delta t)$ as a function of the time interval Δt . Via spectral optimization, the sample autocorrelations were smoothly reduced at all time intervals from pure diffusion to the minimal-autocorrelation regime as predicted. Shaded error bands represent s.e.m. ($n = 100$). **g**, Just after the turn, the diffusion propagator sweeps ahead of the sampled rodent position. **h**, Just after the turn, the minimal-autocorrelation propagator switches the sequence generation process to the other arm.

high-level hierarchical decomposition of the environment that we refer to as the dominant spectral component (DSC) (Fig. 5c).

Despite the fact that the minimally autocorrelating propagator (Fig. 5d) samples states in both arms at the junction similar to a diffusive propagator, minimally autocorrelated sequence generation subsequently deviated radically from diffusion. Sequences generated under minimally autocorrelated spectral modulation (orange, Fig. 5e) were strongly reminiscent of the reported generative cycling phenomenon in that successive state samples were repeatedly drawn from the opposing arm in the maze¹². These stood in stark contrast to diffusive sequences generated both before and after the junction (red, Fig. 5e). Diffusively propagating into one of the arms means that it is relatively likely that a sequence will remain in that arm for a long time thus increasing the generative autocorrelation. This would indicate that the rodent's internal simulation does not have sufficient diversity (in particular, it has not sampled the other arm) to make an informed decision. As predicted, estimated autocorrelations (Fig. 5f) were lower in generatively cycling sequences (orange) compared to the diffusive (red) or superdiffusive (blue) regimes. In contrast to diffusive propagation (Fig. 5g), minimally autocorrelated sampling leverages the hierarchical structure of the environment to generate sequences that efficiently sample across dynamically

divergent components of the environment based on repulsive propagation (Fig. 5h).

Diffusive hippocampal reactivations for structure consolidation.

In contrast to the superdiffusive sequences observed during random and goal-directed foraging, offline hippocampal reactivations exhibit a diffusive operational mode during rest³. In an experiment where rats foraged for randomly dropped food pellets, spatial trajectories were decoded from SWRs during a postexploration rest period ('sleep SWRs') as well as during immobile pauses in active exploration ('wake SWRs'). The SWR sequence generation regime is statistically identified by estimating the mean displacement function (MD) $MD(t) = \langle \|x_t - x_0\| \rangle$ from the decoded position sequences $\mathbf{x} = (x_0, \dots, x_t)$ (Methods). On a log-log plot, the MDs of diffusions and superdiffusions are linearly related to time with different slopes α^{-1} determined by stability²⁴. By studying the slope of estimated MDs, it can therefore be concluded that while wake SWRs were superdiffusive ($\alpha < 1$, consistent with rodent movements), sleep SWRs encoded random walks ($\alpha = 1$). Furthermore, sleep SWRs were recorded over a range of velocities as parametrized by tempo τ in our model (with τ decreasing as velocity increases). Notably, SWR trajectory velocity was uniquely related to fast gamma

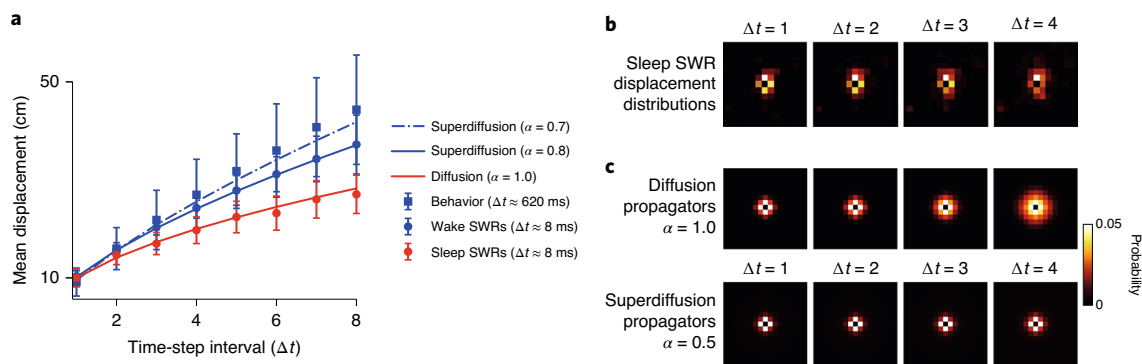


Fig. 6 | Spatial trajectories encoded in SWRs exhibit distinct stability modulation between wake and sleep phases. **a**, Blue and red dots correspond to the recorded MDs of spatial trajectories decoded from SWRs in the wake phase and sleep phase, respectively (Stella et al.⁸). There was a difference between the MD slopes (Methods), indicating that wake SWRs were generated in the superdiffusive regime ($\alpha = 0.8$, blue line is the predicted MD curve) while sleep SWRs were diffusive ($\alpha = 1$, red line is the predicted MD curve). Each of these distinct sampling regimes can be generated by subtly varying the power spectrum s_{ω} applied to a common underlying environment representation. Note that the physical trajectories of the rodents ('behavior', blue square markers) were superdiffusive ($\alpha = 0.7$), consistent with the idea that SWRs may supply candidate exploratory trajectories. This graph is log-log plotted in Extended Data Fig. 6 where the linearity of MD as a function of time can be observed. Error bars show \pm s.e.m. **b**, Propagators estimated from reactivated hippocampal trajectories⁸. These resemble those predicted in the diffusive regime, but not the superdiffusive regime, that lack the increasingly wide-spreading activation surrounding the initial position in the center (**c**). **c**, Propagators as a function of time interval Δt in the diffusive and superdiffusive regimes.

power suggesting that it may vary as a function of MEC input^{8,38}. We reconcile these two SWR modes within our model from two perspectives. Mechanistically, we show that spectral modulation can interpolate between the distinct statistical regimes of SWRs associated with sleep versus wake by fitting the empirical MD measurements as a function of stability α (Fig. 6a) and that the empirical sleep SWRs step distributions (Fig. 6b) are well-approximated by diffusive propagators (Fig. 6c).

Furthermore, in support of our computational hypothesis that the statistical dynamics of sequential hippocampal reactivations are flexibly altered between exploration or planning in the wake state and memory consolidation during rest, we studied performance metrics for exploration, learning and Monte Carlo estimation algorithms based on the sequence generation of input state sequences (Fig. 7a–d). We simulated a learning process whereby an environment representation, formalized as the successor representation (SR), is acquired through error-driven learning based on state sequences^{28,39}. The SR is a predictive state representation whereby the representation of each state encodes the rate at which other states will be visited in the future (Supplementary Math Note). While superdiffusions provide better exploration efficiency (Fig. 7b), and minimally autocorrelated sampling exhibits the best sampling coverage (Fig. 7d), both are conspicuously inferior in terms of the SR consolidation accuracy (Fig. 7c). Diffusive sequence generation resulted in the best learned approximation to the random-walk SR (Fig. 7e–l). Note that this result can be understood theoretically as diffusions embody fundamental spatial biases in their statistical structure (Methods). Although we focused on consolidating the random-walk SR as reflecting a homogeneous predictive map in the absence of salient states such as rewards or landmarks, diffusive sequence generation is also relatively optimized for consolidating directed predictive maps (Extended Data Fig. 4).

Dysregulated entorhinal input degrades spatiotemporal consistency of hippocampal activity. To flexibly shift between different regimes of sequence generation, our model requires that the spectral modulation of MEC activity be coherently balanced across grid modules. Dysregulated entorhinal activity may imbalance spectral modulation, thereby disrupting the spatiotemporal structure of

hippocampal representations. In particular, spectral modulation with stability $\alpha > 1$ defines a pathological regime of sequence generation that we refer to as turbulence. Despite seemingly minor differences in the power spectra from diffusion $\alpha = 1$ to turbulence $\alpha = 2$ (Fig. 8a), the sequences generated differ greatly. Turbulent sequences are highly irregular and fail to reflect the structure of the underlying space (Fig. 8b). In simulation, increasingly turbulent sampling propagates approximately uniformly across states independent of the initial position and propagators may even fail to preserve the state probability density. We suggest that the erroneous modulation of grid cell activity may therefore contribute to psychopathologies in cognitive processes dependent on coherent sampling from internal representations.

In particular, a core positive symptom of schizophrenia is conceptual disorganization. Identifying conceptual representations with nodes in an internal semantic network, this formal thought disorder is indicative of a progressive degradation of the relational structure between nodes during sequence generation consistent with that observed in the generator model as EHC dynamics is shifted deeper into the turbulent regime. The resultant disorder in the hippocampal layer of our model is reminiscent of that observed in mouse models of schizophrenia⁴⁰. In one study, electrophysiological recordings of neural activity in the CA1 hippocampal subregion of forebrain-specific calcineurin knockout mice were acquired as the mice freely explored⁴⁰. Such knockout mice had previously been shown to exhibit several behavioral abnormalities reminiscent of those diagnosed in schizophrenia such as impairments in working memory and latent inhibition. The knockout of plasticity-mediating calcineurin leads to a shift toward potentiation and we sought to highlight a possible mechanism by which the resulting over-excitability may disrupt hippocampal SWRs through the lens of our theory. In particular, we model the effect of calcineurin knockout as an imbalance in the spectral regulation of entorhinal input into the hippocampus. Notably, the relationship between the temporal displacement of spikes between distinct cells given the spatial displacement of their place fields was abolished under calcineurin knockout⁴⁰. This is the key quality of turbulent sequence generation. We made a direct quantitative comparison based on a spiking cross-correlation analysis⁴¹. In the diffusive regime, the cross-correlations between place cells as a function of the distance

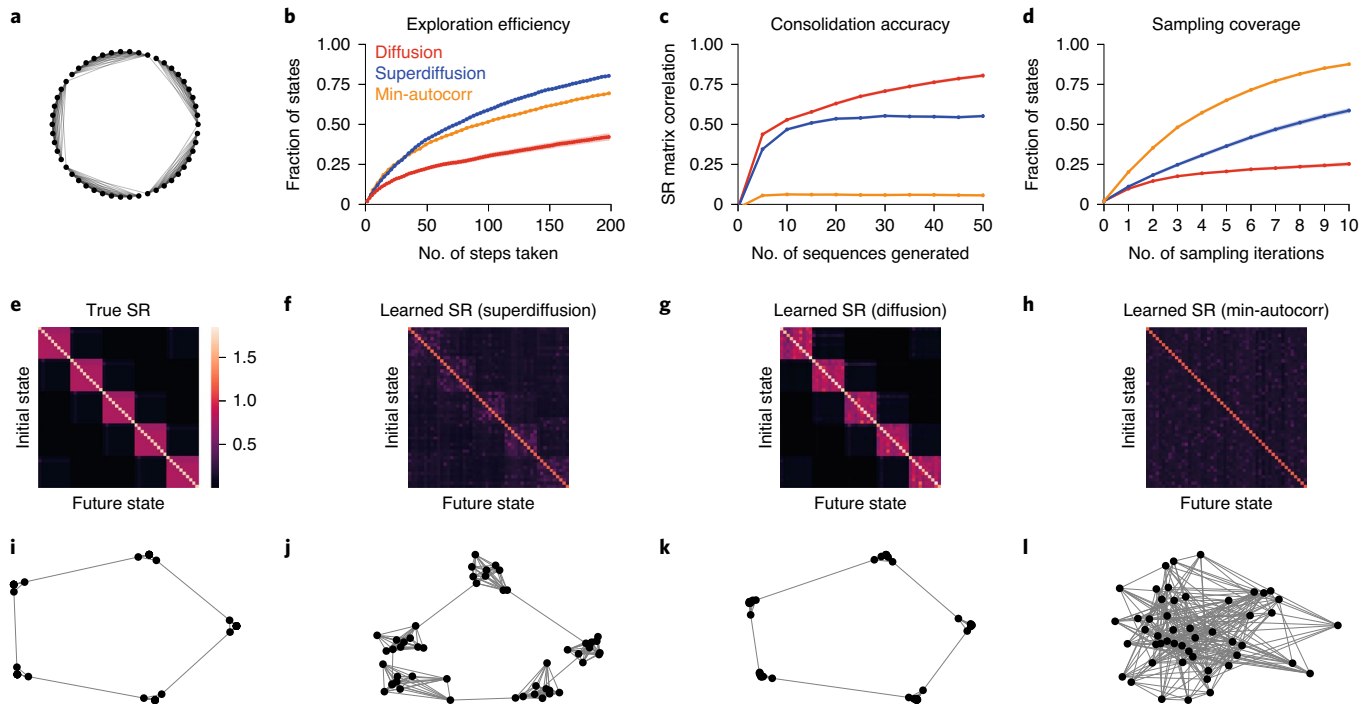


Fig. 7 | The performance of cognitive algorithms depends on the sampling regime. **a**, We consider a random-walk generator on a ‘ring of cliques’ state-space⁵⁰ and compare three different sampling regimes (diffusion, superdiffusion and minimally autocorrelated sampling) with respect to three distinct objectives that quantify the degree to which the sequences generated are optimized for exploration (**b**), learning (**c**) and sampling-based estimation (**d**), respectively. **b**, Exploration efficiency is defined as the fraction of states explored in an environment relative to the cumulative distance traversed. A steeper slope indicates that more states are explored for shorter travel distance. With respect to clustered state-spaces (**a**), the characteristic heavy tail of the superdiffusive propagator indicates that occasionally the next state sampled will be drawn from an alternative clique of states unlike diffusion, which tends to remain within a single clique (Extended Data Fig. 7). Minimally autocorrelated sampling is relatively inefficient since it repeatedly samples states across cliques to avoid the likelihood of resampling nearby states and thus is penalized heavily for traversal distance. **c**, Consolidation accuracy is the matrix correlation between the learned and true SRs. This is plotted as a function of the number of sequences generated. Diffusive replay facilitates the most accurate structure learning of the underlying state-space. **d**, Sampling coverage is the fraction of states sampled across multiple sequences as a function of the number of sampling iterations. Minimally autocorrelated sampling excels, as it is far less likely to resample states within and across sequences. **e**, The true SR matrix for the environment. **f–h**, The learned SRs for superdiffusive, diffusive and minimally autocorrelated sequence generation, respectively. **i–l**, Two-dimensional spectral embeddings of the true (**i**) and learned (**j–l**) SRs. The diffusive regime (**k**) successively recapitulates the geometric structure of the true SR (**i**), being composed of five well-spaced cliques of states. In contrast, the SR learned from superdiffusive replay (**j**) does not separate the cliques as clearly due to Lévy jumps, leading to the erroneous consolidation of illusory long-range transitions. Minimally autocorrelated sequence generation (**l**) corrupts the spatial structure of the state-space. (Sampling optimization, Methods).

between their place fields embodies the expected relationship between spatial and temporal displacements across the population code whereby place cells at a greater distance from one another tend to activate after larger time intervals (Fig. 8c). Performing the same analysis with the same grid code but with a turbulent spectral modulation returns qualitatively distinct results (Fig. 8d) in which the cross-correlation is essentially independent of place field distance consistent with a complete loss of sensitivity to the underlying spatial structure in sequence generation. While the former reflects the characteristic ‘V’ pattern observed in healthy hippocampal activity (Fig. 8e), the latter replicates the key characteristic of disordered hippocampal sequences recorded from this mouse model of schizophrenia⁴⁰ (Fig. 8f).

Discussion

Sequential hippocampal reactivations traversing cognitive maps are viewed as a potential neural substrate of internal simulation at the systems level^{5,13,42}. We have sought to address some of the apparent variability in the structure and statistics of sequence generation in the EHC^{8,9,11,12,36,40,41}. Our contributions are threefold. A principled motivation for activating distinct sequence generation regimes depending on the current cognitive process was established through

simulation and technical arguments. A linear feedback network model of EHC was proposed that implements the new technique of spectrally modulating sequence generation. This provides a mechanistic account of how random-walk processes may be smoothly interpolated with superdiffusive foraging patterns and minimally autocorrelated sampling. Applications of this model across a variety of environment geometries, behavioral states and cognitive functions were shown to explain variations in the spatiotemporal structure of hippocampal sequence generation and behavior across a number of experiments.

An open computational question is how might the optimal sequence generation regime for a particular cognitive algorithm and environmental scenario be identified within the brain. We suggest that this arbitration problem is subsumed within the framework of computational rationality that seeks to understand how the parameters of resource-limited algorithms may be optimized⁴³. Previous analyses within this broad remit have focused on optimizing the distribution from which samples are drawn during decision-making⁴⁴ or how state sequence propagation should be initialized and directed for reinforcement learning¹⁶. Generators may be custom designed to embed such desirable algorithmic features within sequence generation.

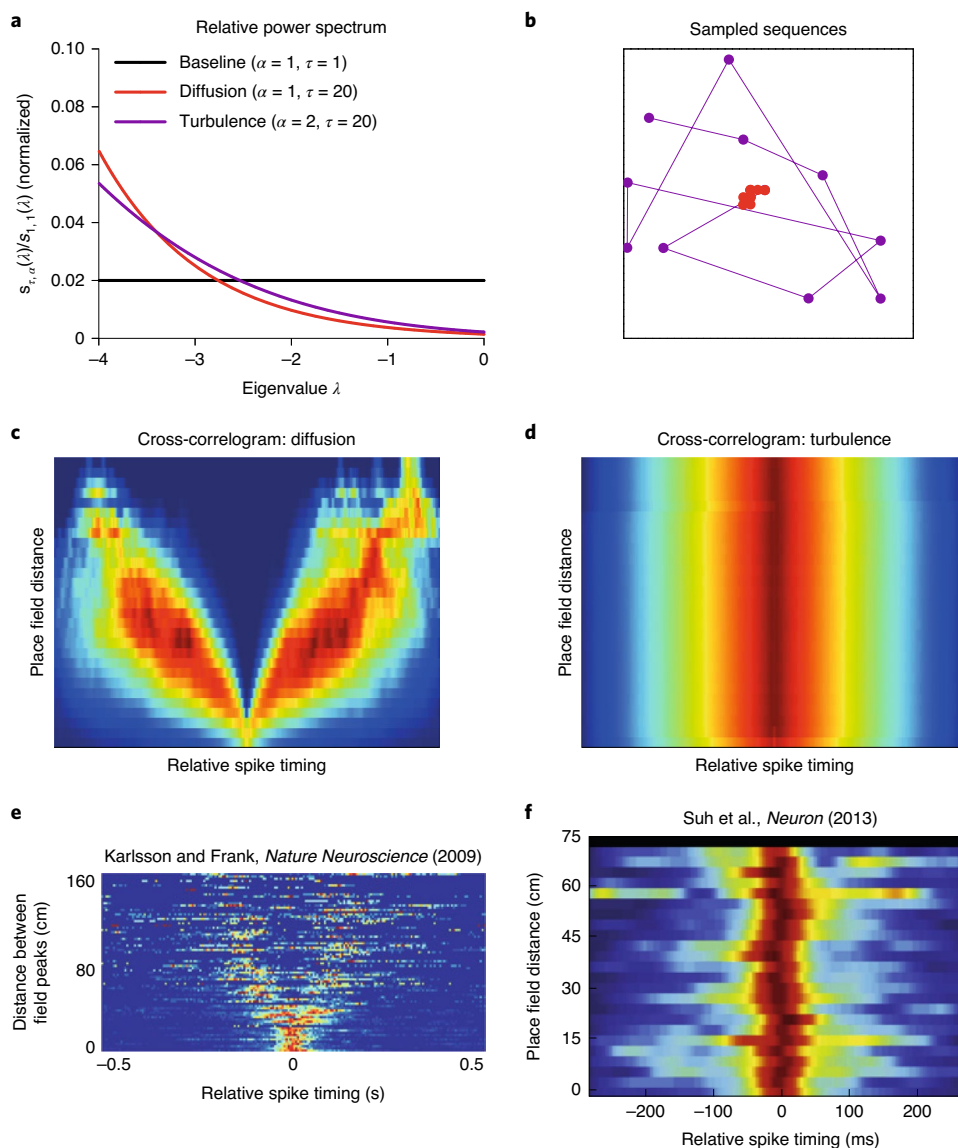


Fig. 8 | Dysregulated spectral modulation cause pathological forms of hippocampal reactivations. **a**, An imbalance in the entorhinal power spectrum (purple line) results in turbulent sequence generation. **b**, As demonstrated in an open-box environment, the resulting trajectories essentially sample randomly and thus are inconsistent with the spatiotemporal structure of the environment. **c,d**, Propagator cross-correlograms organized as a function of place field distance in the diffusion (**c**) and turbulent (**d**) regimes. **e**, The same analysis of electrophysiological recordings in the rodent hippocampus revealed the characteristic V structure as in the structurally consistent diffusion regime (**c**) (Karlsson and Frank, *Nature Neuroscience* (2009); image reproduced from ref. ⁴¹, Nature Publishing Group). Intuitively, the V structure reflects the fact that it takes time for the process to diffuse over space as a function of the distance between locations. **f**, In a genetic mouse model of schizophrenia (Suh et al., *Neuron* (2013)⁴⁰), the time interval between place cell activations in mutant mice did not reflect the spatial distance between the corresponding place fields, and the resulting cross-correlograms resembled those computed in the turbulent regime (**d**).

Our neural network model is designed to provide conceptual clarity regarding how entorhinal spectral modulation may explain variations in hippocampal sequence generation. An integration of other anatomical and functional features of the broader entorhinal–hippocampal system into this model is warranted. Inspiration may be drawn from continuous attractor network models of theta sequences and replay⁴⁵ and MEC electrophysiological recordings that have gleaned evidence that MEC input plays a causal role in refining the temporal organization of hippocampal activity^{33,34}. Our simulations are based on sampling from the propagated state distribution on each network iteration assuming a noisy readout to the hippocampal code. Possible extensions of this model include the sampling of multiple steps, or subsequences, on each iteration that

may be regulated by intra-hippocampal processing between CA3 and CA1 (refs. ^{7,34}). Indeed, a more comprehensive generalization of our model would also operationalize the spectral modulation mechanism along the trisynaptic pathway based on the projections from MEC layer II (ref. ⁴⁶). With respect to theta–gamma coupling^{38,47}, sequence generation may be embedded within gamma oscillations alternating between hetero-associative and auto-associative intra-hippocampal dynamics³⁶.

The causal relationship between the power spectrum of MEC activity and the statistical dynamics of hippocampal sequence generation is a core prediction of our model that has not been directly tested to our knowledge. In particular, spectral modulation between diffusive sequence generation (associated with learning and

consolidation during rest) and superdiffusive dynamics (expected during active exploration and goal-directed behavior) specifies distinct profiles of total grid population activity along the MEC dorsoventral axis potentially underpinned by variations in neural gain and grid rescaling. Furthermore, minimally autocorrelated hippocampal sequence generation (appropriate for sampling-based planning and inference) specifies new nonparametric spectral motifs that may be identifiable experimentally.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-021-00831-7>.

Received: 22 April 2020; Accepted: 3 March 2021;

Published online: 12 April 2021

References

- Buzsáki, G. & Moser, E. I. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* **16**, 130–138 (2013).
- Rowland, D. C., Roudi, Y., Moser, M.-B. & Moser, E. I. Ten years of grid cells. *Annu. Rev. Neurosci.* **39**, 19–40 (2016).
- Lisman, J. et al. Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nat. Neurosci.* **20**, 1434–1447 (2017).
- Ólafsdóttir, H. F., Bush, D. & Barry, C. The role of hippocampal replay in memory and planning. *Curr. Biol.* **28**, R37–R50 (2018).
- Foster, D. J. Replay comes of age. *Annu. Rev. Neurosci.* **40**, 581–602 (2017).
- Klinzing, J. G., Niethard, N. & Born, J. Mechanisms of systems memory consolidation during sleep. *Nat. Neurosci.* **22**, 1598–1610 (2019).
- Pfeiffer, B. E. The content of hippocampal ‘replay’. *Hippocampus* **30**, 6–18 (2020).
- Stella, F., Baracska, P., O’Neill, J. & Csicsvari, J. Hippocampal reactivation of random trajectories resembling Brownian diffusion. *Neuron* **102**, 450–461 (2019).
- Pfeiffer, B. E. & Foster, D. J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74–79 (2013).
- Johnson, A. & Redish, A. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* **27**, 12176–12189 (2007).
- Wikenheiser, A. M. & Redish, D. Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* **18**, 289–294 (2015).
- Kay, K. et al. Constant sub-second cycling between representations of possible futures in the hippocampus. *Cell* **180**, 552–567 (2020).
- Buzsáki, G. & Tingley, D. Space and time: the hippocampus as a sequence generator. *Trends Cogn. Sci.* **22**, 853–869 (2018).
- Káli, S. & Dayan, P. Off-line replay maintains declarative memories in a model of hippocampal-neocortical interactions. *Nat. Neurosci.* **7**, 286–294 (2004).
- Sanborn, A. & Chater, N. Bayesian brains without probabilities. *Trends Cogn. Sci.* **20**, 883–893 (2016).
- Matter, M. G. & Daw, N. D. Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* **21**, 1609–1617 (2018).
- Sokal, A. in *Functional Integration* Vol. 361 (eds DeWitt-Morette, C. et al.) 131–192 (Springer, 1997).
- McNaughton, B. L. et al. Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J. Exp. Biol.* **199**, 173–185 (1996).
- Burak, Y. & Fiete, I. R. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput. Biol.* **5**, e1000291 (2009).
- Ocko, S., Hardcastle, K., Giocomo, L. & Ganguli, S. Emergent elasticity in the neural code for space. *Proc. Natl Acad. Sci. USA* **115**, E11798–E11806 (2018).
- Bush, D., Barry, C., Manson, D. & Burgess, N. Using grid cells for navigation. *Neuron* **87**, 507–520 (2015).
- Behrens, T. et al. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018).
- Norris, J. *Markov Chains* (Cambridge Univ. Press, 1997).
- Klages, R., Radons, G. & Sokolov, I. M. *Anomalous Transport: Foundations and Applications* (Wiley, 2008).
- McNamee, D. & Wolpert, D. M. Internal models in biological control. *Annu. Rev. Control Robot. Auton. Syst.* **2**, 339–364 (2019).
- Weber, M. F. & Frey, E. Master equations and the theory of stochastic path integrals. *Rep. Prog. Phys.* **80**, 046601 (2017).
- McNamee, D., Wolpert, D. & Lengyel, M. Efficient state-space modularization for planning: theory, behavioral and neural signatures. *Adv. Neural Inf. Process. Syst.* **29**, 4511–4519 (2016).
- Stachenfeld, K., Botvinick, M. & Gershman, S. The hippocampus as a predictive map. *Nat. Neurosci.* **20**, 1643–1653 (2017).
- Michelmann, S., Staresina, B. P., Bowman, H. & Hanslmayr, S. Speed of time-compressed forward replay flexibly changes in human episodic memory. *Nat. Hum. Behav.* **3**, 143–154 (2019).
- Stensola, H. et al. The entorhinal grid map is discretized. *Nature* **492**, 72–78 (2012).
- Barry, C., Ginzberg, L. L., O’Keefe, J. & Burgess, N. Grid cell firing patterns signal environmental novelty by expansion. *Proc. Natl Acad. Sci. USA* **109**, 17687–17692 (2012).
- Mallory, C., Hardcastle, K., Bant, J. & Giocomo, L. Grid scale drives the scale and long-term stability of place maps. *Nat. Neurosci.* **21**, 270–282 (2018).
- Schlesiger, M. et al. The medial entorhinal cortex is necessary for temporal organization of hippocampal neuronal activity. *Nat. Neurosci.* **18**, 1123–1132 (2015).
- Yamamoto, J. & Tonegawa, S. Direct medial entorhinal cortex input to hippocampal CA1 is crucial for extended quiet awake replay. *Neuron* **96**, 217–227.e4 (2017).
- Viswanathan, G. M. et al. Optimizing the success of random searches. *Nature* **401**, 911–914 (1999).
- Pfeiffer, B. E. & Foster, D. J. Autoassociative dynamics in the generation of sequences of hippocampal place cells. *Science* **349**, 180–184 (2015).
- Browne, C. B. et al. A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games* **4**, 1–43 (2012).
- Colgin, L. et al. Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature* **462**, 353–357 (2009).
- Dayan, P. Improving generalization for temporal difference learning: the successor representation. *Neural Comput.* **5**, 613–624 (1993).
- Suh, J., Foster, D., Davoudi, H., Wilson, M. & Tonegawa, S. Impaired hippocampal ripple-associated replay in a mouse model of schizophrenia. *Neuron* **80**, 484–493 (2013).
- Karlsson, M. & Frank, L. Awake replay of remote experiences in the hippocampus. *Nat. Neurosci.* **12**, 913–918 (2009).
- Wikenheiser, A. M. & Redish, A. D. Decoding the cognitive map: ensemble hippocampal sequences and decision making. *Curr. Opin. Neurobiol.* **32**, 8–15 (2015).
- Gershman, S. J., Horvitz, E. J. & Tenenbaum, J. B. Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. *Science* **349**, 273–278 (2015).
- Lieder, F., Griffiths, T. & Hsu, M. Overrepresentation of extreme events in decision making reflects rational use of cognitive resources. *Psychological Rev.* **125**, 1–32 (2018).
- Kang, L. & DeWeese, M. Replay as wavefronts and theta sequences as bump oscillations in a grid cell attractor network. *eLife* **8**, e46351 (2019).
- Witter, M. P., Canto, C. B., Couey, J. J., Koganezawa, N. & O’Reilly, K. C. Architecture of spatial circuits in the hippocampal region. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20120515 (2014).
- Lisman, J. E. & Jensen, O. The theta-gamma neural code. *Neuron* **77**, 1002–1016 (2013).
- Hills, T. et al. Exploration versus exploitation in space, mind, and society. *Trends Cogn. Sci.* **19**, 46–54 (2015).
- Botvinick, M. M., Niv, Y. & Barto, A. C. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* **113**, 262–280 (2009).
- Schapiro, A., Rogers, T., Cordova, N., Turk-Browne, N. & Botvinick, M. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

Simulation details. *Linear track simulation.* Propagator simulations were based on a generator defining random walks on a discretized linear track state-space composed of ten states. Propagation densities were interpolated between states. The (τ, α) -modulated power spectrum $s_{\tau, \alpha}(\lambda)$ is defined as

$$s_{\tau, \alpha}(\lambda) = e^{-\tau^{-1}|\lambda|^\alpha} \quad (10)$$

and the relative (τ, α) -modulated power spectrum $r_{\tau, \alpha}(\lambda)$ is then

$$r_{\tau, \alpha}(\lambda) = \frac{s_{\tau, \alpha}(\lambda)}{s_{1, 1}(\lambda)} \quad (11)$$

The latter is normalized for each propagator to facilitate comparisons across propagators

$$\bar{r}_{\tau, \alpha}(\lambda) = \frac{r_{\tau, \alpha}(\lambda)}{\sum_k r_{\tau, \alpha}(\lambda_k)} \quad (12)$$

This measures the fraction of the total power spectrum accounted for by each spectral component relative to the baseline ($\tau = 1, \alpha = 1$) (Fig. 2).

Pfeiffer and Foster, Science (2015). A square enclosure containing no obstacles was represented by a 50×50 lattice of states. The stability parameter was set to $\alpha = 1$ and $\alpha = 0.5$ for diffusive and superdiffusive exploration, respectively. The tempo parameter was set to $\tau = 1$ in both cases. The simulation results reflect 20 sequences composed of 75 samples each. Superdiffusive exploration was more efficient than diffusive across a range of parameter values and changes in the size of the environment. The performance advantage of superdiffusion diminished as the environment size was reduced as all states became locally accessible to diffusive exploration and global repositioning was rendered unnecessary. In Fig. 2k, the step size was computed using the Euclidean distance between states embedded within an ambient continuous space.

Pfeiffer and Foster, Nature (2013). The square arena containing no obstacles was discretely represented by a 25×25 lattice of states (Fig. 3). The random-walk generator O^w (Supplementary Math Note) on this graph was modified to embed the rewarded home location by scaling the generator transition rates at home states according to

$$O_{h_i} \leftarrow v^{-1} O_{h_i} \quad (13)$$

where h indexes home states $x_{h_i} \in \mathcal{X}$ and v is a scalar quantity specifying the motivational value of home states. This leverages a unique property of the generator formalism whereby the temporal localization of states can be controlled independently of spatial structure (Supplementary Math Note).

In our simulations, the stability parameter was set to $\alpha = 1$ and $\alpha = 0.5$ for diffusive and superdiffusive sequence generation, respectively. The tempo parameter was set to $\tau = 5$ in both cases and the motivational value parameter $v = 100$. Trajectory events were simulated via sequence generation emanating from a home location ('home events') and four away locations ('away events'). The home location was specified as four adjacent states close to the center of the environment and the four away locations were placed in each corner. Twenty home events and ten away events per away location, each composed of ten samples, were generated. Sampling density was estimated from these trajectory events as the number of samples of each state divided by the total number of samples. This function was computed for each combination of condition (home and away) and sampling regime (diffusive and superdiffusive). The results were robust to variations in simulation parameters and the number of trajectory events.

Wikenheiser and Redish, Nature Neuroscience (2015). The circular track was discretized into 24 states with three further states connected at regular intervals on the outside of the circle serving as the goal locations (Fig. 4). Consistent with the behaviors acquired by the rats during task training, the generator used reflected a goal-directed behavioral bias in favor of anticlockwise movement around the track and turn-offs to goal locations. For sequence generation targeting a particular goal, the generator reflected a bias toward staying at the goal location once it was reached to reflect the time taken for the reward to be delivered, consummatory behaviors and trial termination. In the multi-goal scenario (Extended Data Fig. 3), the generator transition structure facilitated a return to the main track after goal arrival thus enabling multiple goal locations to be visited. Simulations of theta sequences (for both the individual sequence plots and look-ahead estimations) were composed of five iterations to match the number of decoded positions in the figure panels from ref. 11. As usual, stabilities $\alpha = 1$ and $\alpha = 0.5$ were used for diffusive and superdiffusive sequence generation, respectively, with $\tau = 1$ in both. Look-ahead distances were computed as the distance from the initial position to the furthest sampled position. The initial position was the start location marked by an X in the 'trajectory initiation' simulations (Fig. 4g). In the 'goal arrival' simulations, the initial position for a particular targeted goal was located in the center of the preceding track segment. The mean and standard error of

the look-ahead distances were estimated from 50 samples (an arbitrary number ensuring a stable estimate). The spatial scale (that is, the numeric distance between two adjacent states) was set to roughly match that of the circular track used in the experiments.

Kay et al., Cell (2020). We assumed that the rodent's decision-making policy reflected a directed run to the junction, a random choice to turn left or right, followed by a directed run to the end of the maze arm where a reward might be available (Fig. 5). A small transition rate opposing the directed transitions was added to ensure reversible and aperiodic propagation dynamics and numerically stable analyses. We initialized the spectrum in the diffusion regime and applied standard minimization routines (scipy.minimize) to solve the constrained optimization problem (equation 84, Supplementary Math Note). We minimized the integrated autocorrelation time summed over nine lags. Autocorrelations were estimated from 100 generated sequences of 20 iterations each.

Note that the planning objective (equation (8)) is based on the average-reward, infinite-horizon formulation of the Markov decision process⁵¹. This facilitates a simpler exposition and a more direct analogy to MCMC methods. For episodic and discounted MDPs, a similar but more complicated analysis may be pursued. In this case, the expected cumulative reward forms the value function objective requiring Monte Carlo estimation over sampled sequences. This may be accomplished via MCMC over the joint distribution over states across time points, or else by applying sequential Monte Carlo methods. The integrated autocorrelation time Δt_{ac} remains the key objective for sampling optimization since it is also reflected in the sample variance of the expected cumulative reward value function estimator.

Stella et al., Neuron (2019). We hypothesized that, although the rodents' physical experiences conformed to superdiffusive trajectories, the EHC recapitulated the environment experiences in the form of diffusive trajectories to facilitate accurate spatial consolidation, reconsolidation and maintenance processes during sleep. This is because diffusive replay embodies fundamental inductive biases regarding space, namely that space has a smooth, localized and isotropic structure. In contrast, superdiffusive trajectories are superior for foraging as shown in simulation (Fig. 7) and theoretical studies³⁵. Therefore, we also conjectured that, in contrast to sleep SWRs, SWRs occurring during immobile periods interleaved with active foraging (that is, 'wake SWRs') may reflect nonlocal, superdiffusive spatial trajectories to leverage a cognitive map of the environment in support of exploration. In particular, a superdiffusive sequence of positions may be generated during pauses, which the rodent can then follow physically to search for the food pellets.

Although rodents were familiar with the foraging environment at the time of the electrophysiological recordings and thus only a modest amount of new information may be required to be consolidated, theoretical and computational studies have shown that offline replay is crucial to maintaining precise storage and retrieval mechanisms for previously learned information¹⁴. This need is particularly pressing to avoid interference due to ongoing cortical plasticity as well as the consolidation of new information such as that drawn from experiences in new environments as is the case in the protocol of the current experiment. Furthermore, from a spatial cognition perspective, offline replay may contribute toward 'map refinement'⁸. That is, increasing the accuracy and resolution, and reducing any residual uncertainty, with respect to the structure of cognitive maps that may already be consolidated to a degree.

To show that our model supports the switching of the desired hippocampal sampling regime between diffusion for consolidation to superdiffusion for exploration based on monosynaptic input into the CA1 region, we modeled the MD curves of generated sequences under diffusion and superdiffusion. These curves were plotted as a function of the time interval Δt and compared to the empirical MDs for sleep and wake SWRs, respectively (Fig. 6a). The MD curves of sequence generation under the generator model were analytically computed (equation (17) and Methods) as a function of the stability parameters α and diffusion constants K estimated in ref. 8 (Fig. 6a). To focus on stability modulation, which specifically distinguishes between diffusion and superdiffusion, we averaged the empirical data over trajectory velocities (parametrized by the τ parameter in our model). In addition to SWR trajectories, we also plot the MD curve for the median velocity behavioral trajectories ('behavior', Fig. 6a). The spatial scale parameter σ was set to four to approximately match the ratio between the spatial (y axis) and temporal (x axis) scales. This parameter can be thought of as defining what constitutes a centimeter in our simulation and does not affect the relative slopes in MD.

Sampling optimization. A 'ring-of-cliques' state-space with five cliques composed of ten states each was studied (Fig. 7a). The definitive feature of this state-space is that it exhibits a community structure. Within each clique, states are densely connected, however, only sparse connections are available between cliques. Rings of cliques are commonly studied in the field of social psychology as a proxy for social networks, and it is known that human structure learning is sensitive to such community architectures⁵⁰. Propagators were based on the random-walk generator on this state-space. The power spectrum was parametrized such that the propagator shape was approximately equalized across conditions. This was accomplished by equating

the modal probability (that is, the 'height') of the propagator distributions. In particular, the stability parameter was set to $\alpha = 1$ for diffusion and $\alpha = 0.3$ for superdiffusion. These settings were paired with tempo parameters of $\tau = 20.7$ and $\tau = 3.1$, respectively. The integrated autocorrelation time over nine lags was minimized for the minimal-autocorrelation propagator. See Extended Data Fig. 7 for a visualization of sample sequences.

For exploration, a single sequence of 100 samples was generated and interpolated into a complete behavioral trajectory. For SR consolidation, sequences of 50 states were generated. The learned SRs (Fig. 7f–h) reflect the estimated SR after 500 sequences. For sampling, ten sequences of ten states each were generated (resulting in 100 states sampled in total). These are conceptualized as chains running in parallel as is commonly implemented in MCMC algorithms and thus the sampling coverage was computed by integrating over the ten sequences generated. For exploration and sampling, all sequences were initialized at a fixed state presumed to correspond to the position of the agent in the environment while, for learning, sequences were randomly initialized according to the stationary distribution as if randomly sampling from memory during sleep. All curves reflect the mean and standard error over 50 simulations.

The performance rank order across generative regimes for each measure was robust with respect to variations in τ and α , the number of sequences and sampling iterations, and the parametrization of the environment (that is, changes to the number of cliques or the number of states per clique). With regards to the latter, it seemed that the graph diameter was an important factor in determining the gain in exploration efficiency for the superdiffusive regime relative to the minimal-autocorrelation regime (Fig. 7b). The graph diameter measures the minimum distance the agent must travel in the worst-case scenario. In the ring-of-cliques model, the graph diameter scales with the number of cliques. As the graph diameter increased and, consequently, large nonlocal 'jumps' were more heavily penalized for distance, superdiffusions tended to excel. Therefore, one would expect that superdiffusive dynamics may be less important in small-world networks that are characterized by connectivity structures admitting a short path between almost all pairs of nodes. Such networks are inconsistent with Euclidean spaces but are notably over-represented in social and transport networks. We also observed that running multiple chains in parallel increased the minimal-autocorrelation performance in sampling coverage relative to superdiffusions (Fig. 7d). This suggests that, consistent with MCMC intuitions, there is a cumulative effect of sampling diversity when autocorrelations are minimized. That is, the evolution of each chain increasingly diverges from other chains oversampling iterations. More formally, the sample cross-correlation across chains is low due to the generative minimum autocorrelation property.

The structure consolidation results can be understood theoretically. Diffusion is a random process with zero drift and finite scale (measured as propagation variance as a function of time). Therefore, it encapsulates inductive biases regarding spatial structure that space is isotropic and has a scale that distinguishes between local and global structure. Moreover, from an information-theoretical point of view, it is the simplest such random process, being the unique maximum entropy process with these properties⁵². This suggests that diffusive sequences implicitly encode isotropy and scale, and nothing else. In contrast, the distinguishing feature of superdiffusive Lévy flight processes in unbounded Euclidean spaces is the infinite variance of its heavy-tailed propagation distribution²⁴. This means that extraordinarily large jumps may be generated and so the downstream learning process is unable to distinguish between local structure (learned from individual small jumps within short time periods) and global structure (learned from the accumulation of many small jumps over a period of time).

Karlsson and Frank, *Nature Neuroscience* (2009) and Suh et al., *Neuron* (2013). The diffusion and turbulence regimes were parametrized by $\alpha = 1$ and $\alpha = 2$, respectively, with τ set to 20 in both cases. In a linear track environment, 100 state sequences with 100 steps were generated in each of these stochastic regimes. The average spiking activity of each place cell (each encoding a distinct state) was taken to be the output propagator density. The Euclidean distance between states (assuming the underlying graph is embedded in an ambient continuous space) was taken to be the distance between place fields. The relative spike timings were offsets in the number of steps within the generated sequences. Given these modeling assumptions, we computed normalized cross-correlograms of place cell activity as a function of the distance between place fields⁴¹. The key qualitative distinction between the cross-correlograms was robustly observed regardless of modifications to the other model parameters (for example, τ , number of sequences, number of steps).

In Fig. 8e, we present a cross-correlogram of SWRs recorded in healthy mice, which clearly shows the predicted V structure consisted with a systematic relationship between temporal and spatial displacements during sequence generation. This plot comes from ref. 41. Note that, in ref. 40, the cross-correlogram for the littermate controls is also presented (Fig. 4c, ref. 40) but the V structure is less clear (presumably, due to fewer SWRs and a smaller environment). Therefore, we include the V structure from the prototypical analyses in ref. 41 since it is easier to interpret on viewing for the first time.

Evaluating sequence generation across exploration, learning and planning. The central computational motivation in conceptualizing MEC as a modulator

of hippocampal sequence generation is that distinct cognitive algorithms have fundamentally different requirements as to the statistical and dynamical structure of the input sequences received from EHC. To establish this empirically, we evaluated sequences generated with respect to metrics sensitive to the performance of exploration, learning and sampling-based planning algorithms. The three metrics are exploration efficiency, consolidation accuracy and sampling coverage, and are defined in the following three subsections.

Exploration efficiency. To evaluate how appropriate a generative sampling regime is for exploration, we quantified exploration efficiency as the fraction of environment states or positions visited relative to the cumulative distance traversed. This definition is a simple adaptation to graph structures of the standard definition commonly studied in the foraging literature⁵³. Exploration efficiency is plotted as a function of the number of states sampled in Fig. 7b for diffusive, superdiffusive and minimally autocorrelated sequence generation. Note that, if the generator model samples successive nonadjacent states, then it is assumed that the agent visits all states along the shortest path between the sampled states as if physically traversing the environment. In a discrete state-space, the distance is taken to be the number of steps taken while in a continuous domain, it is the Euclidean distance between positions.

Propagators may repeatedly sample the same state on multiple iterations leading to a period of immobility for the agent. We assume that the cost of remaining in a state is zero since the agent has not traversed any distance. However, resampling the same state adds to the total time cost of exploration, which can be defined as the total time taken for sampling subsequent states and moving to those states. Assuming that hippocampal state sampling is embedded within theta⁴¹ or slow-gamma³⁶ cycles, the rate at which new states are sampled is on the order of centiseconds. In contrast, a rodent requires time periods on the order of seconds to move between positions sampled in a typical open-box environment. Thus, the contribution of sampling time to total exploration time is negligible and is therefore not reflected in the exploration efficiency measure. Note that, in this regard, the definition of exploration efficiency contrasts sharply with the sampling coverage measure that specifically penalizes for time and not distance.

Consolidation accuracy. We quantify how well a structural representation of an environment (in particular, a SR³⁹) can be learned from replayed state sequences. As a measure of consolidation accuracy, we use the Spearman correlation between the true and learned state-space SRs as a function of the number of sequences generated. This measure is plotted as a function of the number of sequences generated in Fig. 7c for each modulatory regime. Before learning commences, the previous SR is taken to be that generated by a fully connected graph. This suggests that the future expected rate of occupancy is completely homogeneous across initial and successor states a priori. The SR is learned via a standard temporal difference learning rule^{28,39}.

Sampling coverage. An important objective in sample-based inference (for example, in planning²⁷ and reinforcement learning²⁴) is to generate a diversity of sampled states as quickly as possible as this leads to robust and rapid estimators. Sampling coverage is defined as the fraction of environment states or positions visited relative to the number of sampling iterations (Fig. 7d). Note that, unlike the exploration efficiency measure, sampling coverage does not penalize for the distance traversed between states, as it is presumed to reflect an internal sampling mechanism that is completed before any contingent behavior. The sampling coverage measure specifically penalizes for repeatedly sampling the same state since the fraction of states sampled remains the same despite the number of sampling iterations increasing. Both exploration efficiency and sampling coverage are sensitive to the fraction of states visited or sampled, respectively. The major distinction between the two is that the former isolates the cost of distance traversed by the agent, while the latter penalizes the time taken to generate the sequence.

Mean-squared displacement (MSD) and MD as diagnostic measures of the sequence generation regime. MSD, $\text{MSD}(t) = \langle |d(t)|^2 \rangle$, is a measure of the scaling law of a stochastic process and can be estimated from trajectory samples^{8,24}. It is a time-dependent function that characterizes the relationship between the spatial and temporal displacements. For any time period t , $\text{MSD}(t)$ measures how far away from the initial state at $t=0$ that sequence generation is expected to sample as a function of time. Given a trajectory $\mathbf{x} = (x_0, \dots, x_t)$ in Euclidean spaces, the measure of spatial displacement is taken to be the squared Euclidean distance $d(t)$ with respect to the initial state:

$$d(t) = \|\mathbf{x}_t - \mathbf{x}_0\| = \sqrt{\sum_{i=1}^D (x_{t,i} - x_{0,i})^2} \quad (14)$$

where $x_{t,i}$ is the i th position coordinate of the process at time t and D the dimensionality of the state-space. It can be shown analytically²⁴ that, for diffusive-type processes, MSD scales according to

$$\text{MSD}(t) = \langle d(t)^2 \rangle \propto K (\tau^{-1}t)^{\alpha-1} \quad (15)$$

In log space

$$\log \langle d(t)^2 \rangle \propto \log K + \alpha^{-1} \log(\tau^{-1}t) \quad (16)$$

and so the intercept on a log–log plot is the log of the diffusion constant K (which sets the rate of diffusion independent of τ and α modulation) and the slope is the inverse of the stability parameter α . If the slope of the MSD graph equals 1 then the process is a diffusion. A slope greater than 1 indicates supralinear displacements and therefore the process is known as a superdiffusion ($\alpha < 1$). This is the genesis of the term superdiffusion. Therefore, the stability regime of sequence generation can be inferred from sample trajectories by estimating $\text{MSD}(t)$. The key implication of a supralinear MSD is that a superdiffusion is expected to sample further away from its origin over time compared to diffusion. When computing $\text{MSD}(t)$ for sequences generated by our model, we consider the Euclidean distance d induced by embedding the state-space within an ambient Euclidean space.

The MD is analyzed in Fig. 6 and ref. ⁸. This is similar to the MSD formula but without squaring the distance function:

$$\text{MD}(t) = \langle d(t) \rangle \propto K (\tau^{-1}t)^{(2\alpha)^{-1}} \quad (17)$$

therefore, in log space, we have

$$\log \text{MD}(t) \propto \log K + (2\alpha)^{-1} \log(\tau^{-1}t). \quad (18)$$

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

No data was acquired for this study.

Code availability

Simulation code was written using open source packages in the Python 3 programming environment and can be found at <https://github.com/dmcnamee/FlexModEHC>.

References

51. Sutton, R. & Barto, A. *Reinforcement Learning: An Introduction* (MIT Press, 2018).

52. Cover, T. & Thomas, J. *Elements of Information Theory* (Wiley, 2006).

53. Viswanathan, G. M., Da Luz, M. G., Raposo, E. P. & Stanley, H. E. *The Physics of Foraging: An Introduction to Random Searches and Biological Encounters* (Cambridge Univ. Press, 2011).

54. Jinnai, Y., Park, J. W., Abel, D. & Konidaris, G. Discovering options for exploration by minimizing cover time. Preprint at arXiv:1903.00606 (2019).

Acknowledgements

This work was supported by the Center for Brains, Minds and Machines (funded by NSF STC award no. CCF-1231216) and the Wellcome Trust (Sir Henry Wellcome Fellowship grant no. 110257/Z/15/Z to D.C.M.). For the purpose of Open Access, we have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. We thank A. Bernacchia, Q. Huys, Z. Kurth-Nelson, M. Nour, E. Russek, H. Sheahan, H. Stojic, S. Tootoonian and E. Wimmer for discussions and feedback. We thank F. Stella and J. Csicsvari (IST Austria) for sharing their data and for performing additional analyses. We thank B. Sorscher and colleagues (Stanford University) for sharing their code.

Author contributions

D.C.M., K.L.S., M.M.B. and S.J.G. contributed toward model conceptualization. D.C.M. performed the simulations. D.C.M. prepared a draft of the manuscript. D.C.M., K.L.S., M.M.B. and S.J.G. reviewed the manuscript contributing edits and comments.

Competing interests

D.C.M. and S.J.G. declare no competing interests. K.L.S. and M.M.B. are employed by Google DeepMind.

Additional information

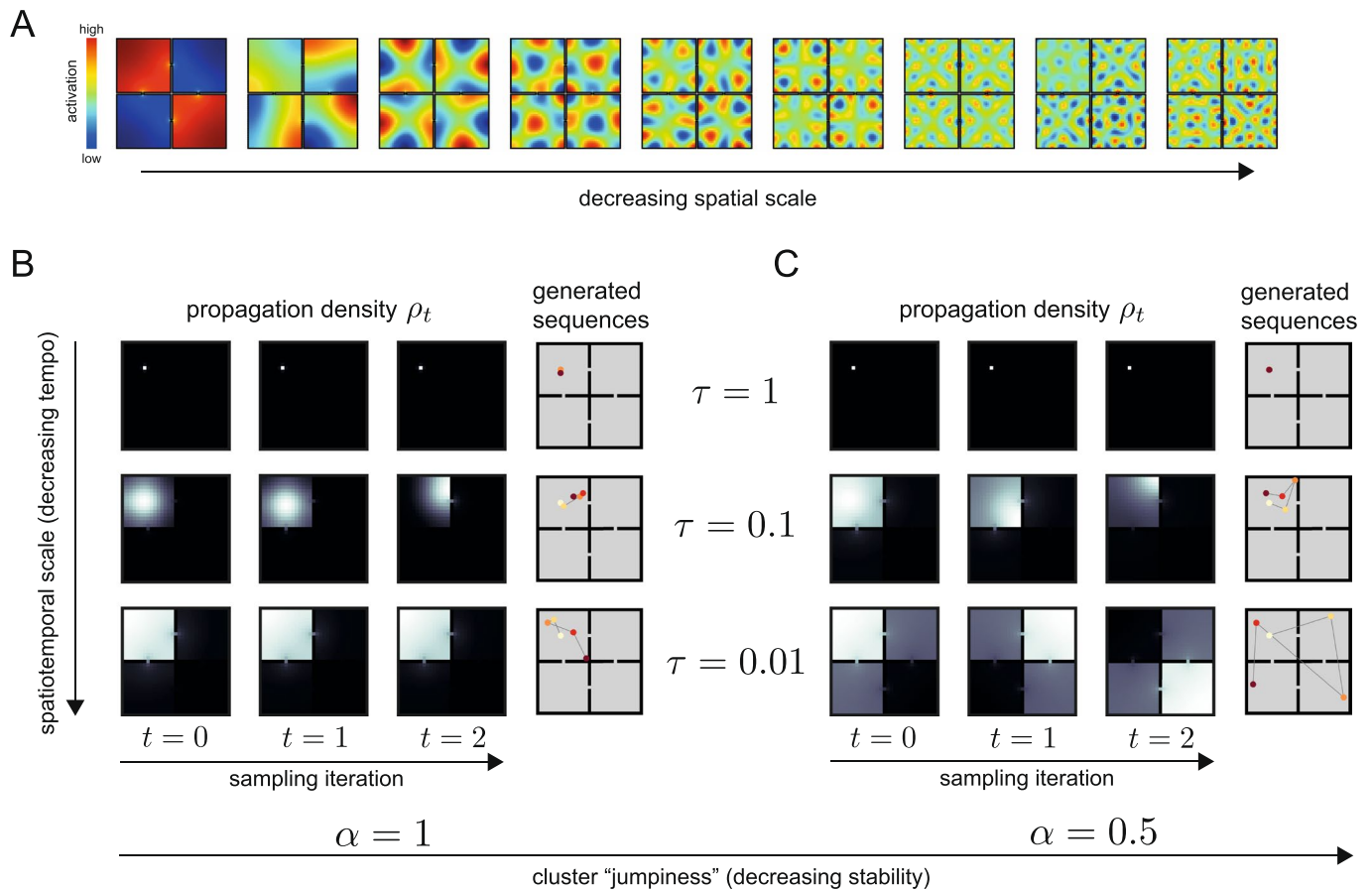
Extended data is available for this paper at <https://doi.org/10.1038/s41593-021-00831-7>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41593-021-00831-7>.

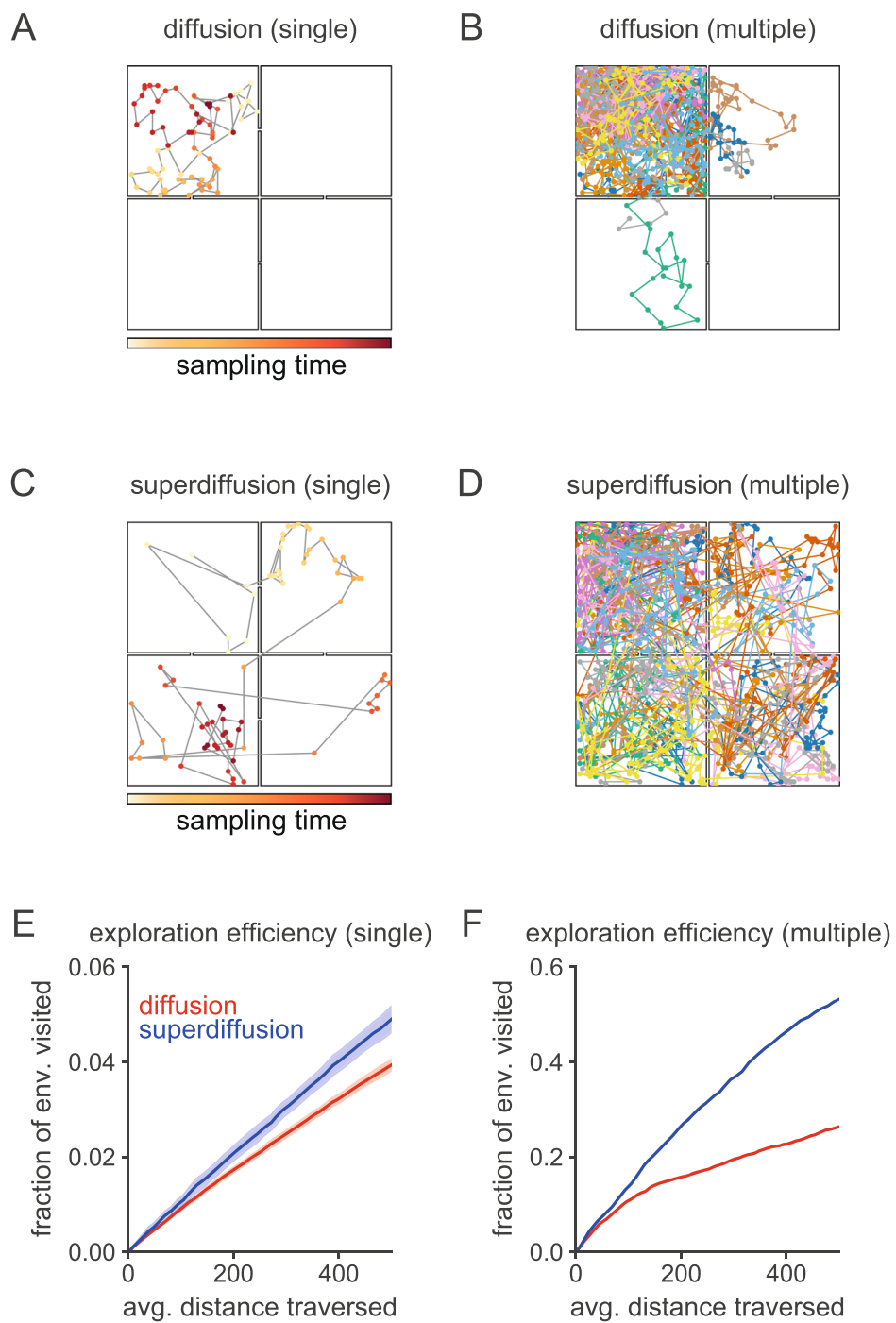
Correspondence and requests for materials should be addressed to D.C.M.

Peer review information *Nature Neuroscience* thanks Sen Cheng, Kiah Hardcastle and Martin Stemmler for their contribution to the peer review of this work.

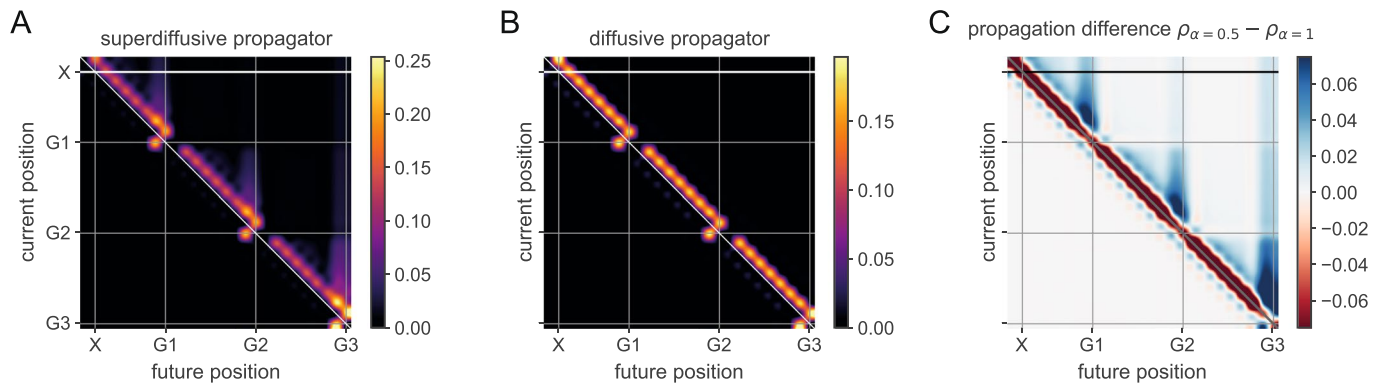
Reprints and permissions information is available at www.nature.com/reprints.



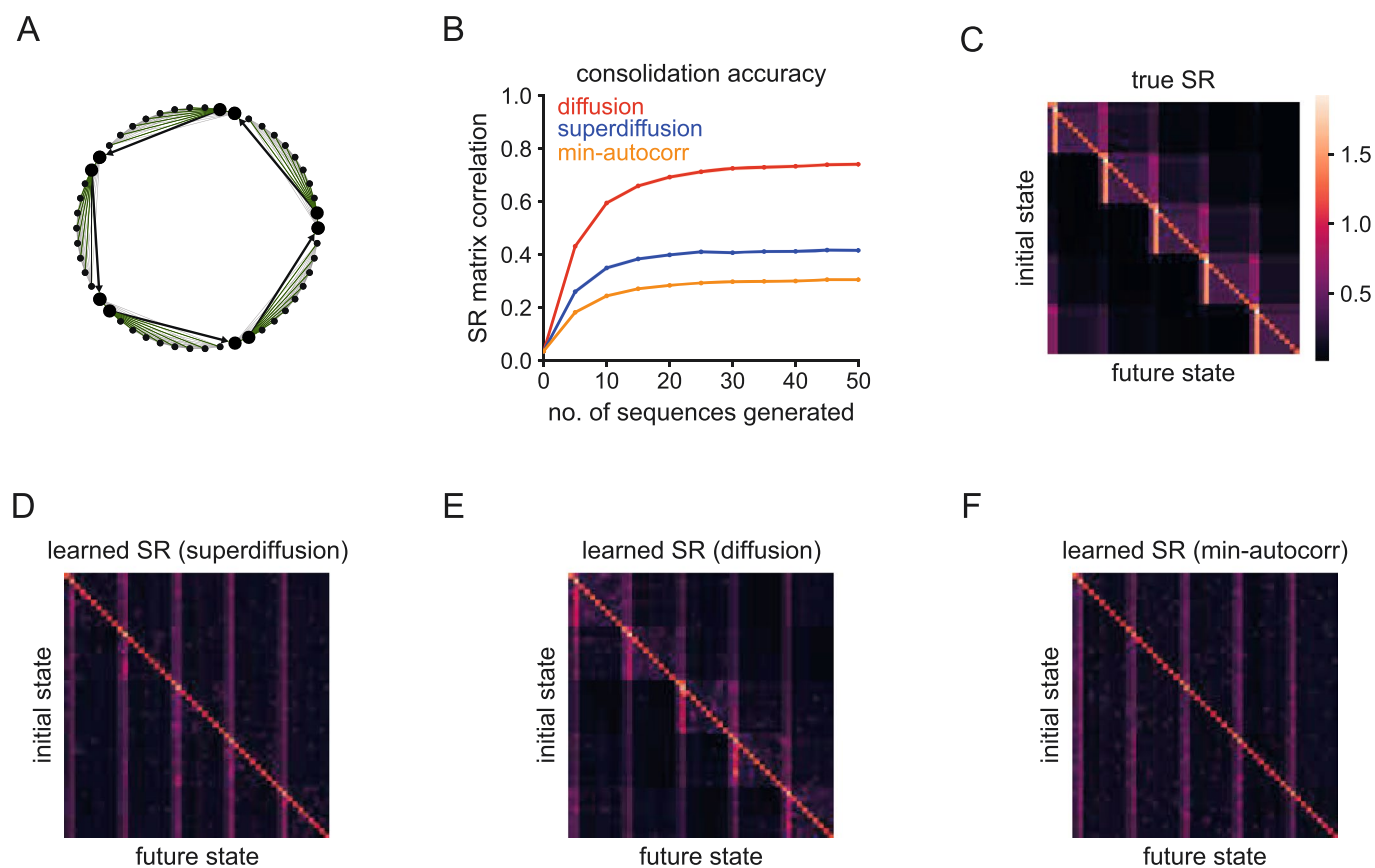
Extended Data Fig. 1 | Sequence generation in an inhomogeneous environment with tempo τ and stability α modulation. **a**, Spectral components of the four-room grid world are presented as a function of spatial scale. Generator eigenvectors coincide with the eigenvectors of the corresponding SR. Therefore, many examples of generator eigenvectors may be observed in ref. ²⁸. **b**, As the tempo parameter τ is reduced (with the stability parameter held at $\alpha = 1$), the power spectrum $s_{\tau, \alpha}(\lambda)$ reweights the spectral components leading to changes in the spatial scale of the propagation densities ρ_t . **c**, The generated sequences vary in the extent of their spatiotemporal traversal of the environment.



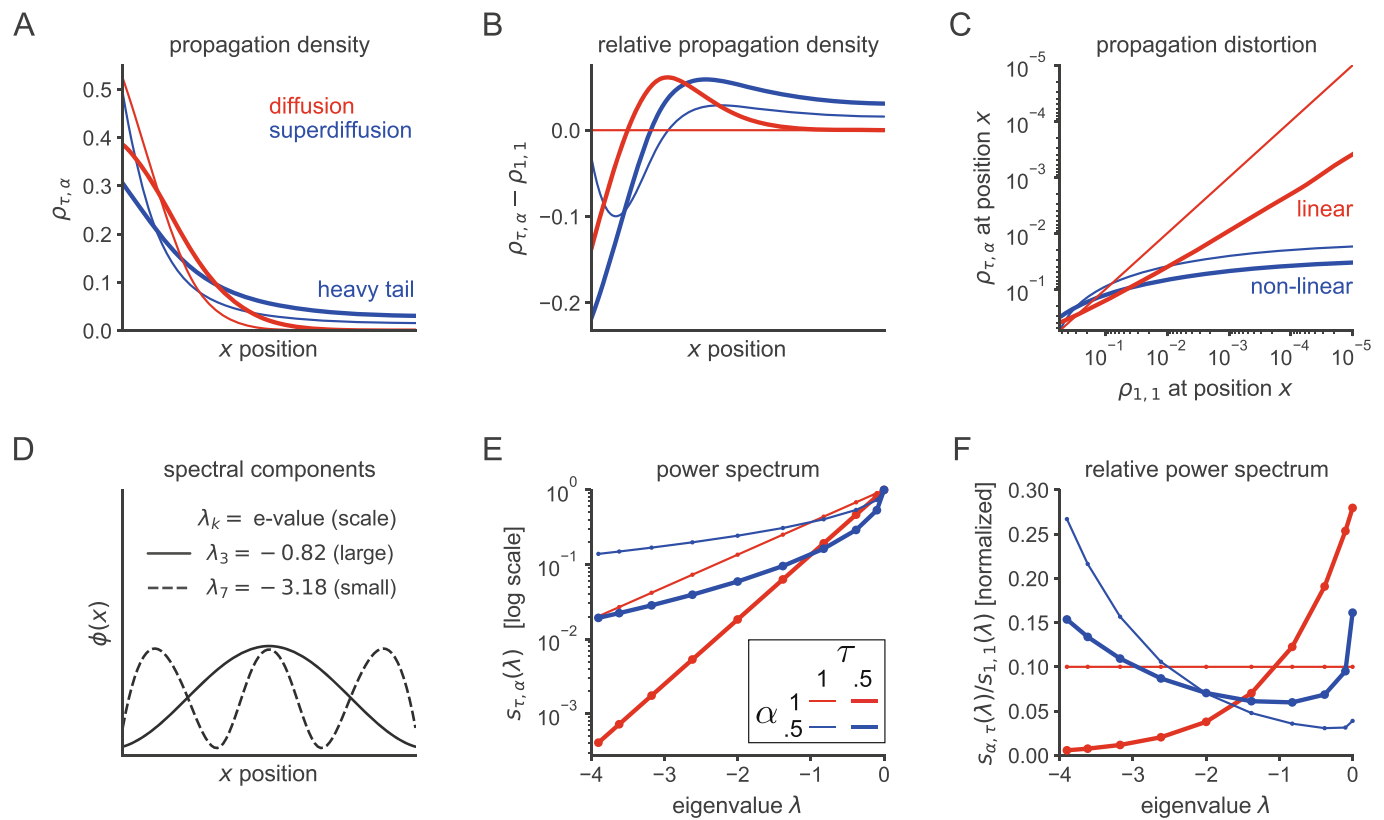
Extended Data Fig. 2 | Exploration efficiency in a structured environment. The exploration simulation in Fig. 2 is repeated here but for an environment decomposed into four areas connected by bottlenecks. Each area consisted of 50×50 states. Simulation parameters were unchanged from Fig. 2 (Methods). **e**, Error bands \pm s.e.m. for $n=20$ trajectories.



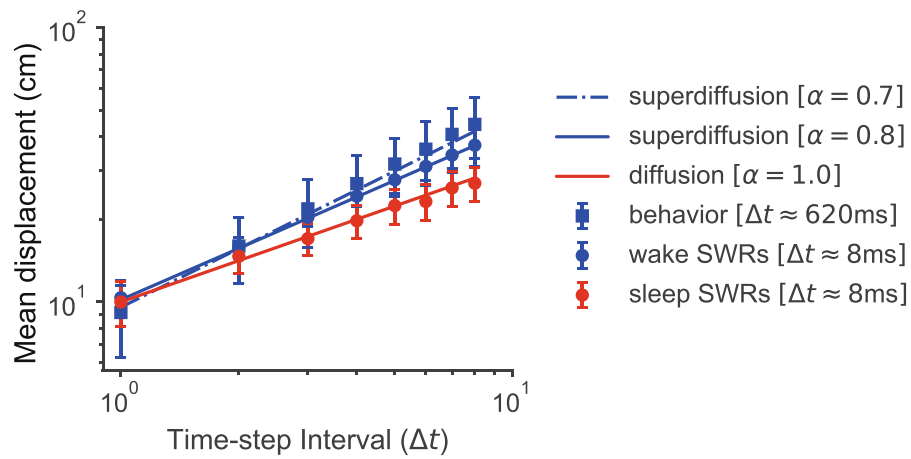
Extended Data Fig. 3 | Propagator analysis across all possible current and future positions for the circular track task. a, b. We plot the superdiffusive (panel A, $\alpha=0.5$) and diffusive propagators (panel B, $\alpha=1$) for the circular track task¹¹ for all current positions (y-axis) with all goals activated. The y-axis (x-axis) indexes the current (future) position. Each row reports the probability of each position being sampled given the current position (the row index). **c.** The difference in diffusive and superdiffusive propagators is plotted. For example, the row corresponding to the initial X position (marked by the thick black line), shows that the local activations are less likely (red) under superdiffusions compared to diffusion (though they are still possible) while remote goal locations specifically are more likely (blue).



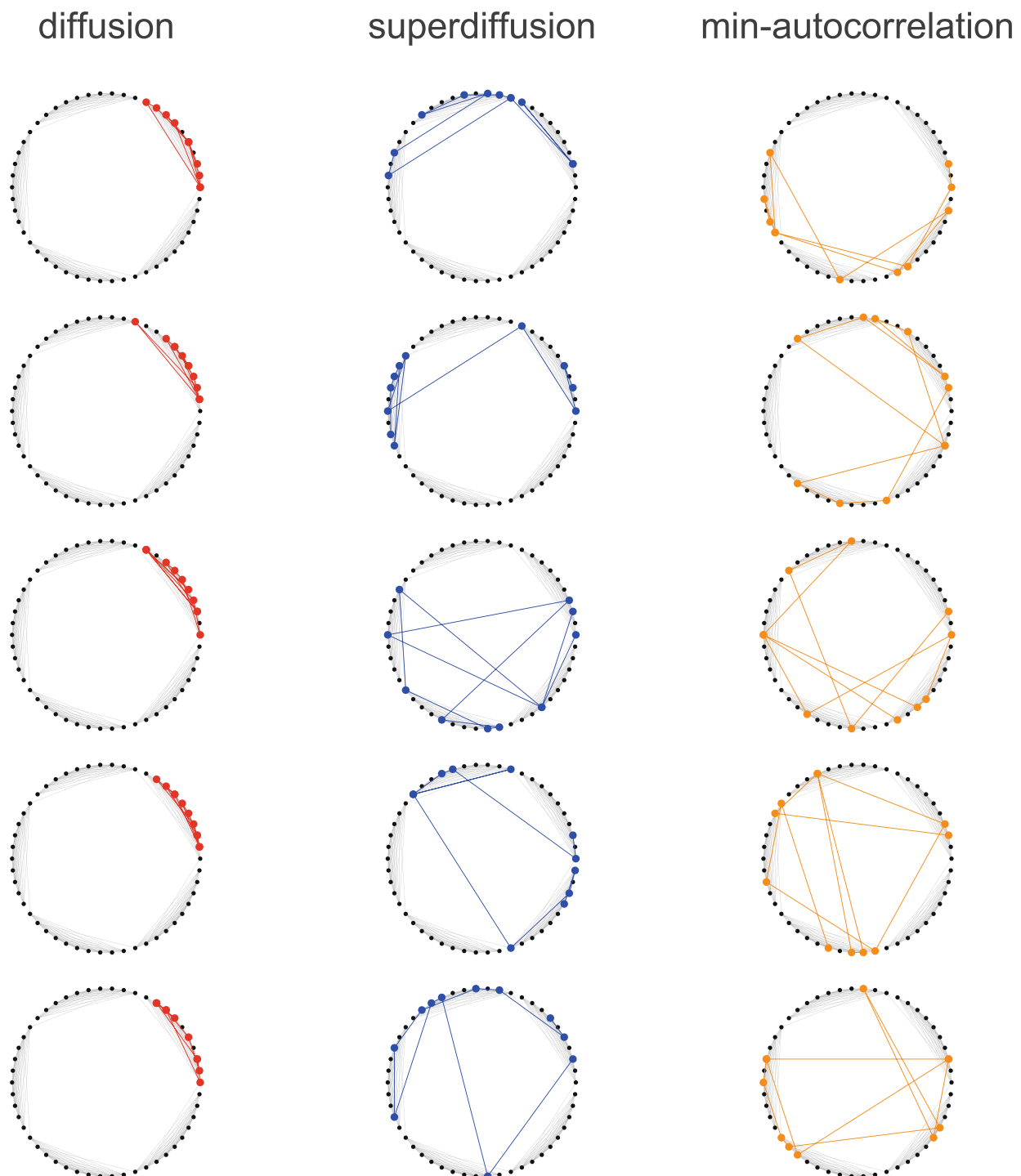
Extended Data Fig. 4 | Diffusions are optimized for consolidating directed transition structures. Three regimes of sequence generation are compared in terms of consolidating a directed policy on the same graph used in Fig. 7 where the policy was assumed to be undirected (that is a random walk). **a**, The directed policy is dominated by high probability anticlockwise transitions (heavy black arrows) between the clique bottleneck states (highlighted by increased size). Within each clique, transitions to the bottleneck state facilitating an anti-clockwise transition to the next clique are relatively likely to occur. **b**, The same rank order as in Fig. 7c is observed with the diffusive regime leading to the most accurate consolidation of the directed policy successor representation (SR). **c**, The true directed policy SR. **d–f**, The SRs learned via superdiffusion, diffusion, and minimally autocorrelated sampling respectively. Note that the superdiffusive and minimal autocorrelation regimes identify the the dominant anti-clockwise transition structure but fail to reflect the clustering of states based on their clique membership. This simulation indicates that the diffusive regime is optimal for consolidation regardless of whether the target transition structure corresponds to a random walk or not.



Extended Data Fig. 5 | Detailed analyses of spectrally modulated propagation in the linear track environment. **a**, Copy of Fig. 2d for context and comparative visualization. **b**, The propagation distributions in panel A are re-expressed as the difference with respect to the baseline ($\tau=1, \alpha=1$). **c**, Each propagation density is plotted against the baseline propagation density ($\tau=1, \alpha=1$) on log scales. Note that only superdiffusion results in a non-linear warping of the propagation density. This non-linearity underpins the heavy-tail in the propagation density. **d**, Small-scale (dashed line) and large-scale (full line) spectral components (that is, generator eigenvectors) are plotted as a function of linear track position. Eigenvalues take a maximum value of 0 and are ordered according to the spatial scale of the corresponding spectral component. For example, eigenvalues close to zero correspond to large-scale spectral components with the spatial scale of the corresponding spectral component decreasing with the eigenvalue which are always less than or equal to zero. Therefore, the scale of a spectral component ϕ_k decreases as the absolute value $|\lambda_k|$ of its corresponding eigenvalue λ_k increases. **e**, Spectral power is plotted as a function of eigenvalue on a log scale. As the tempo τ decreases towards zero, the spatiotemporal horizon increases and the propagation density converges on the stationary state distribution. This corresponds to all spectral components with non-zero eigenvalues decaying to zero. Note that diffusion linearly changes the slope of the power spectrum while superdiffusion imposes a non-linear reweighting. This effect underpins the nonlinear time warping in the propagation densities in panel C. **f**, The characteristic modulation of the power spectrum under diffusion and superdiffusion is demonstrated by the relative power spectrum $\frac{s_{\tau,\alpha}(\lambda)}{s_{1,1}(\lambda)}$ computed as the power ratio with respect to the baseline propagator. Note that the long-range diffusion propagator ($\alpha=1, \tau=0.5$, thick red line) downweights small-scale components and upweights large-scale components while superdiffusion ($\alpha=0.5, \tau=1$, thick blue line) relatively upweights both small- and large-scale components but suppresses at medium scales. This enables a superdiffusive propagator to jump to remote positions without traversing intermediate locations.



Extended Data Fig. 6 | Mean displacement scales linearly on a log-log plot. The data and simulations from Fig. 6 are plotted on a logarithmic scales demonstrating the linear relationship between mean spatial displacement and time intervals. Mean displacement (MD) measures the ratio between the spatial displacement and time interval of sequence generation and can be used to infer whether a generative sampling process is operating in a diffusive or superdiffusive regime based on sampled sequences. In particular, the MD function has a slope proportional to α^{-1} with respect to the time interval Δt (Methods). Error bars \pm s.e.m..



Extended Data Fig. 7 | Sampled sequences exhibit distinctive patterns across spectral modulation regimes. Each column shows five sample sequences generated in one of the three sampling regimes studied in Fig. 7. Note that diffusions (left column) tend to remain within a single clique. Superdiffusions (middle column) interleave local diffusions with occasional jumps to other cliques. Jumps are observed as lines crossing the center of the ring which do not conform to any transition step in the underlying state-space structure (gray lines). Minimally autocorrelating propagators (right column) repeatedly generate large jumps between cliques.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Simulation code was written using open source packages in the Python 3 programming environment. Code available at <https://github.com/dmcnamee/FlexModEHC>.

Data analysis Analysis code was written using open source packages in the Python 3 programming environment. Code available at <https://github.com/dmcnamee/FlexModEHC>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

No data was acquired for this study.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="N/A"/>
Data exclusions	<input type="text" value="N/A"/>
Replication	<input type="text" value="N/A"/>
Randomization	<input type="text" value="N/A"/>
Blinding	<input type="text" value="N/A"/>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

- | n/a | Involvement in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |