PNAS Nexus, 2024, 3, pgae525
https://doi.org/10.1093/pnasnexus/pgae525
Advance access publication 22 November 2024
Research Report

Continuous Quasi-Attractors dissolve with too much – or too little – variability

Francesca Schönsberg [a,b,*], Rémi Monasson [a] and Alessandro Treves [b,c,*]

[a]Laboratory of Physics of the Ecole Normale Supérieure, PSL and CNRS UMR8023, Sorbonne Université, Paris 75005, France
[b]SISSA, Scuola Internazionale Superiore di Studi Avanzati, Cognitive Neuroscience, Trieste 34136, Italy
[c]DSyNC lab, University of Agder, Kristiansand 4604, Norway
*To whom correspondence should be addressed: Email: francesca.schonsberg@phys.ens.fr (F.S.); ale@sissa.it (A.T.)
Edited By Ivet Bahar

Abstract

Recent research involving bats flying in long tunnels has confirmed that hippocampal place cells can be active at multiple locations, with considerable variability in place field size and peak rate. With self-organizing recurrent networks, variability implies inhomogeneity in the synaptic weights, impeding the establishment of a continuous manifold of fixed points. Are continuous attractor neural networks still valid models for understanding spatial memory in the hippocampus, given such variability? Here, we ask what are the noise limits, in terms of an experimentally inspired parametrization of the irregularity of a single map, beyond which the notion of continuous attractor is no longer relevant. Through numerical simulations we show that (i) a continuous attractor can be approximated even when neural dynamics ultimately converge onto very few fixed points, since a quasi-attractive continuous manifold supports dynamically localized activity; (ii) excess irregularity in field size however disrupts the continuity of the manifold, while too little irregularity, with multiple fields, surprisingly prevents localized activity; and (iii) the boundaries in parameter space among these three regimes, extracted from simulations, are well matched by analytical estimates. These results lead to predict that there will be a maximum size of a 1D environment which can be retained in memory, and that the replay of spatial activity during sleep or quiet wakefulness will be for short segments of the environment.

Keywords: spatial memory, recurrent networks, associative plasticity, attractor dynamics, place cells

Significance Statement

Remembering one's position on a spatial continuum seemed to require a smooth orderly representation inside the brain, arising from a continuous attractor neural network dynamics. Yet, place cells in the hippocampus have been shown to be quite disorderly, when recorded in bats flying in a long tunnel. Does it mean they cannot remember where they were? No, disorder still enables a continuous quasi-attractive line of dynamical states to emerge, but up to a limit. In fact, simulations show that there is a minimum degree of disorder as well as a maximum one enabling spatial memory retrieval. These two values, which we can estimate analytically, coalesce when the environment gets larger, implying a maximum size that can be stored in memory.

Introduction

The study of spatial representations in the hippocampus has been extensively conducted by analyzing recordings of place cells in simple laboratory environments. This body of research has consolidated the notion that a typical place cell exhibits activity within a single defined place field (1). However, this perspective has been increasingly questioned in recent years. Studies conducted in larger environments (2–6) have demonstrated the presence of multiple fields per place cell, while earlier multifield recordings, particularly in the dentate gyrus, had already hinted at such complexity (7–9). A pivotal contribution to this ongoing debate was made by Eliav et al. (10). Their recent study quantified the distribution of the fields expressed by individual CA1 place cells in bats flying in a 200 m long tunnel. Place cells were shown to have up to >20 different fields, with huge variability in the peak firing rate and in the width. The latter was reported to be well fit by a log normal distribution, which allows for small fields from under a meter wide, up to large ones of tens of meters. Further, experimental evidence both in bats (10) and rats (5) indicates that the same place cell can exhibit a single field in small environments and multiple fields in larger ones, challenging the notion that some place cells are strictly uni-field while others are multifield. For years, associative memory network models (11–13) have explored the hypothesis that the hippocampal representation of space might be comprised of place cells expressing predominantly single fields of standard size. Theoretically, within a recurrent neural network this assumption facilitates modeling

Competing Interest: The authors declare no competing interests.
Received: May 31, 2024. **Accepted:** November 12, 2024

place field storage as continuous attractors being established through unsupervised Hebbian learning. In these models, each position in the spatial environment, whether in one, two, or three dimensions—or, equivalently, in abstract continuous spaces of any dimensionality—maps to a fixed point on an attractive continuous manifold. The fixed points, however, are only marginally stable, susceptible to even minimal forces acting along the manifold, and then true continuity only arises in mathematical limit conditions. Consequently, since the introduction of such (approximately) continuous attractor neural networks (cANN) in a wider neuroscience context (14), a significant effort has been dedicated to understanding the factors that may drive in practice their neural dynamics and determine their stability. "Bumps" of neural activity can slide along the manifold due to various factors, such as fast neural noise (14–24), firing rate adaptation (25, 26), or quenched disorder, which itself can be induced by the storage of multiple maps (27–30), by random noise (24, 31–36), by the encoding of additional variables (37–40), or by a consistent direction of motion in the synaptic weights (24, 41). All these considerations stretch the notion of continuous attractor networks, as they are in practice endowed with only a discrete number of true fixed points rather than a genuinely continuous attractive manifold. Yet, the precise noise limits at which such a notion ceases to be valid remain undefined. This poses a significant challenge to considering cANNs as effective models for cognitive maps with irregular place field statistics. The presence of such irregularity, indeed, implies pronounced quenched noise in synaptic weights within the network, leading to a small residual number of fixed points. Motivated by the recent experiments in bats, we have addressed this question through a detailed numerical study grounded in the recorded place field statistics, followed by a first, heuristic mathematical analysis.

We examine three dimensions of variability for place fields: their number per cell, their size, and their peak firing rate, following the distributions reported in Ref. (10), each parameterized by a single variable. Next, we encode this variability into the synaptic weights of a recurrent neural network through Hebbian learning, effectively introducing quenched noise. We then analyze networks dynamics at different locations in this 3D space, where standard cANNs would correspond to the zero-variability point (0,0,0), which exhibits a semi-continuous manifold of fixed points that approaches continuity as the number of cells $N \to \infty$. Our study reveals three distinct regions within this "phase diagram," characterized by specific dynamical properties. These regions are delimited by abrupt transitions in dynamical behavior, akin at least numerically to phase transitions. Importantly, we demonstrate that the coordinates defining the experimental recordings fall within the same dynamical region as standard cANNs, the only one where memory retrieval can effectively take place.

While a multifield multiscale neural code has been suggested to be advantageous in terms of decoding error (10), whether a disorderly arrangement of place fields could be the basis for a stable memory representation of an environment had remained unclear. Here, we show that highly irregular place fields can indeed be effectively stored and retrieved within a continuous quasi-attractor (CQA) neural network. We identify, however, specific boundaries of quenched disorder, for which we provide analytical heuristic estimates, that delimit the region of existence for such a CQA regime. We note the challenging nature of the analytical evaluation, due in part to the out-of-equilibrium character of the dynamical phenomena we find numerically. Conventional analysis methods based on the study of fixed points are not sufficient, and thus we hope to stimulate future endeavors aimed at achieving a more rigorous analytical understanding of these boundaries.

Results

Model definition

We first introduce a neural network model incorporating the activity statistics observed in bats by (10). The network includes N pyramidal cells, labeled $i = 1, \ldots, N$ and modeled as threshold-linear units (42). Recurrent connections among these N units are taken to be dense (as they are known to be in CA3, not in CA1), for simplicity all-to-all, and to be endowed with Hebbian plasticity, through which the network is assumed to have stored a representation of a tunnel of length L—and only of that tunnel.

With these assumptions, as a result of a learning phase not explicitly modeled here, and discretizing the tunnel of length L into spatial bins of width $W = L/S$ centered in $s_u = u \times W$, with $u = 1, \ldots, S$, the connectivity matrix between the neurons is given by the Hebbian covariance rule:

$$J_{ij} = \frac{1}{NS}\sum_{u=1}^{S}\left[\frac{\eta_i(s_u)}{\langle\eta\rangle} - 1\right]\left[\frac{\eta_j(s_u)}{\langle\eta\rangle} - 1\right], \tag{1}$$

where $\eta_i(s)$ is the reproduced recorded activity of cell i when the bat is at position s in the tunnel. Here, $\langle\eta\rangle$ denotes the average activity over all cells and all positions.

The distribution of activity $\{\eta_i(s)\}$ is chosen to replicate the statistics observed in bats flying in a long tunnel. The recorded place maps, of which we report two samples in Fig. 1A, were indeed shown to have an overall high variability as reported in the observed distributions, shown in Fig. 1B–F. We focus on reproducing the variability in the number of fields, reported in Fig. 1B, the one in the width, reported in Fig. 1C and D, the correlation between width and peak firing rate, reported in Fig. 1E, and the one of the maximal peak firing rate, reported in Fig. 1F. To model the observed sources of variability, we assume that (1) Each unit is taken to have, along the tunnel, a variable number $M \geq 1$ of place fields. We consider, in agreement with experimental data, that the distribution of M is exponential: $P(M) \propto \exp(-M/\zeta)$ where $\zeta > 0$ controls the average number of fields $\langle M\rangle$, see Fig. S1 for details. The case of a single field per neuron, i.e. $P(M) = \delta(M, 1)$, is formally associated with $\zeta = 0$. (2) Each field k is centered at a random location uniformly distributed along the tunnel and has a Gaussian shape of width $d_k = 2\sigma$, truncated at $\pm\sigma$, and peak rate p_k; with both parameters drawn from *log normal* distributions, which fit quite well the experimental data. In detail, for each k, we draw d_k such that $\ln(d_k)$ is normally distributed with mean μ_d and variance σ_d and then p_k such that $\ln(p_k)$ is normally distributed with mean $\tilde{\mu}_p(d_k) \equiv \mu_p + \gamma \ln(d_k/\langle d_k\rangle)$ and variance σ_p. A factor $\gamma > 0$ introduces a correlation between field widths and peak rates, with $\gamma = 0.5$ reproducing the observed correlation, as shown in Fig. 1E&M. We will see at the end of the analytical section estimating the quasi-circular boundary that taking into account this correlation is important. Furthermore, Burak and colleagues were recently able to show that precisely this degree of correlation would arise naturally in a model in which the place fields result from random Gaussian processes (43). This construction procedure for the fields is summarized in Fig. S2 and leads to firing rate profiles and distributions similar to those observed experimentally, as shown in Fig. 1G–N; For additional details, see SI Section 1. The network evolves under its recurrent connectivity in Eq. 1 only (no external input after the initial cue, and learning is taken to have been consolidated). We call $V_i(t)$ the activity of cell i at time t, during such recurrent dynamics. The discrete time evolution equation for neural activity reads

$$V_i(t+1) = e^{-1/\tau}V_i(t) + (1 - e^{-1/\tau})g[h_i(t)]^+, \tag{2}$$

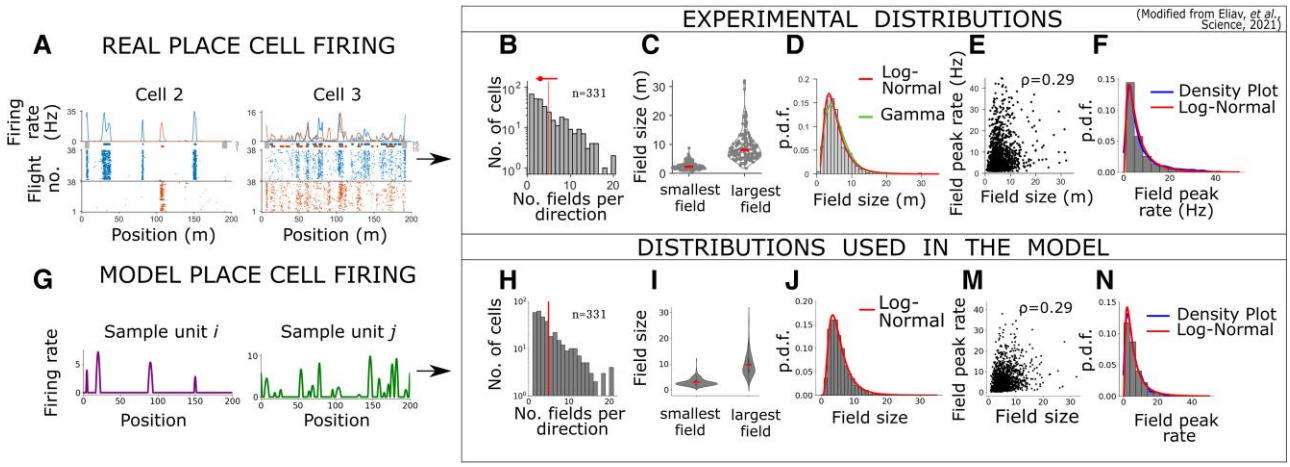

Fig. 1. Experimental and model place field profiles and distributions Top (A–E): Experimental results of (10), reproduced with permission; F) was obtained from the experimental data in E), kindly given to us by the authors. Bottom (G–N): analogous plots obtained with the field generating algorithm we design, introduced in the main text. A) Firing rate profiles of two neurons in CA1, different colors represent different flying directions (bottom: raster plots across multiple back and forth flights). B) Distribution of the number of place fields in one direction. The bar at 20 includes all values above 20, and the average number of fields is $\langle M \rangle = 4.9$. C) Distribution of the smallest and largest field sizes per neuron (including those with at least 2 fields). D) Distribution of fields sizes, the parameters of the log normal fit, in red, are $e^{\mu_d} = 4.8$ m, $\sigma_d = 0.575$. E) Scatter plots of field size versus peak firing rate for each field. The Spearman correlation coefficient is $\rho = 0.29$. F) Distribution of peak firing rates, the parameters of the log normal fit, in red, are $e^{\mu_p} = 4.7$ Hz, $\sigma_p = 0.884$. G) Two sample activity profiles of two different units, as obtained from our algorithm. We model only one flying direction. H–N same as B–F but for the field distributions sampled from a random realization of the field generating algorithm for $N = 331$ units (with parameters: $\zeta = 4.7$, $\mu_d = 1.57$, $\mu_p = 1.549$, $\sigma_d = 0.575$, $\sigma_p = 0.884$, $\gamma = 0.5$).

where $[h]^+ = \max(0, h)$, g is a fixed gain parameter and τ is the time scale for collective firing rate dynamics, in units of the discrete simulation time steps. In recurrent networks of spiking units, τ is, according to mean-field theory, related to the time needed for synaptic conductances to close (44), and can be smaller close to a discrete attractor state (45, 46). Although the case of continuous attractors is less clear and likely depends on their smoothness, it can be argued that τ should correspond to at most 10 ms of real time. In the simulations, we use $\tau \geq 9.5$ time steps, which implies that a time step corresponds to at most 1 ms of real time. Finally,

$$h_i(t) = \sum_{j \neq i} J_{ij} V_j(t) - T\left(\frac{1}{N}\sum_j V_j(t)\right) \tag{3}$$

is the input current to unit i at time t, relative to a threshold T incorporating fast activity-dependent inhibition. We choose $T(v) = 4\omega(v - v_0)^3$, where v_0 is the target mean activity, which we keep constant, and ω the strength of this inhibitory feedback, also a constant. See the Methods section for the setting of simulation parameters. Note that these dynamics are associated with an energy (Lyapunov) function, as described in SI Section 2.

Numerical results: three different regimes

To assess whether the network activity $\vec{V}(t)$ at time t, evolving without external inputs, reinstates the population vector $\vec{\eta}(s)$ encoding during learning a specific position s, we initialize it at some location s_0, i.e. $\vec{V}(t=0) = \vec{\eta}(s_0)$, acting as an initial cue, and then measure the cosine similarity

$$O(\vec{\eta}(s), \vec{V}(t)) = \frac{\sum_i^N \eta_i(s) \cdot V_i(t)}{\sqrt{\sum_i^N (\eta_i(s))^2 \cdot \sum_i^N (V_i(t))^2}}, \tag{4}$$

for all positions along the tunnel.

The set of overlaps $O(\vec{\eta}(s), \vec{V}(t))$ for all s defines the overlap profile at a given time t. If it is a bump-like profile peaked in s^*, we can say that the network is reactivating or retrieving the code for this position. Conversely, when the profile is not spatially tuned, network activity is not representing any position along the tunnel—it has left its attractive manifold. Monitoring this profile over time, we therefore characterize the spatial memory expressed in high-dimensional neural dynamics.

Three scenarios are observed, depending on the parameters:

- *Continuous quasi-attractor (CQA)*: the profile shows a bump-like shape, with width small compared to the tunnel length, at all times, Figs. 2A and B. The bump slides with time until a fixed point is reached (Fig. 2A).
- *Fragmented Manifold (FM)*: the bump deforms, gradually spreads out, see Figs. 2C and D. After some time, the bump relocalizes, as if teleported elsewhere, until it reaches a fixed point (Fig. 2D). Informally speaking, the neural activity transiently leaves the manifold.
- *Nonlocalized (NL)*: the bump rapidly grows in width and ceases to be localized (Fig. 2E) until a spatially noninformative fixed point is reached, spread over the whole environment (Fig. 2F).

By spanning wide ranges of values for the quenched parameters ζ, σ_p, σ_d, as depicted in Fig. 3A (with a sample realization in Fig. 3B), we locate the boundaries of regions associated to the CQA, FM, and NL regimes. To do so, we consider four estimators of the bump-like nature of the activity, see Methods:

- the percentage of vanished manifold, quantified as the percentage of initial positions from which the neural population activity leaves the manifold;
- the tangent overlap $\langle O_{\text{tang}} \rangle$, characterizing the alignment of the direction of instability of the fixed points with the direction of the presumed 1D manifold;
- the bump width at the fixed points, indicating whether the retrieved stable activity is localized on the manifold.
- the number of fixed points, indicating the number of different positions on the manifold encoded as fixed points.

These estimators vary with the parameters ζ, σ_p, σ_d.

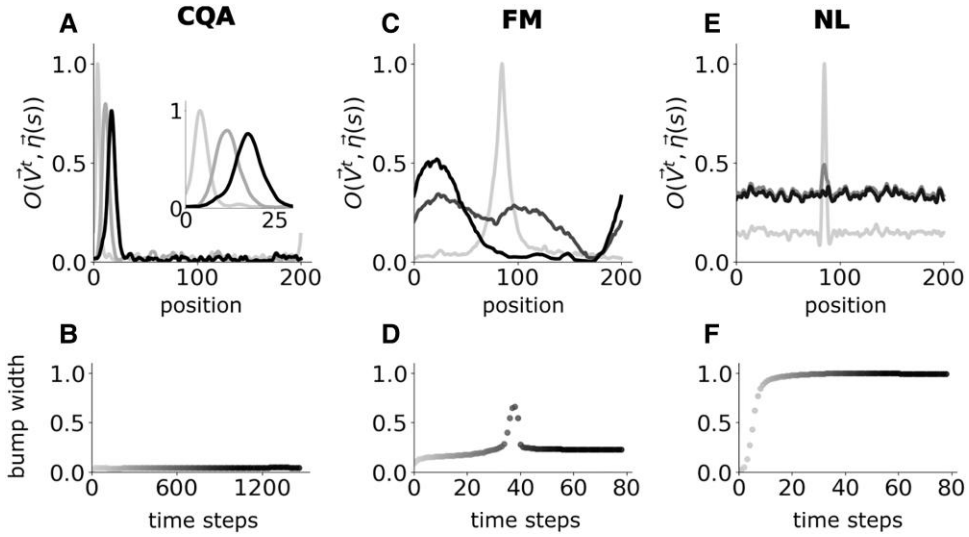

Fig. 2. The three dynamical regimes of the neural activity and their spatial correlates. Top row: overlap profiles $O(\eta(s), V(t))$ vs. position s at different steps t of the dynamics; selected steps t are indicated by gray levels ranging from light gray (initial condition) to black (fixed point). Bottom row: bump width (see Methods) of the overlap profile as a function of the number t of steps; same gray level code as in the top row. A, B) Continuous Quasi-Attractor regime: A) the bump slides on the manifold—the inset shows a zoom-in—B) maintaining a constant width at all times, and stops at a fixed point, which reflects the initial condition. Dark gray in A) corresponds to the profile at the 200th time step. C, D) Fragmented Manifold regime: C) the activity "jumps" outside the manifolds and re-enters at a different location elsewhere, accordingly D) the bump width transiently increases. Dark gray in C) corresponds to the profile at the 38th time step. E, F) Non Localized regime: E) the initial bump rapidly vanishes and the activity is not clearly related to any position in space, F) the smeared overlap is reflected in a bump width ≈ 1 . Dark gray in A) corresponds to the profile at the 30th time step. Parameters: $N = 8000$, $g = 2.5$, $\tau = 9.5$; A, B) $\zeta = 1$, $\sigma_d = \sigma_p = 0.4$; C, D) $\zeta = 1$, $\sigma_d = \sigma_p = 0.9$; E, F) $\zeta = 4.7$, $\sigma_d = \sigma_p = 0.2$.

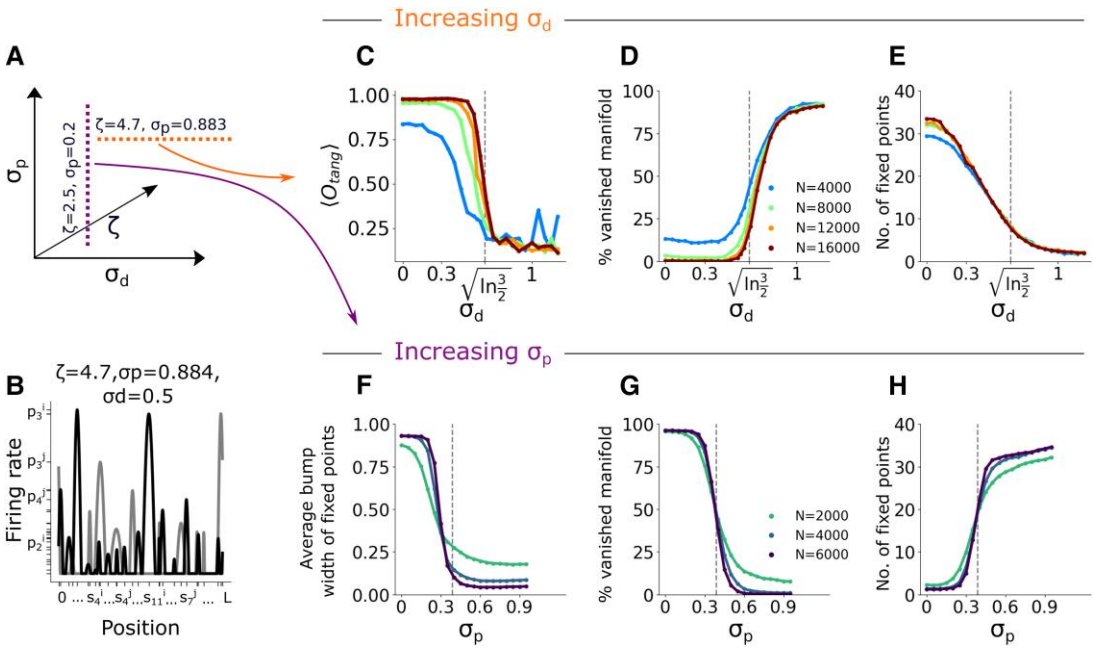

Fig. 3. Dynamical changes between regimes. A) The distribution of the fields is defined by three quenched order parameters, as described in the main text. The vertical and horizontal dashed lines represent the values explored in the subplots C, D) and F–H) respectively. B) Place fields profiles of two sample units (gray/black), as generated from our procedure, for the labeled value of ζ, σ_p, σ_d. C) $\langle O_{\text{tang}} \rangle$ profile, the curves indicate the 0.25 quantile (75% of the overall data lie above the line). D) Percentage of vanished manifold and E) Average number of fixed points for increasing σ_d when maintaining ζ and σ_p fixed at their experimental values, as labeled, for increasing N. F) Averaged bump width of the fixed points G) Percentage of vanished manifold and H) Average number of fixed points for increasing σ_p maintain ζ and σ_d fixed. The vertical lines in C–E) and F–H) correspond to the analytical predictions, $\sigma_d = \sqrt{\ln(3/2)}$ and σ_p extracted from the numerical solution of Eq. 10, respectively. For further comparison refer to Fig. S3; see Methods for details over the measurements and model parameters. Each data point is averaged over 20–50 different quenched realizations of the network, each probed with 50 different runs initialized from equidistant $\eta(s)$ along S.

In particular, we observe that when the dispersion σ_d of place field widths increases, at some point, as shown in Fig. 3C–E, the percentage of vanished manifold abruptly increases and the direction of instability of the fixed point also gets abruptly orthogonal to the direction of the manifold, while the number of fixed points decreases only smoothly. This large-σ_d effect is thus visible in the *dynamical behavior* towards fixed points. Further, we observe that when the dispersion σ_p of place field peak firing rates increases,

with $\zeta > 0$ and σ_d fixed, at some point, as shown in Fig. 3F–H, the stable states start to be localized on the manifold, the percentage of vanished manifold abruptly decreases and the number of fixed points abruptly increases. This effect of σ_p is thus visible in the fixed points themselves. The simultaneity and the abruptness of both these types of change are compatible with sharp changes of regime, reminiscent of phase transitions in physical systems. For more comparisons, refer to Fig. S3.

The results of an exhaustive exploration of parameter space are summarized in Fig. 4 in the form of empirical phase diagrams, locating the regions associated with the CQA, FM, and NL regimes. We note some general features of this summary of the simulation results, see also Fig. 3:

- The CQA regime is found in a central region of the (σ_d, σ_p) plane, when $\zeta > 0$.
- For σ_d above a critical value that appears to be almost independent of σ_p and ζ, the network abruptly transitions to the FM regime, becoming effectively unable to represent, based on memory, almost any position along the tunnel (Fig. 4A–C).
- When units have multiple fields, and both σ_d and σ_p are small, approximately within a circular boundary (which depends on ζ), population dynamics always delocalizes, and the network can be said to have entered, again abruptly, the NL regime (Fig. 4D–F).

The boundary between the CQA and FM regimes seems to be at a value of σ_d above, but close to the experimental value from CA1 recordings, $\sigma_d = 0.575$. We emphasize that its exact location is not sensitive to the precise form of the initial conditions, see Fig. S4, nor to the parameters regulating, in our network model, its dynamical evolution, such as the gain or the global inhibition term, nor to those related to the discretization of the tunnel length, see Figs. S5 and S6. Note the contrast between the sharp increase with σ_d in the proportion of starting positions from which activity is eventually teleported and the smooth decrease in the number of fixed points (see Fig. 3D and E), which in itself would not suggest that the network enters a distinct phase.

Analytical estimates

A rigorous analytical treatment of the model, estimating the dynamical evolution of the network, as a function of the connections, could provide solid explanations regarding the nature of the transitions between the regimes we described numerically. Yet, a rigorous approach is challenging, due to the nonindependent distribution of the interaction terms J_{ij} and, for the CQA-FM boundary, due to the nonstationarity of the phenomenon. Nevertheless, we can derive precise estimates of the boundaries, using signal-to-noise analyses, reported below.

Quasi-circular boundary of nonlocalization.

When units have multiple fields and limited variability in their width and peak rate, the network activity gets "smeared" over the entire length of the tunnel, as shown in Fig. 2E, Fig. 3F and in Fig. 4D–F. This indicates that the competition among the fields fails to produce a winning location. Intriguingly, adding variability in the distribution of width and peak rate promotes localization, as if noise allows the network to better differentiate among the different fields, restricting the retrieved activity to a fraction of the units. To derive the boundary we make three concatenated ansatzes:

1. that the "strength" of each field k of each unit is proportional to the product of its width and peak rate

$$s_k = d_k \cdot p_k; \tag{5}$$

2. that per each unit only one "strongest" field, say $k^* = 1$, effectively matters to create the signal while all others contribute as noise

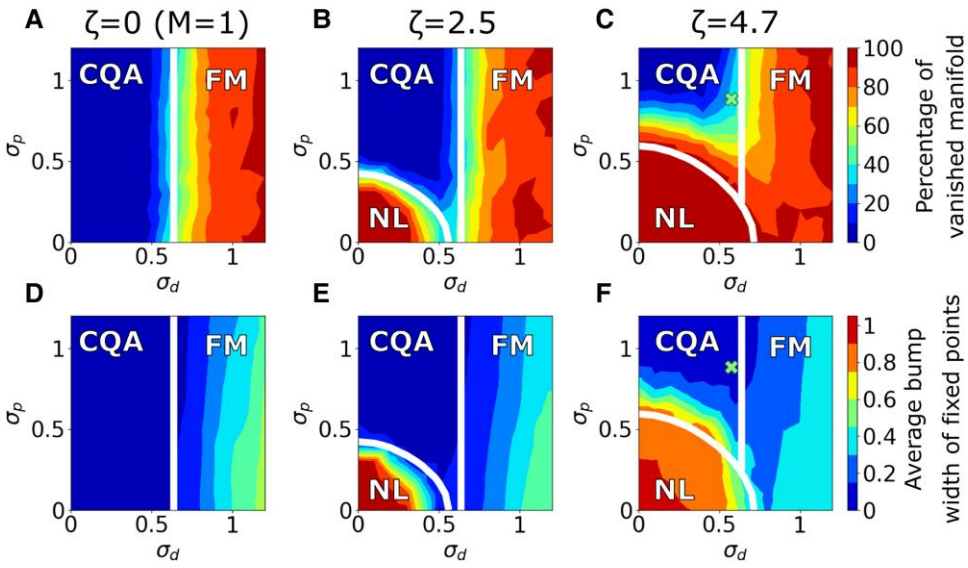

Fig. 4. Phase diagrams in the σ_p-σ_d plane. Both numerical and analytical results are shown for three $\sigma_p - \sigma_d$ sections at increasing average number of fields ($\zeta = 0$, i.e. $M \equiv 1$ for A–D); $\zeta = 2.85$, $\langle M \rangle \approx 3.4$ for B–E); $\zeta = 4.7$, $\langle M \rangle \approx 4.9$ akin to the experimentally observed value for C–F)). A–C) show the percentage of vanished manifold D–F) show the average bump width of the fixed points (see Methods for details about the measures). Green crosses in C–F) indicate the σ_p, σ_d values giving rise to distributions as the experimental results. White curves are derived analytically, as reported in the analytical estimates subsection. Parameters: plots where obtained by interpolation (with 13×13 data points). A data point is averaged over 3–5 different quenched realizations of the network, each probed with 50 different runs initialized from equidistant $\eta(s)$ along S, $N = 5{,}000$.

$$\mathcal{S}^{NL}_{ignal} \propto \langle d_1 \cdot p_1 \rangle \ , \tag{6}$$

$$\mathcal{N}^{NL}_{oise} \propto \left\langle \sum_{k>1} d_k \cdot p_k \right\rangle. \tag{7}$$

3. that localization can only occur when the overall signal is larger than the overall noise scaled by some quantity $C(\sigma_p, \sigma_d)$, i.e. that the boundary occurs when

$$\frac{\mathcal{S}^{NL}_{ignal}}{\mathcal{N}^{NL}_{oise}} = C(\sigma_p, \sigma_d). \tag{8}$$

The analytical steps to evaluate $\mathcal{S}^{NL}_{ignal}$ and $\mathcal{N}^{NL}_{oise}$ are reported in SI Section 4, and they relate, in practice, to the statistics of the population vectors $\eta(s)$, leading to a closed set of equations for units having the same number of M fields. We then perform the weighted average $\langle \rangle_M$ for both quantities at each specific ζ. We find heuristically that the quantity $C(\sigma_d, \sigma_p)$ can be approximated by

$$C(\sigma_d, \sigma_p) = \frac{\langle p_k \rangle}{A\sqrt{Var(p_k)}}. \tag{9}$$

The equation for the boundary, as derived in the SI Section 4 writes then as

$$\frac{\langle \mathcal{S}^{NL}_{ignal}(M) \rangle_M}{\langle \mathcal{N}^{NL}_{oise}(M) \rangle_M} = \frac{1}{A}\frac{\langle p_k \rangle}{\sqrt{Var(p_k)}} = \frac{1}{A\sqrt{\exp\left[\gamma^2\sigma_d^2 + \sigma_p^2\right] - 1}}, \tag{10}$$

Whichever value one chooses for the factor A, it can be verified numerically that this equation produces a quasi-circular boundary for $\gamma = 0.5$, see Fig. S7; and its effective radius increases with ζ, Fig. S8. The boundary, for $A = 3$, is reported as a white quasi-circular line in the subplots of Fig. 4 and as dashed vertical lines in Fig. 3F–H. Numerically, this solution is indeed close to the critical values estimated from the simulations.

As a validation of our analysis we performed simulations with other correlation values, $\gamma \neq 0.5$, between peaks and diameters. This leads to a different shape of the region of nonlocalization, whose boundary is still defined by Eq. 10, see e.g. Fig. S7 for $\gamma = 0$.

Vertical boundary of the Fragmented-Manifold regime

The location of the vertical boundary, as shown in Fig. 4A–C, appears to be independent of the value of ζ and nearly independent of σ_p. Based on this observation, we introduce a simplified model: each unit is characterized by a single field ($M_i = 1, \forall i$), the fields have purely Gaussian shapes and the different fields are uniformly distributed across the environment. As we consider $\sigma_p = 0$, i.e. the same peak rate $p = 1/\sqrt{2\pi}$ for each field, the only source of variability among the units lies in the widths of their fields, which follow a log normal distribution as in the realistic model reproducing experimental data. In Fig. 5A, we show a few sample activity profile of different units for a given σ_d for this reduced model. As reported in Fig. 5B–D, to be compared with Fig. 3C–E, simulations of this reduced model reveal a qualitatively and also quantitatively similar dynamical behavior as that expressed by the realistic model, giving us the possibility to study analytically the reduced version in order to get an understanding of the realistic one.

What makes the network activity suddenly "jump elsewhere" during the dynamics, for $\sigma_d > \sigma_d^c$? The cause must be the disparity in the width of the fields. When some of the widest fields happen to be centered far away from the current position s, but still get activated, they can start suppressing the narrower units centered around s, and the self-reinforcing process can lead to a jump. Critical to the self-reinforcing nature of the process is that the largest fields exert a disproportionate influence with respect to the narrow ones; this is because in the Hebbian plasticity rule (Eq. 1) the normalization of both pre- and post-synaptic factors is by the average activity level across units, $\langle \eta \rangle$. To find the boundary we make, again, a set of ansatzes:

1. given a unit, the average incoming signal from all other units, as a function of the distance d of their fields, is proportional to the average of the corresponding connectivity weight, and the noise square to their variance, i.e.

$$\mathcal{S}^{FM}_{ignal}(d) \propto \langle J(d) \rangle, \tag{11}$$

$$[\mathcal{N}^{FM}_{oise}(d)]^2 \propto \langle (J(d))^2 \rangle - (\langle J(d) \rangle)^2. \tag{12}$$

2. the activity bump can remain localized only if the incoming signal from the furthest unit is larger than the incoming noise scaled by some quantity $D(\sigma_d)$, namely assuming $d \to \infty$, the boundary corresponds to the values at which

$$\frac{\lim_{d\to\infty} \mathcal{S}^{FM}_{ignal}(d)}{\sqrt{\lim_{d\to\infty} [\mathcal{N}^{FM}_{oise}(d)]^2}} = D(\sigma_d). \tag{13}$$

The calculation of the above quantities is reported in SI Section 5, and it reflects, in practice, only the statistics of the connections J. We find, also in this case heuristically, that the quantity $D(\sigma_d)$ is actually a constant, namely

$$D(\sigma_d) \equiv D = 1. \tag{14}$$

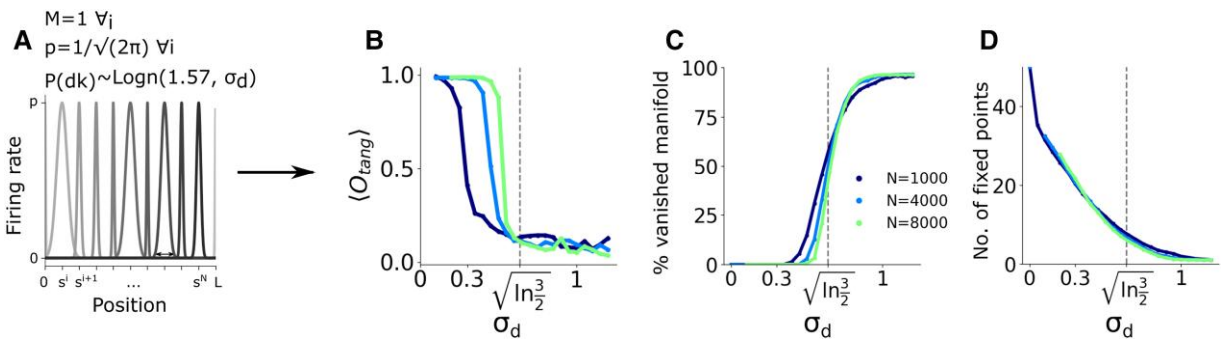

Fig. 5. Dynamical change between CQA and FM regimes, with $M \equiv 1$ and $\sigma_p = 0$. A) Single fields of 10 sample units obtained from the reduced model introduced in the text, with $\sigma_d = 0.87$ (See SI Section 1.4 for details). B–D) Same as Fig. 3 B–D).

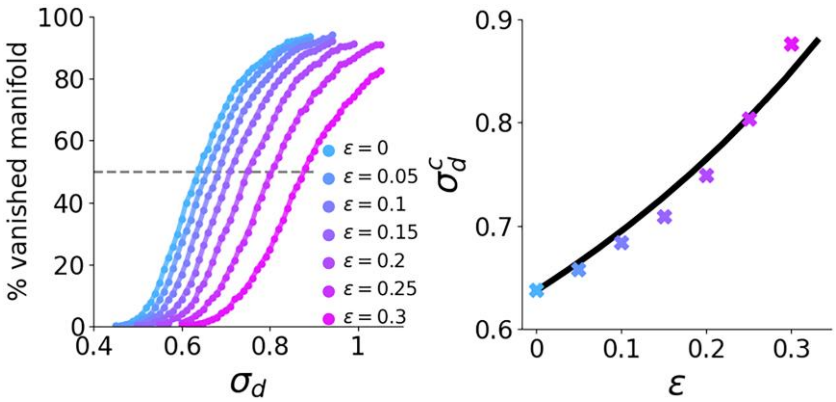

Fig. 6. Numerical and analytical results for the ϵ expansion of $\mathbf{J_{ij}}$. (Left) Percentage of vanished manifold for simulations with J as in Eq. 16 and ϵ as labeled. Horizontal dashed line indicates 50% of the curves, which we take as an estimate of $\sigma_d^c(\epsilon)$. Right) Crosses represent $\sigma_d^c(\epsilon)$ calculated from the numerical simulations on the left, black line corresponds to the analytical estimate introduced in the text (see SI Section 5). Parameters: $N = 4{,}000$. Each data point is averaged over 20–50 different quenched realizations of the network, each probed with 20 different runs initialized from equidistant $\eta(s)$ along S.

Setting the equality in 13 the value corresponding to the estimate for the boundary location is then:

$$\sigma_d^c = \sqrt{\ln(3/2)} \simeq 0.637. \tag{15}$$

This value is reported as a white vertical line in the subplots of Fig. 4 and as dashed vertical lines in Fig. 3C–E and Fig. 5B–D. Also in this case, numerically, this solution is indeed close to the critical value estimated from the simulations.

As a validation of our analysis we show that its predictive power extends beyond the standard Hebbian rule and is able to locate the boundary along the σ_d-axis for a class of plasticity rules. To do so we design an additional model, identical to the reduced model introduced above except for the couplings, given instead by

$$J_{ij} = \frac{1}{NS}\sum_{u=1}^{S}\left[\frac{\eta_i(s_u)}{\langle\eta_i\rangle^{\epsilon}\langle\eta\rangle^{1-\epsilon}} - 1\right]\left[\frac{\eta_j(s_u)}{\langle\eta_j\rangle^{\epsilon}\langle\eta\rangle^{1-\epsilon}} - 1\right], \tag{16}$$

where $0 \leq \epsilon \leq 1$. For $\epsilon = 0$, the original Hebbian rule in Eq. 1 is recovered. For $\epsilon \neq 0$, instead, the normalization of the coupling J_{ij} depends on the neurons i and j through their average activities $\langle\eta_i\rangle$ and $\langle\eta_j\rangle$. We run numerical analysis using this model and, as reported in Fig. 6A, we show that the critical level of noise characterizing the passage between the QCA and FM regime increases with ϵ. We consider small positive values for ϵ. The analytical calculation of the critical $\sigma_d^c(\epsilon)$ is reported in SI Section 5 and the result is indicated by a black line in Fig. 6B. We consider the agreement between this analytical estimate and the crosses in Fig. 6B, corresponding to the numerically estimated level of noise at which 50% of the manifold has disappeared, as a validation of our analysis.

Discussion

It has long been understood that storing multiple regular place maps within a single connectivity matrix creates quenched disorder, which roughens each continuous attractor—i.e. each map—without completely erasing it, up to a certain capacity limit (11, 21, 27, 29, 30, 47). Instead, the impact of quenched noise originating from irregularity, even within a single map, had never been explored, leaving unclear the extent to which the continuous attractor model is relevant for spatial memory, now that significant irregularity is reported, in particular in experiments conducted in large, semi-ecological environments. Inspired by the bat hippocampal recordings in a long tunnel, in Ref. (10), we have characterized numerically the dynamical behavior of a recurrent neural network in which connections self-organize, through Hebbian learning, from realistically irregular place fields along the tunnel. We identify three dynamical behaviors of the isolated network, not driven by external inputs, depending on the level of irregularity. In two of them, NL and FM, activity either delocalizes or dynamically violates, by jumping, the topology of the tunnel. Only in one, CQA, prevailing in a region of parameters which includes the irregularity reported in Ref. (10), the activity is localized at the fixed points as well as throughout the dynamical evolution. We have derived numerically a 3D phase diagram, sketched in Fig. 7, spanning the variability in the number, width and peak firing of the fields, which shows the regions expressing each of the three behaviors. Further, we have developed an initial analytical approximation of the boundaries between these regions.

Assessing the implications of our study for spatial memory in the hippocampus requires some qualifications. First, although inspired by recordings in CA1, the recurrent model we have considered is only, if at all, appropriate to the CA3 network, noted for its extensive collateral connectivity. One critical assumption, then, is that forthcoming recordings from CA3 will show qualitatively, if not quantitatively similar irregularity to those from CA1, supporting the applicability of the same model. We are also assuming that the details of the connectivity (which in the real CA3 is far from the all-to-all scheme considered in the model), the actual biophysics, the operation of inhibitory circuits, etc., do not alter much the scenario with the three dynamical regimes indicated by our simplified model. In order to obtain general insights into the dynamical properties expressed by the hippocampus, we also assume that the recordings we took under consideration in the model are representative of other mammalian species for which neural data on place fields are not available.

Second, the three dynamical behaviors are expressed, in the model, when the recurrent network is only initially driven by spatially selective external inputs and then is isolated. One of them, the CQA regime, implies effective spatial memory retrieval also when external inputs are initially incomplete, conflicting or noisy, and later subside or are suppressed relative to reverberations along the recurrent connections, which structure attractor dynamics. Indeed, in the CQA regime, although only a limited

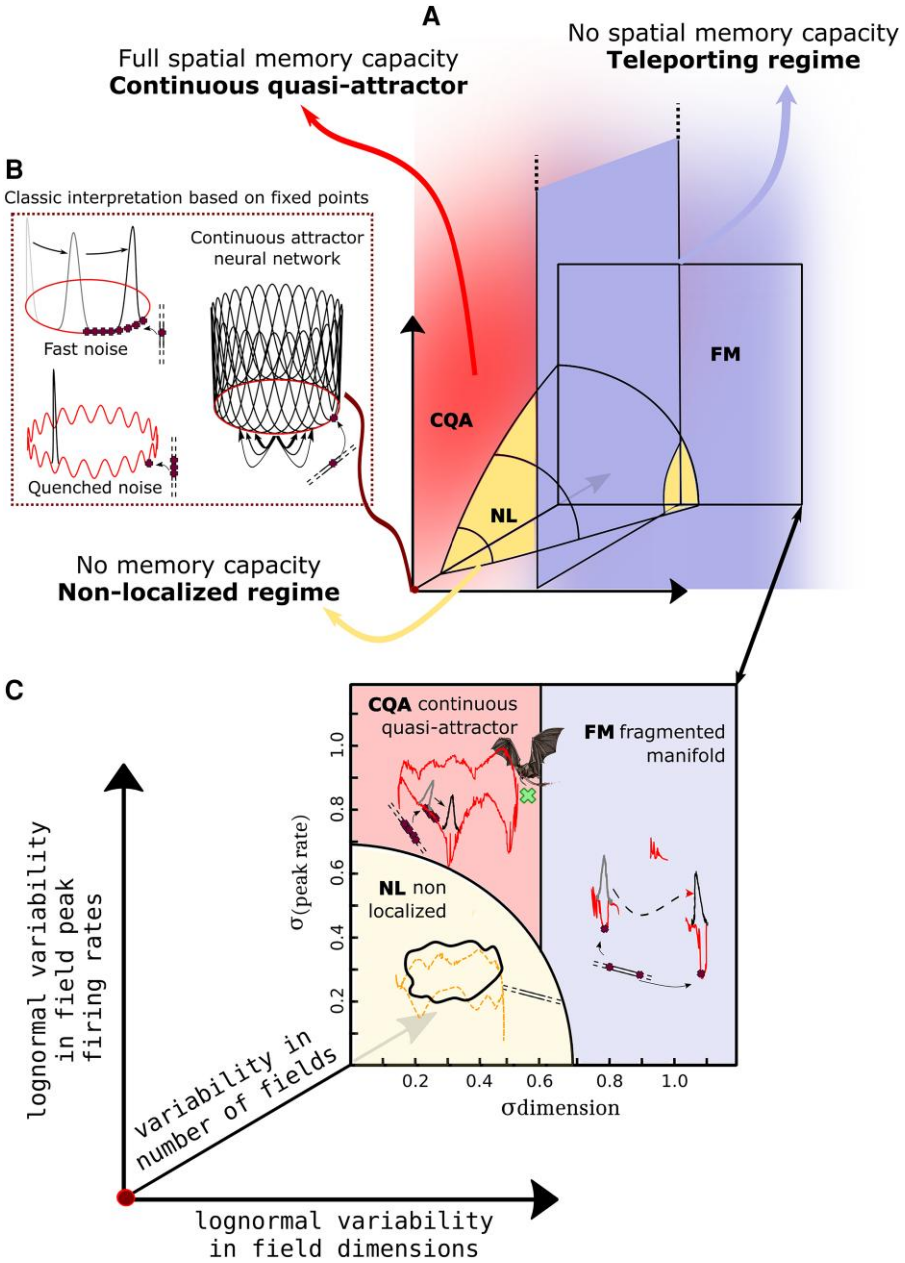

Fig. 7. Sketch of the 3D phase diagram and its insights into memory. A) Rough sketch of the 3D phase diagram, with axes labels as in C), schematizing the results reported in Fig. 4. Three regimes are separated by two boundaries. In the continuous quasi-attractor regime (red), the manifold of solutions, representing different locations along the tunnel, is strongly attractive to all directions but one: following an external cue network states are attracted to the manifold, and spatial memory is expressed as a bump of activity, perhaps sliding a bit along the nonattractive direction. In the teleportation or Fragmented Manifold regime (purple), the manifold effectively vanishes: external cues can only drive the dynamics towards a few residual fixed points, unable to represent space, as dynamics do not smoothly flow along the manifold. In the nonlocalized phase (yellow), the fixed points are not localized: hence no cue can retrieve spatial memory. B) Sketch of the idealized continuous attractor (right) which can only emerge, in the continuous quasi-attractive phase, at the origin, i.e. in the unrealistic condition of single, equal and regularly positioned fields. The parallel lines on the bottom right of the cANN symbolize a 1D track: each fixed point in the continuous attractor neural network (dark red cross) represents a memory of one position in the environment (matching dark red cross in the track). The presence of mild noise (its effects on the cANN are schematized in the two left sketches) downgrades the precision of spatial memory. C) Sketch of the three regimes in a plane corresponding to the distribution in the number of fields observed in the recording in bats (10). Memory retrieval is preserved as long as the dynamical flow is aligned with the manifold. This occurs in the red region, which includes also the firing rate and field width statistics observed in the recordings (green cross). Memory capability suddenly deteriorates beyond each boundary (determining the inability to retrieve anything, or possibly only a few locations). In C) and B), the closed curves (red/yellow) are intended to sketch the energy of the quasi-attractive continuous manifold. Gray–black bumps indicate the overlap profile which one can calculate with the $\{\eta(s)\}$ at each step of the dynamics. The yellow dashed manifold, instead, represent a phantom manifold which the dynamics cannot reach.

number of fixed points are present, the entire trajectories that link these fixed points exhibit quasi-attractive behavior, meaning they are unstable solely along the direction of the trajectory itself. Therefore, we can expect mild external inputs to suffice to keep neural activity localized at many arbitrary locations along the quasi-attractive trajectory; or at worst, if they are too mild, to let it flow to a nearby fixed point, which would come to represent a segment of the trajectory. In rodents, but not in bats, a partial

alternation between externally driven and attractor-driven dynamics may be paced by the Theta rhythm (48). In the NL and FM regimes, instead, either neural activity is not localized in the isolated network, or the manifold is broken into fragments, with jumps between them, which wash away positional information; and thus we expect spatial memory to be poor. In this context, the closeness of the parameter values corresponding to experimental observations in bats (green cross in Fig. 7) to the boundary separating the CQA and FM regimes seems to suggest that the operation point of the network tries to be at a safe distance from the NL boundary, which expands in large environments and would imply a disabled spatial memory. Alternatively, one might speculate that the closeness of the operation point to the FM boundary suggest that the network implements a trade-off between the variability intrinsic to a random process and the need for effective spatial memories.

In addition, following this reasoning, we expect in the near total absence of external inputs—during sleep or quiet wakefulness—the actual network operating in the CQA regime to *replay* only those fixed points that exist, and the close surrounding regions in their basin of attraction—the segments. Only segments of the tunnel adjacent to the fixed points—those more stably stored in memory—might be reached during spontaneous, rambling dynamics. The phenomenon of a segmented replay of very large environments has in fact been reported in the analysis of CA1 resting state recordings (49).

The analytical estimates presented here, although preliminary, yields some initial insights into the neural mechanisms underlying the storage of individual irregular maps and illustrate some computational constraints. With respect to the NL-CQA boundary, our analysis highlights the effect of a dominant field associated to each unit. Such a field always exist, given nonzero variability, but when the average dominance decreases below a critical value—the quasi-circular boundary—suddenly it is unable to lead to localized activity. Turning to the CQA-FM boundary, our simulations and analysis indicate that the "teleportation" is a dynamical effect due to fluctuations beyond the mean-field description. When, during the initial widening of the activity bump, a few units are activated with very wide fields, as often occurs with large values of the field width variability parameter σ_d, they can reinforce each other irrespective of the original position of the bump, maintained by units with narrow fields that give now a diminishing contribution. The latter are eventually suppressed, and the bump is repositioned at the location that best matches the wide fields of the units winning the competition. Apparently, such self-reinforcing amplification, with respect to the position signal, of the noise due to the units with wide fields is guaranteed to occur somewhere, in a system with many units, whenever σ_d exceeds the critical level σ_d^c. While this effect warrants further investigation through analyses of out-of-equilibrium dynamics, and it remains to be examined whether it generalizes to maps in higher dimension, e.g. in 2D, it is remarkable that our analytical estimate for σ_d^c is a pure number, $\sigma_d^c \simeq 0.637$, independent of any parameter. With all the necessary caveats, including those mentioned above, this implies that we can extrapolate a maximum length of a tunnel, or of a general 1D environment, that can be stored in memory by the recurrent network. The variability in the number of fields (ζ) is in fact expected to scale linearly with the length of the tunnel, or in general the size of the environment, whether in CA1 or in CA3. As a result, the NL-CQA boundary expands away from the origin and for any given level of peak rate variability σ_p (and correlation parameter γ) there is a value ζ where this boundary meets the CQA-FM boundary. This would imply, if σ_p is set e.g. by biophysical constraints, the disappearance of the viable CQA region and hence a maximum size of the environment that can be stored in memory. Indeed, the relation

$$\ln(3/2) + \sigma_p^2 = (r_\sigma^c(\zeta))^2 \tag{17}$$

defines when the two boundaries converge and the "acceptable" range for σ_d diminishes to zero. Using parameters from CA1 recordings in (10), this yields $\langle M \rangle \simeq 30$, see SI Section 4.1 for details, corresponding to $L \simeq 1.4$ km, indicating a surprisingly low limit for the length of a memorable tunnel.

The validity of this result remains to be verified, particularly in relation to applying the network model to infer conclusions beyond the specific conditions in which it was developed. First and foremost whether the model is really relevant for the CA3 network should be assessed on the basis of measurements of neural activity and its variability in the CA3 region, and also in light of quantitative estimates of the connectivity in CA3, in this or that species (50), which could be incorporated in an extended model. Second, the analysis should be applied to the case of 2D environments, and perhaps to higher dimensions and nonstandard geometries, which are relevant for different experimental settings than that of bats flying in a long narrow tunnel; this should be straightforward, if maintaining a simple mathematically clean structure, that ideally corresponds to the empty, featureless environment we have assumed in our analysis. Experimental research is however moving towards analyzing neural activity and memory behavior in more ecologically plausible environments, in which objects, landmarks, and other features are known to "attract" place fields (51, 52), and the effects of their presence on the different dynamical regimes we have studied here, and whether there is still a capacity limit on the size of the environment that can be stored in a recurrent network, remain an avenue for future research.

Methods

Assessment of the quality of the attractor

We consider the following estimators:

Percentage of vanished manifold

For a given realization of the network we run ≥ 50 neural dynamics, each starting from a "position" s_0, i.e. $\vec{V}(t=0) = \vec{\eta}(s_0)$, spanning regularly the tunnel. We locate, at each dynamical step t, the center s_c^t of the bump from the overlap profile (Eq. 4). If the bump has moved between steps t and $t+1$, say, to $s_c^{t+1} > s_c^t$, we focus on the overlap profile $O(\vec{\eta}(s), V(t+1))$ for $s \in [s_c^t, s_c^{t+1}]$. If this quantity varies monotonically with s or undergoes nonmonotonic changes only on a short sub-interval of less than $5L/S$, we consider that the dynamics has stayed within the neural manifold; otherwise, the dynamics is considered to have jumped outside the manifold. Note that choosing the particular value $5L/S$ does not really affect the result, as shown in Fig. S5C. We repeat this procedure for every time step t up to the convergence to the fixed point. The "percentage of vanished manifold" is defined as the percentage of different runs across various initial conditions and realizations undergoing at least one jump.

Tangent overlap $\langle \mathbf{O_{tang}} \rangle$

Consider a given realization of the network. For each fixed point of the dynamics $\vec{V}^{FP} = \vec{V}(s_c)$, we first estimate the center of the bump s_c maximizing the overlap profile. We then run new dynamics, starting from an initial condition $\vec{V}(t=0) = \vec{\eta}(\tilde{s})$ where $\tilde{s}$ is close to, but distinct from s_c. As the dynamics converge towards $\vec{V}(s_c)$, in order to obtain a good approximation of $\vec{V}(s_c \pm 1)$ we identify

the intermediate step t at which the overlap between $\vec{V}(t)$ and $\vec{\eta}(s_c \pm 1)$ is maximal and take an average (similar results are obtained by taking the last value $\vec{V}(t)$ having maximal overlap with $\vec{\eta}(s_c \pm 1)$ before the maximal overlap shifts to the position of the fixed point s_c). The direction of the neural manifold around the fixed point is then defined as $\vec{D} = \vec{V}(s_c \pm 1) - \vec{V}(s_c)$ or equally as $\vec{D} = \vec{V}(s_c + 1) - \vec{V}(s_c - 1)$.

In addition, we calculate the eigenvector $\vec{E}_{\min}$ corresponding to the smallest eigenvalue of the Hessian matrix (SI Eq. (6), restricted to active neurons) at the fixed point. The overlap between $\vec{E}_{\min}$ and $\vec{D}$ (restricted to active neurons) gives a measure of the alignment between the direction of instability of the fixed point and the putative manifold. This overlaps varies between 1 (the two vectors are aligned) and 0 (they are orthogonal). This quantity is then averaged over all fixed points and realizations of the network, with the result denoted by $\langle O_{tang} \rangle$.

Bump width

Given a configuration $\vec{V}(t)$ of size N we define for all discretized positions s the overlap $O_s \equiv O(\vec{\eta}(s), \vec{V}(t))$ as in 4. We thus obtain the overlap profile: a vector $\vec{O}$ of entries between 0–1 of size S, describing how similar is $V(t)$ to the different configurations of activity along the tunnel during the learning phase. We compute the dispersion of the center of mass of the overlap profile, keeping in mind periodic boundary conditions as follows: (a) we set all values smaller than 0.2 to zero to remove noise; (b) we change variables to polar coordinates and calculate the cosine

$$co = \frac{\sum_s^S O_d \cdot \cos\left(\frac{2\pi s}{S}\right)}{\sum_{s=1}^S O_s} \tag{18}$$

and sine

$$si = \frac{\sum_{s=1}^S O_s \cdot \sin\left(\frac{2\pi s}{S}\right)}{\sum_{s=1}^S O_s} \tag{19}$$

coefficients of the overlap profile; (c) we calculate the angle that represents the orientation of the center of mass using the computed cosine and sine coefficients as

$$\vartheta = \arccos\left(\frac{co}{\sqrt{co^2 + si^2}}\right) \tag{20}$$

and if the si is negative we take $\vartheta = 2\pi - \vartheta$; (d) we calculate the position of the center of mass along the length of the profile based on the angle as $cm = S\vartheta/2\pi$; (e) we find the difference $\vec{d}$ between each discretized position index and the one corresponding to the center of mass, keeping in mind periodic boundary conditions; (f) we finally calculate the dispersion as this mean square difference weighted by the corresponding overlap

$$\text{Bump width} = \sqrt{\frac{\sum_s O_s d_s^2}{\frac{S^2}{12} \sum_s O_s}}. \tag{21}$$

Note that we normalize the dispersion by dividing by the sum of the overlaps and scaling by an "angular momentum" factor of $\frac{1}{12}$. Bump widths close to 0 correspond to localized configurations of activity, while widths close to 1 indicate that the activity is spread over the manifold.

Number of fixed points

For a given realization of the network we run ≥50 neural dynamics, each starting from a "position" s_0, i.e. $\vec{V}(t = 0) = \vec{\eta}(s_0)$, with s_0 spanning regularly the tunnel. We let each dynamics to convergence to the fixed point, evaluate the overlap profile of the fixed point with $\{\eta(s)\}$ and estimate the location corresponding to the maximum of the overlap profile. The number of fixed points is the average across simulations of the number of different locations in the tunnel corresponding to a fixed point.

Simulations: parameters and implementation

In all simulations, unless otherwise specified, we have used the numerical values $\mu_d = 1.570$ and $\mu_p = 1.549$ extracted from the experimental distributions (10), which correspond to a typical field width of $\exp(\mu_d) = 4.7$ m along the $L = 200$ m long tunnel, and to a typical peak rate of $\exp(\mu_p) = 4.8$ Hz. We then discretize the length of the tunnel into $S = 1{,}000$ bins. The gain g is set to 17, the threshold constant ω to 300, and the target mean activity υ_0 to be equal to the average of the learned profiles $\langle \eta \rangle$; these three values were chosen to be in a (rather broad) regime where neural activity does not diverge. We show in Fig. S5 that the particular choice (like the values of L and S) does not influence the dynamical results we report. The neural time scale $\tau > 9.5$ was sometimes "increased" (which means we used shorter time steps) to facilitate reaching fixed points (in Fig. 2, $\tau \equiv 9.5$). All simulations where left to run until they reached the fixed points, i.e. when $V(t+1) - V(t) < 10^{-8}$.

Acknowledgments

We are grateful to Nachum Ulanovsky, Tamir Eliav, and their colleagues for extensively sharing and discussing their findings, as well as to other participants in the "M-Gate" collaboration. Kwang Il Ryom helped with numerics, and recently Yoram Burak clarified a point about their model of random spatial codes. The authors thank the anonymous reviewers for their valuable suggestions.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

The bulk of this work was funded by the EU Marie Skłodowska-Curie Training Network 765549 M-Gate with contributions from the Human Frontier Science Program RGP0057/2016 and the Italian MUR PRIN grant 20174TPEFJ "TRIPS." F.S. is supported by the QBio Junior Research Chair program of the QBio initiative of ENS-PSL and PariSanté Campus. R.M. is supported by the project MEMNET ANR-22-CE16-0005.

Author Contributions

F.S. and A.T. conceived the study, and F.S. designed, ran, and refined multiple layers of simulations. F.S., R.M., and A.T. made repeated efforts at interpreting the results and describing them analytically, and all co-authors jointly wrote the article.

Preprints

A preprint of this article is published at https://doi.org/10.1101/2023.08.16.553619.

Data Availability

The code used to obtain the results presented in this article is available at https://osf.io/qkajx/.

References

1 O'Keefe J. 1976. Place units in the hippocampus of the freely moving rat. *Exp Neurol*. 51(1):78–109.
2 Fenton AA, *et al.* 2008. Unmasking the CA1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. *J Neurosci*. 28(44):11250–11262.
3 Henriksen EJ, *et al.* 2010. Spatial representation along the proximodistal axis of CA1. *Neuron*. 68(1):127–137.
4 Lee JS, Briguglio JJ, Cohen JD, Romani S, Lee AK. 2020. The statistical structure of the hippocampal code for space as a function of time, context, and value. *Cell*. 183(3):620–635.
5 Park EH, Dvorak D, Fenton AA. 2011. Ensemble place codes in hippocampus: CA1, CA3, and dentate gyrus place cells have multiple place fields in large environments. *PLoS One*. 6(7):e22349.
6 Rich PD, Liaw HP, Lee AK. 2014. Large environments reveal the statistical structure governing hippocampal representations. *Science*. 345(6198):814–817.
7 Jung MW, McNaughton BL. 1993. Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus*. 3(2):165–182.
8 Leutgeb JK, Leutgeb S, Moser M-B, Moser EI. 2007. Pattern separation in the dentate gyrus and ca3 of the hippocampus. *Science*. 315(5814):961–966.
9 Muller RU, Kubie JL, Ranck JB. 1987. Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. *J Neurosci*. 7(7):1935–1950.
10 Eliav T, *et al.* 2021. Multiscale representation of very large environments in the hippocampus of flying bats. *Science*. 372(6545): eabg4020.
11 Battaglia FP, Treves A. 1998. Attractor neural networks storing multiple space representations: a model for hippocampal place fields. *Phys Rev E*. 58(6):7738.
12 Samsonovich A, McNaughton BL. 1997. Path integration and cognitive mapping in a continuous attractor neural network model. *J Neurosci*. 17(15):5900–5920.
13 Tsodyks M, Sejnowski T. 1995. Associative memory and hippocampal place cells. *Int J Neural Syst*. 6:81–86.
14 Amari S-I. 1977. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol Cybern*. 27(2):77–87.
15 Ben-Yishai R, Bar-Or RL, Sompolinsky H. 1995. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A*. 92(9): 3844–3848.
16 Burak Y, Fiete IR. 2012. Fundamental limits on persistent activity in networks of noisy neurons. *Proc Natl Acad Sci U S A*. 109(43): 17645–17650.
17 Camperi M, Wang X-J. 1998. A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *J Comput Neurosci*. 5(4):383–405.
18 Compte A, Brunel N, Goldman-Rakic PS, Wang X-J. 2000. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex*. 10(9): 910–923.
19 Fung CCA, Wong KYM, Wu S. 2010. A moving bump in a continuous manifold: a comprehensive study of the tracking dynamics of continuous attractor neural networks. *Neural Comput*. 22(3): 752–792.
20 Lee DD, Reis BY, Seung HS, Tank DW. 1997. Nonlinear network models of the oculomotor integrator. *Computational neuroscience: Trends in Research*,371–377.
21 Monasson R, Rosay S. 2013. Crosstalk and transitions between multiple spatial maps in an attractor neural network model of the hippocampus: phase diagram. *Phys Rev E*. 87(6):062813.
22 Seung HS. 1996. How the brain keeps the eyes still. *Proc Natl Acad Sci U S A*. 93(23):13339–13344.
23 Wu S, Hamaguchi K, Amari SI. 2008. Dynamics and computation of continuous attractors. *Neural Comput*. 20(4):994–1025.
24 Zhang K. 1996. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci*. 16(6):2112–2126.
25 Chu T., Ji Z., Zuo J., Mi Y., Zhang W. H., *et al.* 2024. Firing rate adaptation affords place cell theta sweeps, phase precession and procession. *Elife 12*. RP87055.
26 Treves A. 2004. Computational constraints between retrieving the past and predicting the future, and the CA3-CA1 differentiation. *Hippocampus*. 14(5):539–556.
27 Cerasti E, Treves A. 2013. The spatial representations acquired in CA3 by self-organizing recurrent connections. *Front Cell Neurosci*. 7:112.
28 Hopfield JJ. 2010. Neurodynamics of mental exploration. *Proc Natl Acad Sci U S A*. 107(4):1648–1653.
29 Monasson R, Rosay S. 2014. Crosstalk and transitions between multiple spatial maps in an attractor neural network model of the hippocampus: collective motion of the activity. *Phys Rev E*. 89(3):032803.
30 Papp G, Witter MP, Treves A. 2007. The CA3 network as a memory store for spatial representations. *Learn Mem*. 14(11):732–744.
31 Hamaguchi K, Hatchett JPL, Okada M. 2006. Analytic solution of neural network with disordered lateral inhibition. *Phys Rev E*. 73(5):051104.
32 Itskov V, Hansel D, Tsodyks M. 2011. Short-term facilitation may stabilize parametric working memory trace. *Front Comput Neurosci*. 5:40.
33 Renart A, Song P, Wang X-J. 2003. Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron*. 38(3):473–485.
34 Stringer SM, Rolls ET, Trappenberg TP, De Araujo IET. 2002. Self-organizing continuous attractor networks and path integration: two-dimensional models of place cells. *Netw Comput Neural Syst*. 13(4):429–446.
35 Stringer SM, Trappenberg TP, Rolls ET, De Araujo IET. 2002. Self-organizing continuous attractor networks and path integration: one-dimensional models of head direction cells. *Netw Comput Neural Syst*. 13(2):217–242.
36 Zhong W, Lu Z, Schwab DJ, Murugan A. 2020. Nonequilibrium statistical mechanics of continuous attractors. *Neural Comput*. 32(6):1033–1068.
37 Roudi Y, Latham PE. 2007. A balanced memory network. *PLoS Comput Biol*. 3(9):e141.
38 Roudi Y, Treves A. 2004. An associative network with spatially organized connectivity. *J Stat Mech*. 2004(07):P07010.
39 Roudi Y, Treves A. 2006. Localized activity profiles and storage capacity of rate-based autoassociative networks. *Phys Rev E*. 73(6):061904.
40 Treves A. 2003. Computational constraints that may have favoured the lamination of sensory cortex. *J Comput Neurosci*. 14(3): 271–282.
41 Spalla D., Cornacchia I.M., Treves A. 2021. Continuous attractors for dynamic memories. *Elife 10*. e69499.
42 Treves A. 1990. Threshold-linear formal neurons in autoassociative nets. *J Phys A Math Gen*. 23(12):2631.
43 Mainali N, da Silveira RA, Burak Y. 2024. Universal statistics of hippocampal place fields across species and dimensionalities. *bioRxiv*.
44 Treves A. 1993. Mean-field analysis of neuronal spike dynamics. *Netw Comput Neural Syst*. 4(3):259.

45 Battaglia FP, Treves A. 1998. Stable and rapid recurrent processing in realistic autoassociative memories. *Neural Comput*. 10(2): 431–450.
46 Tsodyks MV, Sejnowski T. 1995. Rapid state switching in balanced cortical network models. *Netw Comput Neural Syst*. 6(2):111.
47 Monasson R, Rosay S. 2015. Transitions between spatial attractors in place-cell models. *Phys Rev Lett*. 115(9):098101.
48 Jezek K, Henriksen EJ, Treves A, Moser EI, Moser M-B. 2011. Theta-paced flickering between place-cell maps in the hippocampus. *Nature*. 478(7368):246–249.
49 Eliav T, *et al*. 2022. Fragmented replay of very large environments in the hippocampus of bats. *FENS Conference Poster*, Paris (9–13 July 2022).
50 Amaral DG, Ishizuka N, Claiborne B. 1990. Chapter neurons, numbers and the hippocampal network. *Prog Brain Res*. 83:1–11.
51 Boccara CN, Nardin M, Stella F, O'Neill J, Csicsvari J. 2019. The entorhinal cognitive map is attracted to goals. *Science*. 363(6434): 1443–1447.
52 Hollup SA, Molden S, Donnett JG, Moser M-B, Moser EI. 2001. Accumulation of hippocampal place fields at the goal location in an annular Watermaze task. *J Neurosci*. 21(5):1635–1644.