

How Informative Are Spatial CA3 Representations Established by the Dentate Gyrus?

Erika Cerasti^{1*}, Alessandro Treves^{1,2}

1 SISSA, Cognitive Neuroscience Sector, Trieste, Italy, **2** Kavli Institute for Systems Neuroscience and Centre for the Biology of Memory, NTNU, Trondheim, Norway

Abstract

In the mammalian hippocampus, the dentate gyrus (DG) is characterized by sparse and powerful unidirectional projections to CA3 pyramidal cells, the so-called mossy fibers. Mossy fiber synapses appear to duplicate, in terms of the information they convey, what CA3 cells already receive from entorhinal cortex layer II cells, which project both to the dentate gyrus and to CA3. Computational models of episodic memory have hypothesized that the function of the mossy fibers is to enforce a new, well-separated pattern of activity onto CA3 cells, to represent a new memory, prevailing over the interference produced by the traces of older memories already stored on CA3 recurrent collateral connections. Can this hypothesis apply also to spatial representations, as described by recent neurophysiological recordings in rats? To address this issue quantitatively, we estimate the amount of information DG can impart on a new CA3 pattern of spatial activity, using both mathematical analysis and computer simulations of a simplified model. We confirm that, also in the spatial case, the observed sparse connectivity and level of activity are most appropriate for driving memory storage—and not to initiate retrieval. Surprisingly, the model also indicates that even when DG codes just for space, much of the information it passes on to CA3 acquires a non-spatial and episodic character, akin to that of a random number generator. It is suggested that further hippocampal processing is required to make full spatial use of DG inputs.

Citation: Cerasti E, Treves A (2010) How Informative Are Spatial CA3 Representations Established by the Dentate Gyrus? PLoS Comput Biol 6(4): e1000759. doi:10.1371/journal.pcbi.1000759

Editor: Lyle J. Graham, Université Paris Descartes, Centre National de la Recherche Scientifique, France

Received: July 14, 2009; **Accepted:** March 24, 2010; **Published:** April 29, 2010

Copyright: © 2010 Cerasti, Treves. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partially supported by the EU Spacebrain grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: cerasti@sisssa.it

Introduction

The hippocampus presents the same organization across mammals, and distinct ones in reptiles and in birds. A most prominent and intriguing feature of the mammalian hippocampus is the dentate gyrus (DG). As reviewed in [2], the dentate gyrus is positioned as a sort of intermediate station in the information flow between the entorhinal cortex and the CA3 region of the hippocampus proper. Since CA3 receives also direct, *perforant path* connections from entorhinal cortex, the DG inputs to CA3, called *mossy fibers*, appear to essentially duplicate the information that CA3 can already receive directly from the source. What may be the function of such a duplication?

Within the view that the *recurrent* CA3 network operates as an autoassociative memory [3], [4], it has been suggested that the mossy fibers (MF) inputs are those that drive the storage of new representations, whereas the perforant path (PP) inputs relay the cue that initiates the retrieval of a previously stored representation, through *attractor dynamics*, due largely to recurrent connections (RC). Such a proposal is supported by a mathematical model which allows a rough estimate of the amount of information, in bits, that different inputs may impart to a new CA3 representation [5]. That model, however, is formulated in the Marr [6] framework of *discrete* memory states, each of which is represented by a single activity configuration or firing pattern.

Conversely, the prediction that MF inputs may be important for storage and not for retrieval has received tentative experimental

support from experiments with spatial tasks, either the Morris water maze [7] or a dry maze [8]. Two-dimensional spatial representations, to be compatible with the attractor dynamics scenario, require a multiplicity of memory states, which approximate a 2D continuous manifold, isomorphic to the spatial environment to be represented. Moreover, there has to be of course a multiplicity of manifolds, to represent distinct environments with complete remapping from one to the other [9]. Attractor dynamics then occurs along the dimensions locally orthogonal to each manifold, as in the simplified “multi-chart” model [10], [11], whereas tangentially one expects marginal stability, allowing for small signals related to the movement of the animal, reflecting changing sensory cues as well as path integration, to displace a “bump” of activity on the manifold, as appropriate [10], [12].

Although the notion of a really continuous attractor manifold appears as a limit case, which can only be approximated by a network of finite size [13], [14], [15], [16], even the limit case raises the issue of how a 2D attractor manifold can be established. In the rodent hippocampus, the above theoretical suggestion and experimental evidence point at a dominant role of the dentate gyrus, but it has remained unclear how the dentate gyrus, with its MF projections to CA3, can drive the establishment not just of a discrete pattern of activity, as envisaged by [5], but of an entire spatial representation, in its full 2D glory. This paper reports the analysis of a simplified mathematical model aimed at addressing this issue in a quantitative, information theoretical fashion.

Author Summary

The CA3 region at the core of the hippocampus, a structure crucial to memory formation, presents one striking anatomical feature. Its neurons receive many thousands of weak inputs from other sources, but only a few tens of very strong inputs from the neurons in the directly preceding region, the dentate gyrus. It had been proposed that such sparse connectivity helps the dentate gyrus to drive CA3 activity during the storage of new memories, but why it needed to be so sparse had remained unclear. Recent recordings of neuronal activity in the dentate gyrus [1] show the firing maps of granule cells of rodents engaged in exploration: the few cells active in a given environment, about 3% of the total, present multiple firing fields. Following these findings, we could now construct a network model that addresses the question quantitatively. Both mathematical analysis and computer simulations of the model show that, while the memory system would function also otherwise, connections as sparse as those observed make it function optimally, in terms of the bits of information new memories contain. Much of this information, we show, is encoded however in a difficult format, suggesting that other regions of the hippocampus, until now with no clear role, may contribute to decode it.

Such an analysis would have been difficult even only a few years ago, before the experimental discoveries that largely clarified, in the rodent, the nature of the spatial representations in the regions that feed into CA3. First, roughly half of the entorhinal PP inputs, those coming from layer II of the *medial* portion of entorhinal cortex, were found to be often in the form of *grid cells*, i.e. units that are activated when the animal is in one of multiple regions, arranged on a regular triangular grid [17]. Second, the sparse activity earlier described in DG granule cells [18] was found to be concentrated on cells also with multiple fields, but irregularly arranged in the environment [1]. These discoveries can now inform a simplified mathematical model, which would have earlier been based on ill-defined assumptions. Third, over the last decade *neurogenesis* in the adult dentate gyrus has been established as a quantitatively constrained but still significant phenomenon, stimulating novel ideas about its functional role [19]. The first and third of these phenomena will be considered in extended versions of our model, to be analysed elsewhere; here, we focus on the role of the multiple DG place fields in establishing novel CA3 representations.

A simplified mathematical model

The complete model considers the firing rate of a CA3 pyramidal cell, η_i , to be determined by the firing rates $\{\eta\}$ of other cells in CA3, which influence it through RC connections; by the firing rates $\{\beta\}$ of DG granule cells, which feed into it through MF connections; by the firing rates $\{\varphi\}$ of layer II pyramidal cells in entorhinal cortex (medial and lateral), which project to CA3 through PP axons; and by various feedforward and feedback inhibitory units. A most important simplification is that the fine temporal dynamics, e.g. on theta and gamma time scales, is neglected altogether, so that with “firing rate” we mean an average over a time of order the theta period, a hundred *msec* or so. Very recent evidence indicates, in fact, that only one of two competing spatial representations tends to be active in CA3 within each theta period [Jezeq et al, SfN abstract, 2009]. Information coding over shorter time scales would require anyway a more complex analysis, which is left to future refinements of the model.

For the different systems of connections, we assume the existence of anatomical synapses between any two cells to be represented by fixed binary matrices $\{c^{PP}\}, \{c^{MF}\}, \{c^{RC}\}$ taking 0 or 1 values, whereas the efficacy of those synapses to be described by matrices $\{J^{PP}\}, \{J^{MF}\}, \{J^{RC}\}$. Since they have been argued to have a minor influence on coding properties and storage capacity [20], consistent with the diffuse spatial firing of inhibitory interneurons [21], the effect of inhibition and of the current threshold for activating a cell are summarized into a subtractive term, of which we denote with \bar{T} the mean value across CA3 cells, and with $\tilde{\delta}_i$ the deviation from the mean for a particular cell i .

Assuming finally a simple threshold-linear activation function [22] for the relation between the activating current and the output firing rate, we write

$$\eta_i(\vec{x}) = g \left[\sum_l c_{il}^{PP} J_{il}^{PP} \varphi_l(\vec{x}) + \sum_j c_{ij}^{MF} J_{ij}^{MF} \beta_j(\vec{x}) + \sum_k c_{ik}^{RC} J_{ik}^{RC} \eta_k(\vec{x}) + \tilde{\delta}_i - \bar{T} \right]^+ \quad (1)$$

where $[]^+$ indicates taking the sum inside the brackets if positive in value, and zero if negative, and g is a gain factor. The firing rates of the various populations are all assumed to depend on the position \vec{x} of the animal, and the notation is chosen to minimize differences with our previous analyses of other components of the hippocampal system (e.g. [22], [23]).

The storage of a new representation

When the animal is exposed to a new environment, we make the drastic modelling assumption that the new CA3 representation be driven solely by MF inputs, while PP and RC inputs provide interfering information, reflecting the storage of previous representations on those synaptic systems, i.e., noise. Such “noise” can in fact act as an undesired signal and bring about the retrieval of a previous, “wrong” representation, an interesting process which is not however analysed here. We reabsorb the mean of such noise into the mean of the “threshold+inhibition” term \bar{T} and similarly for the deviation from the mean. We use the same symbols for the new variables incorporating RC and PP interference, but removing in both cases the “ \sim ” sign, thus writing

$$\eta_i(\vec{x}) = \left[\sum_j c_{ij}^{MF} J_{ij}^{MF} \beta_j(\vec{x}) + \delta_i - T \right]^+ \quad (2)$$

where the gain has been set to $g=1$, without loss of generality, by an appropriate choice of the units in which to measure $\{c^{MF}\}, \{J^{MF}\}$ (pure numbers) and δ_i, T (s^{-1}).

As for the MF inputs, we consider a couple of simplified models that capture the essential finding by [1], of the irregularly arranged multiple fields, as well as the observed low activity level of DG granule cells [24], while retaining the mathematical simplicity that favours an analytical treatment. We thus assume that only a randomly selected fraction p_{DG} of the granule cells are active in a new environment, of size A , and that those units are active in a variable number Q_j of locations, with Q_j drawn from a distribution with mean q . In model A, which we take as our reference, the distribution is taken to be Poisson (the data reported by Leutgeb et al [1] are fit very well by a Poisson distribution with $q=1.7$, but their sampling is limited). In model B, which we use as a variant, the distribution is taken to be exponential (this better

describes the results of the simulations in [25], though that simple model may well be inappropriate). Therefore, in either model, the firing rate $\beta_j(\vec{x})$ of DG unit j is a combination of Q_j gaussian ‘‘bumps’’, or fields, of equal effective size $(\sigma_f)^2$ and equal height β_0 , centered at random points \vec{x}_{jk} in the new environment

$$\beta_j(\vec{x}) = \sum_{k=0}^{Q_j} \beta_0 e^{-\frac{(\vec{x}-\vec{x}_{jk})^2}{2\sigma_f^2}}. \quad (3)$$

The informative inputs driving the firing of a CA3 pyramidal cell, during storage of a new representation, result therefore from a combination of three distributions, in the model. The first, Poisson but close to normal, determines the MF connectivity, that is how it is that each CA3 unit receives only a few tens of connections out of $N_{DG} \simeq 10^6$ granule cells (in the rat), whereby $\{c_{ij}^{MF}\} = 0,1$ with $P(c_{ij}^{MF} = 1) = C_{MF}/N_{DG} \equiv c^{MF}$. The second, Poisson, determines which of the DG units presynaptic to a CA3 unit is active in the new environment, with $P(\text{unit } j \text{ is active}) = p_{DG}$. The third, either Poisson or exponential (and see model C below), determines how many fields an active DG unit has in the new environment. Note that in the rat $C_{MF} \simeq 46$ [26] whereas $p_{DG} \approx 0.02 - 0.05$, even when considering presumed newborn neurons [24]. As a result, the total number of active DG units presynaptic to a given CA3 unit, $p_{DG}C_{MF} \equiv \alpha$, is of order one, $\alpha \sim 1-2$, so that the second Poisson distribution effectively dominates over the first, and the number of active MF impinging on a CA3 unit can approximately be taken to be itself a Poisson variable with mean α . As a qualification to such an approximation, one has to consider that different CA3 pyramidal cells, among the $N_{CA3} \simeq 3 \times 10^5$ present in the rat (on each side), occasionally receive inputs from the *same* active DG granule cells, but rarely, as $N_{DG} \simeq 10^6$, hence the pool of active units $p_{DG}N_{DG}$ is only one order of magnitude smaller than the population of receiving units N_{CA3} .

In a further simplification, we consider the MF synaptic weights to be uniform in value, $J_{ij}^{MF} \equiv J$. This assumption, like those of equal height and width of the DG firing fields, is convenient for the analytical treatment but not necessary for the simulations. It will be relaxed later, in the computer simulations addressing the effect of MF synaptic plasticity.

The new representation is therefore taken to be established by an informative signal coming from the dentate gyrus

$$\bar{\eta}_i(\vec{x}) = J \sum_j c_{ij}^{MF} \beta_j(\vec{x}) - T \quad (4)$$

modulated, independently for each CA3 unit, by a noise term δ_i , reflecting recurrent and perforant path inputs as well as other sources of variability, and which we take to be normally distributed with zero mean and standard deviation δ .

The position \vec{x} of the animal determines the firing $\{\beta\}$ of DG units, which in turn determine the probability distribution for the firing rate of any given CA3 pyramidal unit

$$P(\eta_i|\vec{x}) = \delta(\eta_i) \Phi\left(-\frac{\bar{\eta}_i(\vec{x})}{\delta}\right) + \Theta(\eta_i) e^{-\frac{(\eta_i - \bar{\eta}_i(\vec{x}))^2}{2\delta^2}} / (\sqrt{2\pi}\delta)$$

where

$$\Phi(\rho(\vec{x})) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\rho(\vec{x})} e^{-t^2/2} dt$$

$$\psi(\vec{x} - \vec{x}_{jk}) \equiv \left(\beta_0 e^{-\frac{(\vec{x} - \vec{x}_{jk})^2}{2\sigma_f^2}} \right).$$

is the integral of the gaussian noise up to given signal-to-noise ratio

$$\rho(\vec{x}) \equiv \bar{\eta}(\vec{x})/\delta,$$

and $\Theta(\eta)$ is Heaviside’s function vanishing for negative values of its argument. The first term, multiplying Dirac’s $\delta(\eta_i)$, expresses the fact that negative activation values result in zero firing rates, rather than negative rates.

Note that the resulting sparsity, i.e. how many of the CA3 units end up firing significantly at each position, which is a main factor affecting memory storage [22], is determined by the threshold T , once the other parameters have been set. The approach taken here is to assume that the system requires the new representation to be sparse and regulates the threshold accordingly. We therefore set the sparsity parameter $a_{CA3} = 0.1$, in broad agreement with experimental data [15], and adjust T (as shown, for the mathematical analysis, in the third section of the Methods).

The distribution of fields per DG unit is given in model A by the Poisson form

$$P_A(Q) = \frac{q^Q}{Q!} e^{-q}$$

in model B by the exponential form

$$P_B(Q) = \frac{1}{1+q} \left(\frac{q}{1+q} \right)^Q$$

and we also consider, as another variant, model C, where each DG unit has one and only one field

$$P_C(Q) = \delta_{1Q}.$$

Assessing spatial information content

In the model, spatial position \vec{x} is represented by CA3 units, whose activity is informed about position by the activity of DG units. The activity of each DG unit is determined independently of others by its place fields

$$P(\{\beta(\vec{x})\}) = \prod_j P(\beta_j(\vec{x}))$$

with

$$P(\beta_j(\vec{x})) = (1 - p_{DG}) \delta(\beta_j) + p_{DG} \times$$

$$\sum_{Q_j=0}^{\infty} P_{A,B \text{ or } C}(Q_j) \delta\left(\beta_j - \sum_{k=1}^{Q_j} \psi(\vec{x} - \vec{x}_{jk})\right)$$

where each contributing field is a gaussian bump

The Mutual Information $I(\vec{x}, \{\eta_i\})$ quantifies the efficiency with which CA3 activity codes for position, on average, as

$$\langle I(\vec{x}, \{\eta_i\}) \rangle = \langle H_1(\{\eta_i\}) \rangle - \langle \langle H_2(\{\eta_i\} | \vec{x}) \rangle_{\vec{x}} \rangle \quad (5)$$

where the outer brackets $\langle \cdot \rangle$ indicate that the average is not just over the noise δ , as usual in the estimation of mutual information, but also, in our case, over the *quenched*, i.e. constant but unknown values of the microscopic quantities c_{ij} , the connectivity matrix, Q_j , the number of fields per active unit, and \vec{x}_{jk} , their centers. For given values of the quenched variables, the total entropy H_1 and the (average) equivocation H_2 are defined as

$$H_1(\{\eta_i\}) = - \int \prod_i d\eta_i P(\{\eta_i\}) \log(P(\{\eta_i\})) \quad (6)$$

$$\langle H_2(\{\eta_i\} | \vec{x}) \rangle_{\vec{x}} = - \int (d\vec{x}/A) \prod_i d\eta_i P(\{\eta_i\} | \vec{x}) \times \log(P(\{\eta_i\} | \vec{x})) \quad (7)$$

where A is the area of the given environment; the logs are intended in base 2, to yield information values in bits.

The estimation of the mutual information can be approached analytically directly from these formulas, using the replica trick (see [27]), as shown by [28] and [29], and briefly described in the first section of the Methods. As in those two studies, however, here too we are only able to complete the derivation in the limit of low signal-to-noise, or more precisely of limited variation, across space, of the signal-to-noise around its mean, that is $\langle (\rho_i(\vec{x}) - \langle \rho_i(\vec{x}) \rangle_{\vec{x}})^2 \rangle_{\vec{x}} \rightarrow 0$. In this case we obtain, to first order in $N \equiv N_{CA3}$, an expression that can be shown to be equivalent to

$$\begin{aligned} \langle I(\vec{x}, \{\eta_i\}) \rangle &= \frac{N}{\ln 2} \left\langle \int \frac{d\vec{x}}{A} \left\{ \Phi(-\rho_i(\vec{x})) \ln \Phi(-\rho_i(\vec{x})) - \Phi(-\rho_i(\vec{x})) \right. \right. \\ &\quad \left. \left. \ln \int \frac{d\vec{x}'}{A} \Phi(-\rho_i(\vec{x}')) \right\} + \int \frac{d\vec{x}}{A} \frac{d\vec{x}'}{A} \left\{ \frac{\Phi(\rho_i(\vec{x}))}{2} \right. \right. \\ &\quad \left. \left. [\rho_i(\vec{x}) - \rho_i(\vec{x}')]^2 + [\rho_i(\vec{x}) - \rho_i(\vec{x}')]\sigma(\rho_i(\vec{x})) \right\} \right. \\ &\quad \left. - \int \frac{d\vec{x}}{A} \frac{d\vec{x}'}{A} \frac{d\vec{x}''}{A} \frac{\Phi(\rho_i(\vec{x}))}{4} [\rho_i(\vec{x}') - \rho_i(\vec{x}'')]^2 \right\rangle \end{aligned} \quad (8)$$

where we use the notation $\sigma(\rho) = (1/\sqrt{2\pi}) \exp -\rho^2/2$ (cp. [29], Eqs.17, 45).

Being limited to the first order in N , the expression above can be obtained in a straightforward manner by directly expanding the logarithms, in the large noise limit $\delta \rightarrow \infty$, in the simpler formula quantifying the information conveyed by a single CA3 unit

$$\begin{aligned} \langle I(\vec{x}, \{\eta_i\}) \rangle &= \frac{1}{\ln 2} \left\langle \int \frac{d\vec{x}}{A} \left\{ \Phi(-\rho_i(\vec{x})) \ln \Phi(-\rho_i(\vec{x})) - \Phi(-\rho_i(\vec{x})) \right. \right. \\ &\quad \left. \left. \ln \int \frac{d\vec{x}'}{A} \Phi(-\rho_i(\vec{x}')) \right\} - \int \frac{d\vec{x}}{A} \int \frac{d\eta}{\sqrt{2\pi}\delta} e^{-\frac{(\eta - \bar{\eta}(x))^2}{2\delta^2}} \right. \\ &\quad \left. \ln \left[\int \frac{d\vec{y}}{A} e^{-\frac{\bar{\eta}^2(x) - \bar{\eta}^2(y) - 2\eta(\bar{\eta}(x) - \bar{\eta}(y))}{2\delta^2}} \right] \right\rangle \end{aligned} \quad (9)$$

This single-unit formula cannot quantify the higher-order contributions in N , which decrease the information conveyed by a population in which some of the units inevitably convey some of the same information. The replica derivation, instead, in principle would allow one to take into proper account such correlated selectivity, which ultimately results in the information conveyed by large CA3 populations not scaling up linearly with N , and saturating instead once enough CA3 units have been sampled, as shown in related models by [28], [29]. In our case however the calculation of e.g. the second order terms in N is further complicated by the fact that different CA3 units receive inputs coming from partially overlapping subsets of DG units. This may cause saturation at a lower level, once all DG units have been effectively sampled. The interested reader can follow the derivation sketched in the Methods.

Having to take, in any case, the large noise limit implies that the resulting formula is not really applicable to neuronally plausible values of the parameters, but only to the uninteresting case in which DG units impart very little information onto CA3 units. Therefore we use only the single-unit formula, and resort to computer simulations to assess the effects of correlated DG inputs. The second and third sections of the Methods indicate how to obtain numerical results by evaluating the expression in Eq. 9.

Computer simulations can be used to estimate the information present in samples of CA3 units of arbitrary size, and at arbitrary levels of noise, but at the price of an indirect *decoding* procedure. A decoding step is required because the dimensionality of the space spanned by the CA3 activity $\{\eta_i\}$ is too high. It increases in fact exponentially with the number N of neurons sampled, as M^N , where M is the number of possible responses of each neuron. The decoding method we use, described in the fourth section of the Methods, leads to two different types of information estimates, based on either the full or reduced localization matrix. The difference between the two, and between them and the analytical estimate, is illustrated under Results and further discussed at the end of the paper.

Results

The essential mechanism described by the model is very simple, as illustrated in Fig. 1. CA3 units which happen to receive a few DG overlapping fields combine them in a resulting field of their own, that can survive thresholding. The devil is in the quantitative details: what proportion of CA3 cells express place fields, how large are the fields, and how strong are the fields compared with the noise, all factors that determine the information contained in the spatial representation. Note that a given CA3 unit can express multiple fields.

It is convenient to discuss such quantitative details with reference to a standard set of parameters. Our model of reference is a network of DG units with fields represented by Gaussian-like functions of space, with the number of fields per each DG units given by a Poisson distribution with mean value q , and parameters as specified in Table 1.

In general, the stronger the mean DG input, the more it dominates over the noise, and also the higher the threshold has to be set in CA3 to make the pattern of activity as sparse as required, by fixing $a_{CA3} = 0.1$. To control for the trivial advantage of a higher signal-to-noise, we perform comparisons in which it is kept fixed, by adjusting e.g. the MF synaptic strength J .

Multiple input cells vs. multiple fields per cell

The first parameter we considered is q , the average number of fields for each DG unit, in light of the recent finding that DG units

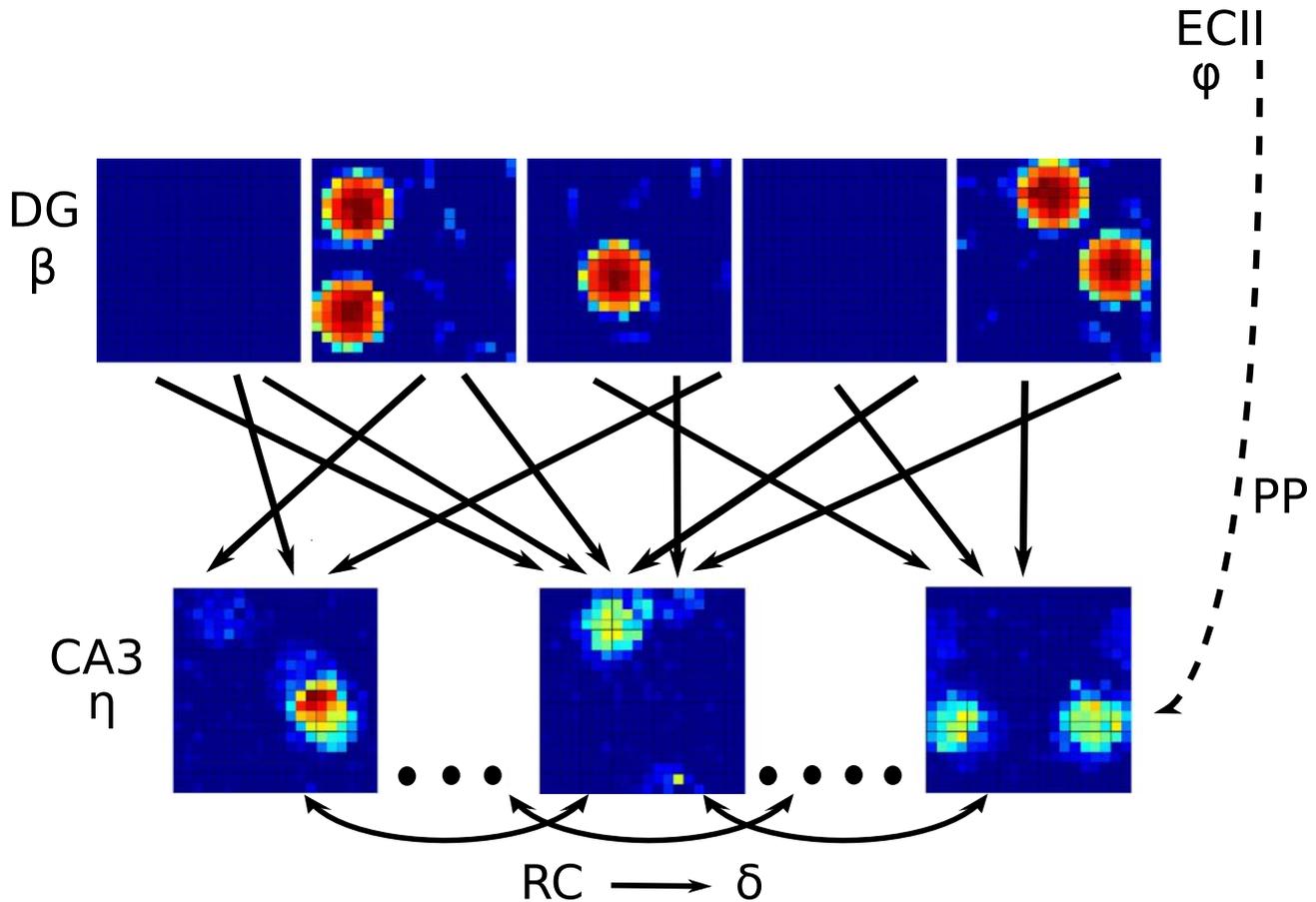


Figure 1. Network scheme. The DG-CA3 system indicating examples of the fields attributed to DG units and of those resulting in CA3 units, the connectivity between the two populations, and the noise δ that replaces, in the model, also the effect of recurrent connections in CA3. doi:10.1371/journal.pcbi.1000759.g001

active in a restricted environment are more likely to have multiple fields than CA3 units, and much more often than expected, given their weak probability of being active [1]. We wondered whether receiving multiple fields from the same input units would be advantageous for CA3, and if so whether there is an optimal q value. We therefore estimated the mutual information when q varies and μ , the total mean number of DG fields that each CA3 cell receives as input, is kept fixed, by varying C_{MF} correspondingly.

As shown in Fig. 2, varying q in this manner makes very little difference in the bits conveyed by each CA3 cell. This figure reports the results of computer simulations, that illustrate also the dependence of the mutual information on N_{CA3} , the number of cells sampled. The dependence is sub-linear, but rather smooth, with significant fluctuations from sample-to-sample which are largely averaged out in the graph. The different lines correspond to different distributions of the input DG fields among active DG cells projecting to CA3, that is different combinations of values for q and $C_{MF} = \mu / (q p_{DG})$, with μ kept constant; these different distributions do not affect much the information in the representation.

The analytical estimate of the information per CA3 unit confirms that there is no dependence on q (Fig. 2, inset). This is not a trivial result, as it would be if only the parameter μ entered the analytical expression. Instead, the second section of the Methods shows that the parameters C_m of the m -field decomposition depend separately on q and $\alpha \equiv p_{DG} C_{MF}$, so the fact that the two separate dependencies almost cancel out in a single dependence on their product, μ , is remarkable. Moreover, such analytical estimate of the information conveyed by one unit does not match the first datapoints, for $N_{CA3} = 1$, extracted from the computer simulation; it is not higher, as might have been expected considering that the simulation requires an additional information losing decoding step, but lower, by over a factor of 2. The finding that the analytical estimate differs from, and is in fact much lower than, the slope parameter extracted from the simulations, after the decoding

Table 1. Parameters.

parameter	symbol	standard value
probability a DG unit is active in one environment	p_{DG}	0.033
number of DG inputs to a CA3 unit	C_{MF}	50
mean number of fields per active DG unit	q	1.7
mean number of fields activating a CA3 unit	μ	$C_{MF} p_{DG} q = 2.833$
strength of MF inputs	J	1, otherwise $2.833/\mu$
noise affecting CA3 activity	δ	1 (in units in which $\beta_0 = 2.02$)
sparsity of CA3 activity	a_{CA3}	0.1

Values used in the standard version of the model.

doi:10.1371/journal.pcbi.1000759.t001

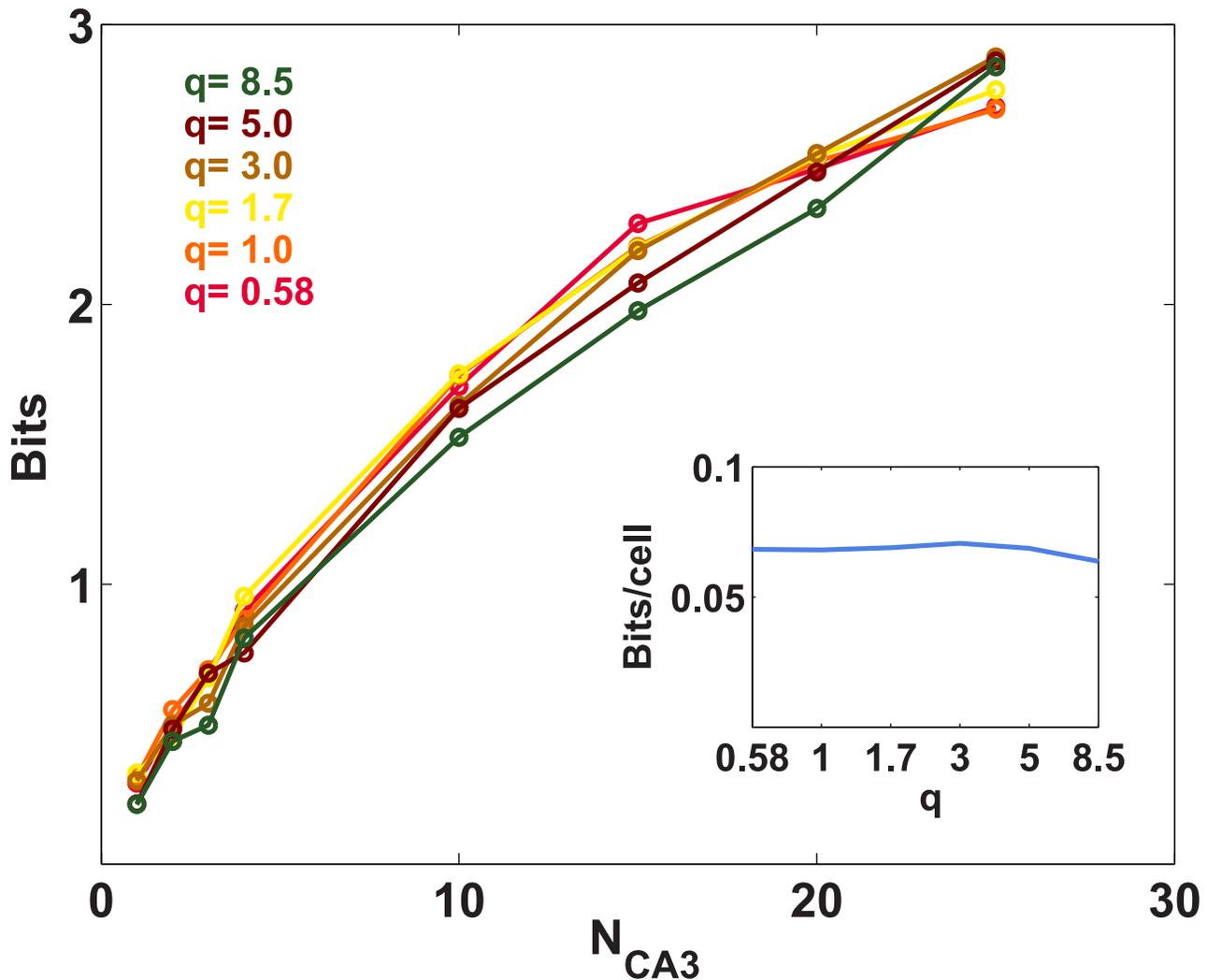


Figure 2. The exact multiplicity of fields in DG units is irrelevant. Information about position plotted versus the number of CA3 units, N_{CA3} from which it is decoded, with the mean number of fields in the input to each CA3 unit constant at the value $\mu = 2.833$. Different lines correspond to a different mean number of fields per DG input units, balanced by different mean number of input units per CA3 unit. Inset: analytical estimate of the information per CA3 unit, from numerically integrating Eq. 9. doi:10.1371/journal.pcbi.1000759.g002

step, is further discussed below. Despite their incongruity in absolute values, neither the estimate derived from the simulations nor the analytical estimate have separate dependencies on q and α , as shown in Fig. 2.

More MF connections, but weaker

Motivated by the striking sparsity of MF connections, compared to the thousands of RC and PP synaptic connections impinging on CA3 cells in the rat, we have then tested the effect of changing C_{MF} without changing q . In order to vary the mean number of DG units that project to a single CA3 unit, while keeping constant the total mean input strength, assumed to be an independent biophysically constrained parameter, we varied inversely to C_{MF} the synaptic strength parameter J . As shown in Fig. 3, the information presents a maximum at some intermediate value $C_{MF} \approx 20 - 30$, which is observed both in simulations and in the analytical estimate, despite the fact that again they differ by more than a factor of two.

Again we find that the analytical estimate differs from, and is in fact much lower than, the slope parameter extracted from the simulations, after the decoding step. Both measures, however, show

that the standard model is not indifferent to how sparse are the MF connections. If they are very sparse, most CA3 units receive no inputs from active DG units, and the competition induced by the sparsity constraint tends to be won, at any point in space, by those few CA3 units that are receiving input from just one active DG unit. The resulting mapping is effectively one-to-one, unit-to-unit, and this is not optimal information-wise, because too few CA3 units are active – many of them in fact have multiple fields (Fig. 4, right), reflecting the multiple fields of their “parent” units in DG. As C_{MF} increases (with a corresponding decrease in MF synaptic weight), the units that win the competition tend to be those that summate inputs from two or more concurrently active DG units. The mapping ceases to be one-to-one, and this increases the amount of information, up to a point. When C_{MF} is large enough that CA3 units begin to sample more effectively DG activity, those that win the competition tend to be the “happy few” that happen to summate several active DG inputs, and this tends to occur at only one place in the environment. As a result, an ever smaller fraction of CA3 units have place fields, and those tend to have just one, often very irregular, as shown in Fig. 4, right. From

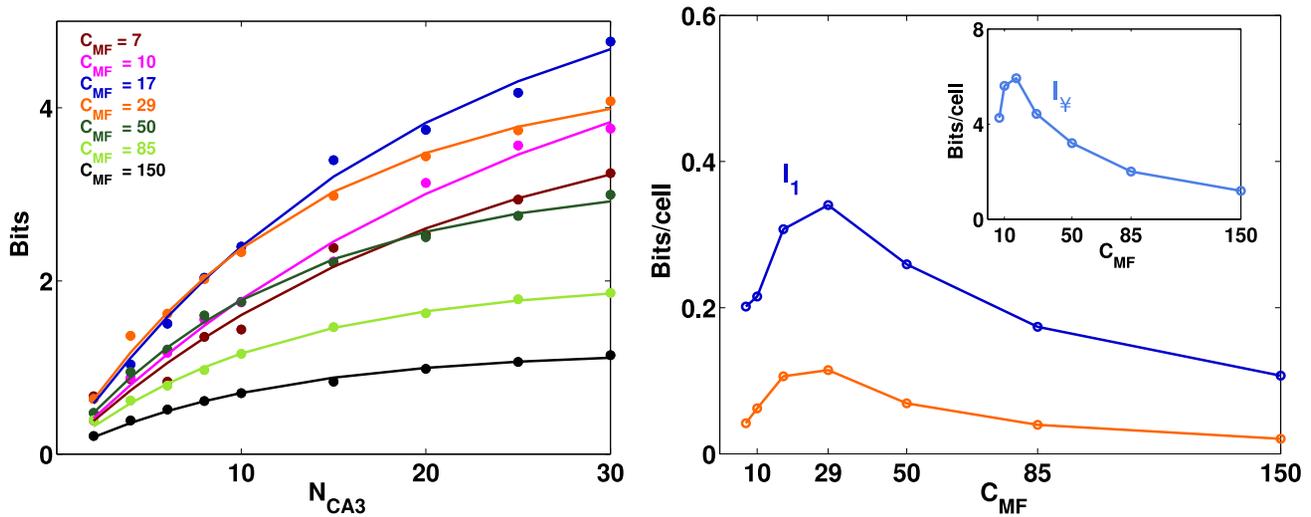


Figure 3. A sparse MF connectivity is optimal, but not too sparse. Left: information plotted versus the number of CA3 cells, with different colors for different values of C_{MF} . Dots represent information values obtained from simulations, while curves are exponentially saturating fits to the data points, as described in Methods. Right: plot of the two parameters of the fit curves. Main figure: slope parameter describing the slope of the linear part of the curve (for low N_{CA3}), contrasted with the analytical estimate of the term proportional to N_{CA3} (Eq.9); inset: total information parameter, describing the saturation level reached by the curve. doi:10.1371/journal.pcbi.1000759.g003

that point on, the information in the representation decreases monotonically. The optimal MF connectivity is then in the range which maximizes the fraction of CA3 units that have a field in the newly learned environment, at a value, roughly one third, broadly consistent with experimental data (see e.g. [30]).

It is important to emphasize that what we are reporting is a quantitative effect: the underlying mechanism is always the same, the random summation of inputs from active DG units. DG in the model effectively operates as a sort of random number generator, whatever the values of the various parameters. How informative are the CA3 representations established by that random number generator, however, depends on the values of the parameters.

Other DG field distribution models

We repeated the simulations using other models for the DG fields distribution, the exponential (model B) and the single field one (model C), and the results are similar to those obtained for model A: the information has a maximum when varying C_{MF} on its own, and is instead roughly constant if the parameter μ is held constant (by varying q inversely to C_{MF}). Fig. 5 reports the comparison, as C_{MF} varies, between models A and B, with $q = 1.7$, and model C, where $q \equiv 1$, so that in this latter case the inputs are $1/1.7$ times weaker (we did not compensate by multiplying J by 1.7). Information measures are obtained by decoding several samples of 10 units, averaging and dividing by 10, and not by

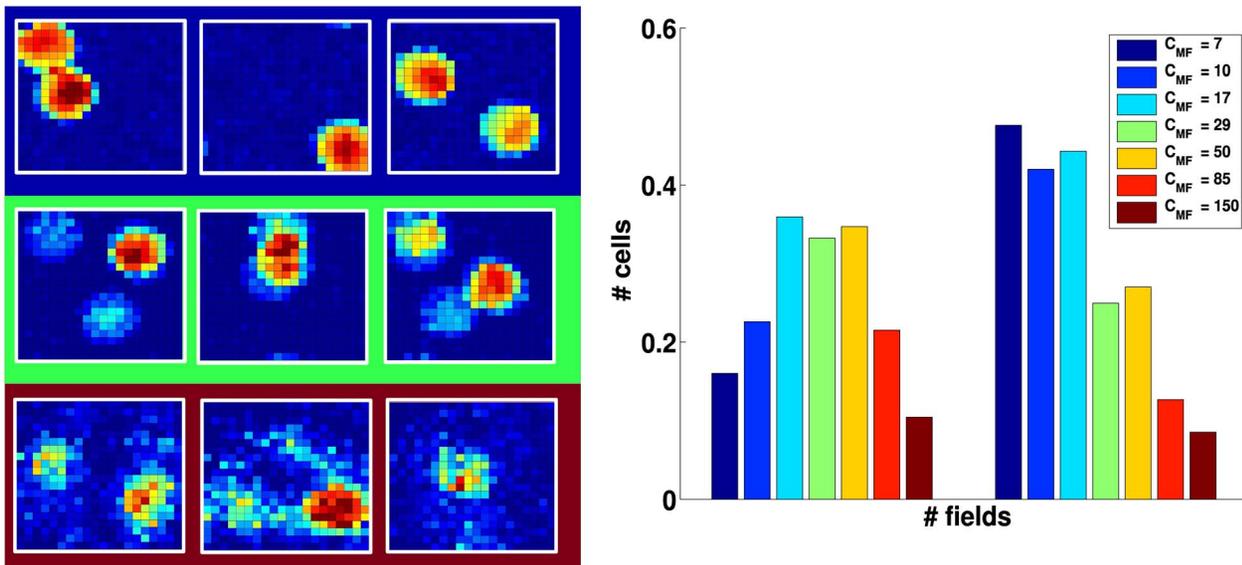


Figure 4. Information vs. connectivity. Left: Examples of CA3 firing rate maps for $C_{MF} = 7$ (top row); $C_{MF} = 29$ (middle) and $C_{MF} = 150$ (bottom); Right: Histogram that shows the fraction of CA3 units active somewhere in the environment, left, and the fraction of active CA3 units with more than one field, right, for different C_{MF} values. doi:10.1371/journal.pcbi.1000759.g004

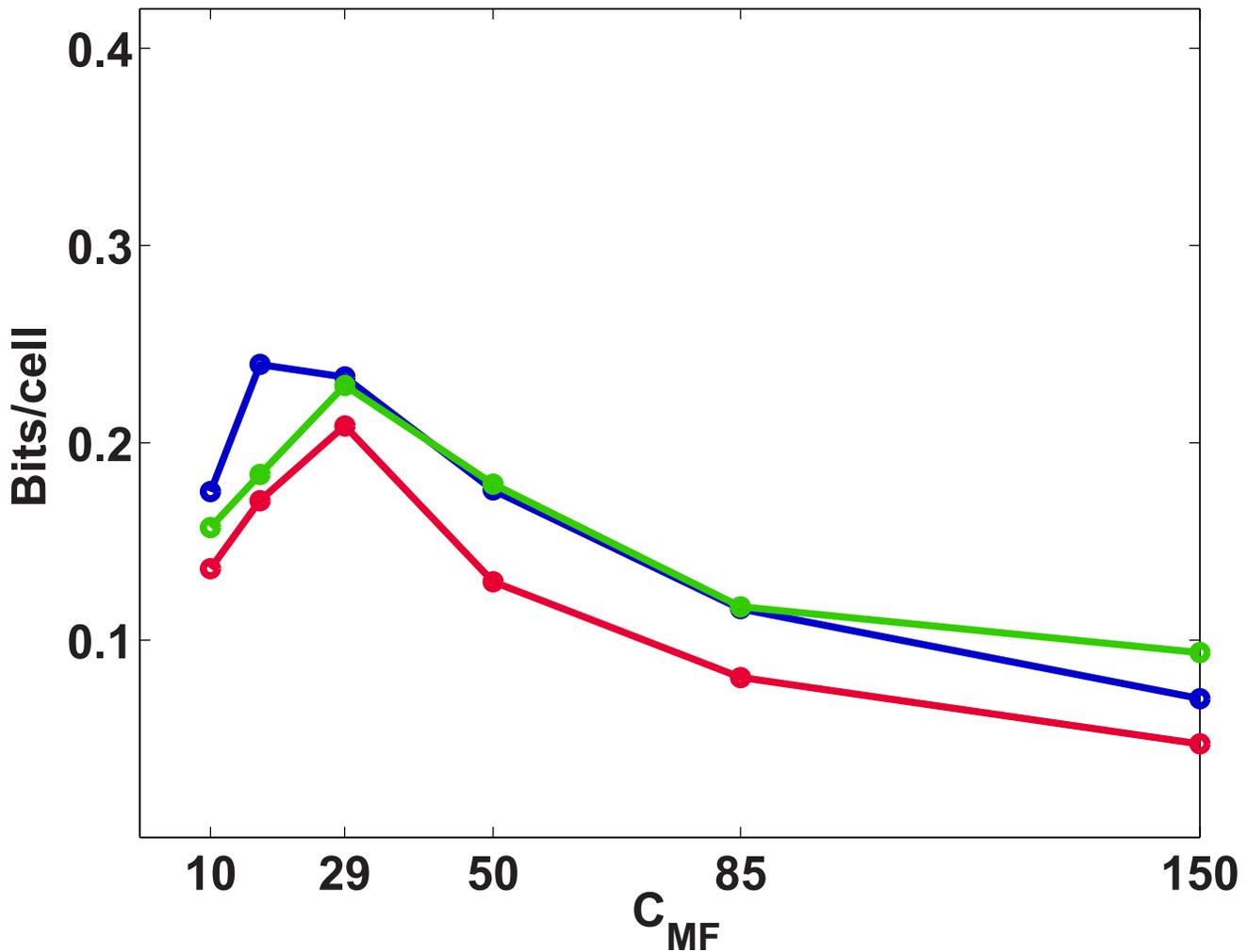


Figure 5. Information vs. connectivity. Information plotted versus different values of connectivity between DG and CA3. Solid lines are all from simulations (localization information from samples of 10 units, divided by 10), as follows: for the blue line, the distribution defining the number of fields in DG cells is Poisson (model A); for the green line, it is exponential (model B); and for the red line, each DG active unit has one field only (model C).

doi:10.1371/journal.pcbi.1000759.g005

extracting the fit parameters. As one can see, the lower mean input for model C leads to lower information values, but the trend with C_{MF} is the same in all three models. This further indicates that the multiplicity of fields in DG units, as well as its exact distribution, is of no major consequence, if comparisons are made keeping constant the mean number of fields in the input to a CA3 unit.

Sparsity of DG activity

We study also how the level of DG activity affects the information flow. We choose different values for the probability p_{DG} that a single DG unit fires in the given environment, and again we adjust the synaptic weight J to keep the mean DG input per CA3 cell constant across the comparisons.

Results (Fig. 6) are similar to those obtained varying the sparsity of the MF connections. Indeed, the analytical estimate in the two conditions would be exactly the same, within the approximation with which we compute it, because the two parameters p_{DG} and C_{MF} enter the calculation in equivalent form, as a product. The actual difference between the two parameters stems from the fact that increasing C_{MF} , CA3 units end up sampling more and more the same limited population of active DG units, while increasing

p_{DG} this population increases in size. This difference can only be appreciated from the simulations, which however show that the main effect remains the same: an information maximum for rather sparse DG activity (and sparse MF connections). The subtle difference between varying the two parameters can be seen better in the saturation information value: with reference to the standard case, in the center of the graph in the inset, to the right increasing p_{DG} leads to more information than increasing C_{MF} , while to the left the opposite is the case, as expected.

Full and simplified decoding procedures

As noted above, we find that the analytical estimate of the information per unit is always considerably lower than the slope parameter of the fit to the measures extracted from the simulations, contrary to expectations, since the latter require an additional decoding step, which implies some loss of information. We also find, however, that the measures of mutual information that we extract from the simulations are strongly dependent on the method used, in the decoding step, to construct the “localization matrix”, i.e. the matrix which compiles the frequency with which the virtual rat was decoded as being in position \vec{x} when it was

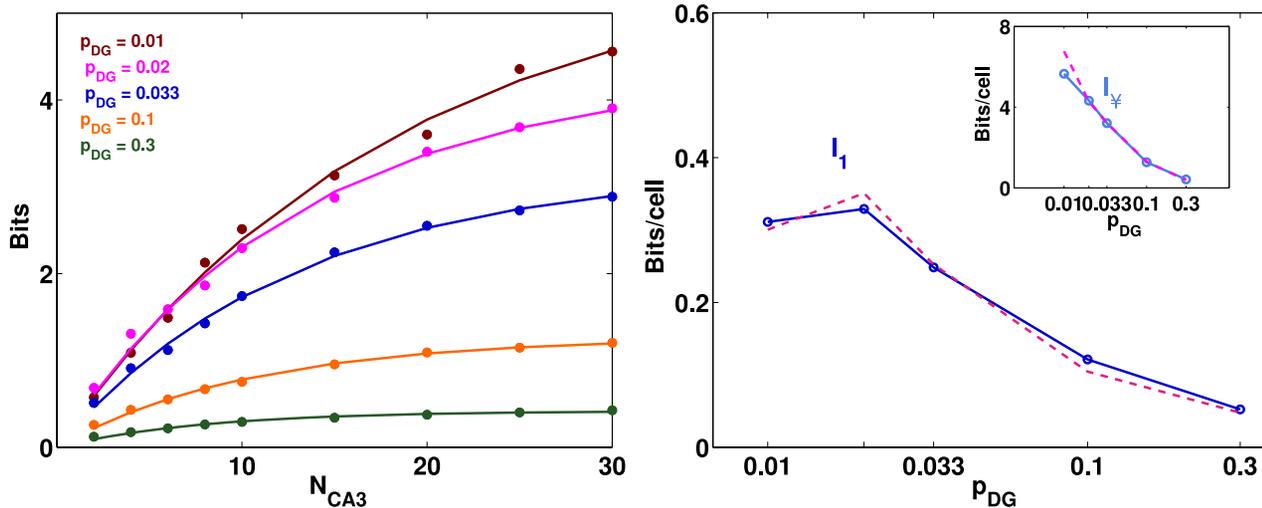


Figure 6. Sparse DG activity is effective at driving CA3. Left: Information plotted versus the number of CA3 units, different colors correspond to different values for p_{DG} . Dots represent information values obtained from simulations, while the curves are exponentially saturating fits to the data points, as described in Methods. Right: Plot of the two parameters of the fits. Main figure: slope parameter describing the slope of the linear part of the information curve (for low N_{CA3}); inset: total information parameter describing the saturation level reached by the information - both are contrasted with the corresponding measures (dashed lines) obtained varying C_{MF} instead of p_{DG} . doi:10.1371/journal.pcbi.1000759.g006

actually in position \vec{x} . All measures reported so far, from simulations, are obtained constructing what we call the *full* localization matrix $Q(\vec{x}, \vec{x}')$ which, if the square environment is discretized into 20×20 spatial bins, is a large 400×400 matrix, which requires of order 160,000 decoding events to be effectively sampled. We run simulations with trajectories of 400,000 steps, and additionally corrected the information measures to avoid the limited sampling bias [31].

An alternative, that allows extracting unbiased measures from much shorter simulations, is to construct a simplified matrix $Q(\vec{x} - \vec{x}')$, which averages over decoding events with the same vector displacement between actual and decoded positions. $Q(\vec{x} - \vec{x}')$ is easily constructed on the torus we used in all simulations, and being a much smaller 20×20 matrix it is effectively sampled in just a few thousand steps.

The two decoding procedures, given that the simplified matrix is the shifted average of the rows of the full matrix, might be expected to yield similar measures, but they do not, as shown in Fig. 7. The simplified matrix, by assuming translation invariance of the errors in decoding, is unable to quantify the information implicitly present in the full distribution of errors around each actual position. Such errors are of an “episodic” nature: the local view from position \vec{x} might happen to be similar to that from position \vec{x}' , hence neural activity reflecting in part local views might lead to confuse the two positions, but this does not imply that another position \vec{z} has anything in common with $\vec{z} + (\vec{x}' - \vec{x})$. Our little network model captures this discrepancy, in showing, in Fig. 7, that for any actual position there are a few selected position that are likely to be erroneously decoded from the activity of a given sample of units; when constructing instead the translationally invariant simplified matrix, all average errors are distributed smoothly around the correct position (zero error), in a roughly Gaussian bell. The upper right panel in Fig. 7 shows that such episodic information always prevails, whatever the connectivity, i.e. in all three parameter regimes illustrated in Fig. 4. The lower right panel in Fig. 7 compares, instead, the entropies of the decoded positions with the two matrices, conditioned on the actual position - that is, the *equivocation* values. Unlike the mutual

information, such equivocation is much higher for the simplified matrix; for this matrix, it is simply a measure of how widely displaced are decoded positions, with respect to the actual positions, represented at the center of the square; and for small samples of units, which are not very informative, the “displacement” entropy approaches that of a flat distribution of decoded positions, i.e. $\log_2(400) \approx 8.64$ bits. For larger samples, which enable better localization, the simplified localization matrix begins to be clustered in a Gaussian bell around zero displacement, so that the equivocation gradually decreases (the list of displacements, with their frequencies, is computed for each sample, and it is the equivocation, not the list itself, which is averaged across samples). In contrast, the entropy of each row of the full localization matrix, i.e. the entropy of decoded positions conditioned on any actual position, is lower, and also decreasing more steeply with sample size; it differs from the full entropy, in fact, by the mutual information between decoded and actual positions, which *increases* with sample size. The two equivocation measures therefore both add up to the two mutual information measures to yield the same full entropy of about 8.64 bits (a bit less in the case of the full matrix, where the sampling is more limited), and thus serve as controls that the difference in mutual information is not due, for example, to inaccuracy. As a third crucial control, we calculated also the average conditional entropy of the full localization matrix, when the matrix is averaged across samples of a given size: the resulting entropy is virtually identical to the displacement entropy (which implies instead an average of the full matrix across rows, i.e. across actual positions). This indicates that different samples of units express distinct episodic content at each location, such that averaging across samples is equivalent to averaging across locations.

Apparently, also the analytical estimate is unable to capture the spatial information implicit in such “episodic” errors, as its values are well below those obtained with the full matrix, and somewhat above those obtained with the simplified matrix (consistent with some loss with decoding). One may wonder how can the information from the full localization matrix (which also requires a decoding step) be higher than the decoding-free analytical

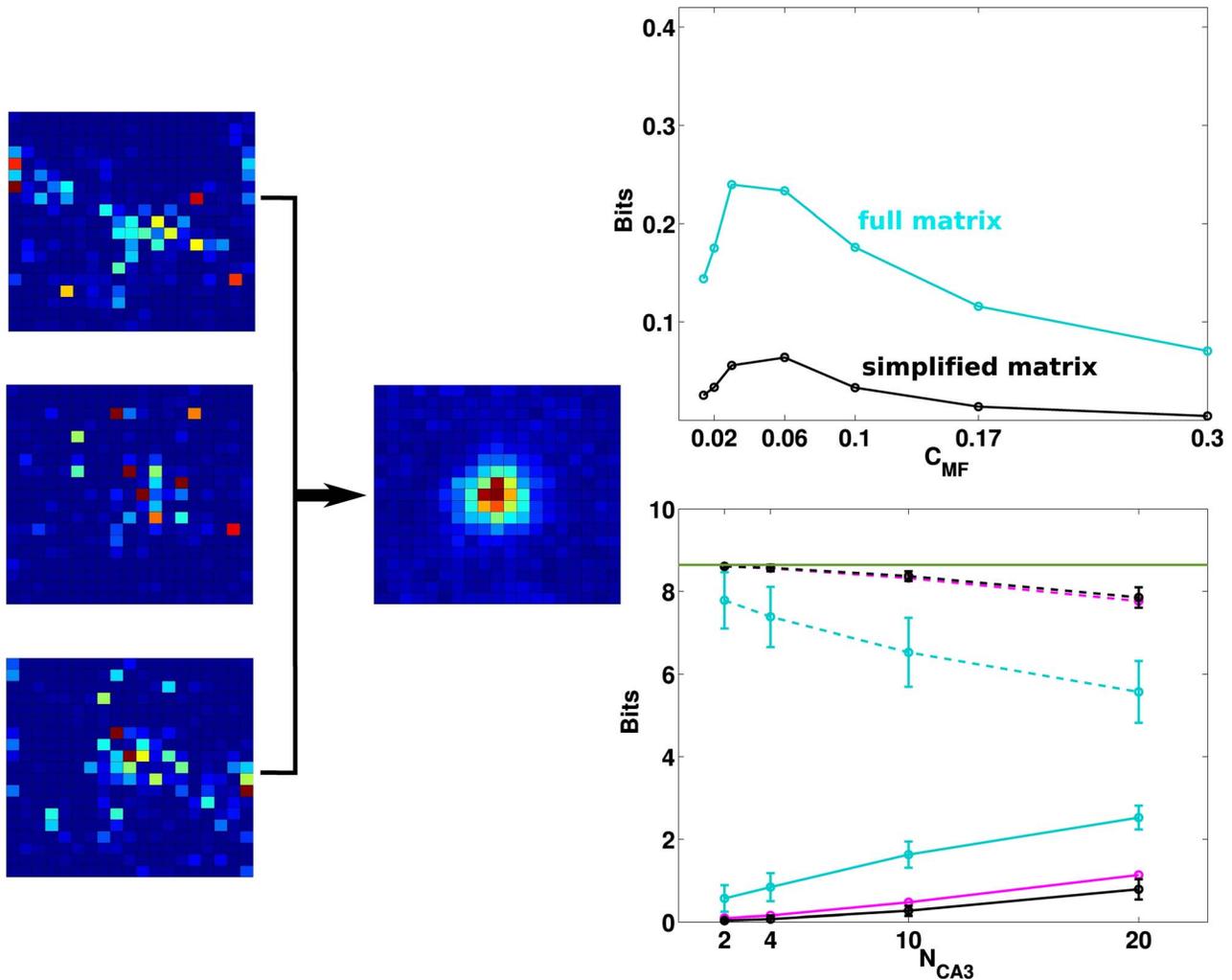


Figure 7. Localization matrices. Left: the rows of the full matrix represent the actual positions of the virtual rat while its columns represent decoded positions (the full matrix is actually 400×400); three examples of rows are shown, rendered here as 20×20 squares, all from decoding by a given sample of 10 units. The simplified matrix is a single 20×20 matrix obtained (from the same sample) as the average of the full matrix taking into account translation invariance. Right, top: the two procedures lead to large quantitative differences in information (here, the measures from samples of 10 units, divided by 10, from the full matrix, cyan, and from the simplified matrix, black), but with the same dependence on C_{MF} . Right, bottom: The conditional entropies of the full and simplified localization matrices (cyan and black, dashed) in both cases add up to the respective mutual information measure (cyan and black, solid) to give the full entropy of $\log(400) \approx 8.64$ bits (green line). The conditional entropy calculated from the full matrix averaged across samples (red, dashed) is equivalent to that calculated from the displacements, for each sample (black, dashed).
doi:10.1371/journal.pcbi.1000759.g007

estimate, without violating the basic information processing theorem. The solution to the riddle, as we understand it, is subtle: when decoding, one takes essentially a maximum likelihood estimate, assigning a unique decoded position per trial, or time step. This leads to a “quantized” localization matrix, which in general tends to have substantially higher information content than the “smoothed” matrix based on probabilities [32]. In the analytical derivation there is no concept of trial, time step or maximal likelihood, and the matrix expresses smoothly varying probabilities. The more technical implications are discussed further at the end of the Methods. These differences do not alter the other results of our study, since they affect the height of the curves, not their shape, however they have important implications. The simplified matrix has the advantage of requiring much less data, i.e. less simulation time, but also less real data if applied to neurophysiological recordings, than the full matrix, and in most situations it might be the only feasible measure of spatial

information (the analytical estimate is not available of course for real data). So in most cases it is only practical to measure spatial information with methods that, our model suggests, miss out much of the information present in neuronal activity, what we may refer to as “dark information”, not easily revealed. One might conjecture that the prevalence of dark information is linked to the random nature of the spatial code established by DG inputs. It might be that additional stages of hippocampal processing, either with the refinement of recurrent CA3 connections or in CA1, are instrumental in making dark information more transparent.

Effect of learning on the Mossy Fibers

While the results reported this far assume that MF weights are fixed, $J = 1$, we have also conducted a preliminary analysis of how the amount of spatial information in CA3 might change as a consequence of plasticity on the mossy fibers. In an extension of the standard model, we allow the weights of the connections

between DG and CA3 to change with a model “Hebbian” rule. This is not an attempt to capture the nature of MF plasticity, which is not NMDA-dependent and might not be associative [33], but only the adoption of a simple plasticity model that we use in other simulations. At each time step (that corresponds to a different place in space) weights are taken to change as follows:

$$\Delta J_{ij}^{MF}(t) = \gamma_{MF} \eta_i(\vec{x}(t)) (\beta_j(\vec{x}(t)) - \langle \beta(\vec{x}(t)) \rangle) \quad (10)$$

where γ_{MF} is a plasticity factor that regulates the amount of learning. Modifying in this way the MF weights has the general effect of increasing information values, so that they approach saturation levels for lower number of CA3 cells; in particular this is true for the information extracted from both full and simplified matrices. In Fig. 8, the effect of such “learning” is shown for different values of the parameter γ_{MF} , as a function of connectivity.

We see that allowing for this type of plasticity on mossy fibers leads to shift the maximum of information as a function of the

connectivity level. The structuring of the weights effectively results in the selection of favorite input connections, for each CA3 unit, among a pool of available ones; the remaining strong connections are a subset of those “anatomically” present originally. It is logical, then, that starting with a larger pool of connections, among which to pick the “right” ones, leads to more information than starting with few connections, which further decrease in effective number with plasticity. We expect better models of the details of MF plasticity to preserve this main effect.

A further effect of learning, along with the disappearance of some CA3 fields and the strengthening of others, is the refinement of their shape, as illustrated in Fig. 9. It is likely that also this effect will be observed even when using more biologically accurate models of MF plasticity.

Retrieval abilities

Finally, all simulations reported so far involved a full complement of DG inputs at each time step in the simulation. We have also tested the ability of the MF network to *retrieve* a

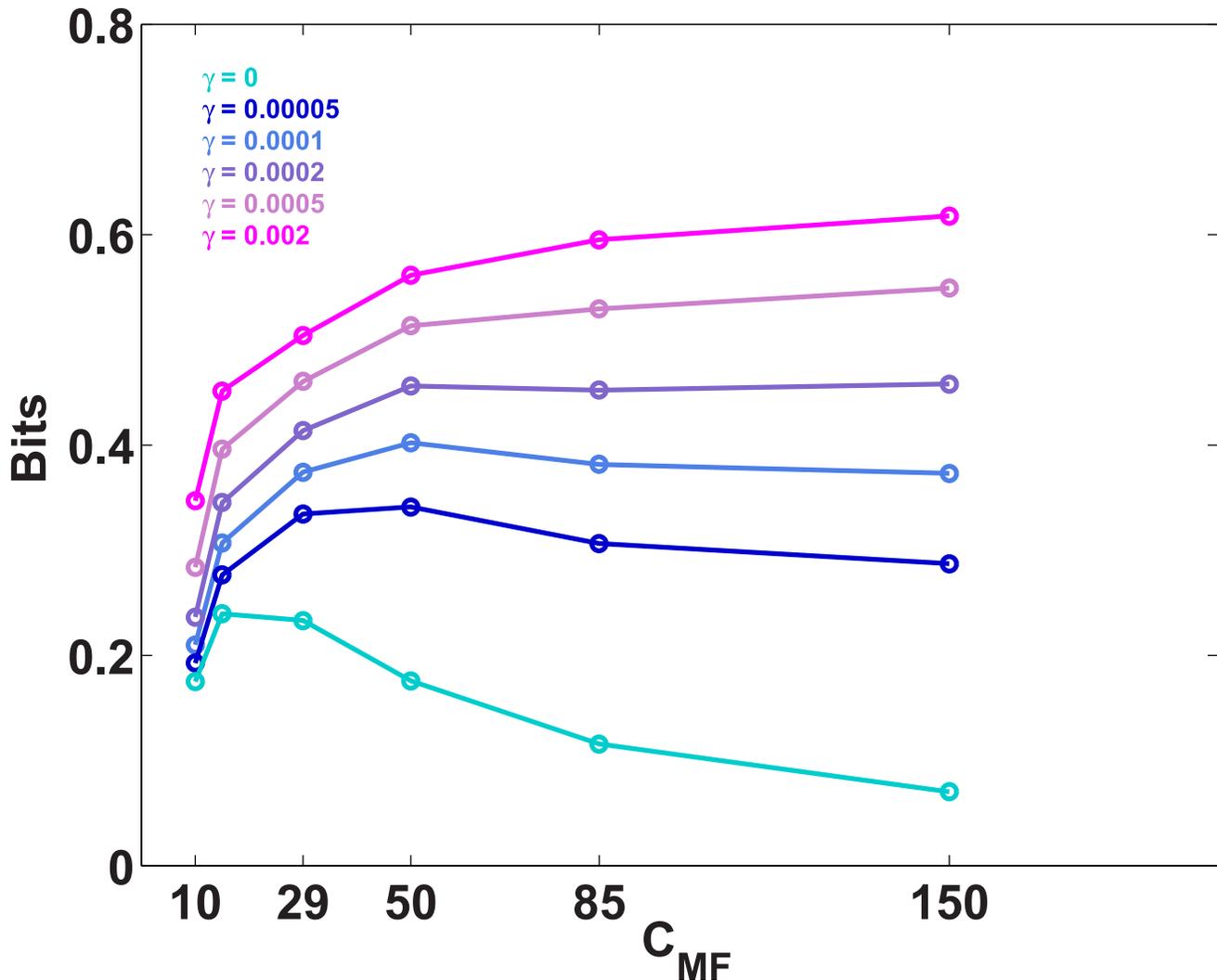


Figure 8. Information vs. connectivity for different levels of learning. Information is plotted as a function of the connectivity level between DG and CA3, different colors correspond to different values of the learning factor γ_{MF} . Simulations run for 100,000 training steps, during a fraction $\sim a_{CA3} = 0.1$ of which each postsynaptic units is strongly activated, and its incoming weights liable to be modified. The γ values tested hence span the range from minor modification of the existing weight, for $\gamma = 0.00005$, to major restructuring of all available weights for $\gamma = 0.002$. doi:10.1371/journal.pcbi.1000759.g008

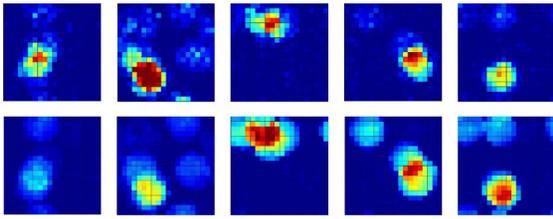


Figure 9. MF plasticity can suppress, enlarge and in general refine CA3 place fields. The place fields of five example units are shown before plasticity is turned on (top row) and after 100,000 steps with a large plasticity factor $\gamma_{MF}=0.0001$ (bottom row). The rounding and regularization of the fields was observed also for several other units in the simulation.
doi:10.1371/journal.pcbi.1000759.g009

spatial representation when fed with a degraded input signal, with and without MF plasticity. The input is degraded, in our simulation, simply by turning on only a given fraction, randomly selected, of the DG units that would normally be active in the environment. The information extracted after decoding by a sample of units (in Fig. 10, 10 units) is then contrasted with the size

of the cue itself. In the absence of MF plasticity, there is obviously no real retrieval process to talk about, and the DG-CA3 network simply relays partial information. When Hebbian plasticity is turned on, the expectation from similar network models (see e.g. [34], Fig. 9) is that there would be some pattern completion, i.e. some tendency for the network to express nearly complete output information when the input is partial, resulting in a more sigmoidal input-output curve (the exact shape of the curve depends of course also on the particular measure used).

It is apparent from Fig. 10 that while, in the absence of plasticity, both parameters characterizing the information that can be extracted from CA3 grow roughly linearly with the size of the cue, with plasticity the growth is supralinear. This amounts to the statement that the beneficial effects of plasticity require a full cue to be felt – the conceptual opposite to pattern completion, the process of integrating a partial cue using information stored on modified synaptic weights. This result suggests that the sparse MF connectivity is sub-optimal for the associative storage that leads to pattern completion, a role that current perspectives ascribe instead to perforant path and recurrent connections to CA3. The role of the mossy fibers, even if plastic, may be limited to the establishment of new spatial representations.

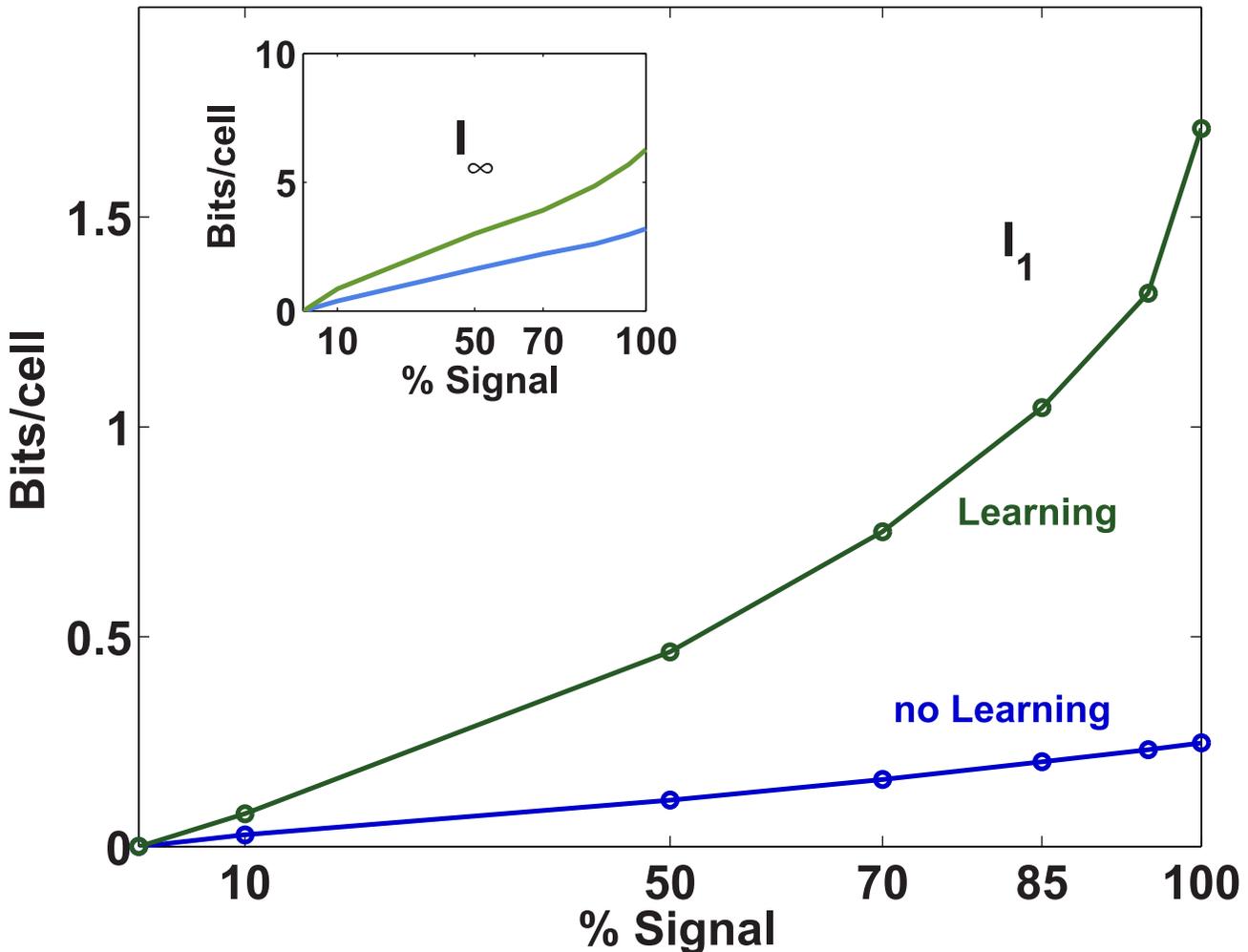


Figure 10. Information reconstructed from a degraded input signal. Slope parameter I_1 of the information curve as a function of the percentage of the DG input that CA3 receives. Inset: the same plot for the total information parameter I_∞ . The same training protocol was run as for Figs. 8–9.
doi:10.1371/journal.pcbi.1000759.g010

Discussion

Ours is a minimal model, which by design overlooks several of the elements likely to play an important role in the functions of the dentate gyrus - perhaps foremost, neurogenesis [35]. Nevertheless, by virtue of its simplicity, the model helps clarify a number of quantitative issues that are important in refining a theoretical perspective of how the dentate gyrus may work.

First, the model indicates that the recently discovered multiplicity of place fields by active dentate granule cells [1] might be just a “fact of life”, with no major computational implications for dentate information processing. Still, requiring that active granule cells express multiple fields seems to lead, in another simple network model (of how dentate activity may result from entorhinal cortex input [25]), to the necessity of inputs coming from *lateral* EC, as well as from *medial* EC. The lateral EC inputs need not carry any spatial information but help to select the DG cells active in one environment. Thus the multiplicity of DG fields refines the computational constraints on the operation of hippocampal circuits.

Second, the model shows that, assuming a fixed total MF input strength on CA3 units, it is beneficial in information terms for the MF connectivity to be very sparse; but not vanishingly sparse. The optimal number of anatomical MF connections on CA3 units, designated as C_{MF} in the model, depends somewhat on the various parameters (the noise in the system, how sparse is the activity in DG and CA3, etc.) and it may increase slightly when taking MF plasticity into account, but it appears within the range of the number, 46, reported for the rat by [26]. It will be interesting to see whether future measures of MF connectivity in other species correspond to those “predicted” by our model once the appropriate values of the other parameters are also experimentally measured and inserted into the model. A similar set of consideration applies to the fraction of granule cells active in a given environment, p_{DG} , which in the model plays a similar, though not completely identical, role to C_{MF} in determining information content.

Third, the model confirms that the sparse MF connections, even when endowed with associative plasticity, are not appropriate as devices to store associations between input and output patterns of activity – they are just too sparse. This reinforces the earlier theoretical view [3], [5], which was not based however on an analysis of spatial representations, that the role of the dentate gyrus is in establishing new CA3 representations and not in associating them to representations expressed elsewhere in the system. Availing itself of more precise experimental parameters, and based on the spatial analysis, the current model can refine the earlier theoretical view and correct, for example, the notion that “detonator” synapses, firing CA3 cells on a one-to-one basis, would be optimal for the mossy fiber system. The optimal situation turns out to be the one in which CA3 units are fired by the combination of a couple of DG input units, although this is only a statistical statement. Whatever the exact distribution of the number of coincident inputs to CA3, DG can be seen as a sort of *random pattern generator*, that sets up a CA3 pattern of activity without any structure that can be related to its anatomical lay-out [36], or to the identity of the entorhinal cortex units that have activated the dentate gyrus. As with random number generators in digital computers, once the product has been spit out, the exact process that led to it can be forgotten. This is consistent with experimental evidence that inactivating MF transmission or lesioning the DG does not lead to hippocampal memory impairments once the information has already been stored, but leads to impairments in the storage of new information [7], [8].

The inability of MF connection to subserve pattern completion is also consistent with suggestive evidence from imaging studies with human subjects [37].

Fourth, and more novel, our findings imply that a substantial fraction of the information content of a spatial CA3 representation, over half when sampling limited subsets of CA3 units, can neither be extracted through the simplified method which assumes translation invariance, nor assessed through the analytical method (which anyway requires an underlying model of neuronal firing, and is hence only indirectly applicable to real neuronal data). This large fraction of the information content is only extracted through the time-consuming construction of the full localization matrix. To avoid the limited sampling bias [38] this would require, in our hands, the equivalent of a ten hour session of recording from a running rat (!), with a square box sampled in 20×20 spatial bins. We have hence labeled this large fraction as *dark information*, which requires a special effort to reveal. Although we know little of how the real system decodes its own activity, e.g. in downstream neuronal populations, we may hypothesize that the difficulty at extracting dark information affects the real system as well, and that successive stages of hippocampal processing have evolved to address this issue. If so, qualitatively this could be characterized as the representation established in CA3 being *episodic*, i.e. based on an effectively random process that is functionally forgotten once completed, and later processing, e.g. in CA1, may be thought to gradually endow the representations with their appropriate continuous spatial character. Another network model, intended to elucidate how CA1 could operate in this respect, is the object of our on-going analysis.

The model analysed here does not include neurogenesis, a most striking dentate phenomenon, and thus it cannot comment on several intriguing models that have been put forward about the role of neurogenesis in the adult mammalian hippocampus [39], [40], [41]. Nevertheless, presenting a simple and readily expandable model of dentate operation can facilitate the development of further models that address neurogenesis, and help interpret puzzling experimental observations. For example, the idea that once matured newborn cells may temporally “label” memories of episodes occurring over a few weeks [42], [43], [44], [45] has been weakened by the observation that apparently even young adult-born cells, which are not that many [46], [47], [45], are very sparsely active, perhaps only a factor of two or so more active than older granule cells [24]. Maybe such skepticism should be reconsidered, and the issue reanalysed using a quantitative model like ours. One could then investigate the notion that the new cells link together, rather than separating, patterns of activity with common elements (such as the temporal label). To do that clearly requires extending the model to include a description not only of neurogenesis, but also of plasticity within DG itself [48] and of its role in the establishment of successive representations one after the other.

Methods

Replica calculation

Estimation of the equivocation. Calculating the equivocation from its definition in Eq.7 is straightforward, thanks to the simplifying assumption of independent noise in CA3 units. We get

$$\langle H_2(\{\eta_i\}|\bar{x}) \rangle_{\bar{x}} = \frac{N}{\ln 2} \int \frac{d\bar{x}}{A} \left\{ -\Phi(-\rho_i(\bar{x})) \ln \Phi(-\rho_i(\bar{x})) + \Phi(\rho_i(\bar{x})) \left[\frac{1}{2} + \ln(\sqrt{2\pi}\delta) \right] - \frac{1}{2} \rho_i(\bar{x}) \sigma(\rho_i(\bar{x})) \right\} \quad (11)$$

where

$$\rho_i = \frac{\bar{\eta}_i}{\delta}$$

$$\sigma(\rho) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\rho^2}{2}},$$

although the spatial integral remains to be carried out.

Estimation of the entropy. For the entropy, Eq.6, the calculation is more complicated. Starting from

$$H_1(\{\eta_i\}) = -\frac{1}{\ln 2} \int \frac{d\vec{x}}{A} \prod_i d\eta_i P(\{\eta_i\}|\vec{x}) \ln \left[\int \frac{d\vec{x}'}{A} P(\{\eta_i\}|\vec{x}') \right]$$

we remove the logarithm using the replica trick (see [27])

$$H_1(\{\eta_i\}) = -\frac{1}{\ln 2} \int \frac{d\vec{x}}{A} \prod_i d\eta_i P(\{\eta_i\}|\vec{x}) \lim_{n \rightarrow 0} \frac{1}{n} \left\{ \left[\int \frac{d\vec{x}'}{A} P(\{\eta_i\}|\vec{x}') \right]^n - 1 \right\} \quad (12)$$

which can be rewritten (Nadal and Parga [49] have shown how to use the replica trick in the $n \rightarrow 1$ limit, a suggestion used in [50] to analyse information transfer in the CA3-CA1 system)

$$H_1(\{\eta_i\}) = -\frac{1}{\ln 2} \lim_{n \rightarrow 1} \frac{1}{(n-1)} [\tilde{H}_1(n) - \tilde{H}_1(1)] \quad (13)$$

using the spatial averages, defined for an arbitrary real-valued number n of replicas

$$\tilde{H}_1(n) \equiv \int \prod_i d\eta_i \left[\int \frac{d\vec{x}}{A} P(\{\eta_i\}|\vec{x}) \right]^n$$

$$= \int \frac{d\vec{x}_1}{A} \dots \frac{d\vec{x}_\beta}{A} \dots \frac{d\vec{x}_n}{A} \prod_i h_i(\{\vec{x}_\beta\}, n) \quad (14)$$

where we have defined a quantity dependent on both the number n of replicas and on the position in space, later to be integrated over, of each replica β :

$$h_i(\{\vec{x}_\beta\}, n) \equiv \int d\eta_i \prod_\alpha \left[\delta(\eta_i) \Phi \left(-\frac{\bar{\eta}_i(\vec{x}_\beta)}{\delta} \right) + \frac{1}{(\sqrt{2\pi}\delta)} e^{-\frac{(\eta_i - \bar{\eta}_i(\vec{x}_\beta))^2}{2\delta^2}} \Theta(\eta_i) \right].$$

We need therefore to carry out integrals over the firing rate of each CA3 unit, η_i , in order to estimate $h_i(\{\vec{x}_\beta\}, n)$, while keeping in mind that in the end we want to take $n \rightarrow 1$. Carrying out the integrals yields a below-threshold and an above-threshold term

$$h_i(\{\vec{x}_\beta\}, n) = \prod_\beta \Phi(-\rho_i(\vec{x}_\beta)) + e^{-\frac{(n-1)S}{2\delta^2}} \int_{-\infty}^{\eta_0} \frac{e^{-\frac{\eta^2}{2\delta^2}}}{(\sqrt{2\pi}\delta)^n} d\eta \quad (15)$$

where we have defined the quantities

$$S(\{\vec{x}_\beta\}) = - \left[\frac{1}{n(n-1)} \left(\sum_\beta \bar{\eta}_i(\vec{x}_\beta) \right)^2 - \frac{1}{(n-1)} \sum_\beta \bar{\eta}_i^2(\vec{x}_\beta) \right] \quad (16)$$

$$= \frac{1}{2n(n-1)} \sum_{\alpha, \beta} [\bar{\eta}_i(\vec{x}_\alpha) - \bar{\eta}_i(\vec{x}_\beta)]^2$$

and $\eta_0 = (1/n) \sum_\beta \bar{\eta}_i(\vec{x}_\beta)$, while $\hat{\eta} = \eta - \eta_0$.

One might think that $h_i(\{\vec{x}_\beta\}, n \sim 1) \rightarrow 1 + (n-1)h_{i,n}(\{\vec{x}_\beta\}, 1) + O((n-1)^2)$, hence in the product over cells, that defines the entropy $H_1(\{\eta_i\})$, the only terms that survive in the limit $n \rightarrow 1$ would just be the summed single-unit contributions obtained from the first derivatives with respect to n . This is not true, however, as taking the replica limit produces the counter-intuitive effect that replica-tensor products of terms, which individually disappear for $n \rightarrow 1$, only vanish to first order in $n-1$, as shown by [29]. The replica method is therefore able, in principle, to quantify the effect of correlations among units, expressed in entropy terms stemming from the product of h_i across units.

Briefly, one has

$$h_i(\{\vec{x}_\beta\}, n) = \prod_\beta \Phi(-\rho_i(\vec{x}_\beta)) + e^{-\frac{(n-1)S}{2\delta^2}} \left[\frac{1}{\sqrt{n}(\sqrt{2\pi}\delta)^{n-1}} \Phi(\sqrt{n}\eta_0/\delta) \right]$$

$$\simeq \Phi(-\rho_o) + \frac{1}{2} \frac{\sigma^2(-\rho_o)}{\Phi(-\rho_o)} \sum_{\alpha, \beta \neq \alpha} (\rho(\vec{x}_\alpha) - \rho_o)(\rho(\vec{x}_\beta) - \rho_o)$$

$$+ (n-1)\Phi(-\rho_o) \ln \Phi(-\rho_o) + \Phi(\rho_o) + \frac{(n-1)}{2} \sigma(\rho_o)\rho_o \quad (17)$$

$$- (n-1)\Phi(\rho_o) \left[\frac{1}{2} + \ln(\sqrt{2\pi}\delta) + \sum_{\alpha, \beta \neq \alpha} \frac{S(\vec{x}_\alpha, \vec{x}_\beta)}{2\delta^2} \right]$$

$$+ (n-1)\sigma(\rho_o) \frac{\partial \rho_o}{\partial n}$$

where the first two rows come from the term below threshold, and the last two from the one above threshold. Then, following [29],

$$H_1(\{\eta_i\}) \simeq -\frac{1}{\ln 2} \lim_{n \rightarrow 1} \frac{1}{(n-1)} \left[\int \frac{d\vec{x}_\alpha}{A} \frac{d\vec{x}_\beta}{A} \left(1 - (n-1)C + (n-1)\Gamma - \sum_{\alpha, \beta \neq \alpha} G_{\alpha, \beta} \right)^N - 1 \right] \quad (18)$$

where

$$C = \Phi(\rho_o) \left[\frac{1}{2} + \ln(\sqrt{2\pi}\delta) \right] - \frac{1}{2} \sigma(\rho_o)\rho_o - \Phi(\rho_o) \ln \Phi(-\rho_o)$$

$$\Gamma = \sigma(\rho_o) \left[-\rho(\vec{x}_1) + \frac{1}{n-1} \sum_{\beta > 1} \rho(\vec{x}_\beta) \right] \quad (19)$$

$$G_{\alpha, \beta} = \frac{\Phi(\rho_o)}{4n(n-1)} (\rho(\vec{x}_\alpha) - \rho(\vec{x}_\beta))^2 + \frac{\sigma^2(-\rho_o)}{2\Phi(-\rho_o)} (\rho(\vec{x}_\alpha) - \rho_o)(\rho(\vec{x}_\beta) - \rho_o)$$

and where we have considered that in the limit $n \rightarrow 1$ we have $\eta_0/\delta \equiv \rho_o$ appear in all terms of finite weight.

The products between the matrices $G_{\alpha,\beta}$ attached to each CA3 unit generate the higher order terms in N . Calculating them in our case, in which different CA3 units can receive partially overlapping inputs from DG units, is extremely complex (see [51], where information transmission across a network is also considered), and we do not pursue here the analysis of such higher order terms. One can retrieve the result of the TG model in Ref. [29] by taking the further limit $\rho_o \rightarrow 0$, which implies $\Phi(\rho_o) \rightarrow 1/2$ and $\sigma^2(\rho_o) \rightarrow 1/(2\pi)$. A further subtlety is that, in taking the $n \rightarrow 1$ limit, there is a single replica, say \bar{x} , which is counted *once* in the limit, but also several *different* replicas, denoted $\bar{x}', \bar{x}'', \dots$, whose weights vanish, but which remain to determine e.g. the terms proportional to $(n-1)$ emerging from the derivatives. Thus, in the very last term of Eq. 17, one has to derive ρ_o with respect to n to produce the Γ term of Eq. 19, which is absent in [29] because it vanishes with ρ_o . In the off-diagonal terms of the G matrix there are $2(n-1)$ entries dependent on replicas \bar{x} and \bar{x}' , and $(n-1)(n-2)$ entries dependent on replicas \bar{x}' and \bar{x}'' .

Focusing now solely on terms of order N , note that the term S is effectively a *spatial signal*. In the $n \rightarrow 1$ limit it can be rewritten, using \bar{x} for the single surviving replica, as

$$S(\bar{x}, \bar{x}') = [\bar{\eta}_i(\bar{x}) - \bar{\eta}_i(\bar{x}')]^2 - \frac{1}{2} [\bar{\eta}_i(\bar{x}') - \bar{\eta}_i(\bar{x}'')]^2.$$

This allows us to derive, to order N , our result for the spatial information content, Eq. 8.

Note that when the threshold of each unit tends to $-\infty$, and therefore its mean activation $\rho_i \rightarrow \infty$, our units behave as threshold-less linear units with gaussian noise, and the information they convey tends to

$$\langle I(\bar{x}, \{\eta_i\}) \rangle = \frac{N}{4 \ln 2} \left\langle \int \frac{d\bar{x} d\bar{x}'}{A A} [\rho_i(\bar{x}) - \rho_i(\bar{x}')]^2 \right\rangle \quad (20)$$

which is simply expressed in terms of a spatial signal-to-noise ratio, and coincides with the results in Refs. [28], [29].

m -field decomposition

Eqs. 8 and 9 simply sum equivalent average contributions from each CA3 unit. Each such contribution can then be calculated as a series in m , the number of DG fields feeding into the CA3 unit. One can in fact write, for example,

$$\begin{aligned} \langle \Phi(\rho_i(\bar{x})) \rho_i^2(\bar{x}) \rangle = & \\ & \frac{P(1)}{\delta^2} \left\{ \sum_{Q_1=0}^{\infty} P(Q_1) \Phi(\rho_i(\bar{x})) \left[\sum_j J\beta_j(\bar{x}, \{\bar{x}_{jk}\}) - T \right]^2 \right\} \\ & + \frac{P(2)}{\delta^2} \left\{ \sum_{Q_1, Q_2=0}^{\infty} P(Q_1) P(Q_2) \Phi(\rho_i(\bar{x})) \left[\sum_j J\beta_j(\bar{x}, \{\bar{x}_{jk}\}) - T \right]^2 \right\} \\ & + \dots + \frac{P(\gamma)}{\delta^2} \left\{ \sum_{Q_1, Q_2, \dots, Q_\gamma=0}^{\infty} P(Q_1) P(Q_2) \dots P(Q_\gamma) \Phi(\rho_i(\bar{x})) \right. \\ & \left. \left[\sum_j J\beta_j(\bar{x}, \{\bar{x}_{jk}\}) - T \right]^2 \right\} + \dots \end{aligned}$$

where in each term there are γ active DG units, indexed by j , presynaptic to CA3 unit i , and each has Q_j fields (including the

possibility that $Q_j=0$), indexed by k . A similar expansion can be written for the other terms. One then realizes that the spatial component reduces to integrals that depend solely on the total number of fields $m = \sum_j Q_j$, no matter how many DG active units they come from, and the expansion can be rearranged into an expansion in m

$$\langle I(\bar{x}, \{\eta_i\}) \rangle = \frac{N}{2 \ln 2} \sum_{m=0}^{\infty} C_m D_m(T) \quad (21)$$

where one of the components in each term is, for example,

$$D_m(T) = \int \frac{d\bar{x} d\bar{x}' d\bar{x}_1}{A A A} \dots \frac{d\bar{x}_m}{A} \Phi(\rho(\bar{x}, \{\bar{x}_j\})) \rho^2(\bar{x}', \{\bar{x}_j\}) \quad (22)$$

with $\rho(\bar{x}', \{\bar{x}_j\}) = [J \sum_{i=1}^m \psi(\bar{x}' - \bar{x}_j) - T] / \delta$ the mean signal-to-noise at position \bar{x} produced by m fields, from no matter how many DG units. The numerical coefficient C_m , instead, stems from the combination of the distribution for the number of fields for each presynaptic DG unit active in the environment, which differs between models A, B and C, and the Poisson distribution for the number of such units

$$P(\gamma) = \frac{(\alpha)^\gamma}{\gamma!} e^{-\alpha}$$

$$\alpha = p_{DG} C_{MF}.$$

The sum extends in principle to $m \rightarrow \infty$, but in practice it can be truncated after checking that successive terms give vanishing contributions. The appropriate truncation point obviously depends on the mean number of fields q , as well as on the model distribution of fields per unit. Note that the first few terms (e.g. for $m=0, 1, \dots$) may give negative but not necessarily negligible contributions if the effective threshold T is high.

For model A,

$$P_A(Q) = \frac{q^Q}{Q!} e^{-q}$$

and combining the two Poisson series one finds

$$C_m = e^{\alpha(e^{-q}-1)} K_m \frac{q^m}{m!} \quad (23)$$

where $K_0 \equiv 1$ and the other $K_m(\lambda)$ are the polynomials

$$\begin{cases} K_1 &= \lambda \\ K_2 &= \lambda + \lambda^2 \\ K_3 &= \lambda + 3\lambda^2 + \lambda^3 \\ K_4 &= \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4 \\ &\vdots \\ K_m &= \sum_{l=1}^m \Gamma(l, m) \lambda^l \\ &\vdots \end{cases}$$

given by the modified Khayyam-Tartaglia recursion relation

$$\Gamma(l, m) = \Gamma(l-1, m-1) + l \Gamma(l, m-1)$$

and where $\lambda = \alpha e^{-q}$.

For model B,

$$P_B(Q) = \frac{1}{1+q} \left(\frac{q}{1+q} \right)^{Q_i}$$

and combining the Poisson with the exponential series one finds

$$C_m = e^{\alpha \left(\frac{1}{1+q} \right)^{-1}} \tilde{K}_m \left(\frac{q}{1+q} \right)^m \quad (24)$$

where again $\tilde{K}_0 \equiv 1$, while the other $\tilde{K}_m(\tilde{\lambda})$ are the distinct polynomials

$$\begin{cases} \tilde{K}_1 = \tilde{\lambda} \\ \tilde{K}_2 = \tilde{\lambda} + \tilde{\lambda}^2/2! \\ \tilde{K}_3 = \tilde{\lambda} + 2\tilde{\lambda}^2/2! + \tilde{\lambda}^3/3! \\ \tilde{K}_4 = \tilde{\lambda} + 3\tilde{\lambda}^2/2! + 3\tilde{\lambda}^3/3! + \tilde{\lambda}^4/4! \\ \vdots \\ \tilde{K}_m = \sum_{l=1}^m \tilde{\Gamma}(l,m) \tilde{\lambda}^l \\ \vdots \end{cases}$$

given by the further modified Khayyam-Tartaglia recursion relation

$$\tilde{\Gamma}(l,m) = \tilde{\Gamma}(l-1,m-1)/l + \tilde{\Gamma}(l,m-1)$$

and where $\tilde{\lambda} = \alpha/(1+q)$.

For model C,

$$P_C(Q) = \delta_{1Q}$$

there is no parameter q (i.e., $q \equiv 1$), and one simply finds

$$C_m = e^{-\alpha} \frac{\alpha^m}{m!}. \quad (25)$$

Note that in the limit $q \rightarrow 0$, when the mean input per CA3 unit $\mu = \alpha q$ remains finite, for both models A and B one finds

$$\lim_{q \rightarrow 0} C_m = e^{-\mu} \frac{\mu^m}{m!}$$

which is equivalent to Eq. 25, in line with the fact that both models A and B reduce, in the $q \rightarrow 0$ limit, to single-field distributions, but even units with single fields become vanishingly rare, so formally one has to scale up the mean number of active presynaptic units, α , to keep $\mu \equiv \alpha q$ finite and establish the correct comparison to model C.

Sparsity and threshold

The analytical relation between the threshold T of CA3 units and the sparsity aa_{CA3} of the layer is obtained starting from the formula defining the sparsity a_{CA3} (see below) which can be rewritten

$$a = \left\langle \frac{[\sigma(\rho(\vec{x})) + \rho(\vec{x})\Phi(\rho(\vec{x}))]^2}{[\rho(\vec{x})\sigma(\rho(\vec{x})) + \Phi(\rho(\vec{x}))(1 + \rho^2(\vec{x}))]} \right\rangle \quad (26)$$

Since in the analytical calculation we have T as parameter, this equation can be taken as a relation $aa_{CA3}(T)$ which has to be inverted to allow a comparison with the simulations, which are run controlling the sparsity level at a predefined level (in our case $aa_{CA3} = 0.1$) and adjusting the threshold parameter accordingly. The inversion requires using the m -field decomposition and numerical integration. A graphical example of the numerical relation is given in Fig. 11.

Simulations

The mathematical model described above was simulated with a network of 15000 DG cells and 500 CA3 cells. A virtual rat explores a continuous two dimensional space, intended to represent a $1sqm$ square environment but realized as a torus, with periodic boundary conditions. For the numerical estimation of mutual information, the environment is discretized in a grid of 20×20 locations, whereas trajectories are in continuous space, but in discretized time steps. In each time step (intended to correspond to roughly $62.5ms$, half a theta cycle) the virtual rat moves half a grid unit ($2.5cm$) in a direction similar to the direction of the previous time step, with a small amount of noise. To allow construction of a full localization matrix with good statistics, simulations are run for typically 400,000 time steps (while for the simplified translationally invariant matrix 5,000 steps would be sufficient). The space has periodic boundary conditions, as in a torus, to avoid border effects; the longest possible distance between any two locations is hence equal to 14.1 grid units, or $70cm$.

DG place fields. After assigning a number of firing fields for each DG units, according to the distributions of models A, B and C, we assign to each field a randomly chosen center. The shape of the field is then given by a Gaussian bell with that center. The tails of the Gaussian function are truncated to zero when the distance

from the center is larger than a fixed radius $r = \sqrt{\frac{fA}{\pi}}$, with $f = 0.1$ the ratio between the area of the field and the environment area A . In the standard model, only about 3 percent of the DG units on average are active in a given environment, in agreement with experimental findings [24]; i.e. the DG firing probability is $p_{DG} = 0.033$. The firing of DG units is not affected by noise, nor by any further threshold. Peak firing is conventionally set, in the center of the field, at the value $\frac{r^2}{2\pi} = 2.02$, but DG units can fire at higher levels if they are assigned two or more overlapping fields.

CA3 activation. CA3 units fire according to Eq. 2: the firing of a CA3 unit is a linear function of the total incoming DG input, distorted by a noise term. This term is taken from a gaussian distribution centered on zero, with variance $\delta = 1$, and it changes for each unit and each time step. A threshold is imposed in the simulations to model the action of inhibition, hypothesizing that it serves to adjust the sparsity a_{CA3} of CA3 activity to its required value. The sparsity is defined as

$$a_{CA3} = \frac{\langle \eta_i(\vec{x}) \rangle^2}{\langle \eta_i^2(\vec{x}) \rangle}$$

and it is fixed to $a_{CA3} = 0.1$. This implies that the activity of the CA3 cells population is under tight inhibitory control.

The decoding procedure and information extraction. At each time step, the firing vector of a set of CA3 units is compared to all the average vectors recorded at each position in the 20×20 grid, for the same sample, in a test trial (these are called template vectors). The comparison is made calculating the Euclidean distance between the current vector and each template, and the position of the closest template is taken to be the decoded position at that time step, for that sample. This procedure has been termed

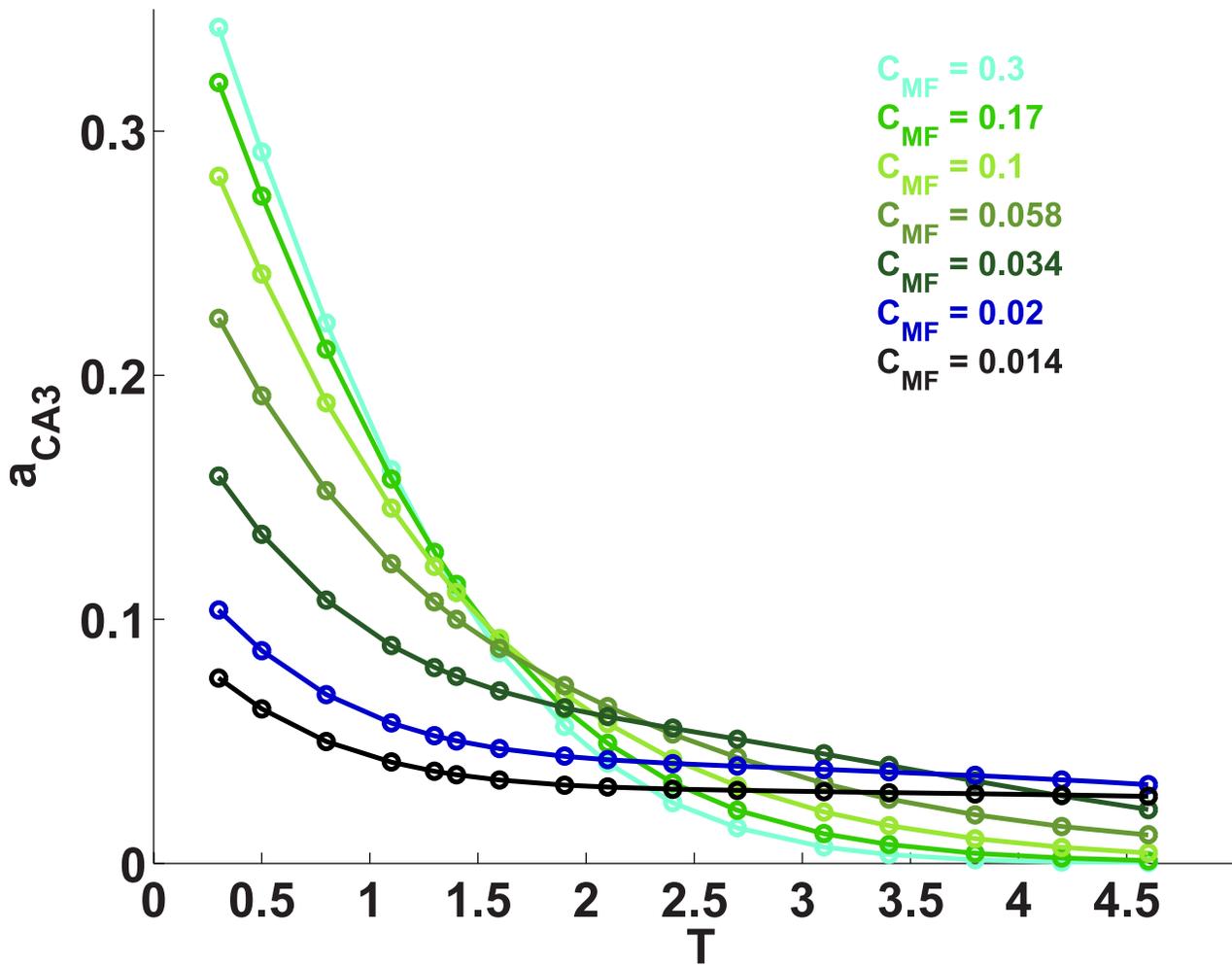


Figure 11. Sparsity-Threshold relation. The sparsity a of CA3 layer vs. the threshold T of CA3 units, from the numerical integration of Eq. 26. Different lines correspond to different degrees of connectivity between DG and CA3. doi:10.1371/journal.pcbi.1000759.g011

maximum likelihood Euclidean distance decoding [32]. The frequency of each pair of decoded and real positions are compiled in a so-called “confusion matrix”, or localization matrix, that reflects the ensemble of conditional probabilities $\{P(\{\eta_i\}|\vec{x})\}$ for that set of units. Should decoding “work” in a perfect manner, in the sense of always detecting the correct position in space of the virtual rat, the confusion matrix would be the identity matrix. From the confusion matrix obtained at the end of the simulation, the amount of information is extracted, and plotted versus the number of CA3 units present in the set. We averaged extensively over CA3 samples, as there are large fluctuations from sample to sample, i.e. for each given number of CA3 units we randomly picked several different groups of CA3 units and then averaged the mutual information values obtained. In all the results reported we averaged also over 3–4 simulation run with a different random number generator, i.e. over different trajectories. The same procedure leading to the information curve was repeated for different values of the parameters. In all the information measures we reported, we also corrected for the limited sampling bias, as discussed by [31]. In our case of spatial information, the bias is essentially determined by the spatial binning we used (20×20) and by the decoding method [52].

One should note the maximum likelihood decoding procedure to better understand the discrepancy between the information

estimated from simulations (with the procedure based on the full matrix) and that calculated analytically. The analytical calculation distinguishes in a clear-cut manner so called *annealed* variables, which are interpreted as “fast” noise and are averaged in computing the relation between position and neuronal activity, and so called *quenched* variables, which are interpreted as frozen disorder and are averaged over only later, in computing average the entropy, free-energy or mutual information [27]. In using maximum likelihood decoding, instead, the localization matrix that relates actual and decoding position effectively averages only trial-to-trial variability, i.e. the noise that occurs on intermediate time scales. The variability on genuinely fast time scales is suppressed, in fact, by the maximum likelihood operation, which acts as a sort of temporal low pass filter with a cut-off time equal to one time step. This suppression of part of the annealed noise leads to larger information values extracted from the simulations, and hence to the notion of “dark” information. In the real system, the spiking nature of neuronal activity may induce a similar cut-off, although its quantitative relation to the one-time-step cut-off in the simulations (here intended to be half a theta cycle) remains to be firmly established.

Fitting. We fit the information curves obtained in simulations to exponentially saturating curves as a function of N in order to get the values of the two most relevant parameter that describe

their shape: the initial slope I_1 (i.e. the average information conveyed by the activity of individual units) and the total amount of information I_∞ (i.e. the asymptotic saturation value). The function we used for the fit is the following

$$F(N) = I_\infty \left(1 - e^{-N \frac{I_1}{I_\infty}} \right) \quad (27)$$

In most cases the fit was in excellent agreement with individual data points, as expected on the basis of previous analyses [28].

References

- Leutgeb JK, Leutgeb S, Moser MB, Moser EI (2007) Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315: 91–966.
- Treves A, Tashiro A, Witter MP, Moser EI (2008) What is the mammalian dentate gyrus good for? *Neuroscience* 154: 155–1172.
- McNaughton BL, Morris RG (1987) Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences* 10: 48–415.
- Rolls E (1989) Functions of neuronal networks in the hippocampus and neocortex in memory. In: Byrne JH, Berry WO, eds. *Neural Models of Plasticity: Experimental and Theoretical Approaches*. San Diego: Academic Press. pp 240–265.
- Treves A, Rolls ET (1992) Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2: 19–199.
- Marr D (1971) Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society (London)* B 262: 2–31.
- Lassalle JM, Bataille T, Halley H (2000) Reversible inactivation of the hippocampal mossy fiber synapses in mice impairs spatial learning, but neither consolidation nor memory retrieval, in the Morris navigation task. *Neurobiology of Learning and Memory* 73: 23–257.
- Lee I, Kesner RP (2004) Encoding versus retrieval of spatial memory: Double dissociation between the dentate gyrus and the perforant path inputs into CA3 in the dorsal hippocampus. *Hippocampus* 14: 6–76.
- Leutgeb S, Leutgeb JK, Barnes CA, Moser EI, McNaughton BL, et al. (2005) Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science*. pp 619–623.
- Samsonovich A, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience* 17: 500–5920.
- Battaglia F, Treves A (1998) Attractor neural networks storing multiple space representations: A model for hippocampal place fields. *Phys Rev E* 58: 7738–7753.
- Stringer MS, Rolls ET (2002) Invariant object recognition in the visual system with novel views of 3D objects. *Neural Comput* 14: 285–2596.
- Tsodyks M, Sejnowski T (1995) Associative memory and hippocampal place cells. *International Journal of Neural Systems* 6: 8–86.
- Hamaguchi K, Hattett JP (2006) Analytic solution of neural network with disordered lateral inhibition. *Phys Rev E* 73.
- Papp G, Witter MP, Treves A (2007) The CA3 network as a memory store for spatial representations. *Learning and Memory* 14: 72–744.
- Roudi Y, Treves A (2008) Representing where along with what information in a model of a cortical patch. *PLoS Computational Biology* 4.
- Hafting T, Fyhn M, Molden S, Moser MB, et al. (2005) Microstructure of a spatial map in the entorhinal cortex. *Science* 436: 81–806.
- Jung MW, Wiener SI, McNaughton BL (1994) Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *Journal of Neuroscience* 14: 747–7356.
- Aimone JB, Wiskott L (2008) Computational modeling of neurogenesis. *Adult Neurogenesis* 52: 43–481.
- Treves A, Rolls ET (1991) What determines the capacity of autoassociative memories in the brain? *Network: Computation in Neural Systems* 2: 31–397.
- Wilson MA, BL M (1993) Dynamics of the hippocampal ensemble code for space. *Science* 261: 1055–1058.
- Treves A (1990) Graded-response neurons and information encodings in autoassociative memories. *Physical review A, General physics* 42: 218–2430.
- Kropff E, Treves A (2008) The emergence of grid cells: Intelligent design or just adaptation? *Hippocampus* 18: 156–1269.
- Chawla MK, Guzowski JF, Ramirez-Amaya V, Lipa P, Hoffman KL, et al. (2005) Sparse, environmentally selective expression of arc rna in the upper blade of the rodent fascia dentata by brief spatial experience. *Hippocampus* 15: 59–586.
- Si B, Treves A (2009) The role of competitive learning in the generation of DG fields from EC inputs. *Cognitive Neurodynamics* 3: 19–187.
- Amaral DG, Ishizuka N, Claiborne B (1990) Neurons, numbers and the hippocampal network. *Progress in brain research* 83: 1–11.
- Mezard M, Parisi G, Virasoro (1986) *Spin Glasses and Beyond* World Scientific.
- Samengo I, Treves A (2000) Representational capacity of a set of independent neurons. *Phys Rev E* 63: art. 011910.
- DelPrete, Treves A (2001) Theoretical model of neuronal population coding of stimuli with both continuous and discrete dimensions. *Phys Rev E* 64: 021912.
- Leutgeb S, Leutgeb JK, Treves A, Moser MB, et al. (2004) Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* 305: 195–1298.
- Treves A, Panzeri S (1995) The upward bias in measures of information derived from limited data samples. *Neural Comput* 2: 39–407.
- Rolls ET, Treves A, Tovee MJ (1997) The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. *Exp Brain Res*. pp 149–162.
- Nicoll RA, Schmitz D (2005) Synaptic plasticity at hippocampal mossy fibre synapses. *Nature Reviews Neuroscience* 6: 83–876.
- Treves A (2004) Computational constraints between retrieving the past and predicting the future, and the CA3-CA1 differentiation. *Hippocampus* 14: 539–556.
- Kuhn HG, Dickinson-Anson H, Gage FH (1996) Neurogenesis in the dentate gyrus of the adult rat: Age-related decrease of neuronal progenitor proliferation. *Journal of Neuroscience* 16: 227–2033.
- Redish AD, Battaglia FP, Chawla MK, Ekstrom AD, Gerrard JL, et al. (2001) Independence of firing correlates of anatomically proximate hippocampal pyramidal cell. *The Journal of Neuroscience* 21.
- Bakker A, Kirwan CB, Miller M, Stark CE (2008) Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*. pp 1640–1642.
- Panzeri S, Treves A (1996) Analytical estimates of limited sampling biases in different information measures. *Network: Computation in Neural Systems* 7: 8–107.
- Aimone JB, Wiles J, Gage FH (2006) Potential role for adult neurogenesis in the encoding of time in new memories. *Nature Neuroscience* 9: 73–727.
- Becker S (2005) A computational principle for hippocampal learning and neurogenesis. *Hippocampus* 15: 72–738.
- Wiskott L, Rasch MJ, Kempermann G (2006) A functional hypothesis for adult hippocampal neurogenesis: Avoidance of catastrophic interference in the dentate gyrus. *Hippocampus* 16: 39–343.
- Kee N, Teixeira CM, Wang AH, Frankland PW (2007) Preferential incorporation of adult-generated granule cells into spatial memory networks in the dentate gyrus. *Nature Neuroscience* 10: 35–362.
- Ge S, Yang C, Hsu K, Ming G, et al. (2007) A critical period for enhanced synaptic plasticity in newly generated neurons of the adult brain. *Neuron*. pp 559–566.
- Buzzetti RA, Marrone DF, Schaner MJ, Chawla MK, Bohanick JD, et al. (2007) Do dentate gyrus granule cells tag time-specific experiences? *Soc Neurosci Abstr* 744: 16.
- Tashiro A, Makino H, Gage FH (2007) Experience-specific functional modification of the dentate gyrus through adult neurogenesis: A critical period during an immature stage. *The Journal of Neuroscience* 27: 352–3259.
- Cameron HA, Mckay RD (2001) Adult neurogenesis produces a large pool of new granule cells in the dentate gyrus. *The Journal of Comparative Neurology* 435: 406–417.
- McDonald HY, Wojtowicz JM (2005) Dynamics of neurogenesis in the dentate gyrus of adult rats. *Neuroscience Letters* 385: 7–75.
- McHugh TJ, Jones MW, Quinn JJ, Balthasar N, Coppari R, et al. (2007) Dentate gyrus nmda receptors mediate rapid pattern separation in the hippocampal network. *Science* 317: 9–99.
- Nadal J, Parga N (1993) Information processing by a perceptron in an unsupervised learning task. *Network: Computation in Neural Systems*. pp 295–312.
- Treves A (1995) Quantitative estimate of the information relayed by the schaffer collaterals. *Journal of Computational Neuroscience*. pp 259–272.
- DelPrete, Treves A (2002) Replica symmetric evaluation of the information transfer in a two-layer network in the presence of continuous and discrete stimuli. *Phys Rev E* 65: 041918.
- Panzeri S, Treves A, Schultz S, Rolls ET (1999) On decoding the responses of a population of neurons from short time windows. *Neural Comput* 11: 153–1577.

Acknowledgements

We had valuable discussion with Jill Leutgeb, Bailu Si and Federico Stella.

Author Contributions

Analyzed the data: EC. Wrote the paper: EC AT. Carried out the mathematical analysis of the model and the computer simulations: EC. Designed the model, participated in its analysis and supervised the simulations: AT.