Resolution of Nested Neuronal Representations Can Be Exponential in the Number of Neurons

Alexander Mathis, Andreas V. M. Herz, and Martin B. Stemmler

*Bernstein Center for Computational Neuroscience, Graduate School of Systemic Neurosciences,
Ludwig-Maximilians-Universität München, Grosshadernerstrasse 2, D-82152 Martinsried, Germany*
(Received 20 September 2011; published 6 July 2012)

Collective computation is typically polynomial in the number of computational elements, such as transistors or neurons, whether one considers the storage capacity of a memory device or the number of floating-point operations per second of a CPU. However, we show here that the capacity of a computational network to resolve real-valued signals of arbitrary dimensions can be exponential in N, even if the individual elements are noisy and unreliable. Nested, modular codes that achieve such high resolutions mirror the properties of grid cells in vertebrates, which underlie spatial navigation.

The brain encodes many stimulus variables, such as the position of an object, the orientation of an edge, or the frequency of a sound, with high precision. Multidimensional stimulus spaces are represented by joint activity across neurons, each of which fires noisy, unreliable spikes. Yet, despite the stochastic nature of neuronal discharge, the nervous system achieves a highly efficient representation of the outside world [1].

The simplest representation of a continuous stimulus variable is a one-to-one map onto a neuronal firing rate. Given unreliable spikes, however, a labeled line code across different neurons is more robust and efficient [2]. The place code for spatial position in the hippocampus is an instance of such a labeled line code. One drawback of such a code is that the resolution only scales linearly in the number of neurons [2–4].

A place code can be improved upon by using a cascade of self-similar, periodic representations at different scales, as depicted in Fig. 1. Each successive level refines the representation at the previous coarser scale, such that the overall resolution scales exponentially in the number of neurons, as we show in this Letter.

Neuronal coding of sensory information at multiple scales occurs in many brain areas [5] and arises naturally in the theory of sparse coding [6]; we show here how the dense coding at multiple scales found in the entorhinal cortex and related areas, where each neuron has multiple firing fields, can be highly efficient, even though the firing rate of a single neuron no longer maps onto a single stimulus, but to many possible stimuli.

We consider N statistically independent neurons encoding a compact but possibly high-dimensional stimulus space, normalized to $[0, 1]^D$. For simplicity, each stimulus is assumed to be equally likely. Each neuron's response is characterized by the number of spikes k_i emitted within a time τ after stimulus onset. The neuron's mean firing rate depends on the stimulus x through its tuning curve $\Omega_i(x)$. The number of spikes is stochastic, so that observing a response $K = (k_1, \ldots, k_N)$ across the population of neurons has a probability

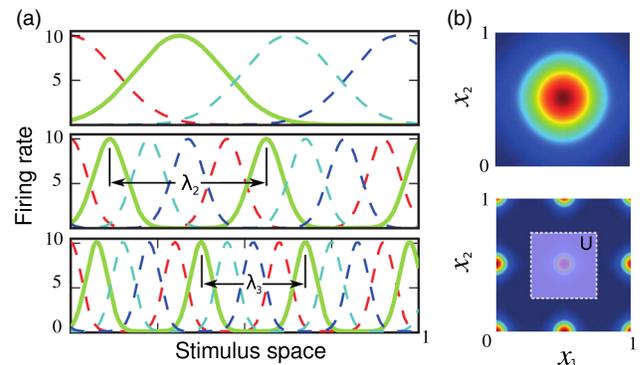$$P(K|x) = \prod_{i=1}^{N} P_i(k_i|\tau\Omega_i(x)). \tag{1}$$

FIG. 1 (color online). Example of nested modules. (a) All modules, except for the coarsest one, have periodic tuning curves $\Omega_i(x - c_i)$. A module consists of a set of tuning curves with the same period but different phases c_i. The spatial period for modules 2 and 3 are $\lambda_2 = 0.45$ and $\lambda_3 = 0.3$, respectively. In each module, we highlight a single tuning curve by a solid line to show the period. Shifted but otherwise identical tuning curves are dashed. Nested modules successively refine the representation of the stimulus. Periodicity implies that the map from stimulus to population response is not one-to-one within a single module. Only the ensemble response provides a unique representation of x. (b) A unimodal tuning curve Ω in two dimensions, shown at the top, can be rescaled and periodically extended using Eqs. (5) and (6). The periodic tuning curves Ω_Γ in the lower panel is based on a rectangular lattice Γ spanned by $v_1 = \lambda(1, 0)'$ and $v_2 = \lambda(0, 1)'$, with $\lambda = 1/2$. For this lattice, a fundamental domain U is depicted.

We ask how accurately an ideal observer can deduce the stimulus x from K. For this purpose, we use the Fisher information matrix $\boldsymbol{J}$ with components

$$\boldsymbol{J}_{\alpha\beta}(x) = \int \frac{\partial \ln P(K, x)}{\partial x_\alpha} \frac{\partial \ln P(K, x)}{\partial x_\beta} P(K, x) dK, \quad (2)$$

where $\alpha, \beta \in \{1, \cdots, D\}$. The Cramér-Rao inequality [7] relates the inverse of $\boldsymbol{J}$ to the covariance matrix Σ of any unbiased estimator $\hat{x}$ of the encoded stimulus x,

$$\Sigma(\hat{x}|x) \geq \boldsymbol{J}(x)^{-1}. \quad (3)$$

Let us first consider a population of neurons with tuning curves $\Omega_i(x)$ such that each has a single peak of width σ. Suppose the tuning curves differ only in the center c_i of the peak, so that $\Omega_i(x) = \Omega(x - c_i)$. A single tuning curve has Fisher information $\boldsymbol{J}_{\Omega_i}(x)$. For the stochastic population model in Eq. (1), the Fisher information $\overline{\boldsymbol{J}_\Omega(x)}$ for the population is simply the sum

$$\overline{\boldsymbol{J}_\Omega(x)} \overset{\text{def}}{=} \sum_{i=1}^N \boldsymbol{J}_{\Omega_i}(x) = \int_{[0,1]^D} \boldsymbol{J}_\Omega(x - \varphi)\rho(\varphi)d\varphi,$$

where $\rho(\varphi) = \sum_{i=1}^N \delta(\varphi - c_i)$ is the density of the centers.

Assume that the centers c_i are uniformly distributed in $[0, 1]^D$, so that the tuning curves cover the entire stimulus space. As the centers become increasingly dense with increasing N, the Fisher information $\overline{\boldsymbol{J}_\Omega(x)}$ becomes independent of the specific stimulus x and scales linearly in the number of neurons

$$\overline{\boldsymbol{J}_\Omega} = N K_\Omega(\sigma, \tau, D) \quad (4)$$

for some function K_Ω [3,8,9]. If the stimulus space were not compact, but instead encompassed all of $\mathbb{R}^D$, then $K_\Omega \sim \sigma^{D-2}$ [3].

The linear scaling of Eq. (4) in N can be dramatically improved by switching to a nested modular code, in which each module consists of a subpopulation of M periodic tuning curves, as in Fig. 1. Each module is associated with a unique spatial period.

We first compute the Fisher information for a single module that has periodic tuning curves, such that $M = N$. Any unimodal tuning curve $\Omega(x)$ on $[0, 1]^D$ can be periodically extended. Let $\Gamma \subset \mathbb{R}^D$ be a nondegenerate, affine point lattice [10]

$$\Gamma = u + \sum_{\alpha=1}^D k_\alpha v_\alpha \quad \text{for } k_\alpha \in \mathbb{Z}, \quad u, \quad v_\alpha \in \mathbb{R}^D, \quad (5)$$

such that $(v_\alpha)_{1 \leq \alpha \leq D}$ is a basis for $\mathbb{R}^D$ and u a center. Let $U \subset \mathbb{R}^D$ be a fundamental domain of this lattice. Then there is a canonical coordinate transformation $\Phi: U \rightarrow [0, 1)^D$ in terms of an invertible matrix $\boldsymbol{T}$ and a vector w, such that $\Phi = -w + \boldsymbol{T}$. One defines the periodic extension of Ω as

$$\Omega_\Gamma: \mathbb{R}^D \rightarrow \mathbb{R}^+, \qquad x \mapsto \Omega \circ \Phi(x \bmod \Gamma). \quad (6)$$

This definition is illustrated in Fig. 1(b) and is independent of the particular representation (or translational shift) for U.

A family of shifted, periodic tuning curves on the lattice Γ constitutes a module and is associated with Fisher information $\overline{\boldsymbol{J}_{\Omega,\Gamma}}$. We now relate $\overline{\boldsymbol{J}_{\Omega,\Gamma}}$ to the original $\overline{\boldsymbol{J}_\Omega}$. Under the inverse map Φ^{-1}, the transformed tuning curves have centers $c_i' = \Phi^{-1}(c_i) \in U$. Therefore,

$$\overline{\boldsymbol{J}_{\Omega,\Gamma}(x)} \overset{\text{def}}{=} \sum_{i=1}^N \boldsymbol{J}_{\Omega_\Gamma}(x + c_i') = \int_U \boldsymbol{J}_{\Omega_\Gamma}(x + \varphi)\rho'(\varphi)d\varphi,$$

where $\rho'(\varphi) = \sum_{i=1}^N \delta(\varphi - c_i')$. By changing the variables in the Fisher information [7] and using the periodicity of $\Phi((x + \varphi) \bmod \Gamma)$ on U, we get

$$\overline{\boldsymbol{J}_{\Omega,\Gamma}(x)} = \int_U \boldsymbol{T} \boldsymbol{J}_\Omega(\Phi(x + \varphi))\boldsymbol{T}'\rho'(\varphi)d\varphi$$

$$= \boldsymbol{T} \int_U \boldsymbol{J}_\Omega(\Phi(x + \varphi))\rho'(\varphi)d\varphi\boldsymbol{T}'.$$

In the last step, we interchanged the multidimensional integration and matrix multiplication, as T and its transpose T' are independent of φ. Under the map Φ, we obtain

$$\overline{\boldsymbol{J}_{\Omega,\Gamma}(x)} = \boldsymbol{T} \int_{[0,1]^D} \boldsymbol{J}_\Omega(x + y) \cdot \rho(y)dy\boldsymbol{T}'$$

$$= \boldsymbol{T}\left(\sum_{i=1}^N \boldsymbol{J}_\Omega(x + c_i)\right)\boldsymbol{T}' = \boldsymbol{T}\overline{\boldsymbol{J}_\Omega(x)}\boldsymbol{T}'. \quad (7)$$

Thus, we have derived the following rule: $\overline{\boldsymbol{J}_{\Omega,\Gamma}(x)} = \boldsymbol{T}\overline{\boldsymbol{J}_\Omega(x)}\boldsymbol{T}'$. For an orthogonal lattice Γ defined by $v_\alpha = \lambda_\alpha e_\alpha$ on the canonical basis $(e_\alpha)_{1 \leq \alpha \leq D}$ of $\mathbb{R}^D$, each entry in the Fisher information matrix is rescaled by $(\lambda_\alpha \lambda_\beta)^{-1}$ under this transformation. If the original Fisher information matrix is diagonal, then so is $\overline{\boldsymbol{J}_{\Omega,\Gamma}(x)}$. Therefore, rescaling the periodic tuning curves by a factor of 2 quadruples the Fisher information, but at the cost of introducing ambiguity—the value of x can only be recovered modulo the lattice Γ. Resolving this ambiguity requires a multiscale representation consisting of multiple modules spanning different spatial periods.

How should these different spatial periods be chosen? Suppose there are L modules with spatial periods $\{\lambda_1, \ldots, \lambda_L\}$, arranged in decreasing order from the longest period to the shortest, as in Fig. 1. Each module has $M = N/L$ tuning curves, and each tuning curve within a module is associated with a different phase shift.

The easiest case to analyze is the one in which the lattices Γ_k are orthogonal [11], the rescaling is uniform in each dimension, and the tuning curve Ω is radially symmetric. In such a case, a module on lattice Γ_k has a Fisher information matrix $\overline{\boldsymbol{J}_{\Omega,\Gamma_k}(x)} = \lambda_k^{-2}\overline{\boldsymbol{J}_\Omega}$, and radial symmetry implies that $\overline{\boldsymbol{J}_\Omega} = J\mathbf{I}$ is diagonal and proportional to the identity matrix $\mathbf{I}$ [3]. To ensure that the first module represents $x \in [0, 1]^D$ unambiguously, we treat this module as a special case and make it aperiodic;
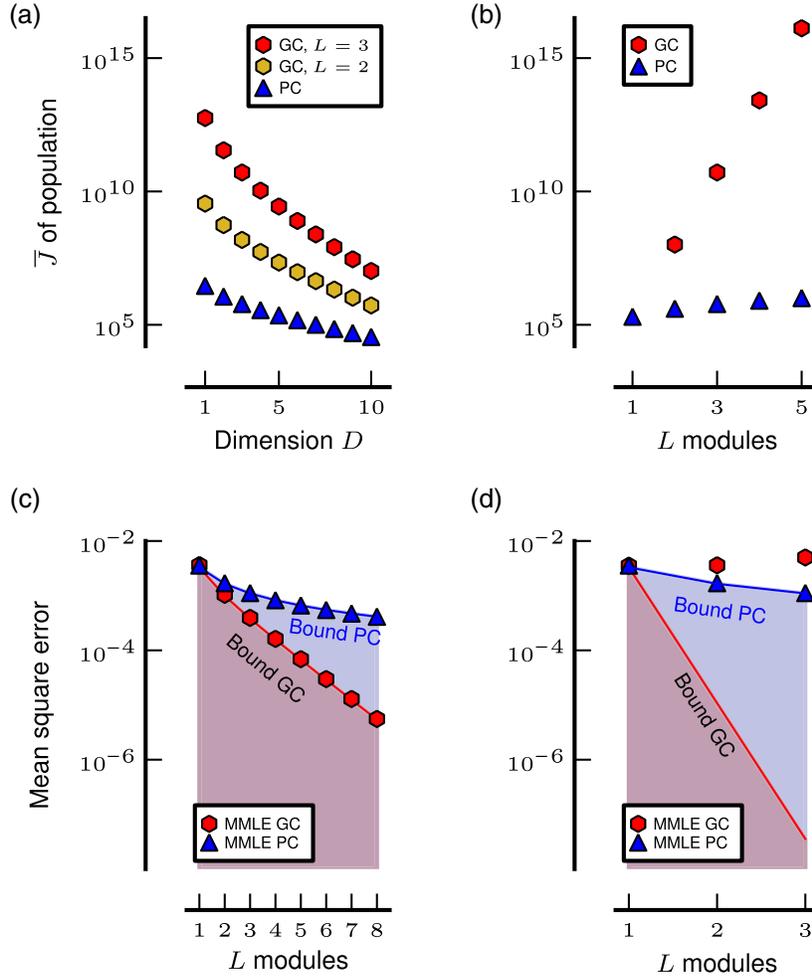
FIG. 2 (color online). (a) Grid codes (GC) outperform place codes (PC), regardless of the number of stimulus dimensions. A population of $N = 3 \times 10^5$ neurons is divided into one, two, or three modules, according to Eq. (8). The neurons' tuning width is fixed as $\sigma = 2$. (b) The Fisher information for a neuronal population with $M = 10^5$ neurons per module and $D = 3$. For a nested modular code with L modules, the Fisher information grows exponentially in $N = L \cdot M$, whereas it is linear in N for a place code. (c) The error of the estimator that minimizes $(x - \hat{x})^2$ for place and grid codes in $D = 3$ dimensions, based on sampling the stochastic response 1200 times. Each module comprises $M = 8^3$ equidistantly spaced cells. The lattice lengths for different modules are scaled according to Eq. (8), with a safety factor $C = 20$. The Cramér-Rao bound of Eq. (10) is tight for both the grid and place codes. (d) When the lattice lengths contract more strongly than allowed by Eq. (8), such that $C = 1$, the error fails to improve in a nested modular code. Although the Fisher information predicts an error even lower than in (c), the uncertainty derived from the first module's response is larger than the lattice length scale of the next finer-grained module, so that adding modules with finer spacing does not improve the resolution.

i.e., the first module is a place code with the same tuning width that a periodic module with $\lambda = 1$ would have.

Within the first module, the expected error in an unbiased estimate of x asymptotically approaches $1/\sqrt{J}$, according to Eq. (3). This error sets a lower bound on the period of the next module that refines the representation of x. Hence, each λ_k should obey

$$\lambda_{k+1} = \frac{C\lambda_k}{\sqrt{J}}, \qquad (8)$$

where C is a safety factor, such that $1 \ll C < \sqrt{J}$. For a C larger than unity, the next module can correct for the error in the previous module. The Fisher

information for the nested population is $\overline{J_{\Omega,\Gamma_1\cdots\Gamma_L}} = \overline{J_{\Omega,\Gamma_1\cdots\Gamma_L}}I$ with

$$\overline{J_{\Omega,\Gamma_1\cdots\Gamma_L}} = \sum_{k=1}^{L} \frac{J}{\lambda_k^2} = \sum_{k=1}^{L} \frac{J^k}{C^{2(k-1)}}.$$

If we take only the last term in the series and substitute $L = N/M$, we see that

$$\overline{J_{\Omega,\Gamma_1\cdots\Gamma_L}} > \frac{J^{N/M}}{C^{2(N/M-1)}}. \qquad (9)$$

For fixed module size M, the Fisher information scales exponentially in the number of neurons N. Such a coding scheme, therefore, outperforms a single module that only

scales linearly in N. Note that this scaling is independent of the dimension D.

As a concrete example, we consider a set of tuning curves with Poisson noise, centers c_i, tuning width σ, and period λ, as given by $\Omega_i(x) = f_{max} \exp(\frac{1}{\sigma^2} \sum_{\alpha=1}^{D} \{\cos[2\pi(x_\alpha - c_{i,\alpha})/\lambda] - 1\})$ [12]. If the number of neurons per period λ is $M \gg 1$ and the centers are uniformly distributed, the module's average Fisher information is given by

$$\overline{J} = \frac{M 4\pi^2 f_{max}\tau}{\lambda^2 \sigma^2} K_1(\sigma^2) K_0^{D-1}(\sigma^2) \mathbf{I} \qquad (10)$$

where $K_n(x) = \exp(-1/x)I_n(-1/x)$ and $I_n(x)$ is the nth order modified Bessel function of the first kind. In Fig. 2(a), the Fisher information for a large population of $N = 300\,000$ (place) cells is plotted for stimulus dimensions $D = 1$ to 10. Dividing the population into separate grid modules according to Eq. (10) with $C = 20$ leads to a much higher Fisher information—orders of magnitude, irrespective of the dimension D of the stimulus space. Figure 2(b) underscores the key finding of this Letter: the Fisher information grows exponentially in the number N of encoding neurons. The place code, in contrast, is linear in N. To corroborate these analytical results, we sampled the response K and estimated the minimum mean square error based on the posterior probability distribution $p(x|K)$ [13], as shown in Figs. 2(c) and 2(d). These simulations show that the Cramér-Rao bound is tight as long as the grid codes obey the constraint in Eq. (8). For place codes, it is known that on short time scales [14], or for low numbers of neurons [9], the Cramér-Rao bound will not be tight, so that the Fisher information underestimates the error in decoding the signal. The same will hold true for grid codes. However, when $N > M \gg 1$ and the expected number of spikes at the center of each tuning curve is appreciable, a nested modular code leads to an error that scales as $M^{-N/M}$.

Note that we assume that the firing of neurons is uncorrelated. Whether this assumption holds in cortex is a matter of fierce debate [15]. The Fisher information deteriorates with increasing noise correlations, but its scaling in N does not, at least not for the correlation strengths measured by Ecker *et al.* in cortex.

Periodic tuning curves have been found in entorhinal cortex of rodents—coined grid cells ([16]; see Supplemental Material [17]). This unexpected discovery has inspired theorists to explore the combinatorial capacity of modular periodic codes and how they might be used in the brain [18]. In some cases, the stimulus space is intrinsically periodic—orientation and color hue are but two examples. But when the space of stimulus x is infinite instead of periodic, different spatial periods can be combined to encode a much larger range of x uniquely than would otherwise be possible [19]. Indeed, the exponential range that can result confers a relative precision that is also exponential [20]. This Letter, in contrast, shows that the

absolute precision in x can be exponential in N. Precision is of paramount importance for path integration, for which the mammalian brain is thought to use grid cells [21]. Interestingly, the periodic lattices for neighboring grid cells share similar spatial periods and orientations, but are spatially translated relative to each other [16]. Moreover, along the dorso-ventral axis of the entorhinal cortex, the typical spatial period of the lattice grows from roughly 20 cm to several meters, while the ratio of grid field width to spatial period remains constant [22]. Our theoretical analysis indicates that these grid cell properties may endow the brain with a highly accurate representation of space; the same principles might be used for representing other continuous, high-dimensional stimuli.

[1] H. B. Barlow, *The Mechanisation of Thought Processes* (Her Majesty's Stationery Office, London, 1959), p. 535; J. J. Atick, Network **3**, 213 (1992).

[2] M. Bethge and K. Pawelzik, Neurocomputing **44-46**, 323 (2002).

[3] K. Zhang and T. Sejnowski, Neural Comput. **11**, 75 (1999).

[4] If the width of each neuron's tuning curve is simultaneously scaled as N^{-1}, the net resolution scales as N^2.

[5] D. H. Hubel and T. N. Wiesel, J. Physiol. **195**, 215 (1968); R. L. De Valois, D. G. Albrecht, and L. G. Thorell, Vision Res. **22**, 545 (1982); M. S. Lewicki, Nat. Neurosci. **5**, 356 (2002).

[6] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, IEEE Trans. Inf. Theory **38**, 587 (1992); B. Olshausen and D. Field, Vision Res. **37**, 3311 (1997).

[7] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (Springer, New York, 1998), 2nd ed.

[8] W. M. Brown and A. Bäcker, Neural Comput. **18**, 1511 (2006).

[9] C. W. Eurich and S. D. Wilke, Neural Comput. **12**, 1519 (2000).

[10] We use the more generalized concept of an affine point lattice, in which u is an arbitrary affine translation.

[11] The rescaling rule of Eq. (8) becomes a matrix equation in the general case. Let $\mathbf{\Lambda}_i$ be a diagonal matrix of scaling factors for the point lattice Γ. If the tuning curves form a complete cover for the unit hypercube $[0, 1]^D$, then the Fisher information is a positive-definite symmetric matrix with no zero eigenvalues. Therefore, there is an orthonormal matrix $\mathbf{O}$ and a positive definite diagonal matrix $\mathbf{D}$ such that $\mathbf{J} = \mathbf{ODO}'$. By the Cramér-Rao bound, $\Sigma \geq \mathbf{J}^{-1}$. Hence, we define the following refinement rule: $\mathbf{\Lambda}_{k+1} = C(\mathbf{OD}^{-1/2}\mathbf{O}')\mathbf{\Lambda}_k$.

[12] Similar tuning curves were used by A. Montemurro and S. Panzeri, Neural Comput. **18**, 1555 (2006).

[13] S. Yaeli and R. Meir, Front. Comput. Neurosci. **4**, 130 (2010).

[14] M. Bethge, D. Rotermund, and K. Pawelzik, Neural Comput. **14**, 2317 (2002); P. Berens, A. S. Ecker, S. Gerwinn, A. S. Tolias, and M. Bethge, Proc. Natl. Acad. Sci. U.S.A. **108**, 4423 (2011).

[15] M. A. Smith and A. Kohn, J. Neurosci. **28**, 12591 (2008);
A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K.
Logothetis, and A. S. Tolias, Science **327**, 584
(2010).

[16] T. Hafting, M. Fyhn, S. Molden, M-B Moser, and E. I.
Moser, Nature (London) **436**, 801 (2005); F. Sargolini, M.
Fyhn, T. Hafting, B. L. McNaughton, M. P. Witter, M-B
Moser, and E. I. Moser, Science **312**, 758 (2006); C. N.
Boccara, F. Sargolini, V. H. Thoresen, T. Solstad, M. P.
Witter, E. I. Moser, and M-B Moser, Nat. Neurosci. **13**,
987 (2010).

[17] See Supplemental Material at http://link.aps.org/
supplemental/10.1103/PhysRevLett.109.018103 for a
more detailed discussion of grid cells.

[18] A. Guanella and P. Verschure, J. Integ. Neurosci. **6**, 433
(2007).

[19] I. R. Fiete, Y. Burak, and T. Brookings, J. Neurosci. **28**,
6858 (2008).

[20] S. Sreenivasan and I. Fiete, Nat. Neurosci. **14**, 1330
(2011).

[21] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser,
and M-B Moser, Nat. Rev. Neurosci. **7**, 663 (2006).

[22] M. Fyhn, S. Molden, M. P. Witter, E. I. Moser, and M-B
Moser, Science **305**, 1258 (2004); L. M. Giocomo, E. A.
Zilli, E. Fransén, and M. E. Hasselmo, Science **315**, 1719
(2007); V. H. Brun, T. Solstad, K. B. Kjelstrup, M. Fyhn,
M. P. Witter, E. I. Moser, and M-B Moser, Hippocampus
18, 1200 (2008).