Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Neurobiology

Periodic population codes: From a single circular variable to higher dimensions, multiple nested scales, and conceptual spaces

Andreas VM Herz[1], Alexander Mathis[2,3] and Martin Stemmler[1]

Across the nervous system, neurons often encode circular stimuli using tuning curves that are not sine or cosine functions, but that belong to the richer class of von Mises functions, which are periodic variants of Gaussians. For a population of neurons encoding a single circular variable with such canonical tuning curves, computing a simple population vector is the *optimal* read-out of the most likely stimulus. We argue that the advantages of population vector read-outs are so compelling that even the neural representation of the outside world's flat Euclidean geometry is curled up into a torus (a circle times a circle), creating the hexagonal activity patterns of mammalian grid cells. Here, the circular scale is not set *a priori*, so the nervous system can use multiple scales and gain fields to overcome the ambiguity inherent in periodic representations of linear variables. We review the experimental evidence for this framework and discuss its testable predictions and generalizations to more abstract grid-like neural representations.

Addresses
[1] Bernstein Center for Computational Neuroscience Munich and Faculty of Biology, Ludwig-Maximilians-Universität München, Grosshadernerstrasse 2, 82152 Planegg-Martinsried, Germany
[2] Department of Molecular and Cellular Biology and Center for Brain Science, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA
[3] Werner Reichardt Centre for Integrative Neuroscience and Institute for Theoretical Physics, University of Tübingen, 72076 Tübingen, Germany

Corresponding author: Herz, Andreas VM (herz@bio.lmu.de)

Current Opinion in Neurobiology 2017, **46**:99–108

This review comes from a themed issue on **Computational neuroscience**

Edited by **Adrienne Fairhall** and **Christian Machens**

http://dx.doi.org/10.1016/j.conb.2017.07.005

Introduction

Angular variables are of key importance for sensory and motor systems — they describe the rotation of a rigid body part around a joint, the orientation of a visual stimulus projected onto the retina, or one's own movement direction relative to some external landmark. Reflecting the periodic nature of angular variables, their neural representation is periodic, too. However, periodic tuning curves may also result from the way neural responses are measured. For instance, the oriented receptive field of a neuron in visual cortex is reduced to a periodic tuning curve when the cell's response is measured as a function of angle. More surprisingly, though, the nervous system uses periodic representations for spatial navigation [1] and conceptual categorization [2••], two tasks involving variables that are not periodic in nature. The observed grid-like codes are fascinating, but what are their representational and computational merits?

The ambiguity inherent to any periodic representation of non-periodic variables, such as spatial position, confounds decoding at the single-neuron level — if multiple positions are mapped onto the same value of the internal coding variable, there is simply no way to recover the one true position. The same holds for 'mixed' neural representations of cognitive task variables [3,4]. Here, several stimulus attributes or even multimodal inputs drive the same measure of the neuronal response, such as the firing rate. Mixed selectivity stands in contrast to the concept of multiplexing, in which a neuron might represent the intensity of a visual stimulus in the firing rate, and another stimulus attribute, such as the stimulus orientation, in the latency of the response.

At the population level, representations of neurons with mixed selectivity can be read out linearly and efficiently: although every neuron carries ambiguous information, different neurons encode different stimulus combinations so that each individual stimulus triggers a unique population-level response [3–5]. Similarly, the non-uniqueness of grid codes at the single-grid scale might be resolved by pooling information across multiple scales. As we will argue in this review, simple readouts of grid codes based on canonical tuning curves are indeed possible. Key ingredients are two mechanisms long known from motor and sensory neuroscience: population-vector decoding [6] and gain modulation [7].

Decoding circular variables in one dimension

Because of noise intrinsic to the nervous system, neurons never respond the same way twice [8–10]. Accordingly, the population response $\boldsymbol{n} = (n_1, n_2, \ldots, n_N)$ of an ensemble of N neurons is statistical in nature. It occurs with a likelihood $P(\boldsymbol{n}|x)$ — the conditional probability of the response $\boldsymbol{n}$ given the input x, which might describe motor actions or sensory signals. Optimal neural inference consists of computing the most likely input from the noisy response, i.e., the 'maximum likelihood estimate' x_{ML}. As $P(\boldsymbol{n}|x)$ depends on the neural firing statistics and tuning curves, determining x_{ML} requires elaborate calculation.

The challenge for down-stream neurons (and external observers) is even more demanding: to infer the input x solely from the noisy population response $\boldsymbol{n}$. Here, the observer or down-stream neural system need to maximize the posterior probability $P(x|\boldsymbol{n})$ over all possible x-values for given $\boldsymbol{n}$ so as to obtain the 'maximum *a posteriori* estimate' x_{MAP}. It is a question of debate how neural systems solve this fundamental challenge [11–14].

For continuous circular variables, on the other hand, there exists an intuitive stimulus estimate, the population vector (PV), which weighs the response of each neuron by its preferred stimulus direction x (Box 1 and Figure 2e). First proposed as a coding mechanism for motor cortex, the PV is linear, robust and computable by a linear network [11,15–17]. Yet the PV remains controversial, even for motor cortex, as it is ill-suited to describe the time-varying kinematics of motor actions [18] and is quite sensitive to the uniformity of the distribution of preferred directions [19] and to the nature of noise correlations in the population [20]. The consensus view holds that PV decoding might be 'good enough', but rarely perfect [12,13,21,22•]. In fact, for a population with cosine or various other unimodal tuning curves and Poisson spike statistics, the PV is strictly suboptimal.

Surprisingly, though, there are canonical tuning curves and conditions under which the PV is the *optimal* decoder for Poisson statistics [23]. These canonical tuning curves are von Mises functions, which are exponentials of a sinusoid (see Box 1). Compared to sine or cosine functions, von Mises functions have an additional parameter that controls the tuning width. Figure 1 shows typical von Mises fits for orientation tuning in V1 [24], head-direction tuning in the anterior thalamic nucleus [25,26], reaching direction in motor cortex [27], and one-dimensional slices through 2D spatial firing fields of grid cells in medial entorhinal cortex [28]. These data and quantitative analyses, e.g., [27], suggest that, apart from their theoretical appeal, von Mises functions readily capture the essence of circular tuning in many cases.

In addition, von Mises functions are highly appealing from a theoretical point of view: For statistically independent Poisson neurons with von Mises tuning (see Box 1 for detailed mathematical definitions), one finds [23] that

(1) the most likely stimulus can be directly read out from the PV, $x_{ML} = x_{PV}$,
(2) the uncertainty in x_{ML} is given by the inverse PV length,
(3) the likelihood $P(\boldsymbol{n}|x)$ is von Mises, too, and uniquely fixed by PV length and direction,
(4) the expected PV length equals the average Fisher information, up to a fixed constant.

If the prior stimulus distribution $P(x)$ is flat, then Bayes' rule states that the posterior probability $P(x|\boldsymbol{n}) \sim P(\boldsymbol{n}|x)$. In this case, the prediction for the maximum *a posteriori* probability stimulus x_{MAP} is identical to the maximum likelihood stimulus x_{ML} so that the first two results apply to x_{MAP}, too, and similarly the third result is also true for $P(x|\boldsymbol{n})$. Von Mises functions are thus not only advantageous when it comes to fitting experimentally measured circular tuning characteristics but also to improve the interpretation of these data within a sound theoretical framework.

Just as each neuron's response is a random variable, so is the PV. As such, the PV fluctuates from trial to trial, so that the uncertainty in x_{ML} varies, too. A reliable population response, though, could simplify downstream processing. The coefficient of variation in the Fisher information (PV length) is smallest when the tuning curve's concentration parameter κ (see Box 1) is around 2.5 (Figure 1e). The minimum is broad, so that values $2 < \kappa < 5$ are close to optimal (Figure 1a,c,d); these values for κ correspond to an orientation tuning width of 30° to 50°. In contrast, maximizing the information predicts narrower tuning widths [29,30].

Decoding linear variables in one dimension by coiling up stimulus space

Many stimulus variables, such as the pitch of a pure sound, the wave-length of a light source or the position of an animal, are not circular but linear variables. Do such stimuli require an entirely different neural representation or could periodic population codes still be used? Figure 2 demonstrates that this is indeed possible. However, there is a price to be paid: the many-to-one mapping (Figure 2a) from a straight line to a circle does not have a unique inverse mapping. The same is true in higher dimensions (Figure 2b). To overcome this fundamental problem, multiple periodic representations with different spatial scales λ need to be combined (Figure 2c). For PV-decoding, multiple neurons are required at each scale, which thereby predicts the existence of distinct *modules*. Within one module, neuronal tuning curves must share the same spatial period but different curves will be phase shifted relative to each other. If no common factor exists that would divide the different modular scales evenly, the coding range is potentially very large [31].

Coiling up stimulus space for the sake of a periodic neural representation may seem costly in terms of neural hardware. Information theoretical analyses show, however,