Representation of uncertainty during hippocampal theta sequences

Balázs B Ujfalussy (ujfalussy.balazs@koki.mta.hu)
Laboratory of Neuronal Signaling, MTA Institute of Experimental Medicine,
Budapest, Hungary

Márton Kis (kis.marton@koki.mta.hu)
Laboratory of Neuronal Signaling, MTA Institute of Experimental Medicine,
Budapest, Hungary

Gergő Orbán (orban.gergo@wigner.mta.hu)
Computational Systems Neuroscience Lab, MTA Wigner Research Center for Physics
Budapest, Hungary

Abstract

Animals are able to perform probabilistic computations implying that the nervous system is capable of the representation and manipulation of probability distributions. However, the way encoded distributions are related to population activity of neurons remains unknown because measures that could dissociate alternative models based on experimental data are remarkably lacking. Here, we focus on hippocampal activity during exploratory behavior, where the place cell activtions outline the trajectory of the animal starting from past towards future positions during each theta cycle (theta sequences). Critically, during a single theta sequence the uncertainty is expected to change systematically, thus providing an opportunity to identify how it is encoded in the population activity. We derived contrasting predictions for four alternative models: (a) encoding the most likely trajectory; (b) sampling from the posterior distribution; (c) standard probabilistic population coding and (d) distributed distributional code. We have started to apply these results to experimental data to identify if and how uncertainty of spatial trajectories are represented in the hippocampus. Our analysis framework is an important step towards elucidating the strategies used by the brain to encode probability distributions and to understand the computational role of neuronal variability.

Keywords: population code; uncertainty; theta sequences; navigation; decoding

Introduction

Representing uncertainty associated with sensory stimuli or behavioral decisions is paramount for efficient behavior and learning when the available information is limited and the world can change rapidly. Several lines of evidence demonstrated that humans and other animals are taking an internal estimate of the uncertainty into account when making decisions (Ernst & Banks, 2002; Kording & Wolpert, 2004; Kepecs, Uchida, Zariwala, & Mainen, 2008), indicating that they are able to perform probabilistic computations and the manipulation of probability distributions (Knill & Pouget, 2004; Kersten & Yuille, 2003).

Coding schemes

In the past decades several competing hypotheses have been proposed regarding the representation of complex probability distributions by the neuronal population activity. First, the population activity can represent a probability distribution by encoding its parameters via a probabilistic population code (PPC; Ma, Beck, Latham, and Pouget (2006)). This coding scheme has particularly useful properties for implementing cue combination or evidence accumulation in the case of relatively simple (i.e. exponential family) distributions but implementing hierarchical inference or learning in a generative model using PPCs is more challenging (Beck, Pouget, & Heller, 2012). A distinctive property of the PPC is that it is a product of experts-type representation where the entropy of the encoded distribution decreases with the firing rate of the population.

Second, complex probably distributions can be represented using a distributed distributional code (DDC; Zemel, Dayan, and Pouget (1998); Sahani and Dayan (2003)) where the activity of individual neurons represent the expectation of a set of nonlinear functions (corresponding to the receptive field of the neurons) under the encoded distribution. Although decoding DDCs is not straight-forward, calculating expectation of arbitrary nonlinear functions has a simple form which is a useful property for both learning and inference (Vértes & Sahani, 2018). The DDC is a mixture of experts representation where large-entropy distributions are encoded in a more heterogenous population activity. Importantly, the full distribution is encoded instantaneously in the population activity in both the DDC and the PPC.

Third, the distribution could be represented by drawing samples from it (Fiser, Berkes, Orban, & Lengyel, 2010), in which case a single snapshot from the population activity represents only a point estimate of the represented variable, but integrating over time can capture the full, potentially complex distribution. Sampling based representations provide a natural explanation for the highly structured trial to trial variability observed in the cortex (Orbán, Berkes, Fiser, & Lengyel, 2016) and can be used for learning generative models in neuronal networks (Hinton, Dayan, Frey, & Neal, 1995).

Which of the above coding schemes is actually used by cor-

tical neuronal populations remains hotly debated mostly because of the lack of the appropriate measures to distinguish them based on experimental data. To successfully identify the coding scheme used by the brain three different requirements has to be fulfilled: a) we need to know what are the variables encoded in the population activity. b) we need to know how the represented uncertainty changes during individual trials. Finally, c) we need to be able to collect a high number of trials to evaluate potential changes in higher order statistics in the population activity. Here we argue, that theta activity during spatial exploration in rodents fulfills all these three criteria.

Theta sequences

During theta sequences, place cells, selective for specific locations in the environment, become sequentially activated in each theta cycle. In this way neurons encoding past, present, and future locations outline the trajectory of the animal (Feng, Silva, & Foster, 2015). We interpret this activity pattern as the result of repeatedly performing probabilistic inference and predictions about possible trajectories in a dynamical generative model, which underlies model-based planning, a task in which the hippocampus has been implicated (Miller, Botvinick, & Brody, 2017). Thus the population activity during thousands of theta cycles in a typical spatial navigation experiment (requirement c), encodes the trajectory of the animal (a). Critically, during a single theta sequence the uncertainty is expected to increase systematically (b), thus providing a chance to identify how probability distributions are encoded in the population activity.

Results

Encoding the posterior by the firing of place cells

We illustrate the workflow of the analysis in a synthetic dataset in which we simulated the motion of a rat in two dimensional open arena and modelled the population activity of hippocampal neurons using the encoding models described in the introduction. Importantly, similar to the real situation, in our simulations the animal does not have access to its true position $\mathbf{x}_t$, but have to infer it from the sensory inputs observed in the past $\mathbf{y}_{0:t}$. We assume that the activity of the place cells is driven by the posterior distribution over the trajectories $P(\vec{\mathbf{x}}|\mathbf{y}_{0:t})$. We used a flexible model, a slightly modified version of the standard Kalman filter with noisy observation on the motor control signal (Bishop, 2006), for the motion of the animal that generates smooth movement and allows efficient dynamical inference of the trajectories.

Fig. 1a-d illustrates a short segment of the true trajectory of the simulated animal along with its inferred past and predicted future positions. The animal's inferred position is characterised by an increasing level of uncertainty as the trajectory moves from the past to the future (beginning versus the end of the theta cycle; illustrated by the increasing diameter of the ellipses in Fig. 1b). We simulated the activity of 100 hippocampal pyramidal neurons with Gaussian place fields of $20 - 80\,\text{cm}$ diameter and $15 - 50\,\text{Hz}$ maximal firing rate.

The firing rate of the hippocampal neurons in our synthetic dataset was driven by trajectories encoded using one of the three alternative encoding models specified in the introduction (PPC, DDC and sampling) and a control model (maximum a posteriory, MAP) in which only the most likely trajectory is encoded. Spikes were generated from an inhomogeneous Poisson process using the firing rates controlled by the posterior over the trajectories encoded by one of the 4 possible models. Importantly, population activity in all of the competing models is largely consistent with many defining characteristics of hippocampal data during exploratory behavior e.g., neurons show place cell activity, phase precession and theta sequences (Fig. 1e).

Identification of the encoding model

Next we derive 3 specific measures of the population activity that are sufficient to distinguish between these four alternative coding strategies. These encoding strategies make different predictions regarding the detailed structure of the population activity. To obtain these measures, we first calculate the difference between the trajectories encoded in two subsequent theta cycles (Fig. 1d):

$$\varepsilon_1 = \sum_\tau (\hat{x}_t - \hat{x}_{t-1})^2 \tag{1}$$

and the difference between the encoded and the true trajectory (Fig. 1d):

$$\varepsilon_2 = \sum_\tau (\vec{x}_t - \hat{x}_t)^2 \tag{2}$$

In the case of the MAP, PPC and DDC encoding $\varepsilon_1 \approx \varepsilon + \zeta$ and $\varepsilon_2 \approx \varepsilon + \vec{\sigma}^2$, where ε is the decoding error (due to the finite number of observed neurons in the population and their stochastic spiking), $\vec{\sigma}^2$ is the variance of the encoded distribution characterising the inference of the position from sensory data and ζ is the change in the encoded distribution between two subsequent theta cycles due to novel sensory input. In the case of sampling $\varepsilon_1 \approx \varepsilon + 2\vec{\sigma}^2 + \zeta$ whereas $\varepsilon_2 \approx \varepsilon + 2\vec{\sigma}^2$ where we assumed that independent samples are drawn in each theta cycle. Further we assume that the incoming sensory input does not change the posterior too much in a theta cycle (100 ms) thus $\vec{\sigma}^2 \gg \zeta \approx 0$. Therefore, our first measure is $\varepsilon_2 - \varepsilon_1 \approx 0$, which is zero only in the case of sampling and it is $\varepsilon_2 - \varepsilon_1 = \vec{\sigma}^2$ otherwise (Fig. 1c-d).

To obtain the second measure, we note that in the case of MAP encoding and sampling the population activity at any given time encodes a single position therefore the precision of the encoding of the trajectory is constant within a theta cycle. Conversely, PPC and DDC instantaneously represent the uncertainty of the distribution which increases within a theta cycle. Thus, our second measure is the variance of the posterior distribution decoded from the population activity which is expected to be constant for sampling and MAP and increase for PPC and DDC.

For the third measure we build on the insight that in the case of the PPC the encoded uncertainty is directly proportional
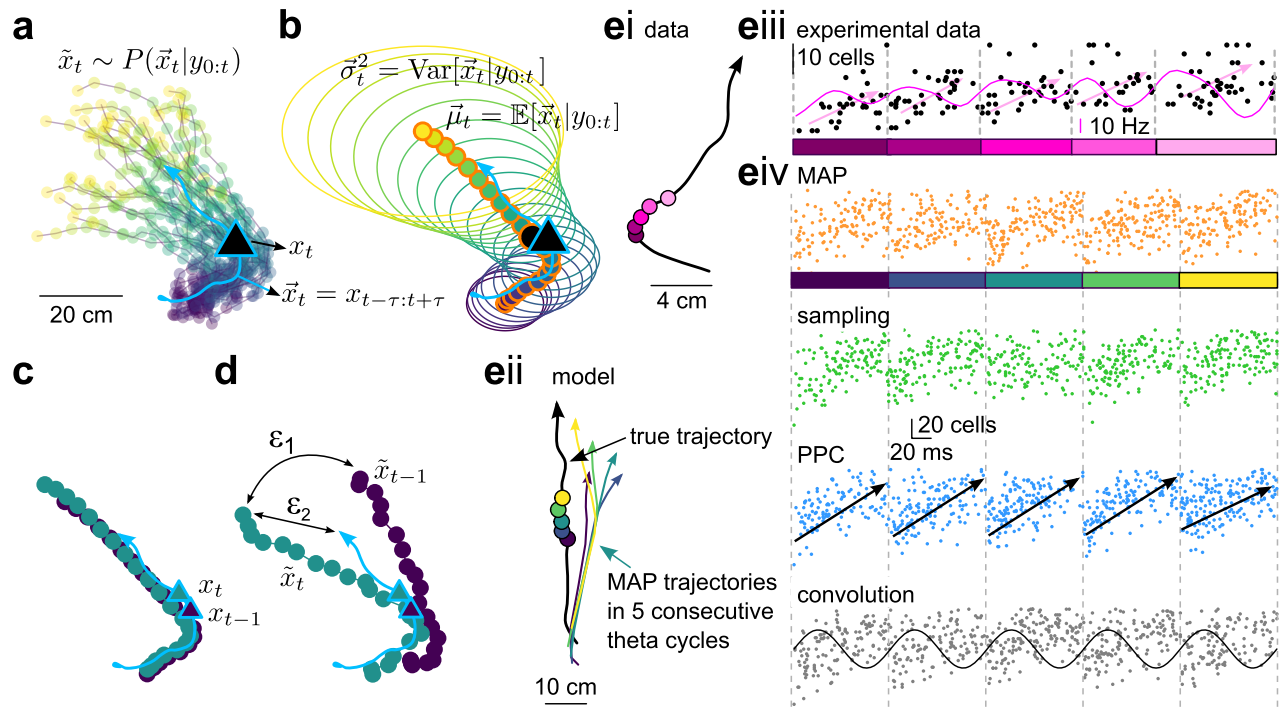
Figure 1: Representation during theta sequences. (**a-b**) Inference in the model. (**a**) A 2 s segment of the true trajectory of the animal $\vec{\mathbf{x}}$ is shown in blue together with the 25 samples, $\tilde{x}_t$, drown from the posterior over the possible trajectories. (**b**) Marginal means ($\vec{\mu}_t$, filled circles) and covariances ($\vec{\sigma}_t^2$, open ellipses) of the trajectories inferred at x_t. Colour changes from purple (past) to yellow (future). (**c**) MAP and (**d**) sampled trajectories during two consecutive theta cycles. The true location at the time of the inference is indicated by the filled triangles. $\varepsilon_1 = \sum_\tau (\vec{x}_t - \hat{x}_t)^2$ is the error between the decoded and the true trajectory. $\varepsilon_2 = \sum_\tau (\hat{x}_t - \hat{x}_{t-1})^2$ the difference between trajectories decoded during subsequent theta cycles. (**e**) Comparison of real and synthetic data. Animal trajectories in experimental data (ei) and in the model (eii) with the position in 5 consecutive theta cycles. Theta sequences in data (eiii) and in the model when the place cells use one of the 4 models to represent uncertainty (eiv). Population activity shown in (eiv) is decoded and analysed to identify the strategy for encoding uncertainty. Experimental data is from Grosmark et al. (2016).

to the response gain. Therefore the population firing rate is expected to decrease within a theta cycle, whereas there is no explicit relationship between these quantities in the other two coding schemes.

Results on synthetic data

We decoded the spike trains of the 4 competing models within each theta cycle separately in three non-overlapping time frames using two different static decoding methods. The first method, which is consistent with the MAP and sampling encoding model, assumed that the population activity encodes a single point at any given time and provided a maximum likelihood estimate for the encoded position ($\hat{x}_t$) which was used to estimate $\varepsilon_2 - \varepsilon_1$. The second method, consistent with the DDC encoding, provided the maximum likelihood estimate for the mean ($\hat{\mu}_t$) and variance parameter ($\hat{\sigma}_t^2$) of the encoded Gaussian distribution which was used to characterise the instantaneous uncertainty encoded in the population activity.

Figure 2a-c shows the three different measures calculated

using data from the 4 different encoding schemes. Our results indicate that the encoding model can be uniquely identified using these measures, as only sampling has $\varepsilon_2 - \varepsilon_1 \approx 0$ during the whole theta cycle (Fig. 2a); PPC and DDC predicts an increase in the represented uncertainty within a theta cycle (Fig. 2b) but only PPC predicts a decrease in the firing rate (Fig. 2c).

Conclusions and future directions

Our analysis using synthetic data demonstrates that datasets typically recorded in rodent navigational experiments can have enough statistical power to identify the hippocampal representation of uncertainty using these three measures. We have started to apply this framework to experimental data to identify if and how uncertainty of spatial trajectories are represented in the hippocampus. Our framework is an important step towards elucidating the strategies used by the brain to encode probability distributions and to understand the computational role of neuronal variability.
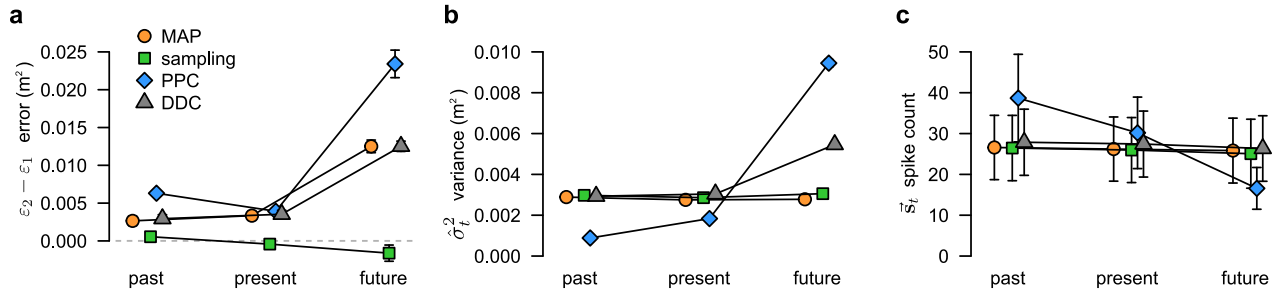
Figure 2: Discriminating between encoding models. (**a**) The error $\varepsilon_1 - \varepsilon_2$ (defined in Fig. 1d) is near 0 only for sampling. Trajectories were decoded assuming single point encoding and the decoding variance was subtracted from both errors. (**b**) The decoded variance, $\hat{\sigma}_t^2 = \mathbb{E}[\vec{\sigma}_t^2 | \vec{s}_t]$ assuming DDC encoding increases for the two models encoding full distributions. (**c**) The spike count within theta cycles decreases only in the PPC encoding. Error bars indicate SE and are often smaller than symbols in (a) and (b).

Acknowledgments

References

Beck, J., Pouget, A., & Heller, K. (2012). Complex inference in neural circuits with probabilistic population codes and topic models. *Advances in Neural Information Processing Systems*.

Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer.

Ernst, M., & Banks, M. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429-433.

Feng, T., Silva, D., & Foster, D. J. (2015). Dissociation between the experience-dependent development of hippocampal theta sequences and single-trial phase precession. *J Neurosci*, *35*(12), 4890-902.

Fiser, J., Berkes, P., Orban, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci*, *14*(3), 119-30.

Grosmark, A., Long, J., & Buzsáki, G. (2016). Recordings from hippocampal area CA1, PRE, during and POST novel spatial learning. *CRCNS.org*.

Hinton, G., Dayan, P., Frey, B., & Neal, R. (1995). The wake--sleep algorithm for unsupervised neural networks. *Science*, *268*, 1158-1161.

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008, Sep). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227-31. doi: 10.1038/nature07200

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Curr Opin Neurobiol*, *13*(2), 150-8.

Knill, D. C., & Pouget, A. (2004, Dec). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci*, *27*(12), 712-9. doi: 10.1016/j.tins.2004.10.007

Kording, K., & Wolpert, D. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*, 244-247.

Ma, W., Beck, J., Latham, P., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.*, *9*, 1432-1438.

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017, Sep). Dorsal hippocampus contributes to model-based planning. *Nat Neurosci*, *20*(9), 1269-1276. doi: 10.1038/nn.4613

Orbán, G., Berkes, P., Fiser, J., & Lengyel, M. (2016, Oct). Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron*, *92*(2), 530-543. doi: 10.1016/j.neuron.2016.09.038

Sahani, M., & Dayan, P. (2003, Oct). Doubly distributional population codes: simultaneous representation of uncertainty and multiplicity. *Neural Comput*, *15*(10), 2255-79. doi: 10.1162/089976603322362356

Vértes, E., & Sahani, M. (2018). Flexible and accurate inference and learning for deep generative models. *Advances in Neural Information Processing Systems*.

Zemel, R., Dayan, P., & Pouget, A. (1998). Probabilistic interpretation of population codes. *Neural Comput.*, *10*, 403-430.