2016 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 13-16, 2016, SALERNO, ITALY

BAYESIAN NONPARAMETRIC METHODS FOR DISCOVERING LATENT STRUCTURES OF RAT HIPPOCAMPAL ENSEMBLE SPIKES

Zhe Chen[1*]

[1]Department of Psychiatry
NYU School of Medicine

Scott W. Linderman[2†], *Matthew A. Wilson*[3‡]

[2]School of Engineering & Applied Sciences, Harvard University
[3]Picower Institute for Learning and Memory, MIT

ABSTRACT

Hippocampal functions are responsible for encoding spatial and temporal dimensions of episodic memory, and hippocampal reactivation of previous awake experiences in sleep is important for learning and memory consolidation. Therefore, uncovering neural representations of hippocampal ensemble spike activity during various behavioral states would provide improved understanding of neural mechanisms of hippocampal-cortical circuits. In this paper, we propose two Bayesian nonparametric methods for this purpose: the Bayesian modeling allows to impose informative priors and constraints into the model, whereas Bayesian nonparametrics allows automatic model selection. We validate these methods to three different hippocampal ensemble recordings under different task behaviors, and provide interpretation and discussion on the derived results.

Index Terms— Bayesian nonparametrics, hidden Markov model, hidden semi-Markov model, population codes

1. INTRODUCTION

Neuronal spikes are the basic codes for representing and transmitting information in the brain. A fundamental goal in computational neuroscience is to understand the representation of neuronal population codes and discover latent structure of spatiotemporal neural data under various brain states. Today, development in neuroscience has enabled us to simultaneously record a large number of neuronal ensemble spike activity in rodents and primates. However, it remains a statistical challenge to fully interpret the ensemble spike activity and link its representation to the animal's behavior, task, or internally-driven computation. In many scenarios, the behavioral measure can be either abstract (non-quantitative) or absent (such as during sleep). In some scenarios, it is desired to infer the latent structure based on the ensemble spike activity alone. In these cases, it is preferred to employ an unsupervised learning paradigm to extract the inherent structure of neuronal ensemble spikes. State-space methods and Bayesian methods are powerful in analyzing spatiotemporal neural data [3, 5].

In this paper, we employ the principle to rat hippocampal ensemble spike activity recorded during different behavioral states, including *spatial navigation* and *sleep*. Uncovering neural representations of hippocampal population codes may reveal important mechanisms of episodic memory of space and time, and memory consolidation during sleep [17, 18, 16]. Specifically, we model the animal's state as a latent Markovian or semi-Markovian process. The state may correspond to animal's spatial location (which is unobserved from the internal brain's perspective, or may be referred to an undefined internal neural process that drives the ensemble spike activity. We develop two Bayesian nonparametric extensions of hidden Markov model (HMM) for characterizing such data. The Bayesian nonparametric approach automatically adapts the model size according to the observed data. We then develop a Markov Chain Monte Carlo (MCMC) inference algorithm to explore the posterior distribution of the unknown state and parameters.

2. BASIC MODEL DESCRIPTION

We used a finite m-state HMM to characterize the temporal population spiking activity from a population of C hippocampal neurons [2, 4]. We assume, first, the latent state process followed a first-order discrete-state Markov chain $\mathcal{S} = S_{1:T} \equiv \{S_t\} \in \{1, \ldots, m\}$, and second, the spike counts of individual place cells at discrete time index t, conditional on the hidden state S_t, followed a Poisson probability with their respective tuning curve functions $\mathbf{\Lambda} = \{\boldsymbol{\lambda}_c\} = \{\lambda_{c,i}\}$. The HMM is summarized as follows:

$$\begin{aligned} p(\boldsymbol{y}_{1:T}, S_{1:T}|\boldsymbol{\pi}, \boldsymbol{P}, \mathbf{\Lambda}) &= p(\mathcal{S}_1|\boldsymbol{\pi}) \prod_{t=2}^{T} p(S_t|S_{t-1}, \boldsymbol{P}) \\ &\quad \times \prod_{t=1}^{T} p(\boldsymbol{y}_t|S_t, \mathbf{\Lambda}), \end{aligned}$$

*Support from NSF-CRCNS grant IIS-130764 from the US National Science Foundation and NIH grant R01-NS100065 from the NINDS.

†Support from the Thomas and Stacey Siebel Foundation.

‡Support from the ONR MURI grant N00014-10-1-0936 and NIH grant TR01-GM10498, and NSF STC award CCF-1231216.

978-1-5090-0746-2/16/$31.00

$$
\begin{aligned}
p(S_1|\boldsymbol{\pi}) &= \texttt{Multinomial}(S_1|\boldsymbol{\pi}), \\
p(S_t|S_{t-1},\boldsymbol{P}) &= \texttt{Multinomial}(S_t|\boldsymbol{P}_{S_{t-1},:}), \\
p(\boldsymbol{y}_t|S_t,\boldsymbol{\Lambda}) &= \prod_{c=1}^{C} \texttt{Poisson}(y_{c,t}|\lambda_{c,S_t}).
\end{aligned}
$$

where $\boldsymbol{P} = \{P_{ij}\}$ denotes an m-by-m state transition matrix, with P_{ij} representing the transition probability from state i to j; $y_{c,t}$ denotes the number of spike counts from cell c within the t-th temporal bin and $\boldsymbol{y}_{1:T} = \{y_{c,t}\}_{C\times T}$ denotes time series of C-dimensional population response vector; and $\texttt{Poisson}(y_{c,t}|\lambda_{c,i})$ defines a Poisson distribution with the rate parameter $\lambda_{c,i}$ when $S_t = i$. Finally, $\log p(\boldsymbol{y}_{1:T}|\mathcal{S},\boldsymbol{\theta})$ defines the observed data log likelihood given the latent state sequence $\mathcal{S}$ and all parameters $\boldsymbol{\theta} = \{\boldsymbol{\pi},\boldsymbol{P},\boldsymbol{\Lambda}\}$ (where $\boldsymbol{\pi} = \{\pi_i\}$ denotes a probability vector for the initial state S_1).

We further introduce the following prior distributions over the parameters [4]:

$$
\begin{aligned}
\boldsymbol{\pi} &\sim \texttt{Dir}(\alpha_0\mathbf{1}), \;\; \boldsymbol{P}_{i,:} \sim \texttt{Dir}(\alpha_0\mathbf{1}), \\
\alpha_0 &\sim \texttt{Gamma}(a_{\alpha_0},1), \;\; \lambda_{c,i} \sim \texttt{Gamma}(a_c^0,b_c^0).
\end{aligned}
$$

where $\texttt{Dir}$ denotes the Dirichlet prior distribution, and $\texttt{Gamma}(a_c^0,b_c^0)$ denotes the gamma prior distribution with shape parameter a_c^0 and scale parameter b_c^0.

3. BAYESIAN NONPARAMETRIC MODELING AND INFERENCE

3.1. HDP-HMM

In our previous work [15], we generalized the finite-state HMM above with a Bayesian nonparametric model known as a hierarchical Dirichlet process (HDP)-HMM. Specifically, we sample a distribution over latent states, G_0, from a DP prior, $G_0 \sim \texttt{DP}(\gamma,H)$, where γ is the concentration parameter and H is the base measure. We also place a prior distribution over the concentration parameter, $\gamma \sim \texttt{Gamma}(a_\gamma,1)$. Given the concentration, we sample from the DP via the "stick-breaking process (STP)": the stick-breaking weights, $\boldsymbol{\beta}$, is drawn from a beta distribution:

$$
\tilde{\beta}_i \sim \texttt{Beta}(1,\gamma), \quad \beta_i = \tilde{\beta}_i \prod_{j=1}^{i-1}(1-\tilde{\beta}_j) \tag{1}
$$

where $\beta_1 = \tilde{\beta}_1$, $\sum_{i=1}^{\infty}\beta_i = 1$, and $\texttt{Beta}(a,b)$ defines a beta distribution with two shape parameters $a > 0$ and $b > 0$. The name "stick-breaking" comes from the interpretation of β_i as the length of the piece of a unit-length stick assigned to the i-th value. After the first $i-1$ values having their portions assigned, the length of the remainder of the stick is broken according to a sample $\tilde{\pi}_i$ from a beta distribution, and $\tilde{\beta}_i$ indicates the portion of the remainder to be assigned to the i-th value. Therefore, the STP defines a DP—the smaller γ, the less of the stick will be left for subsequent values.

After sampling $\boldsymbol{\beta}$, we next sample the latent state variables, in this case $\boldsymbol{\lambda}_c$, from the base measure H. Our draw from the $\text{DP}(\gamma,H)$ prior is then given by $G_0 = \sum_{j=1}^{\infty}\beta_j\delta_{\boldsymbol{\lambda}_c^{(j)}}$. Thus, the stick breaking construction makes clear that draws from a DP are discrete with probability one.

Given a countably infinite set of shared states, we may sample the rows of the transition matrix, $\boldsymbol{P}_{i,:} \sim \text{DP}(\alpha_0,\boldsymbol{\beta})$. We place the same prior over π. The base measure is $\boldsymbol{\beta}$, a countably infinite vector of stick-breaking weights, that serves as the mean of the DP prior over the rows of $\boldsymbol{P}$. The concentration parameter α_0 governs how concentrated the rows are around the mean. The hierarchical prior is required to provide a discrete yet countably infinite set of latent states for the HMM.

3.2. HDP-HSMM

The HSMM extends the standard HMM by assuming that the state transition probability P_{ij} depends that the amount of the sojourn time in state i. In many real-world examples, the latent process is not always Markvoian in that the state durations may follow more specific probability distributions, such as Poisson, negative binomial, lognormal and inverse Gaussian distribution (note that the standard HMM corresponds to a geometric sojourn time distribution such that the probability of staying in state i for d steps is $P_{ii}^d(1-P_{ii})$). One idea is to introduce an explicit-duration semi-Markov modeling for each state. Here, we adapt our HDP-HMM into an HDP-HSMM based on the formulation in [11, 12]. Specifically, we assume that the sojourn duration in state i, denoted by $p(d_t|S_t=i)$, follows a parametric distribution form:

$$
\begin{aligned}
d_t|S_t=i &\sim \texttt{NegBin}(r,p) \\
&= \binom{d+r-2}{d-1}(1-p)^r p^{d-1} \;\; (d=1,2,\cdots)
\end{aligned}
$$

where $\texttt{NegBin}(r,p)$ denotes a negative binomial distribution (discrete analog of the gamma distribution), which reduces to the geometric ($\texttt{Geom}$) distribution when $r=1$ as a special case (i.e., Markovian). We use an "embedding trick" for the negative binomial distribution [12]

$$
d \sim \texttt{NegBin}(r,p) \Longleftrightarrow d = 1 + \sum_{k=1}^{r} z_k
$$

where $z_k \sim \texttt{ShiftedGeom}(1-p)$ are independent and identically distributed (i.i.d.) having a geometric distribution with parameter $1-p$ shifted so that support includes 0.

3.3. Bayesian inference

Previously, we developed variational Bayes (VB) and MCMC inference procedures for HDP-HMM [15]. Specifically, we used a "weak limit" approximation in which the DP prior was

approximated with a symmetric Dirichlet prior:

$$\begin{aligned}
\gamma &\sim \texttt{Gamma}(a_\gamma, 1) \\
\alpha_0 &\sim \texttt{Gamma}(a_{\alpha_0}, 1) \\
\boldsymbol{\beta}|\gamma &\sim \texttt{Dir}(\gamma/M, \ldots, \gamma/M), \\
\boldsymbol{\pi}|\alpha_0, \boldsymbol{\beta} &\sim \texttt{Dir}(\alpha_0\beta_1, \ldots, \alpha_0\beta_M), \\
\boldsymbol{P}_{i,:}|\alpha_0, \boldsymbol{\beta} &\sim \texttt{Dir}(\alpha_0\beta_1, \ldots, \alpha_0\beta_M).
\end{aligned}$$

where M denoted a truncation level for approximating the distribution over the countably infinite number of states. This prior weakly converges to the DP prior as the dimensionality of the Dirichlet distribution approaches infinity [11].

In the case of HDP-HSMM inference, we use the HMM embedding for HSMM with $\texttt{NegBin}(r, p)$. We impose a uniform prior (integer values $\{1, 2, \ldots, r_{\max}\}$) over the shape parameter r and a beta prior over p. The Gibbs sampling procedure follows the procedure in [12]: First, sample the state sequence $S_{1:T}$ using a block sampler and the message passing algorithm; second, sample the negative binomial parameters (r_i, p_i) for each state $i = 1, 2, \ldots, m$ given the state sequences. In total, the time complexity of each Gibbs iteration is reduced from $\mathcal{O}(m^2T + mT^2)$ to $\mathcal{O}(m^2T + mTr)$ (linear in time T).

For Poisson likelihood, we use a Gibbs sampler for parameter $\boldsymbol{\Lambda}$. Since we are using conjugate gamma priors, the posterior can be updated in a closed form

$$\lambda_{c,i}|\boldsymbol{y}, \mathcal{S} \sim \texttt{Gamma}\left(\alpha_c^0 + \sum_{t=1}^{T} y_{c,t}\mathbb{I}[S_t = i],\ \beta_c^0 + \sum_{t=1}^{T}\mathbb{I}[S_t = i]\right).$$

Under the weak limit approximation, the priors on $\boldsymbol{P}_{i,:}$ and $\boldsymbol{\pi}$ reduce to Dirichlet distributions, and we can derive conjugate Gibbs updates for these parameters as follows:

$$\begin{aligned}
\boldsymbol{\pi}|\alpha_0, \boldsymbol{\beta} &\sim \texttt{Dir}\left(\alpha_0\boldsymbol{\beta} + \mathbf{1}_{S_1}\right), \\
\boldsymbol{P}_{i,:}|\alpha_0, \boldsymbol{\beta} &\sim \texttt{Dir}\left(\alpha_0\boldsymbol{\beta} + \boldsymbol{n}_i\right), \\
n_{i,j} &= \sum_{t=1}^{T-1}\mathbb{I}[S_t = i, S_{t+1} = j],
\end{aligned}$$

where $\mathbf{1}_j$ is a unit vector with a one in the j-th entry. Conditioned upon the firing rates, the initial state distribution, and the transition matrix, we can jointly update the latent states using a *forward filtering, backward sampling* algorithm to obtain a full sample from $p(\mathcal{S}|\boldsymbol{P}, \boldsymbol{\pi}, \boldsymbol{\Lambda})$.

Regarding the firing rate hyperparameters $\{\alpha_c^0, \beta_c^0\}$ for cell c, we previously proposed three methods for update [15]: (i) empirical Bayesian (EB), which aims to maximize the marginal likelihood of the spike counts; (ii) Hamiltonian Monte Carlo (HMC) sampling for joint posterior $\{\log\alpha_c^0, \log\beta_c^0\}$; and (iii) sampling the scale hyperparameter β_c^0 (using a gamma prior) while fixing the shape hyperparameter, α_c^0. In practice, we found that the second and third methods worked very well.

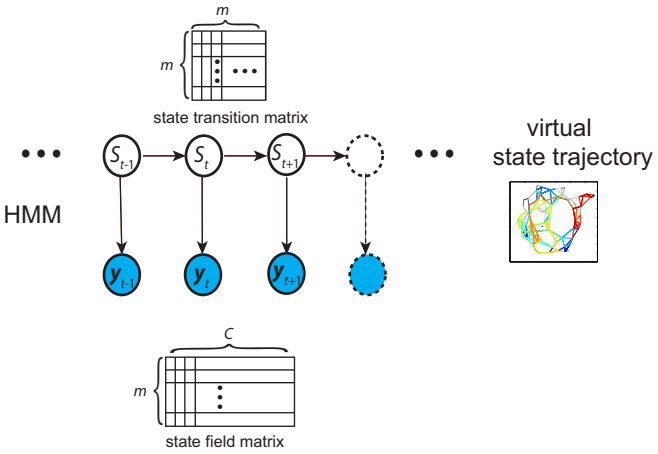

Fig. 1. The m-state HMM (or HDP-HMM or HDP-HSMM) is used to characterize hippocampal neuronal ensemble spikes in time. The virtual spatial topology is derived from the inferred state transition matrix $\boldsymbol{P}$, warm/cold color represents the direction of time evolution along the graph.

In the testing phase, for the test data $\boldsymbol{y}_{test}$ we compute the predictive log likelihood with samples from the posterior distribution generated by our MCMC algorithm:

$$\begin{aligned}
\log p(\boldsymbol{y}_{test}|\boldsymbol{y}_{1:T}) &= \log\sum_{\mathcal{S}}\int p\left(\boldsymbol{y}_{test}, \mathcal{S}_{test}|\boldsymbol{\theta}\right)\ p\left(\boldsymbol{\theta}|\boldsymbol{y}_{1:T}\right)d\boldsymbol{\theta}, \\
&\approx \log\frac{1}{N}\sum_{n=1}^{N}\sum_{\mathcal{S}_{test}} p\left(\boldsymbol{y}_{test}, \mathcal{S}_{test}|\boldsymbol{\theta}_n\right),
\end{aligned}$$

where $\{\boldsymbol{\theta}_n\}_{n=1}^N \sim p(\boldsymbol{\theta}|\boldsymbol{y}_{train})$ denote the Monte Carlo samples from the posterior.

4. RESULT VISUALIZATION AND INTERPRETATION

Upon completion of inference, we obtain the posteriors for latent states $\{S_{1:T}\}$, $\boldsymbol{P}$ and $\boldsymbol{\Lambda}$. Based on the correspondence of state and behavioral measures, we construct a "state space map" [4] for validation: a perfect state space map shall have a one-to-one mapping. To visualize the inferred state transition matrix, we further apply a force-based algorithm to derive a scale-invariant topology graph that defines the connectivity between different states (nodes), which may offer intuitive result interpretation and qualitative assessment. A schematic diagram of the procedure is shown in Fig. 1 [6].

5. EXPERIMENTAL DATA AND RESULTS

5.1. Experimental protocol and recordings

In Dataset 1, Long-Evans rats were freely foraging in a familiar open field arena (Fig. 2, left panel) for about 25 minutes [4]. In Dataset 2, rats were first put in a sleep box for 4 hours,

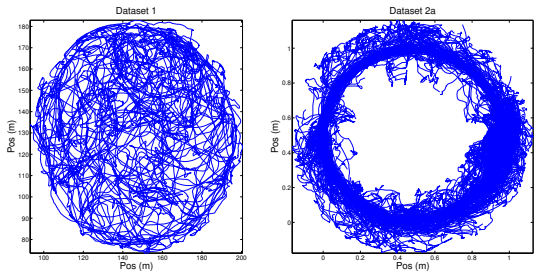

Fig. 2. Rats' run trajectories in the open field (Dataset 1) and circular maze (Dataset 2a).

Table 1. **Statistics of rat hippocampal ensemble recordings.**

Dataset	# cells	Period	Behavior
1	49	24.3 min	free foraging in an open field
2a	77	25.4 min	free foraging in a circular track
2b	77	480 min	rest in a sleep box

and then moved to a familiar circular track (Fig. 2, right panel) for running about 25 minutes, and then put back to the sleep box for another 4 hours [9]. Readers are referred to previous publications for more experimental details [15, 9].

Custom microelectrode drive or silicon probe arrays were implanted in the animal's dorsal hippocampal CA1 area. Extracellular spikes were acquired along local field potentials. We used a custom manual clustering program for spike sorting to obtain well-isolated single units. Putative interneurons were identified based on the spike waveform width and average mean firing rate. Summary of hippocampal recordings is shown in Table 1. All procedures were approved by the Institutional Animal Care and Use Committee and carried out in accordance with the approved guidelines.

5.2. Result on Dataset 1

First, we apply the proposed HDP-HMM and HDP-HSMM algorithms to a previously studied hippocampal neuronal ensemble dataset [15], where the rat was freely foraging in an open field environment (Fig. 2, left panel). One of the objectives in this study is to investigate if we can recover the spatial topological representation of the two-dimensional environment from hippocampal population codes without using their place receptive fields [4]. Note that although the animal's location was recorded, we only use the ensemble spikes alone in the modeling and inference; the animal's location was only used in post-hoc assessment. We test three different hyperparameter optimization (EB/HMC/scale-gamma prior) methods discussed in Section 3.3. We set $r_{\max} = 15$ for HDP-HSMM. We use 80% run-associated (speed filter > 15 cm/s) hippocampal ensemble spikes (temporal bin size: 250 ms) and compute the predictive log-likelihood on the held-out 20% test data. The predictive log-likelihood is further compared to that of a set of independent Poisson processes and normalized by the number of spikes in the test set [15].

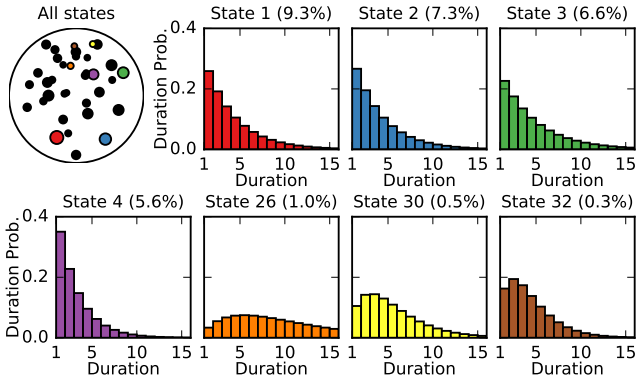

Fig. 3. Inferred state centers (each dot represents one state, and the size is proportional to the state occupant) and duration distributions for 7 representative states.

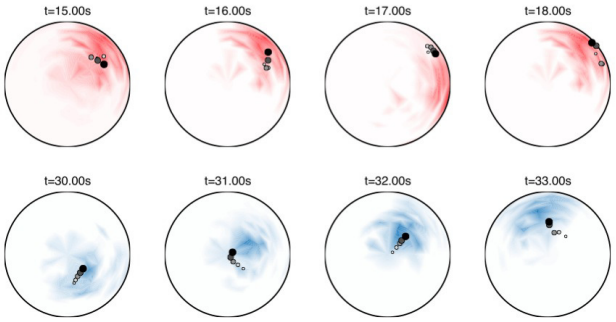

Fig. 4. Snapshots of the inferred state distribution (from HDP-HSMM) of the instantaneous rat's location during the testing period. Two representative sequences show the evolution of instantaneous distributions as the rat explores the environment. The black dots show the rat's trajectory over the past second, with the largest, darkest dot showing the location at time t and the smallest dot showing the location at time $t-1$.

We find that the HDP-HMM produced the best result in predictive log-likelihood (Table 2). To explain why HDP-HSMM did not yield improved predictive performance, we examine the duration distributions of some inferred state (Fig. 3). As seen, the first 4 dominant states have a scale parameter $r = 1$, which implies that $\texttt{NegBin}(1, p)$ is equivalent to a geometric distribution $\texttt{Geom}(1-p)$. In other words, the posterior is concentrated around a special case where the HSMM reduces to the HMM. However, in a general behavioral setting, we expect that the HDP-HSMM would yield better performance than the HDP-HMM. For illustrations, we also show snapshots of the inferred state distribution (Fig. 4) and decoded animal's spatial trajectory (Fig. 5). Specifically, the decoding accuracy of our *unsupervised* HDP-HSMM method is significantly better than that of the *supervised* op-

Table 2. Comparison of predictive log-likelihood (bits/spike) of test data in Dataset 1 (the best result is marked in bold font).

model	predictive log-likelihood
HDP-HMM (EB)	0.579±0.001
HDP-HMM (HMC)	0.647±0.002
HDP-HMM (scale-gamma)	**0.722±0.000**
HDP-HSMM (scale-gamma)	0.712±0.002

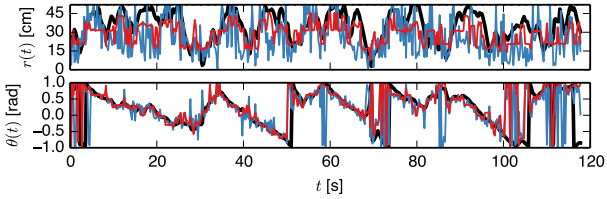

Fig. 5. Decoded rat's two-dimensional location (in polar coordinate $(\rho(t), \theta(t))$) during 2-min recording of test data. *Black curve*: animal's true position; *red*: decoded trajectory from the HDP-HSMM; *magenta*: decoded trajectory from an optimal linear decoder.

timal linear decoder method employing correlations between spike trains [19].

5.3. Result on Dataset 2

Sleep is critical to hippocampus-dependent memory consolidation. Analyzing hippocampal ensemble spike data during both slow-wake sleep (SWS) and rapid-eye-movement (REM) sleep has been an important topic [14, 10]. Typically, hippocampal place cells fire in sequences that span a few seconds as animals run through location-dependent receptive fields. During sleep, the same place cells fire in an orderly manner at a faster timescale within the hippocampal sharp wave (SPW)-ripples, lasting between 50 to 400 milliseconds. Previous findings have suggested that many sequences reflect temporally-compressed spatial sequences of previous experiences. Here, we focus SWS epochs during sleep (sleep scoring was done based on recorded EEG and EMG activities). For screening the candidate events, we use hippocampal LFP ripple band (150-300 Hz) power combined with hippocampal multi-unit activity (threshold>mean+3SD). We also imposed a minimum (>6) cell activation criterion. Upon examining the hippocampal population, we find that the mean firing rates of hippocampal neurons in SWS is significantly lower than in awake state ($P < 10^{-5}$, signed rank test), although the firing rates across states are positively correlated (Pearson's correlation 0.26, $P = 0.023$; Fig. 6).

We apply the HDP-HMM (scale-gamma hyperparameter optimization) to the run-associated (speed filter >10 m/s) hippocampal ensemble spikes (temporal bin: 250 ms), from which we infer the model parameters $\{\boldsymbol{P}, \boldsymbol{\Lambda}\}$ and state space map (Fig. 7, leftmost panel). Next, we apply the estimated

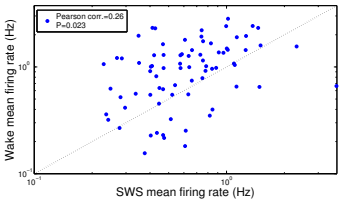

Fig. 6. Comparison of mean firing rates of rat hippocampal neurons (Dataset 2) between wake and SWS.

model to a total of 1519 post-run SWS candidate events (temporal bin: 20 ms; mean±SEM bins: 11.9 ± 0.2) and further infer the unknown state sequences (see an example in Fig. 7). For each detected event, we compute the weighted correlation R and Z-score statistic: $Z = \frac{R - \text{mean of } R_{\text{shuffle}}}{\text{SD of } R_{\text{shuffle}}}$, where R_{shuffle} are derived from 1000 randomly shuffled spike data [9, 6]. A high absolute value (>0.5) of weighted correlation and a high positive Z-score (>1.65) indicates a statistically significant replay event. Interestingly, the proposed unsupervised method has a comparable or higher detection power than the standard (receptive-field based) population decoding method [7]. Detailed results have been reported elsewhere [6].

6. DISCUSSION AND CONCLUSION

Recently, several new parametric and nonparametric methods have been developed to decode rat hippocampal ensemble spike activity [1, 2, 4, 13, 15]. By imposing informative priors and constraints, Bayesian probabilistic modeling offers an appealing method for discovering latent structure in spatiotemporal neural data, and Bayesian nonparametrics allows us to deal with model selection in a principled way. In our specific example, the HDP-HMM produces slightly better predictive log-likelihood than the HDP-HSMM (Table 2). This can be explained by two possible reasons: first, the approximate Markovian nature of the data (Fig. 3); second, the increased model complexity of HDP-HSMM given a relatively small size of training samples. In a general setup, the proposed HDP-HSMM framework is more powerful in characterizing population spike trains. Furthermore, the latent state space analysis may be applied to virtual reality environment, or to temporal dimension of episodic memory— hippocampal "time cells" [8], where animals may engage in non-spatial yet temporally structured task experiences.

Finally, it should be emphasized that although we have used rat hippocampal ensemble spikes as examples for demonstration, our proposed methodology and analysis framework is general and applicable for a wide range of hippocampal-cortical or thalamocortical ensemble spike data.

Acknowledgments

We thank Dr. Matthew J. Johnson (Harvard University) for technical assistance, and thank Dr. Andres Grosmark and

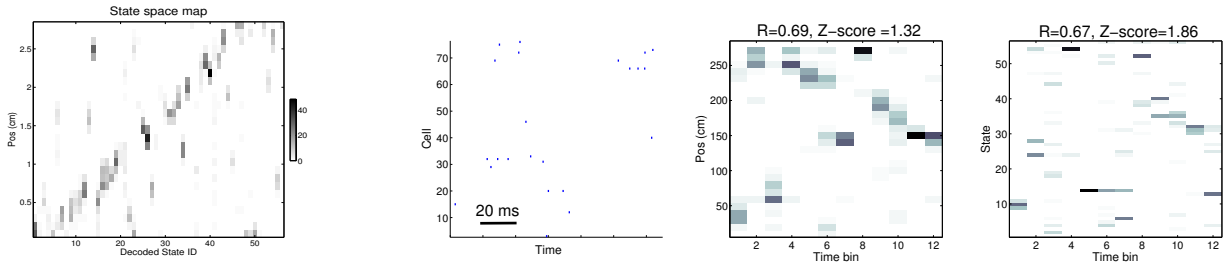

Fig. 7. *1st column*: Inferred state space map for the linearized circular track. *2nd column*: An example of hippocampal ensemble spikes during a post-SWS candidate event detected from hippocampal ripples, and decoded spatial or state sequences from a receptive-field based decoding method (*3rd column*, [7, 9]) and our proposed unsupervised HDP-HMM method (*4th column*). X-axis represents time bin (bin size 20 ms). Weighted correlation R and Z-score are shown at the top of 3rd and 4th panels.

Prof. György Buzsáki (New York University) for providing sleep recordings of hippocampal neurons in this investigation. Parts of results in this paper have been reported in [15, 6].

7. REFERENCES

[1] R. Agarwal, Z. Chen, F. Kloosterman, M. A. Wilson, S. Sarma, "A novel nonparametric approach for neural encoding and decoding models of multimodal receptive fields," *Neural Comput.*, vol. 28, pp. 1356-1387, 2016.

[2] Z. Chen, F. Kloosterman, E. N. Brown and M. A. Wilson, "Uncovering hidden spatial topology represented by hippocampal population neuronal codes," *J. Comput. Neurosci.* , vol. 33, no. 2, pp. 227–255, 2012.

[3] Z. Chen, "An overview of Bayesian methods for neural spike train analysis," *Computational Intelligence and Neuroscience* , volume 2013, Article ID 251905, 2013.

[4] Z. Chen, S. Gomperts, J. Yamamoto and M. A. Wilson, "Neural representation of spatial topology in the rodent hippocampus," *Neural Comput.*, vol. 26, pp. 1–39, 2014.

[5] Z. Chen, editor. *Advanced State Space Methods for Neural and Clinical Data*. Cambridge Univ. Press, 2015.

[6] Z. Chen, A. D. Grosmark, H. Penagos, and M. A. Wilson, "Uncovering representations of sleep-associated hippocampal ensemble spike activity," *Scientific Reports*, vol. 6, p. 32193, 2016.

[7] T. J. Davidson, F. Kloosterman and M. A. Wilson, "Hippocampal replay of extended experience," *Neuron*, vol. 63, no. 4, pp. 497–507, 2009.

[8] H. Eichenbaum, "Time cells in the hippocampus: a new dimension for mapping memories," *Nat. Rev. Neurosci.*, vol. 15, pp. 732-744, 2014.

[9] A. D. Grosmark and G. Buzsáki, "Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences," *Science*, vol. 351, pp. 1440–1443, 2016.

[10] D. Ji and M. A. Wilson, "Coordinated memory replay in the visual cortex and hippocampus during sleep," *Nat. Neurosci.*, vol. 10, no. 1, pp. 100–107, 2007.

[11] M. J. Johnson and A. S. Willsky, "Bayesian nonparametric hidden semi-Markov models," *J. Mach. Learn. Res.*, vol. 14, no. 1, pp. 673–701, 2013.

[12] M. J. Johnson, *Bayesian time series and scalable inference*. PhD thesis, MIT, 2014.

[13] F. Kloosterman, S. Layton, Z. Chen, and M. A. Wilson, "Bayesian decoding of unsorted spikes in the rat hippocampus", *J. Neurophysiol.*, vol. 111, no. 1, pp 217–227, 2014.

[14] A. K. Lee and M. A. Wilson, "Memory of sequential experience in the hippocampus during slow wave sleep," *Neuron*, vol. 36, no. 6, pp. 1183–1194, 2002.

[15] S. W. Linderman, M. J. Johnson, M. A. Wilson and Z. Chen, "A Bayesian nonparametric approach to uncovering rat hippocampal population codes during spatial navigation," *J. Neurosci. Meth.*, vol. 236, pp. 36–47, 2016.

[16] D. K. Roumis and L. M. Frank, "Hippocampal sharp-wave ripples in waking and sleep states," *Curr. Opin. Neurobiol.*, vol. 35, pp. 6–12, 2015.

[17] W. E. Skaggs and B. L. McNaughton, "Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience," *Science*, vol. 271, pp. 1870–1873, 1996.

[18] M. A. Wilson and B. L. McNaughton, "Reactivation of hippocampal ensemble memories during sleep," *Science*, vol. 265, pp. 676–679, 1994.

[19] D. K. Warland, P. Reinagel, and M. Meister, "Decoding visual information from a population of retinal ganglion cells," *J. Neurophysiol.*, vol. 78, pp. 2336–2350, 1997.