

RESEARCH ARTICLE

Inferring oscillatory modulation in neural spike trains

Kensuke Arai^{1,2*}, Robert E. Kass^{1,2,3}

1 Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States of America, **2** Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania, United States of America, **3** Machine Learning Department, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States of America

* kensuke.y.arai@gmail.com



Abstract

Oscillations are observed at various frequency bands in continuous-valued neural recordings like the electroencephalogram (EEG) and local field potential (LFP) in bulk brain matter, and analysis of spike-field coherence reveals that spiking of single neurons often occurs at certain phases of the global oscillation. Oscillatory modulation has been examined in relation to continuous-valued oscillatory signals, and independently from the spike train alone, but behavior or stimulus triggered firing-rate modulation, spiking sparseness, presence of slow modulation not locked to stimuli and irregular oscillations with large variability in oscillatory periods, present challenges to searching for temporal structures present in the spike train. In order to study oscillatory modulation in real data collected under a variety of experimental conditions, we describe a flexible point-process framework we call the Latent Oscillatory Spike Train (LOST) model to decompose the instantaneous firing rate in biologically and behaviorally relevant factors: spiking refractoriness, event-locked firing rate non-stationarity, and trial-to-trial variability accounted for by baseline offset and a stochastic oscillatory modulation. We also extend the LOST model to accommodate changes in the modulatory structure over the duration of the experiment, and thereby discover trial-to-trial variability in the spike-field coherence of a rat primary motor cortical neuron to the LFP theta rhythm. Because LOST incorporates a latent stochastic auto-regressive term, LOST is able to detect oscillations when the firing rate is low, the modulation is weak, and when the modulating oscillation has a broad spectral peak.

OPEN ACCESS

Citation: Arai K, Kass RE (2017) Inferring oscillatory modulation in neural spike trains. *PLoS Comput Biol* 13(10): e1005596. <https://doi.org/10.1371/journal.pcbi.1005596>

Editor: Matthias Bethge, University of Tübingen and Max Planck Institute for Biological Cybernetics, GERMANY

Received: May 14, 2015

Accepted: May 24, 2017

Published: October 6, 2017

Copyright: © 2017 Arai, Kass. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All spike train data files are available from <https://www.github.com/AraiKensuke/PP-AR>.

Funding: This work was supported by the National Institute of Mental Health R01 MH064537 <http://www.nimh.nih.gov/index.shtml>. 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.

Author summary

Oscillatory modulation of neural activity in the brain is widely observed under conditions associated with a variety of cognitive tasks and mental states. Within individual neurons, oscillations may be uncovered in the moment-to-moment variation in neural firing rate. This, however, is often challenging because many factors may affect fluctuations in neural firing rate and, in addition, neurons fire irregular sets of action potentials, or spike trains, due to an unknown combination of meaningful signals and extraneous noise. We have devised a statistical Latent Oscillatory Spike Train (LOST) model with accompanying

model-fitting technology, that is able to detect subtle oscillations in spike trains by taking into account both spiking noise and temporal variation in the oscillation itself. The method couples two techniques developed for other purposes in the literature on Bayesian analysis. Using data simulated from theoretical neurons and real data recorded from cortical motor neurons, we demonstrate the method's ability to track changes in the modulatory structure of the oscillation across experimental trials.

Introduction

Neural oscillations have generated considerable interest for their roles in cognition and as indicators for disease [1–14]. Electroencephalograms (EEGs) and local field potential (LFPs) recordings have revealed transient oscillations in many cortical and subcortical structures, to which neurons both near and far from the LFP recording site show phase preferences in spiking. Groups of neurons that are locked to a common oscillation, and are therefore active in tightly confined temporal windows, may define a cell assembly whose synchronous spiking could select and activate afferent structures [15–17]. Such transient cell assemblies may form and disband as objects are attended to in the visual scene [3, 6–8, 11], as preparation for movements are made [2, 13], or during choice points for reward [9]. Prominent slow oscillations are observed as patients undergo anesthesia [10], and exhibit changes in dynamics during transitions in unconscious brain state [14]. Furthermore, strong oscillatory signals are prominent in neurological disorders such as Parkinson's disease, and can be used to characterize pathophysiology and its relation to behavior [12, 18, 19]. The oscillatory structure of individual neurons can thus provide insight into the dynamics of cognitive processing and motor planning, and their implementation in health and disease. However, identifying oscillations in neural spike trains, particularly within individual experimental trials, can be challenging because they occur in conjunction with both non-oscillatory fluctuations in neural firing rate and extraneous noise. To tease apart multiple factors that affect spiking activity, statisticians have suggested the use of point process models together with modern regression methods associated with generalized linear model (GLM) technology [20, 21]. In this paper we develop a method that can find oscillations in spike trains even when the signal is comparatively weak, and can track changes in oscillatory behavior across trials.

The starting point for our approach is the observation by Smith and Brown [22] that, for many purposes, evolving neural firing rates can be described using state-space models imbedded in point processes. We modify the models used by Smith and Brown, replacing their first-order autoregressive processes with higher-order processes [23] that can fit oscillations found in spiking neurons. This requires careful attention to the form of the higher-order autoregressive process. We take advantage of a Bayesian time-series decomposition introduced by Huerta and West [24], using prior probability distributions to constrain the fit so that it has appropriate power spectral content. We also use a Gibbs sampling method developed recently in a different context [25] to compute posterior distributions efficiently.

Our approach is superficially related to that of Allcroft *et al.* [26], who used autoregressive moving average models to estimate spectral content from censored data, but their method does not accommodate easily the special situation we face with spiking neurons, including history effects that can account for hard and soft refractory periods. Point process models employing the GLM framework [21, 27] have been employed to explain the spiking behavior by taking into account spiking refractoriness, behavioral and stimulus-induced non-stationarities, and spiking of neighboring neurons [20]. A series of investigations [12, 18, 19] have

added long-term history effects to capture oscillation in the history dependence and have also used a state-space smoothing algorithm to track changes in the oscillatory dynamics over time. Eden et al used a long history to capture both, but as we shall see, LOST can better capture the irregularities present in realistic biological oscillations. Our approach is based on what we call the Latent Oscillatory Spike Train (LOST) model, which not only includes terms to describe the spiking refractoriness, event (behavior or stimulus)-locked effects, and trial-to-trial variability in firing rate, but also explicitly models the dynamics of the modulating oscillation itself. However, the latent state cannot capture the discontinuous change in the spiking probability that occurs after every spike, which the aforementioned GLM models do well in characterizing, so we utilize a short-term spiking history to capture the refractory period, while also introducing the latent state to model the oscillatory dynamics.

Oscillatory structure is often considerably degraded when it is observed as a spike train. On the one hand, if the firing rate is low, or the oscillatory modulation is weak, noise associated with erratic spiking will dominate. On the other hand, even if the modulation is strong and the firing rate is high, the oscillatory signal may itself be unsteady, with substantial variation in period from cycle to cycle. LOST accommodates both spiking noise and oscillatory variation by imbedding into the point process intensity function a latent auto-regressive process, and then imposing a suitable soft constraint (in the form of a prior probability distribution) on the oscillatory dynamics. We provide details of the model, along with the posterior sampling scheme and a discussion on how to interpret the posterior samples, and study its effectiveness with leaky integrate-and-fire neural simulations. We then demonstrate the ability of the LOST framework to find interesting trial-to-trial variation in motor cortical neuron during a lever-pulling task.

Model

We seek to extract the underlying oscillation from the binary ($M \times N$) matrix \mathbf{y} representing the binned spiking data from a single neuron during M repetitions of identical behavior or stimulus presentations of time duration N . Because the spiking times are different for each trial, even though the behavior or stimulus are identical, we construct a probabilistic model of the spiking.

A point process models the probability of observing a spike at time t . The conditional intensity function (CIF) completely specifies the point process, and is defined as

$$\lambda[t|\Theta, H(t)] = \lim_{\Delta t \rightarrow 0} \frac{P[N(t + \Delta t) - N(t) = 1|\Theta, H(t)]}{\Delta t}, \tag{1}$$

where $P[\cdot|\cdot]$ is a conditional probability conditioned on the past spiking history $H(t)$ and model parameters Θ [21], making the point process CIF a spike history-dependent generalization of a general inhomogeneous Poisson process. We discretize the problem, and note that $\lambda[t|H(t)]\Delta t$ is the approximate binary probability of observing a single spike in a small time interval Δt . In the discrete time representation, the likelihood of the spike train is then expressed as a product of conditionally independent Bernoulli events

$$P[\mathbf{y}|\Theta, H(t)] = \prod_{mn} p[y_{mn} = 1|\Theta, H(t)]^{y_{mn}} p[y_{mn} = 0|\Theta, H(t)]^{1-y_{mn}} \tag{2}$$

The binary spiking probability is expressed as a sigmoidal function of a sum of several predictors, including a history term λ^R , modeling the refractoriness following a spike. In the simplest case, the predictors will be the shared trial-average effect (TAE) $\mathbf{f} = \boldsymbol{\beta}^T \boldsymbol{\alpha}$, with $\boldsymbol{\beta}$ being the ($K_\beta \times N$) matrix of the K_β B-spline basis vectors and $\boldsymbol{\alpha}$ the K_α weights, the history term $\lambda^R = \mathbf{H}^T \mathbf{v}$,

with \mathbf{H} being the $(K_H \times N)$ matrix of the K_H B-spline basis vectors and \mathbf{v} the K_H weights, the M -dimensional trial-specific baseline offset vector $\boldsymbol{\mu}$, and a time-dependent latent oscillational modulation \mathbf{x} . The TAE is closely related to the peristimulus time histogram (PSTH) often employed in neuroscience. Together, the binary spiking probability for trial m at time n is

$$p(y_{mn} | x_{mn}, H_{n,l}, \Theta) = \frac{(e^{x_{mn} + \mu_m + f_n + \lambda_{l(m,n)}^R})^{y_{mn}}}{1 + e^{x_{mn} + \mu_m + f_n + \lambda_{l(m,n)}^R}}, \tag{3a}$$

$$x_{mn} = \sum_{j=1}^p F_j x_{m,n-j} + \epsilon_{mn}, \tag{3b}$$

where $l(m, n) \equiv n - L(m, n) \in \mathbb{Z}_{\geq 0}$, where $L(m, n)$ is the last spike time before time n of the m th trial. From the data, we infer the parameters and the latent oscillation \mathbf{x} , which is an AR(p) process, the $F_{1:p}$ the p AR coefficients and ϵ_{mn} a Gaussian random variable $\sim N(0, \sigma^2)$, where σ^2 is the innovation variance. In what follows, we will often refer to the latent oscillation as simply oscillation. We also note that the inferred \mathbf{x} may be spectrally broad, or have a very small amplitude, and may not be represent a signal that would commonly be called an oscillation.

The simplest oscillating AR model is of order 2 with complex eigenvalues, and may be a good model for a low dimensional oscillatory dynamical system. However fitting a model to noisy observed data utilizes noisy lagged values. The implied model for such a time series is an ARMA process, and an AR(2) would fail to capture all the structure in the observed data [24], so we approximate the ARMA process with a higher-order AR(p) process.

The spectra of AR(p) models is closely related to its $p = 2C + R$ roots of the characteristic polynomial of $F_{1:p}$, or equivalently to the eigenvalues of the matrix \mathbf{F} in Eq 6. The time-series decomposition [24] approach restricts the AR(p) to those with C imaginary AR(2) components and R real AR(1) components, ie only models that can accommodate the spectral features we believe are present in the data. We work through the rest of the paper using a generic structure that has proven to be a generally reasonable component structure ($C = 4$ and $R = 1$) for a wide range of simulation parameters and for real data, though model order and component structure can also be inferred from the data.

Model fitting

Gibbs sampling [28] and data augmentation [29, 30] are used to infer the model parameters $\Theta = [F_{1:p}, \sigma^2, \boldsymbol{\alpha}, \boldsymbol{\mu}, \mathbf{v}]$ and the oscillation \mathbf{x} . The spike history and TAE are functions of time relative to the last spike of an event or behavior, respectively, and they are estimated by parameterizing a subset of continuous functions with splines. The estimation requires us to choose a set of basis splines, which we choose heuristically, using some prior belief about their general shape to choose the knot locations. In the current work, the number and locations of the knots are not found with the Gibbs sampling procedure, so we briefly describe how they are determined before we describe the Gibbs sampling procedure.

Short-term spike history λ^R . Neurons typically have a refractory period following an action potential, when sodium channels in the membrane become inactive and spiking is strongly suppressed. As the channels reactivate, the neuron may experience a period of rebound excitation, where spiking probability becomes enhanced, before returning to some baseline level. This effect can be modeled by a modulation of probability since the last spike. We employ cubic splines to express the refractory modulation function, with knots placed where we would like to constrain function values, where we believe the function reaches

extremum values, and places in between. The interspike interval (ISI) histogram reflects combined effects of both the spiking history and oscillation, but at the smaller intervals, the refractory inhibition and rebound excitation are rapidly changing, possibly from 0 to well over 1, while a biologically plausible oscillation varies over a more limited range of values centered around 1. The shape of the ISI histogram for smaller intervals should reflect the refractory modulation function more than the oscillation, so we place history spline knots based on features of the ISI near 0. In the case the ISI histogram has a minimum at 0 followed by a maxima, we place knots at $t = \Delta$, one knot at the first maxima, one knot at the mean of the ISIs, and knots at the 70th and 80th percentiles of the ISI distribution. We further have knots at the 97th and at 100ms set to 0, constraining the asymptotic value of the history function. Further, if the bins of the ISI at time Δ are empty, we set $v_0 = -6$ to prevent it from taking unconstrained negative values. If the histogram at 0 is the maximum value, we leave v_0 unconstrained.

TAE spline basis β . To represent a smooth, arbitrary-shaped curve, we use the piecewise cubic B-spline as basis functions. Choosing the appropriate number and locations of basis functions is vital to represent the curve well, and too many points will result in overfitting. We determine the number and locations prior to the Gibbs procedure, by assuming that the shape of the TAE \mathbf{f} in Eq (3a) is similar to the empirical PSTH. We randomly vary the number and locations of the spline knots β and find a set which minimizes the squared error with the PSTH under the constraint that no more than 9 knots are to be used for the data presented in this paper. The shape of the PSTH can be used as a guide to guess the rough number of knots necessary, and the constraint can be changed accordingly for the spike train under investigation.

Posterior simulation. We now introduce the posterior distribution, from which the conditional posteriors necessary for Gibbs sampling, and the generating distributions necessary for data augmentation, are derived. Our model is a state-space model with a point-process observable, and a latent oscillation evolving as an AR(p)-process. The posterior distribution of the parameters Θ is

$$p(\Theta|\mathbf{y}) \propto p(\Theta) \int p(\mathbf{y}, \mathbf{x}|\Theta) d\mathbf{x} = p(\Theta) \int p(\mathbf{y}|\mathbf{x}, \Theta) p(\mathbf{x}|\Theta) d\mathbf{x} \tag{4}$$

For the AR(p) model, $p(\mathbf{x}|\Theta)$ does not have a simple expression. However, if \mathbf{x} were a Markov process, then $p(\mathbf{x}|\Theta)$ could be expressed recursively in terms of the initial probability at time 0. AR(p) can be made a Markov process if we define a new $(N \times M \times p)$ matrix \mathbf{X} , where $\mathbf{X}_{mn} = (x_{m,n-1}, \dots, x_{m,n-p})^T$. Eq (3b) is then

$$\mathbf{X}_{mn} = \mathbf{F}\mathbf{X}_{m,n-1} + \boldsymbol{\epsilon}_{mn}, \tag{5}$$

where

$$\mathbf{F} = \begin{pmatrix} F_1 & F_2 & \dots & F_{p-1} & F_p \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\epsilon}_{mn} = \begin{pmatrix} \epsilon_{mn} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{6}$$

$(N \times M)$ matrix \mathbf{x} is the oscillation we would like to infer, and likely corresponds to oscillations

of the membrane potential due to oscillating synaptic inputs [31, 32]. Eq (4) then becomes

$$p(\Theta|y) \propto p(\Theta) \int_{\mathbf{X}} \prod_{m=1}^M \left[\prod_{n=0}^N p(y_{mn}|x_{mn}, H_{n,l}, \Theta) \prod_{n'=1}^N p(\mathbf{X}_{mn'}|\mathbf{X}_{m,n'-1}, \Theta) p(\mathbf{X}_{m0}|\Theta) \right] d\mathbf{X}. \quad (7)$$

Conditional posteriors and data augmentation for Gibbs sampling

We jointly sample model parameters and latent states from the joint posterior distribution, Eq (4), using Gibbs sampling with data augmentation. The conditional posterior distribution can be read off from the full joint posterior distribution by considering all parameters fixed, except the parameter whose conditional posterior we are interested in. Our problem deviates from the standard Gibbs sampling by the missing data \mathbf{x} in Eq (4), and we utilize the data augmentation strategy [29, 30], a scheme of simplifying analysis by augmenting the observed data with missing values, to generate samples of \mathbf{X} from the predictive distribution $p(\mathbf{X}|y, \Theta)$ in between sampling parameters from the standard conditional posteriors. Even with samples of \mathbf{X} in hand, the posterior distribution does not yield conditional posteriors that can be sampled easily. The recently developed Pólya-Gamma data augmentation scheme [25] allows sampling from simple Gaussian conditional posteriors at the cost of introducing the $M \times N$ matrix of Pólya-Gamma variables ω through the following identity for the spike probability:

$$\begin{aligned} p(y_{mn}|\Theta, H_{n,l}, x_{mn}) &= \frac{(e^{x_{mn} + \mu_m + \lambda_{l(m,n)}^R + f_n})^{y_{mn}}}{1 + e^{x_{mn} + \mu_m + \lambda_{l(m,n)}^R + f_n}} = \frac{1}{2} e^{\kappa_{mn}(x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R)} \times \\ &\int_0^\infty e^{-\omega_{mn}} \frac{(x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R)^2}{2} PG(\omega_{mn}|b = 1, z = 0) d\omega_{mn} \\ &\propto \int_0^\infty N\left(x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R \mid \frac{\kappa_{mn}}{\omega_{mn}}, \frac{1}{\omega_{mn}}\right) PG(\omega_{mn}|b = 1, z = 0) d\omega_{mn} \end{aligned} \quad (8)$$

where $PG(\omega_{mn}|b = 1, z = 0)$ is the Pólya-Gamma distribution with parameters $b = 1$ and $z = 0$, and $\kappa \equiv \mathbf{y} - \frac{1}{2}$.

With the introduction of the Pólya-Gamma variables, the joint posterior Eq (4) becomes

$$\begin{aligned} p(\Theta|y) \propto p(\Theta) \int_{\mathbf{X}} \left[\prod_{m=1}^M \prod_{n'=1}^N p(\mathbf{X}_{mn'}|\mathbf{X}_{m,n'-1}, \Theta) p(\mathbf{X}_{m0}|\Theta) \times \right. \\ \left. \prod_{n=0}^N \int_0^\infty N(x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R \mid \frac{\kappa_{mn}}{\omega_{mn}}, \frac{1}{\omega_{mn}}) PG(\omega_{mn}|b = 1, z = 0) d\omega_{mn} \right] d\mathbf{X}, \end{aligned} \quad (9)$$

from which we derive the conditional posteriors for the parameters and the predictive distributions for the augmented variables. We refer to the set of augmented variables as $\mathbf{V}^+ = \{\mathbf{X}, \omega\}$, and denote by $\Theta_{\setminus x}$ to be the set of all parameters except parameter x .

Sampling μ . The conditional posterior of $\mu \sim N(\mathbf{M}_\mu, \Sigma_\mu)$ is a multivariate Gaussian, and if we constrain μ such that $\sum_m^M \mu_m = 0$, and set $\mu_1 = -\sum_{m=2}^M \mu_m$, it becomes

$$\begin{aligned}
 p(\mu | \Theta_{\mathbf{v}}, \mathbf{V}^+) &\propto \exp \left\{ \sum_{m>1,n}^{M,N} -\frac{\omega_{mn}}{2} \left[(x_{mn} + \mu_m + f_n + \lambda_{l(m,n)}^R) - \frac{\kappa_{mn}}{\omega_{mn}} \right]^2 \right. \\
 &\quad \left. - \sum_n^N \frac{\omega_{1n}}{2} \left[\left(x_{1n} - \sum_{m'>1}^M \mu_{m'} + f_n + O_{1n} \right) - \frac{\kappa_{1n}}{\omega_{1n}} \right]^2 - \frac{1}{2} (\mu - \mu_\mu)^T \mathbf{D}_\mu^{-1} (\mu - \mu_\mu) \right\} \quad (10) \\
 &= \exp \left(-\frac{1}{2} Q_\mu \right),
 \end{aligned}$$

where $O_{mn} \equiv \frac{\kappa_{mn}}{\omega_{mn}} - x_{mn} - f_n - \lambda_{l(m,n)}^R$. The covariance can be read off from the quadratic term,

$$\Sigma_\mu^{-1} = \begin{pmatrix} \sum_n (\omega_{1n} + \omega_{2n}) & \sum_n \omega_{1n} & \dots & \sum_n \omega_{1n} \\ \sum_n \omega_{1n} & \sum_n (\omega_{1n} + \omega_{3n}) & \dots & \sum_n \omega_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_n \omega_{1n} & \sum_n \omega_{1n} & \dots & \sum_n (\omega_{1n} + \omega_{Mn}) \end{pmatrix} \quad (11)$$

The mean \mathbf{M}_μ is at the minimum of Q_μ , which can be found by setting $\frac{\partial Q_\mu}{\partial \mu} = 0$, or $\Sigma_\mu^{-1} \mathbf{M}_\mu = \mathbf{g}$, where $g_i = \sum_n \omega_{1n} O_{1n} - \omega_{in} O_{in}$.

Sampling \mathbf{v} . \mathbf{v} are the vector spline weights for the spike history, and $\{I_H\} = \{1, \dots, K_H\}$ its component indices. We impose our prior belief about the timescale and shape of the spike history term by keeping some of the knots fixed, while re-weighting the remaining ones. The indices of the fixed ones we denote $\{I_f\}$ and the adjustable one $\{I_a\}$, with $\{I_f\} \cup \{I_a\} = \{I_H\}$ and $\{I_f\} \cap \{I_a\} = \emptyset$. We use the shorthand notations \mathbf{a} to be the vector comprised of components $\{I_a\}$ of \mathbf{v} , and $\{I_a\} = \{a(1), a(2), \dots, a(|\{I_a\}|)\}$, its numerical values. Then the conditional posterior of $\mathbf{a} \sim N(\mathbf{M}_a, \Sigma_a)$ is a multivariate Gaussian. Defining $O_{mn} \equiv \left(\frac{\kappa_{mn}}{\omega_{mn}} - x_{mn} - f_n - \mu_m \right)$, it is

$$\begin{aligned}
 p(\mathbf{a} | \Theta_{\mathbf{v}}, \mathbf{V}^+) &\propto \exp \left[-\sum_m^M \left\{ \sum_n^N \frac{\omega_{mn}}{2} (\lambda_{l(m,n)}^R - O_{mn})^2 + \dots \right\} - \frac{1}{2} (\mathbf{v} - \mu_{\mathbf{v}})^T \mathbf{D}^{-1} (\mathbf{v} - \mu_{\mathbf{v}}) \right] \\
 &= \sum_m^M \left\{ \sum_n^N -\frac{\omega_{mn}}{2} \left(\sum_j^{K_H} H_{l(m,n),j} v_j - O_{mn} \right)^2 \right\} - \frac{1}{2} \sum_j \frac{v_j^2 - 2v_j \mu_{v_j} + \mu_{v_j}^2}{\sigma_{v_j}^2} \quad (12) \\
 &= \exp \left(-\frac{1}{2} Q_a \right).
 \end{aligned}$$

The mean \mathbf{M}_a of the adjustable knots is the solution of the set of linear equations whose i th equation, $i \in \{I_a\}$ is

$$\begin{aligned}
 &\sum_{j \in \{I_a\}} 2 \left(\sum_m^M \sum_n^N \omega_{mn} H_{l(m,n),j} H_{l(m,n),i} + \frac{\delta_{ji}}{\sigma_{v_j}^2} \right) v_j = \\
 &= 2 \left(\sum_m^M \sum_n^N \omega_{mn} H_{l(m,n),i} + \frac{\mu_{v_i}}{\sigma_{v_i}^2} \right) - 2 \sum_{j \in \{I_f\}} \left(\sum_m^M \sum_n^N \omega_{mn} H_{l(m,n),j} H_{l(m,n),i} + \frac{\delta_{ji}}{\sigma_{v_j}^2} \right) v_j
 \end{aligned}$$

The covariance can be read off from the quadratic term, and is

$$\Sigma_a^{-1} = \begin{pmatrix} \sum_{mn} \omega_{mn} H_{l(m,n),a(1)}^2 + \frac{1}{\sigma_{v_{a(1)}}^2} & \sum_{mn} \omega_{mn} H_{l(m,n),a(1)} H_{l(m,n),a(2)} & \cdots \\ \sum_{mn} \omega_{mn} H_{l(m,n),a(1)} H_{l(m,n),a(2)} & \sum_{mn} \omega_{mn} H_{l(m,n),a(2)}^2 + \frac{1}{\sigma_{v_{a(2)}}^2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad (13)$$

Sampling α . The conditional posterior of $\alpha \sim N(\mathbf{M}_\alpha, \Sigma_\alpha)$ is multivariate Gaussian,

$$p(\alpha | \Theta_{\alpha}, \mathbf{V}^+) \propto \exp \left\{ \sum_{m \neq 1, n}^{M, N} -\frac{\omega_{mn}}{2} \left[(x_{mn} + \mu_m + f_n + \lambda_{l(m,n)}^R) - \frac{\kappa_{mn}}{\omega_{mn}} \right]^2 \right\} \quad (14)$$

By rearranging, we can easily calculate the mean and variance of the conditionals for α . If n or m is kept fixed in the sum, this is a quadratic form of M or N -dim vectors with a diagonal covariance matrix, and $\alpha \sim N(\mathbf{M}, \Sigma)$, For α ,

$$\begin{aligned} \mathbf{M}_\alpha &= \mu_\alpha + \mathbf{D}_\alpha \boldsymbol{\beta} (\boldsymbol{\beta}^T \mathbf{D}_\alpha \boldsymbol{\beta} + \mathbf{W})^{-1} (\mathbf{O} - \boldsymbol{\beta}^T \mu_\alpha) \\ \Sigma_\alpha^{-1} &= \boldsymbol{\beta} \mathbf{W}^{-1} \boldsymbol{\beta}^T + \mathbf{D}_\alpha^{-1} \\ \mathbf{W}^{-1} &= \sum_m \mathbf{W}_m^{-1}, \text{ where } (\mathbf{W}_m^{-1})_{nn} = \omega_{mn} \\ \mathbf{O} &= \mathbf{W} \sum_m \mathbf{W}_m^{-1} \mathbf{O}_m, \text{ where } (\mathbf{O}_m)_n = \left(\frac{\kappa_{mn}}{\omega_{mn}} - x_{mn} - \lambda_{l(m,n)}^R - \mu_m \right). \end{aligned} \quad (15)$$

μ_α and \mathbf{D}_α are prior mean and covariances for α . Aside from the smoothness imposed by the restriction to the number of knots, we have no prior knowledge about the offsets and knot weights, so we set the means to 0 and the covariances to be broad.

Sampling σ^2 . The conditional posterior distribution of the innovation variance is

$$p(\sigma^2 | \Theta_{\sigma^2, \mathbf{x}, \boldsymbol{\omega}}) \propto \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{M(N-1)} \exp \left(-\frac{1}{2\sigma^2} \sum_{mn} [x_{mn} - (\mathbf{F}\mathbf{X}_{m,n-1})_1]^2 \right), \quad (16)$$

and is in the form of an inverse Gamma distribution. We choose a conjugate inverse Gaussian prior with parameters α_{σ^2} and β_{σ^2} , and sample as

$$\sigma^2 \sim IG \left[\sigma^2 | \alpha_{\sigma^2} + \frac{M(N-1) + 2}{2}, \beta_{\sigma^2} + \frac{1}{2} \sum_{mn} [x_{mn} - (\mathbf{F}\mathbf{X}_{m,n-1})_1]^2 \right] \quad (17)$$

As with the offset and spline weights, we have no *a priori* knowledge about an appropriate range of values for σ^2 . In order to minimize the effect of the prior, we choose a small values for α_{σ^2} (relative to the amount of data), and β_{σ^2} , so that small values of σ^2 can also be sampled.

Sampling $F_{1,p}$. Without further assumptions on the latent dynamics [33], the conditional posterior is a simple multivariate normal distribution. However, for application to spiking data, this naive approach often does not produce desired results for two main reasons. First, the constrained sampling of the coefficients to the stationary regime is inefficient, due to the nontrivial shape of the surface of the stationary manifold [24]. Second, our observations are noisy and sparse, and maximum likelihood using a flat prior may not find a state with an

oscillatory character. The component decomposition [24] approach addresses both of these concerns, and allows us to specifically search for models with an oscillatory character through the imposition of an informative prior.

We seek to sample

$$p(F_{1:p} | \Theta_{\setminus F_{1:p}}, \mathbf{V}^+) \propto \exp\left(-\frac{1}{2\sigma^2} \sum_{mn} [x_{mn} - (\mathbf{F}\mathbf{X}_{m,n-1})_1]^2\right) p(F_{1:p}), \tag{18}$$

with the prior $p(F_{1:p})$ a structural prior which restricts sampling of $F_{1:p}$ to the subspace of coefficients for which the process is stationary and its characteristic polynomial has R real roots and C pairs of imaginary roots. For what follows, we drop the pairs qualifier when referring to one of the pairs of imaginary roots. Instead of sampling the entire $F_{1:p}$ at once, we sample one of the C imaginary or R real roots at a time while keeping the rest of the roots fixed, via Markov Chain Monte Carlo (MCMC) as follows.

The residuals of the AR(p) process are uncorrelated Gaussians, but leaving out one of the roots results in residuals that are correlated. We now outline the case when sampling an imaginary root. Here, we consider the correlated residual to be an AR(2) process. Ignoring the trial index, x_n is related to the p lagged values x_{n-1}, \dots, x_{n-p} through the AR coefficients and an innovation variance. Utilizing the backshift operator, $\hat{B}x_n = x_{n-1}$, eq (3b) can be written as a polynomial in \hat{B} , and the polynomial as a product of its R real roots $\alpha_j, j \in \{1, \dots, R\}$ and C pairs of complex roots $\gamma_j, \gamma_j^*, j \in \{1, \dots, C\}$.

With the roots calculated, the AR process, Eq (3b), can be represented as

$$\epsilon_n = \left[\prod_r^R (1 - \alpha_r \hat{B}) \prod_c^C (1 - \gamma_c \hat{B})(1 - \gamma_c^* \hat{B}) \right] x_n. \tag{19}$$

The coefficient for $\hat{B}^i, i \in \{1, \dots, p\}$, is F_i . Regrouping the terms, we rewrite the residuals when the contribution of the j th complex root is removed, to obtain a new time series ω_{jn} :

$$w_{jn} = (1 - \gamma_j \hat{B})^{-1} (1 - \gamma_j^* \hat{B})^{-1} \epsilon_n = \prod_r^R (1 - \alpha_r \hat{B}) \prod_{c \neq j}^C (1 - \gamma_c \hat{B})(1 - \gamma_c^* \hat{B}) x_n. \tag{20}$$

This equation tells us how to obtain the ω_{jn} , and that we are treating that time series as an AR (2). We note that given a timeseries \mathbf{x}_m and some $F_{1:p}$, we could calculate the residual ϵ , and this would be the same series of N numbers had we first obtained the \mathbf{w}_j and then calculated the $N \epsilon_n = (1 - \gamma_j \hat{B})(1 - \gamma_j^* \hat{B}) w_{jn}$.

Given that its eigenvalues γ_j, γ_j^* have corresponding angle $2\pi/\lambda_j$ and modulus r_j , which control the frequency and the steadiness of the oscillation when these parameters are used to generate a time series. Using Vieta's formula [34], the AR(2) coefficients $\phi_1 = -(\gamma_j + \gamma_j^*)$ and $\phi_2 = \gamma_j \gamma_j^*$ are related to the frequency and modulus as $2r_j \cos(2\pi/\lambda_j)$ and $-r_j^2$, and

$$\phi_{j1}, \phi_{j2} \sim N\left(\mathbf{h}, \frac{\mathbf{H}^{-1}}{\sigma^2}\right) \tag{21}$$

where

$$\mathbf{H}_j^{-1} = \mathbf{D}_j^T \mathbf{D}_j, \quad \mathbf{D}_j^T = \begin{pmatrix} w_{j0} & w_{j1} & \dots & w_{j,N-1} \\ w_{j-1} & w_{j0} & \dots & w_{j,N-2} \end{pmatrix} \quad (22)$$

$$\mathbf{h}_j = \mathbf{H}_j \mathbf{D}_j^T \mathbf{w}'_j, \quad \mathbf{w}'_j{}^T = (w_{j1} \quad w_{j2} \quad \dots \quad w_{jN})$$

The stationary regime is defined by $r_j \leq 1$, so in sampling from the conditional posterior, we must restrict the sampling to regions bounded by the curves

$$\frac{\phi_{j1}^2}{4} + \phi_{j2} = 0, \quad \phi_{j2} = -1 \quad (23)$$

This boundary describes a truncated 2-variate normal distribution.

When the root in question is real, we remove its contribution to obtain the correlated residuals

$$u_{jn} = (1 - \alpha_j \hat{B})^{-1} \epsilon_n = \prod_{\tau \neq j}^R (1 - \alpha_\tau \hat{B}) \prod_c^C (1 - \gamma_c \hat{B}) (1 - \gamma_c^* \hat{B}) x_n. \quad (24)$$

The real roots can be sampled from a truncated univariate normal with mean m_j and variance M_j

$$m_j = \frac{\sum_{n=1}^N u_{jn} u_{j,n-1}}{\sum_{n=1}^N u_{j,n-1}^2}, \quad M_j = \frac{\sigma^2}{\sum_{n=1}^N u_{j,n-1}^2} \quad (25)$$

We further assume a simple informative, uniformly distributed prior of [0.97, 1] for the modulus of the slowest root. Large modulus roots are spectrally peaked, and correspond to oscillations with a relatively well-defined frequency. Our choice of a prior on the modulus is similar to adding regularization terms to the likelihood to address concerns about overfitting when data is limited, an alternative avenue recently investigated by Buesing et al [35], although this requires the structured decomposition of the AR model, which would be difficult to express in terms of regularization functions.

The latent oscillation: X data augmentation. The predictive distribution for \mathbf{X} can be read off from Eq (4), and is

$$\mathbf{X} \sim \prod_m^M \left[\prod_{n=0}^N N \left(x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R \left| \frac{\kappa_{mn}}{\omega_{mn}}, \frac{1}{\omega_{mn}} \right. \right) \prod_{n'=1}^N p(\mathbf{X}_{mn'} | \mathbf{X}_{m,n'-1}, \Theta) p(\mathbf{X}_{m0} | \Theta) \right] \quad (26)$$

The m th observation, which through augmentation by the Pólya-Gamma variables, is no longer the spikes, but a continuous variable $t_{mn} = \kappa_{mn}/\omega_{mn} - f_n - \mu_m - \lambda_{l(m,n)}^R$ with observation noise $r_{mn} = 1/\omega_{mn}$. We generate a sample of \mathbf{X} given observation \mathbf{t} and fixed Θ by employing the FFBS (Forward-filter backward-sample) algorithm [30]. Briefly, FFBS is a recursive algorithm with the same forward filtering step as the Kalman Filter, but on the backward step, instead of obtaining the mean from the smoothing density, we obtain realizable sample paths of the latent state from the smoothing density. The M trials are independent, so for the sake of simple notation, we drop the trial index m in what follows. Forward filtering is done by

recursive application of a prediction and filtering step.

$$\begin{array}{cc}
 \text{prediction} & \text{filter} \\
 \hat{\mathbf{X}}_{m,n|n-1} = \mathbf{F}\hat{\mathbf{X}}_{m,n-1|n-1} & \hat{\mathbf{X}}_{m,n|n} = \hat{\mathbf{X}}_{m,n-1|n-1} + \mathbf{K}_{mn}(t_{mn} - \mathbf{H}\hat{\mathbf{X}}_{m,n-1|n-1}) \\
 \hat{\mathbf{V}}_{m,n|n-1} = \mathbf{F}\hat{\mathbf{V}}_{m,n-1|n-1}\mathbf{F}^T + \mathbf{Q} & \hat{\mathbf{V}}_{m,n|n} = (\mathbf{I} - \mathbf{K}_{mn}\mathbf{H})\hat{\mathbf{V}}_{m,n-1|n-1}
 \end{array} \tag{27}$$

with

$$\mathbf{Q} = \begin{pmatrix} \sigma^2 & 0 \dots \\ 0 & \mathbf{0} \\ \vdots & \end{pmatrix}, \quad \mathbf{H} = (1 \quad 0 \quad \dots) \tag{28}$$

and the Kalman gain, $\mathbf{K}_{mn} = \hat{\mathbf{V}}_{m,n-1|n-1}\mathbf{H}^T(\mathbf{H}\hat{\mathbf{V}}_{m,n-1|n-1}\mathbf{H}^T + r_{mn})^{-1}$. The prediction mean is straightforward given the structure of \mathbf{F} , while the prediction covariance is

$$\hat{\mathbf{V}}_{m,n-1|n-1} = \begin{pmatrix} \sum_{ij} F_i(\hat{\mathbf{V}}_{m,n-1|n-1})_{ij} F_j + \sigma^2 & \sum_i F_i(\hat{\mathbf{V}}_{m,n-1|n-1})_{i,1..p-1} \\ \sum_i F_i(\hat{\mathbf{V}}_{m,n-1|n-1})_{1..p-1,i} & (\hat{\mathbf{V}}_{m,n-1|n-1})_{1..p-1,1..p-1} \end{pmatrix} \tag{29}$$

The Kalman gain becomes

$$\mathbf{K}_{mn} = \frac{1}{(\hat{\mathbf{V}}_{m,n-1|n-1})_{11} + r_{mn}} \begin{pmatrix} (\hat{\mathbf{V}}_{m,n-1|n-1})_{11} \\ \vdots \\ (\hat{\mathbf{V}}_{m,n-1|n-1})_{p1} \end{pmatrix}, \tag{30}$$

while filter mean is $\hat{\mathbf{X}}_{m,n|n} = \hat{\mathbf{X}}_{m,n-1|n-1} + \mathbf{K}_{mn}[t_{mn} - (\hat{\mathbf{X}}_{m,n-1|n-1})_{11}]$ and covariance is

$$(\hat{\mathbf{V}}_{m,n|n})_{ij} = (\hat{\mathbf{V}}_{m,n-1|n-1})_{ij} - \frac{(\hat{\mathbf{V}}_{m,n-1|n-1})_{1i}(\hat{\mathbf{V}}_{m,n-1|n-1})_{1j}}{(\hat{\mathbf{V}}_{m,n-1|n-1})_{11} + r_{mn}} \tag{31}$$

FFBS requires starting values $\hat{\mathbf{X}}_{m,0|0}$ and $\hat{\mathbf{V}}_{m,0|0}$ for the recursion, for which we choose $\mathbf{0}$ and a reasonably broad, diagonal covariance, respectively. The inferred oscillation in the begin of the trials will be of poorer quality, but improves quickly after enough spikes are observed. After the last time point is reached, the latent state is sampled backwards in time. The smoothing distribution can be written

$$p(\mathbf{X}_{m,1..N}|t_{m,1..N}, \Theta) \propto p(\mathbf{X}_N|t_{m,1..N}, \Theta) \prod_{n=0}^{N-1} p(\mathbf{X}_{mn}|\mathbf{X}_{m,n+1}, t_{m,1..n}, \Theta), \tag{32}$$

the first term a normal distribution about $\hat{\mathbf{X}}_{m,N|N}$ with covariance $\hat{\mathbf{V}}_{m,N|N}$, and the term under the product written in terms of known distributions

$$p(\mathbf{X}_{mn}|\mathbf{X}_{m,n+1}, t_{m,1..n}, \Theta) \propto p(\mathbf{X}_{m,n+1}|\mathbf{X}_{mn}, \Theta)p(\mathbf{X}_{mn}|t_{m,1..n}, \Theta), \tag{33}$$

which is a product of the transition density, which is a normal distribution with mean $\mathbf{F}\mathbf{x}_{mn}$ and variance \mathbf{Q} , and a filter density with mean $\hat{\mathbf{X}}_{m,n|n}$ and variance $\hat{\mathbf{V}}_{m,n|n}$. Defining

$\mathbf{A} \equiv \hat{\mathbf{V}}_{m,n|n} \mathbf{F}^T (\mathbf{F} \hat{\mathbf{V}}_{m,n|n} \mathbf{F}^T + \mathbf{Q})^{-1}$, the sampling distribution of $\mathbf{X}_{m,n}$ is a multivariate normal with mean and covariance given by

$$\begin{aligned} \hat{\mathbf{X}}_{m,n|N} &= (\mathbf{I} - \mathbf{A}\mathbf{F})\hat{\mathbf{X}}_{m,n|n} + \mathbf{A}\mathbf{X}_{m,n+1} \\ \hat{\mathbf{V}}_{m,n|N} &= (\mathbf{I} - \mathbf{A}\mathbf{F})\hat{\mathbf{V}}_{m,n|n}. \end{aligned} \tag{34}$$

$\mathbf{X}_{m,n+1}$ is the previously-sampled latent state \mathbf{Q} can be expressed as an outer product of a column vector $\mathbf{u} = (\sigma, 0, 0, \dots)$ with itself, $\mathbf{u}\mathbf{u}^T$. Using the Sherman-Morrison formula [36] for the rank-1 update of an inverse matrix, the moments of the backward-sampling densities may be written as

$$\hat{\mathbf{X}}_{m,n|N} = \begin{pmatrix} (\mathbf{X}_{m,n+1})_2 \\ \vdots \\ (\mathbf{X}_{m,n+1})_{p-1} \\ F_s - \frac{\sigma^2 (\mathbf{F}^{-1})^2 [F_s (\mathbf{V}_{m,n|n})_{pp} + \sum_{i=1}^{p-1} (\mathbf{X}_{m,n+1})_{i+1} (\mathbf{V}_{m,n|n})_{pi} - \sum_i^p (\mathbf{V}_{m,n|n})_{pi} (\hat{\mathbf{X}}_{m,n|n})_i]}{1 + \sigma^2 (\mathbf{V}_{m,n|n})_{pp} (\mathbf{F}^{-1})^2_{p1}} \end{pmatrix} \tag{35}$$

$$\hat{\mathbf{V}}_{m,n|N} = \frac{1}{1 + \sigma^2 (\mathbf{V}_{m,n|n})_{pp} (\mathbf{F}^{-1})^2_{p1}} \begin{pmatrix} 0 & 0 \\ & \vdots \\ 0 & \dots & \sigma^2 (\mathbf{F}^{-1})^2_{p1} \end{pmatrix} \tag{36}$$

where $F_s \equiv \sum_{i=1}^p (\mathbf{F}^{-1})_{pi} (\mathbf{X}_{m,n+1})_i$. The AR(p) model with the extended state-space is only partly dynamic [30], with $\text{rank}(\mathbf{Q}) = 1 < p$, and the computations necessary for the FFBS equations are much less than would be necessary for a full-rank model, as can be seen from the calculations of various moments presented in this section.

Pólya-Gamma data augmentation. The predictive distribution for ω_{mn} , the term under the integral marginalizing out ω_{mn} in Eq (8), is the general PG($\omega_{mn}|b = 1, z \neq 0$) distribution, obtained by an exponential tilting of the PG($\omega_{mn}|b = 1, z = 0$) distribution [25]:

$$\omega_{mn} \sim \text{PG}(\omega_{mn}|1, \psi_{mn}) = 2^{-1} e^{\psi_{mn}} e^{-\frac{\omega_{mn} \psi_{mn}}{2}} \text{PG}(\omega_{mn}|1, 0), \tag{37}$$

with $\psi_{mn} = x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R$. This is sampled using the rejection sampler described in Polson *et al* [25].

Inferring additional latent structures. Oscillations temporally structure the spike train, but it is possible that these oscillations do not occur uniformly throughout all trials. The LOST model allows inferring additional structure present in the neural data, such as modulation of oscillatory strength within a trial, or trends in modulation strength across trials. In the simulation studies, we consider an example where the spike train is significantly modulated in only a fraction of the trials, and demonstrate that the LOST model is able to identify these trials. We also apply this to real data to show the LOST model can predict, using only spike train data, which trials are strongly modulated to a theta rhythm in the LFP, and which are not. For these

applications, we assume the modulation strength takes on 2 distinct values. We fix the modulation strength of unmodulated trials to $s_0 = 0.1$, and parameterize the modulation strength of the modulated trials, s_1 . We introduce the latent indicator variable \mathbf{Z} , which is an $(M \times 2)$ matrix where, for example, $\mathbf{Z}_m = (1, 0)$ if the m th trial is an unmodulated trial, and also the parameter mixture weights $\boldsymbol{\pi}$, a 2 component column vector whose components sum to 1. The likelihood is then

$$\begin{aligned}
 p(\mathbf{y}|\Theta) &= \int_{\mathbf{X}} \prod_m \sum_{\mathbf{Z}_m} p(\mathbf{y}_m, \mathbf{Z}_m | \mathbf{X}_m, \Theta) p(\mathbf{X}_m | \Theta) d\mathbf{X} \\
 &= \int_{\mathbf{X}} \prod_m \sum_{\mathbf{Z}_m} \prod_j [\pi_j p_j(\mathbf{y}_m, | \mathbf{X}_m, \Theta)]^{Z_{mj}} p(\mathbf{X}_m | \Theta) d\mathbf{X} \\
 &= \int_{\mathbf{X}} \prod_m \sum_{\mathbf{Z}_m} \prod_j \left[\pi_j \prod_n \text{Ber}(y_{mn}, | x_{mn}, Z_{mj}, \Theta) \right]^{Z_{mj}} p(\mathbf{X}_m | \Theta) d\mathbf{X} \\
 &= \int_{\mathbf{X}} \prod_m \sum_{\mathbf{Z}_m} \prod_j \left[\pi_j \times \prod_n \int N\left(s_j x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R \middle| \frac{\kappa_{mn}}{\omega_{mn}}, \frac{1}{\omega_{mn}}\right) p(\omega_{mn}) d\omega_{mn} \right]^{Z_{mj}} p(\mathbf{X}_m | \Theta) d\mathbf{X}
 \end{aligned} \tag{38}$$

Define M_L to be all m such that $\mathbf{Z}_m = (1, 0)$ and M_H to be all m such that $\mathbf{Z}_m = (0, 1)$. The likelihood contribution to the conditional posterior for s_1 is Gaussian, $s_1 \sim N\left(s_1 \middle| \frac{B}{2A}, \frac{1}{2A}\right)$, where

$$A \equiv \sum_{m \in M_L, n} \frac{\omega_{mn} x_{mn}^2}{2} \quad B \equiv \sum_{m \in M_L, n} \omega_{mn} x_{mn} \left(\frac{\kappa_{mn}}{\omega_{mn}} - f_n - \mu_m - \lambda_{l(m,n)}^R \right). \tag{39}$$

For the prior distribution for s_1 , we use $p(s_1) = N\left(s_1 \middle| \mu_{s_1}, \sigma_{s_1}^2\right)$, where $\mu_{s_1} = s_0$, and a large variance $\sigma_{s_1}^2$, because we have no prior belief about what s_1 should be. The data augmentation of \mathbf{X} analogous to Eq (26) is carried out by replacing the m th observation and noise by

$$t_{mn} \rightarrow \frac{\kappa_{mn}/\omega_{mn} - f_n - \mu_m - \lambda_{l(m,n)}^R}{\prod_j s_j^{z_{mj}}} \quad r_{mn} \rightarrow \frac{1}{\left(\omega_{mn} \prod_j s_j^{z_{mj}}\right)^2}. \tag{40}$$

The Pólya-Gamma data augmentation analogous to Eq (37) is carried out by replacing the m th ψ_{mn} by

$$\psi_{mn} \rightarrow \prod_j s_j^{z_{mj}} x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R. \tag{41}$$

The data augmentation for Z_{mj} is done by generating from a 2-category multinomial distribution. The probability of the m th trial being in the j th category is proportional

$$p(Z_{mj} = 1 | \Theta, \mathbf{V}^+) \propto \pi_j \prod_n \frac{\left(e^{s_j x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R} \right)^{y_{mn}}}{1 + e^{s_j x_{mn} + f_n + \mu_m + \lambda_{l(m,n)}^R}} \tag{42}$$

The $\boldsymbol{\pi}$ are sampled from a Dirichlet distribution

$$\boldsymbol{\pi} \sim \text{Dir} \left(\alpha_{\pi,0} + \sum_{m \in M_L} Z_{m1}, \alpha_{\pi,1} + \sum_{m \in M_H} Z_{m2} \right). \tag{43}$$

We set the vector $\boldsymbol{\alpha}_{\pi} = (1, 1)$.

Numerical computation

Gibbs sampling was implemented in the Python programming language using the Numpy, SciPy, Matplotlib, StatsModels, and Patsy toolboxes. Numerical routines were written in C/C++ and Cython. Software for the Gibbs sampler and a Python wrapper for the Pólya-Gamma routine based on the original code by Jesse Windle, available at <https://github.com/jwindle>, is available at <https://github.com/AraiKensuke/PP-AR> and <https://github.com/AraiKensuke/pyPG>, respectively.

Results

To illustrate how the LOST model can infer latent oscillations under a variety of realistic conditions, we simulated biologically plausible oscillations, and spike trains modulated with these oscillations using either a simple inhomogeneous renewal process, or the leaky integrate-and-fire (LIF) neuron model. Further, we describe another point process method to infer the CIF, the GLM method, for comparison with the LOST model.

Simulated oscillations and spike trains

The stochastic oscillation for the m th trial \mathbf{w}_m is generated by first generating a stochastic phase \mathbf{t}_m and stochastic amplitude \mathbf{A}_m as

$$t_{m,n+1} = t_{mn} + \Delta t \left(1 + \frac{\xi_{mn}}{C_\xi} \right) \quad w_{mn} = 1 + \frac{A_{mn}}{C_A} \sin(2\pi v t_{mn}). \quad (44)$$

where ξ_m and \mathbf{A}_m are generated by AR(1) processes, and $t_{m0} \in [0, 1]$ a uniform random number. These are then used to generate the irregular oscillation C_ξ and C_A control the size of deviation from uniformity, while the timescale of the AR(1) processes controls the timescale of variation of amplitude and phase in the oscillation. AR(1)s with timescales on the order of the oscillation period $T = 1/v$ produce fluctuations that causes considerable variability of period and amplitude in \mathbf{w} , which we quantify with the oscillator irregularity, the oscillatory coefficient of variation (OCV) of the instantaneous periods T of oscillation, defined as the ratio of the standard deviation to the mean of time intervals between phase 0 crossings of an oscillatory signal. Fig 1 shows 2 example oscillations generated by this procedure, their OCVs and their power spectral densities. Irregular oscillations have a higher OCV and a less peaked spectral density.

To generate spikes, we employ an inhomogeneous renewal process or the LIF model, driven by a stochastic model of oscillation to modulate the firing of spikes. For the inhomogeneous renewal process, for each trial m and time n , if for a uniform random number $r_{mn} \in [0, 1]$, $r_{mn} < \Delta t \left(e^{\mu_m + w_{mn} + \lambda_{l(m,n)}^G} \right)$, we generated a spike, with $\lambda_{l(m,n)}^G$ the user-defined ground truth refractory history function. For the LIF model, spikes were generated using

$$\Delta V_{mn} = \left(-\frac{V_{mn}}{\tau} + f_n + \mu_m + w_{mn} + \chi_{mn} \right) \Delta t, \quad (45)$$

where the Gaussian random variable $\chi_{mn} \sim N(0, \sigma_b^2)$ represents irregular arrival times of afferent spikes, $\tau = 0.2$ the membrane time constant, μ_m a baseline DC current and $\sigma_b^2 = 210$ the amplitude of background fluctuation. Spikes are elicited when $V_{mn} \geq 1$ passes the threshold value of 1, which then causes a reset to $V_{m,n+1} = 0$. LIF have a refractory period that depends on the model parameters, and therefore the CIF naturally is dependent on the last spike time.

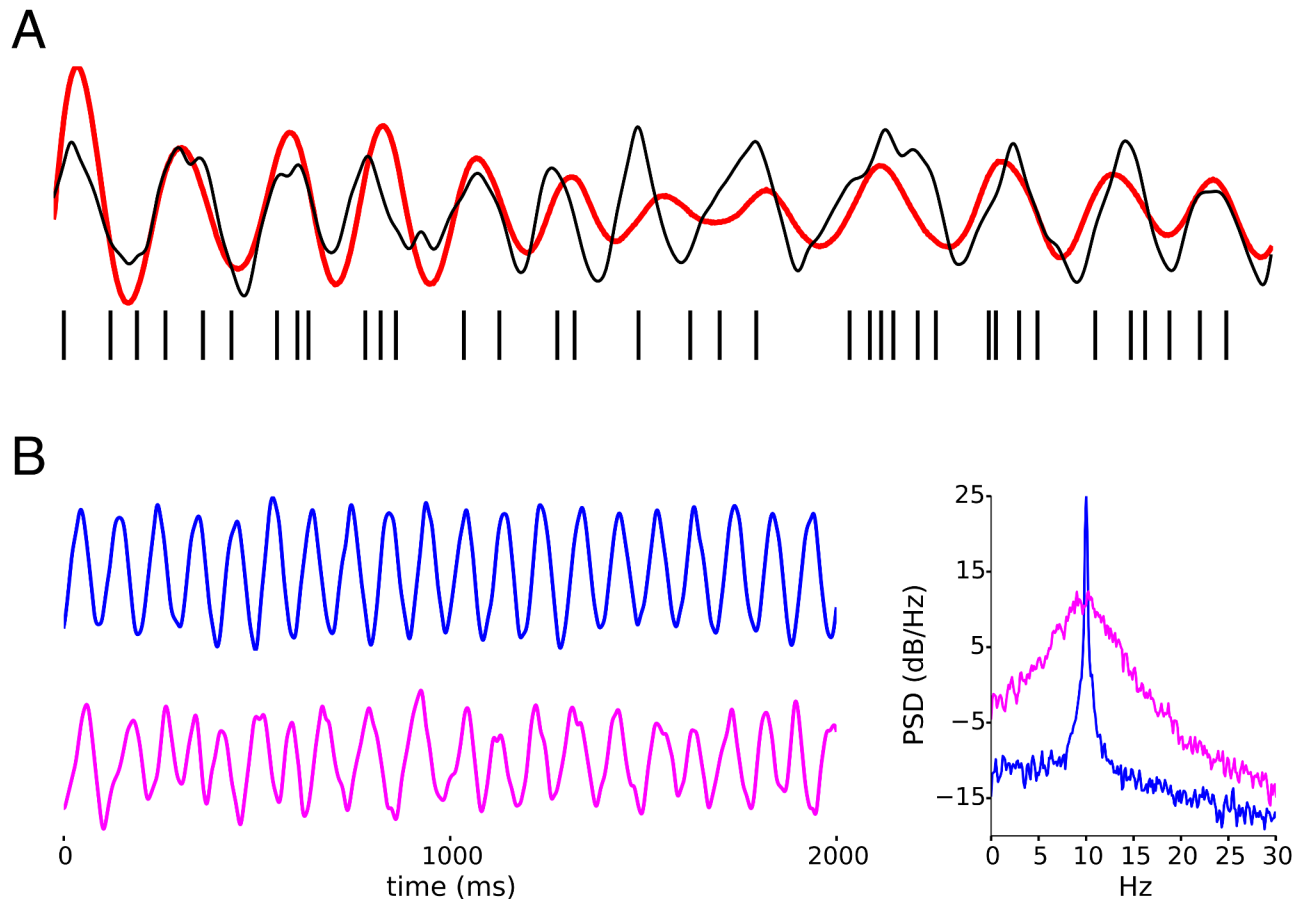


Fig 1. Regular and irregular oscillations. A) LOST infers (red) the oscillatory modulation (black), if any, underlying the spikes (vertical ticks) in a spike train, using only the spiking data. The oscillation amplitude stabilizes after the first few spikes are observed. B) We refer to spectrally narrow (wide) oscillations as regular (irregular), shown by the blue (pink) oscillation with $OCV = 0.02$ (0.22). The correlation between the current and past phase decays slowly (quickly) as a function of lag in regular (irregular) oscillations.

<https://doi.org/10.1371/journal.pcbi.1005596.g001>

LIF neuron with stimulus-triggered change in firing rate. We demonstrate the basic operation of LOST and the inference of the TAE, spike history and oscillation. In the following, we monitor a small subset of the full joint posterior parameters: the modulus r , frequency $2\pi/\lambda$ of the lowest AR(2) component, and the latent state amplitude defined by its standard deviation. Samples of trial offsets or the TAE knots approach their stationary distribution values much more quickly than these parameters representative of the latent state, so monitoring this smaller subset suffices for determining when the posterior LOST draws from has reached its stationary distribution. Fig 2A shows a clear oscillation being inferred, as the parameters show little uncertainty. We took the mean of the last 2000 iterations to calculate the means in Fig 2B, and the oscillation in Fig 2C. In most of the following figures, we use the convention that cyan represents sampled values themselves, their distribution or 5-95% intervals of the sampled values. Red lines represent means of sampled values, and black lines represent ground truth values.

Model identifiability and interpreting the results of LOST. Oscillations and post-spike refractory periods both add temporal structure to spike trains. Because they are presumably separate biological mechanisms and modify the spike train in distinct ways, we model them separately using the latent state and the spike history term. Without imposing constraints on

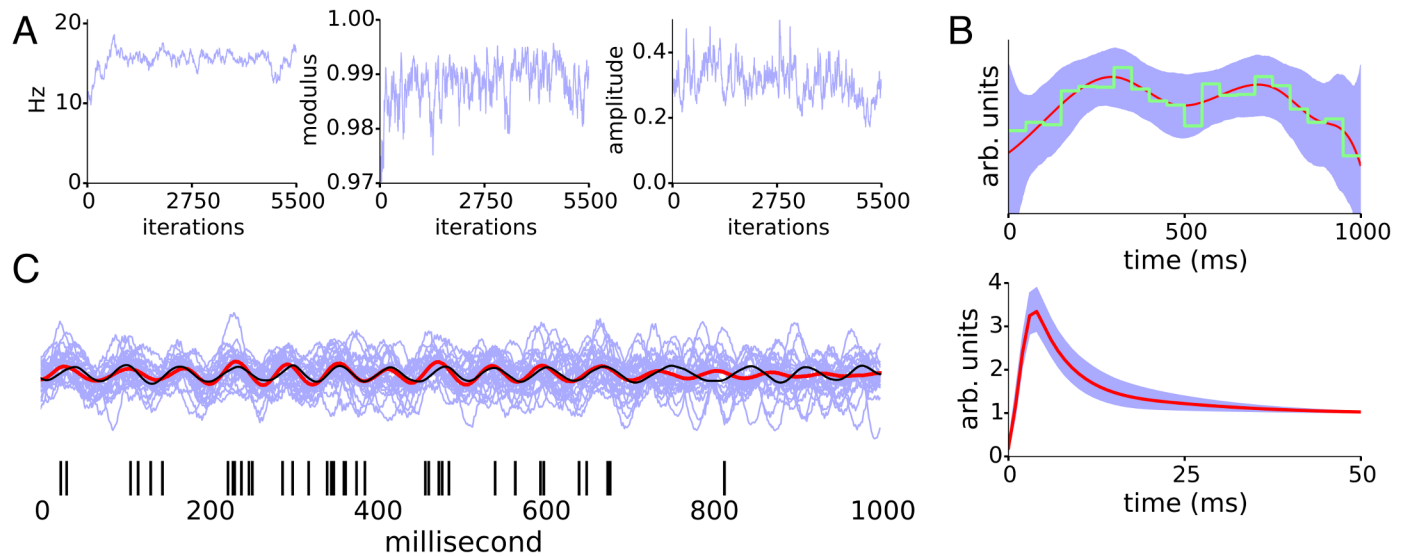


Fig 2. LIF neuron with stimulus-triggered change in firing rate. Phase inference for simulated LIF neuron driven by irregular oscillation with an average firing rate of 40Hz and a 15Hz modulating oscillation with OCV of 0.13. 60 trials of 1 second duration were used for LOST. A) Samples of the representative AR parameters. B) Top, samples of the TAE (cyan), the mean (red), and the empirical PSTH used to initialize TAE knot locations (green), and bottom, samples (cyan) of the spike history and mean (red). C) Example of inferred latent oscillation. The RL of the inferred oscillation is 0.65 ± 0.02 , while the spike train has an R of 0.22 ± 0.01 with the ground truth oscillation.

<https://doi.org/10.1371/journal.pcbi.1005596.g002>

these two terms, the latent state and spike history cannot, in principle be separated. LOST imposes constraints on the latent state and spike history according to our prior beliefs about their properties. The knot locations impose a strong prior on the shapes the spike history can take, so misplacing the knots may strongly affect the inferred latent state, presenting a significant difficulty in interpreting the results. We therefore investigate how mis-specifying the history affects LOST in Fig 3, which shows typical results from mis-specifying the spike history with two simulated datasets. The datasets are spike trains generated by a renewal process, Fig 3A with and Fig 3F without oscillatory modulation.

For the first spike train, the standard spline knot locations allows the successful inference of an oscillation and the history, Fig 3B, as the frequency, modulus and amplitude conditional posteriors are all very certain. The use of a fixed, flat history results in a flat latent state, Fig 3C, as the pronounced silences immediately following spikes are increasingly unlikely under spike train models with the latent state alone, which tends to increase the firing rate immediately following a spike. The latent states take large amplitudes near the spiking frequency of 40Hz when a temporally-compressed version of the ground truth spike history is used, Fig 3D. While the frequency has certainty, the modulus and amplitudes in contrast have large uncertainty, lowering our confidence that this latent state represents an oscillatory modulation. It is likely that a wide range of latent model parameters are nearly interchangeable, leading to a widening of the posterior in the direction of the AR parameters. Stretching the spike history, Fig 3E, results in a similar inferred latent state as Fig 3B, so a mis-specified history does not always lead to spurious results. We analyzed the second spike train 3 times with LOST, twice with an adjustable spike history term but with different history knot locations, and once with a fixed spike history term set to the ground truth, and Fig 3G–3I show samples from their respective posteriors. Fig 3G uses knots chosen by changing the default knot finder parameters, and the amplitudes do not go to 0 while the uncertainty in AR parameters remain high. Fig 3H uses the fixed ground truth history, and the amplitude approaches 0, as it does in Fig 3I

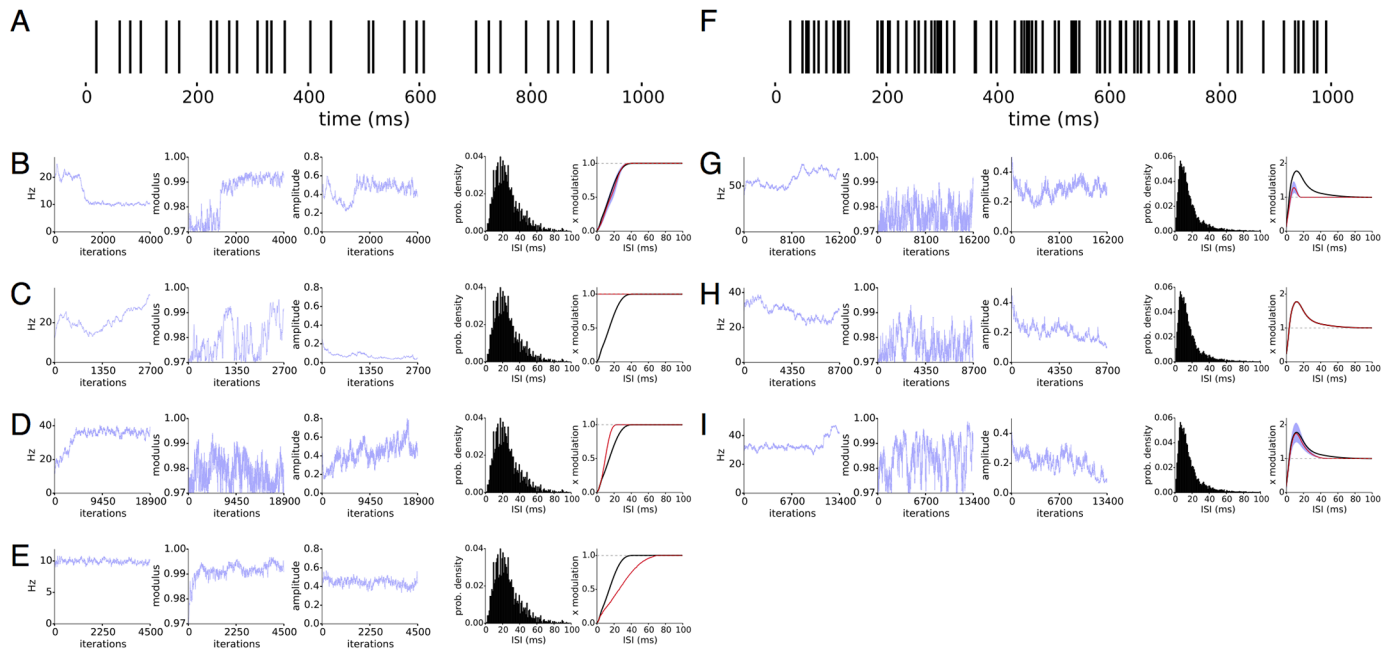


Fig 3. Interaction of spike history and latent state. Two inhomogeneous, history dependent spike trains, with first spike train A) generated with stochastic oscillational modulation at 10Hz and OCV at 0.18, and spiking at 40Hz, and the second spike train F) without modulation, spiking at 60Hz. Gibbs samples from 4 (3) separate LOST for first (second) spike train in B-E (G-I). In B-E and G-I, we show the 3 representative sampled quantities, frequency and modulus of the slowest root and the amplitude of the inferred latent state on the left, and the ISI distribution and spike history term on the right. The ground truth, sampled mean and 5-95% intervals of the sampled spike history term plotted in black, red and cyan, respectively. The differences between the different LOST runs is in how spike history knots were chosen, or if a fixed spike history function was used. For the first spike train A, B uses the standard knot choices, C uses a constant history set to 0 and D, a fixed spike history set to a time-compressed ground-truth history. For the second spike train, G uses the standard knot choice, H the ground truth and I the standard knots moved farther out in time. For the first spike train, misspecification of the history may lead to a false inference of a flat latent state, a latent state with high uncertainty in its parameters or the correct latent state. For the second spike train, a misspecified history can lead to a latent state with significant amplitude but with uncertain parameters, making the inference inconclusive.

<https://doi.org/10.1371/journal.pcbi.1005596.g003>

where knots were placed farther out than in Fig 3G, and the amplitude also approaches 0. We conclude that if the post-spike inhibition is not well characterized, the inference of an oscillation will most likely fail. If the inhibition is well characterized but the rest of the history is not, we find either that an oscillation will still be inferred successfully, or that the posterior will be broad in some of the directions of the AR parameters, making it difficult to conclusively determine that the latent state represents an oscillation.

We now look at spike trains without oscillational modulation, but generated with various post-spike refractory profiles to check that certain refractory profiles are not confounded as oscillations. Fig 4A–4D and 4E–4H are spike trains with average firing rate around 34Hz and 60Hz, respectively. In all cases except Fig 4C and 4F, the latent state amplitude approaches 0 as expected. For Fig 4C and 4F, the amplitude does not approach 0 but fluctuates widely along with the other parameters. In these cases, the large uncertainty in the AR parameters suggests the results of LOST are ambiguous, and we refrain from drawing a conclusion. We find that a reasonable rule of thumb in assessing how much uncertainty is “too much” is that frequency $2\pi/\lambda$ has a standard deviation of around 10% of its mean or less, and that modulus r have a standard deviation of under 0.005. Typically, we also find slow trends over a few thousand Gibbs iterations in the sampled values is commonly observed, consistent with the view that the parameters are nearly interchangeable over a fairly broad region of parameter space and the Gibbs sampler then performs a near random walk over the parameters. A latent state with a

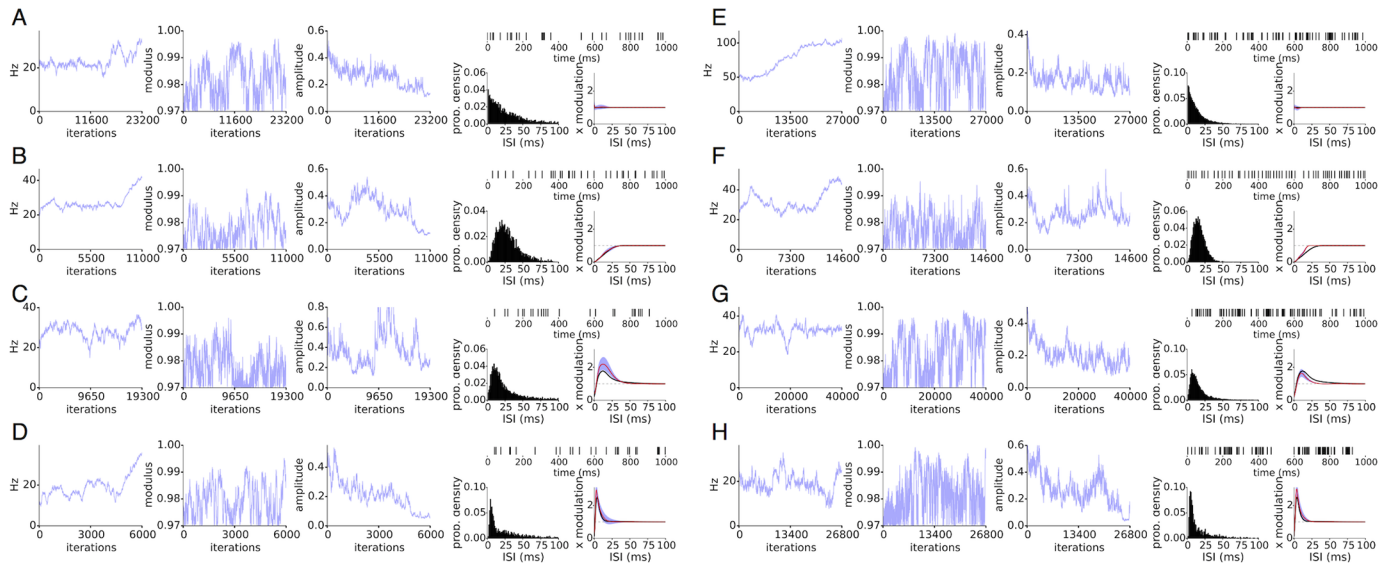


Fig 4. Spike trains without oscillatory modulation. Simulated spike trains at 34Hz (A-D) and 60Hz (E-H), with 4 types of post-spike refractory profiles. Each panel shows Gibbs samples on left and on the right, an example spike train at the top, and the ISI histogram and spike history function as in Fig 1. The expected output of LOST is a nearly flat latent state, which is the case in A, B, D, E, G, H, accompanied by large variability in the other sampled parameters. In C and F, the amplitude of the latent state does not approach 0, but rather continues to fluctuate while taking larger values.

<https://doi.org/10.1371/journal.pcbi.1005596.g004>

large amplitude does not necessarily signify the successful inference of an oscillation, and the uncertainty in the model parameters of the oscillation must also be taken in account in deciding whether to reject the hypothesis that there is an oscillation modulating the spiking. The amplitude of the latent state can inform us of how large the modulation is surrounding the mean firing rate. A latent state amplitude of 0.15 equates to a peak-to-peak fluctuation of about 30% of the average firing rate (for low firing rates < 100Hz), which is quite small a fluctuation considering only a few spikes may be observable for an oscillational cycle, so we use a latent state amplitude of 0.15 as a rule of thumb for when to consider the latent state as practically flat.

Through our examples, we found that the latent state can partially compensate an incorrectly specified spike history, as long as the short-term inhibition is reasonably well characterized by the spike history, or the post-spike inhibition is not too pronounced or long. In the absence of oscillational modulation in the data, this causes the latent state to take on significant amplitude, but the AR parameters themselves also have large uncertainty, making it difficult to conclude that the latent state represents an oscillation. However, in one limiting case, even with a mischaracterized spike history, LOST infers a clear oscillation. A nearly periodic spike train in Fig 5A is shown analyzed with LOST using the adjustable spike history term, Fig 5B and a fixed but modified ground truth spike history term, Fig 5C. When the misspecified history is used, the latent state takes large amplitudes with the posterior distribution of AR parameters having little uncertainty. The frequency of oscillation is nearly the same as the spiking frequency, 62Hz. In the limit of nearly periodic spikes, it may be difficult to conclude what particular combination of latent state and spike history underlies the highly regular spiking.

A flat latent state may signify a lack of oscillatory modulation or a spike history term that poorly captures the post-spike inhibition, or it may also signify that there is not enough data to infer an oscillation. Two spike trains with oscillatory modulation and different refractory profiles are shown in Fig 6A and 6E, with differing amount of trials used in LOST analysis, Fig 6B–6D and 6F–6H. When there is insufficient data to detect an oscillation, the latent state

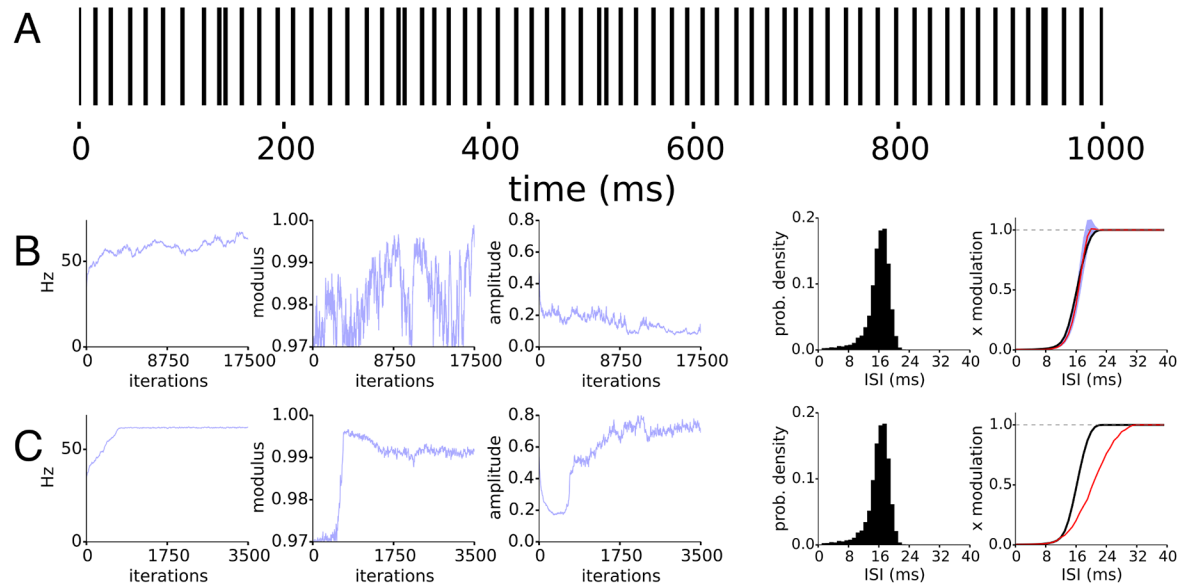


Fig 5. A) limiting case where models are not identifiable. The spike train in A is nearly periodic, but generated without an oscillatory modulation. When using the standard knots, the latent state amplitude approaches 0, B, but when we use a fixed history that is very different than the ground truth, C, the sampled parameters appear to be very certain, and an oscillation at 62Hz, nearly the spiking frequency, 64hz, is inferred. For cases where the spike train itself is nearly periodic, the spike train may be described equally well with various combinations of spike histories and an oscillation at the spiking frequency.

<https://doi.org/10.1371/journal.pcbi.1005596.g005>

amplitude approaches 0, Fig 6B and 6F. In general, the uncertainty in the parameters of the oscillation will decrease as more trials and data are used, Fig 6C, 6D, 6G and 6H, as expected.

Model comparison. We compare the LOST model to another point process model whose CIF is a general linear model (GLM) of past spiking history [12, 18]. Briefly, the CIF for trial m

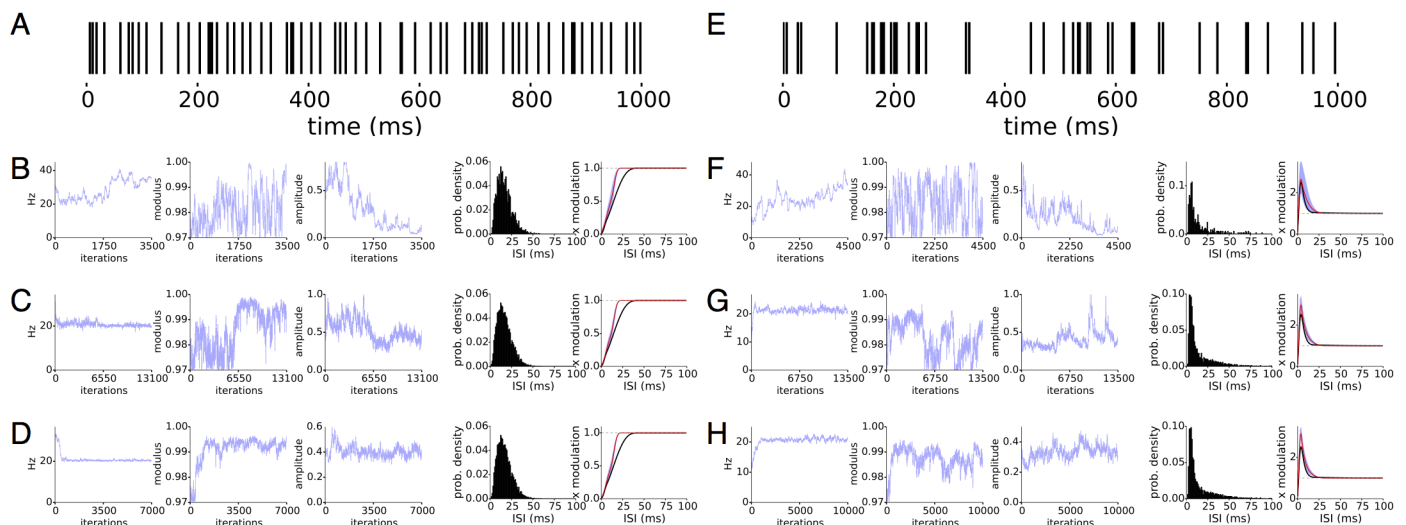


Fig 6. Effect of data size on posterior samples. 2 simulations of spike train with spiking around 60Hz, modulated with a 20Hz oscillation with OCV around 0.14 and 1 second long trials, but with different refractory period profiles, A and E. Layout of the figures is the same as in Fig 3, but with B-D and F-G showing different number of trials used, B) 8, C) 15 trials, and D) 40 for spike train in A and F) 5, G) 60 and H) 80 trials for spike train in E. For the smallest number of trials, the posterior is noticeably more variable, but the amplitude of the latent state approaches 0, meaning the latent state is almost decoupled from the spiking. We can see that the posterior narrows as more trials are used.

<https://doi.org/10.1371/journal.pcbi.1005596.g006>

time n is written

$$\lambda(m, n | H_{m,n}) = \exp \left(\sum_{i=1}^s \bar{a}_i y_{m,n-i} + \sum_{j=1}^o \bar{b}_j C_{mj} \right), \quad C_{mj} = \sum_{k=N_j}^{(S+1)j-1} y_{m,n-k}, \quad (46)$$

and this is easily fit to obtain the $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ coefficients. The $\bar{\mathbf{a}}$ models the short-term refractory self-history and $\bar{\mathbf{b}}$ encapsulates longer-term oscillatory self-history. This model does not take into account the dynamics of oscillation. Rather the spiking activity up to $S(O + 1)$ steps in the past uniquely determine the firing probability at the present, in contrast to our model. We fit the model several times with different initial conditions and values of S and L , using the python function `statsmodels.api.GLM`, and compare their goodness of fit via the time-rescaling theorem [37], and find the best combination of S and O . We then calculate the CIF from that model. Interpreting the fluctuating CIF as an oscillation, we extract the instantaneous phases, and compare to the phases inferred by the LOST model. The GLM model does not separate out the contribution of the short-term refractory self-history and the longer-term oscillatory self-history to the CIF, so we attempted to extract the oscillatory contribution by low-pass filtering the obtained CIF. We refer to this method of inference as GLM hereafter.

We compare the quality of inference of different models by comparing the resultant length of their inference with the ground truth. The resultant length (RL) L is calculated as $Le^{i\Phi_L} = \frac{1}{N} \sum_i e^{i(\phi_j - \psi_j)}$, where the ψ (ϕ) are the ground truth (inferred) phases, and is the circular equivalent to the mean squared error (MSE) of the ground truth and inferred phase. We also use another simple measure of characterizing the strength of oscillatory modulation, the spike phase (SP) histogram and its circular statistic R , calculated as $Re^{i\Phi} = \frac{1}{N} \sum_{n=1}^N e^{i2\pi\phi_{t_n}}$, where t_n is the time of the n th spike. Here R characterizes non-uniformity in the distribution of ϕ_{t_n} . While they are calculated in a similar fashion, the RL directly compares the inferred phase with the ground truth or LFP, while the SP R measures the uniformity of the probability of spiking with respect to ground truth or LFP phase. We caution that a result $L < R$ does not mean that the LOST model performed inference poorly. SP histogram detects modulation, but knowledge that spiking occurs at preferred phases does not also mean we know how to assign instantaneous phases given the spikes in a spike train. The error on R s and L s is calculated by bootstrap sampling with replacement of trials.

Comparison to other models of oscillation. Comparison of the LOST model with the GLM model illustrates how point process noise and oscillatory irregularity pose challenges to inferring an oscillation from a spike train, and how explicitly modeling the dynamics of the oscillation using a latent state with a point process observation model, improves inference when posed with these challenges. Fig 7A and 7B show the dependence of RL on how much spike history is used in the GLM (bottom right), and the history weights \mathbf{b} when 180, 480 or 720ms of history is used. The oscillation OCVs were 0.02 and 0.2 for Fig 7A and 7B, respectively. Temporal phase relationship is preserved for large time lags, as can be seen in the clear periodicity of the history weights \mathbf{b} in Fig 7A, and the GLM model performs as well as the LOST model when long history dependence is used. Shorter history degrades the performance of the GLM model, because we are not averaging out the spiking noise as much when shorter history is used. However, Fig 7B shows that when the oscillation is irregular, the phase information quickly dissipates, as seen by the lack of periodicity in \mathbf{b} at longer lags. Consequently, the spiking noise cannot be averaged out using long history, but rather spikes with large lag contribute only noise as the oscillational irregularity washes out the phase relationship for longer lags. The LOST model suffers less because it models the statistics of the irregularity in oscillation, the observable of the latent oscillation has a point-process observation model that takes

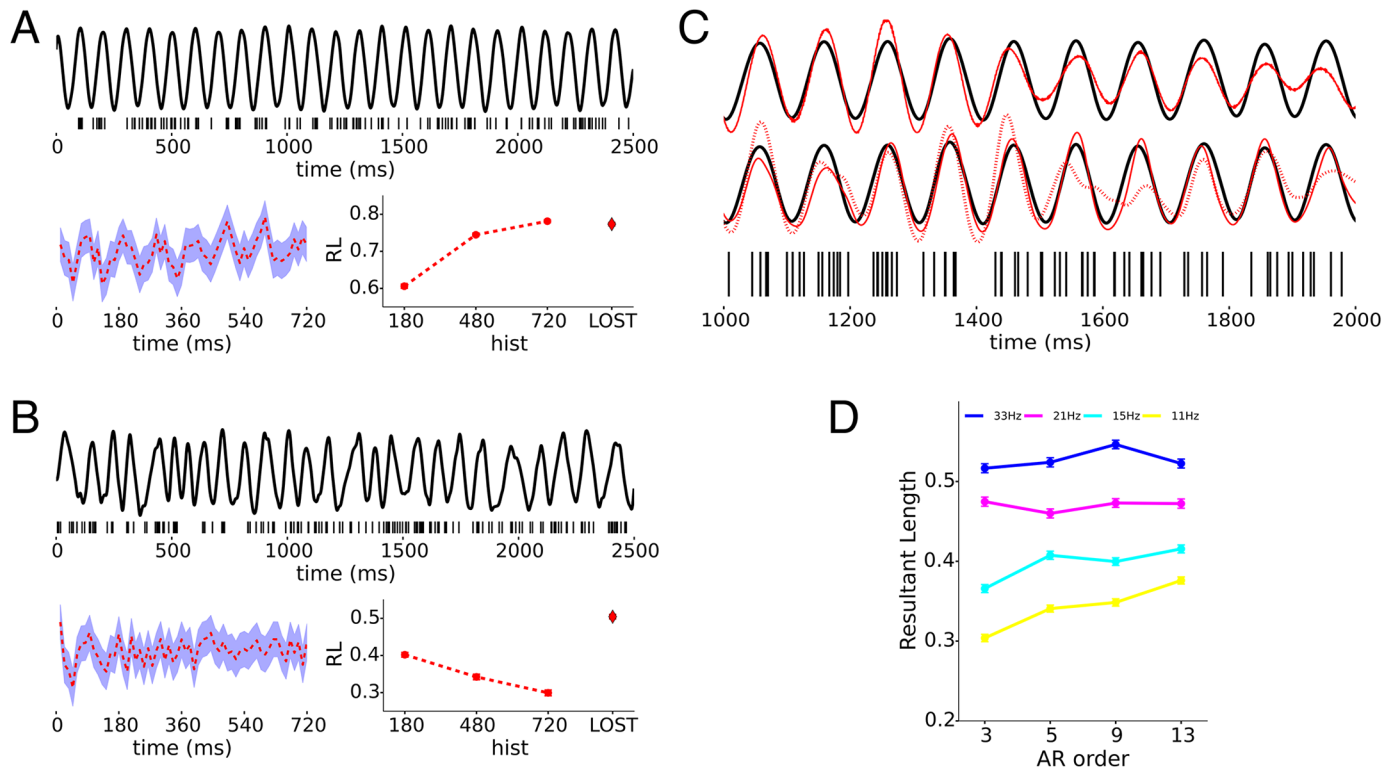


Fig 7. Comparison with other models of spiking and the AR order. Spiking noise and oscillation irregularity affect the quality of phase inference. A, B) Clockwise from top, spike train (ticks) produced by the regular oscillation plotted above, the resultant lengths when oscillation inferred using 3 different history lengths using GLM along with the RL from LOST inference of same data, and the history weights \mathbf{b} (dashed red) and their confidence intervals (purple) for GLM using history length 720ms. Modulating oscillations have OCV = 0.02 and 0.2 in A) and B), respectively. C) Sample CIF inference of data shown in A) using LOST (top trace), and GLM (bottom trace), with black line the duplicated ground truth CIFs. For GLM, the dashed (solid) line is the inference obtained from using 720ms (180ms) of spiking history. D) Comparison of how different AR orders for the LOST model affect inference performance for irregular oscillations.

<https://doi.org/10.1371/journal.pcbi.1005596.g007>

into account spiking noise, and also because the inferred oscillation is conditioned on all the spikes in the trial, Fig 7C.

AR order. Huerta [24] describes a reversible jump MCMC procedure to choose an appropriate AR model structure. We did not implement this due to the complexity of implementation and the relatively high computational cost of the current algorithm. Nonetheless, we investigate the effect of varying the AR order for phase inference with the LOST model. In Fig 7D, we used 4 sets of data with increasing firing rates, and compared the RLs of inference with different AR orders, and the accompanying RL from inference using GLM. The data sets themselves are generated with an oscillation of 10Hz using the LIF model, with OCVs of around 0.2, and have mean firing rates of 11Hz, 15Hz, 21Hz and 33Hz. 140, 130, 100, and 90 trials of 1.2 seconds duration were used in the fitting of the data. The model order comparison in Fig 7D shows that including more components usually resulted in a better inference of instantaneous phase, but qualitatively, even few components can pick out the oscillatory structure present in the spike train. This suggests that a model with the correct order, while important for optimal inference, can still infer valuable information if a sub-optimal order is chosen. For the rest of the paper, we keep the model structure to $4C + 1R$, unless otherwise stated, as it seems to provide a good compromise between good inference and reasonable computation times.

Presence of additional, slower fluctuation. Periods of prolonged silence, followed by prolonged periods of elevated firing rates are often visibly apparent in spike trains [38]. The timescales of these readily visible features are much slower than those characterizing faster neuronal oscillations thought to be relevant to cortical processing. In this example, we seek to demonstrate that the LOST model will give us useful inference of the faster oscillation in the presence of these slow modulations. We generated data with the offset, instead of being constant throughout a trial, to be a randomly generated square wave with a given duty cycle, and state duration generated from an exponential distribution with a minimum duration of 100 ms. Because the square wave is different in each trial, the overall PSTHs are flat. The spike train in Fig 8 have mean firing rates of 22Hz, a 20Hz oscillatory modulation with OCV = 0.27, with 200 trials used, and a duty cycle of 50% for the slow square wave. The resultant length was $L = 0.25 \pm 0.01$. To analyze the data, we begin by inferring the slow time modulation using an AR(1) model as our latent model. We choose an AR(1) model, as we don't expect very much structure in the slow latent state, and also because the time scale of the square waves relative to the duration of each trial ensures that any potential slow oscillatory structure would only be partially sampled, that considering additional structure in the AR is unnecessary. Although an AR(1) model is not the correct model describing these square waves, we are able to roughly infer the shape of the square waves, Fig 8 top.

We account for the slower fluctuation by modifying the CIF, Eq 3a to include the inferred slow fluctuation as a known signal I_{mn} ,

$$p(y_{mn} = 1 | x_{mn}, H_{n,l}, \Theta) = \frac{e^{x_{mn} + \mu_m + f_n + \lambda_{l(m,n)}^R + I_{mn}}}{1 + e^{x_{mn} + \mu_m + f_n + \lambda_{l(m,n)}^R + I_{mn}}} \tag{47}$$

After incorporating this known signal, we are able to infer the phases of the faster oscillation. Without including this step, the oscillatory AR(p) model would infer the slower fluctuation as in Fig 8 top, but inclusion of this known signal allows the faster oscillation to be inferred, Fig 8 bottom.

Neural data often contains more structure than the data analyst is aware of or suspects, and if on initial analysis with the standard LOST the analyst finds a much slower timescale latent state than expected, chances are the user has discovered a real but unknown feature of the data, and that the user may want to infer this slow latent state using an AR(1) as described in this section, and include this signal as a known signal in a second run of LOST to discover the higher-frequency modulation expected.

Mixture of modulated and non-modulated trials. Here we show an example where additional, discrete latent structures beside the oscillation, are present in the data. We motivate this extension to LOST as recent work has shown that brain states during sensory processing and behavior undergo discrete state transitions [39, 40], and because the authors have suspected (unpublished) that the motor cortical neurons analyzed in the Data Analysis section, also exhibit discrete transitions from trial to trial. Fig 9 shows results from 3 simulated spike trains, where 0% (Fig 9A), 25% (Fig 9B) or 40% (Fig 9C) of the of the trials are weakly modulated, and the rest strongly modulated. The mean firing rate for all 3 of the spiketrains are 36Hz, modulated with a 20Hz modulation with OCVs of 0.32. 80 trials were analyzed, each of duration 1.2 seconds.

The LOST model successfully infers the existence of 2 subsets of trials whose modulation strengths are very different in Fig 9B and 9C, and correctly infers that all trials are modulated by similar strengths in Fig 9A. To infer the modulation state, weak or strong, we compiled the posterior mean of the indicators $\langle Z_{m1} \rangle$ for all trials m . Trials that are strongly modulated tend to have small values of $\langle Z_{m1} \rangle$. The mixture coefficient π provides a straightforward threshold

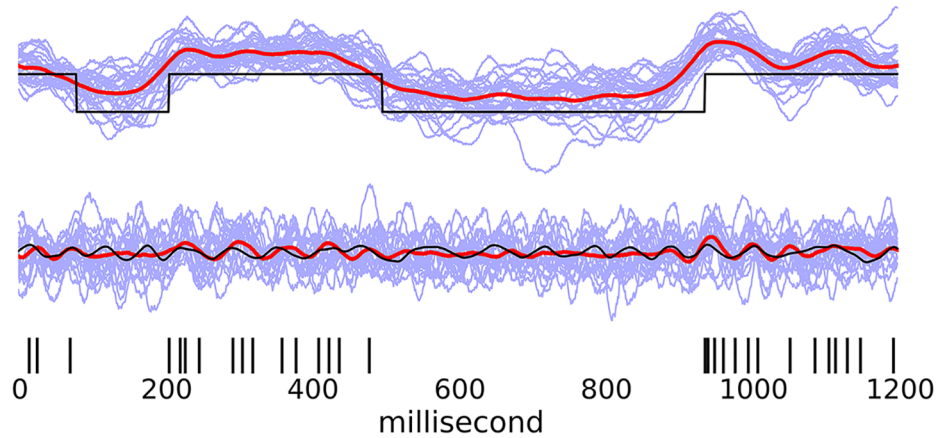


Fig 8. Spike train modulated with fast oscillation and a slower square-wave fluctuation. For a sample trial, top shows inferred, square-wave modulation of spiking (black) and the inferred latent state (red). Bottom shows inference of faster rhythm by taking the slow signal shown above to be a known signal, and re-running LOST.

<https://doi.org/10.1371/journal.pcbi.1005596.g008>

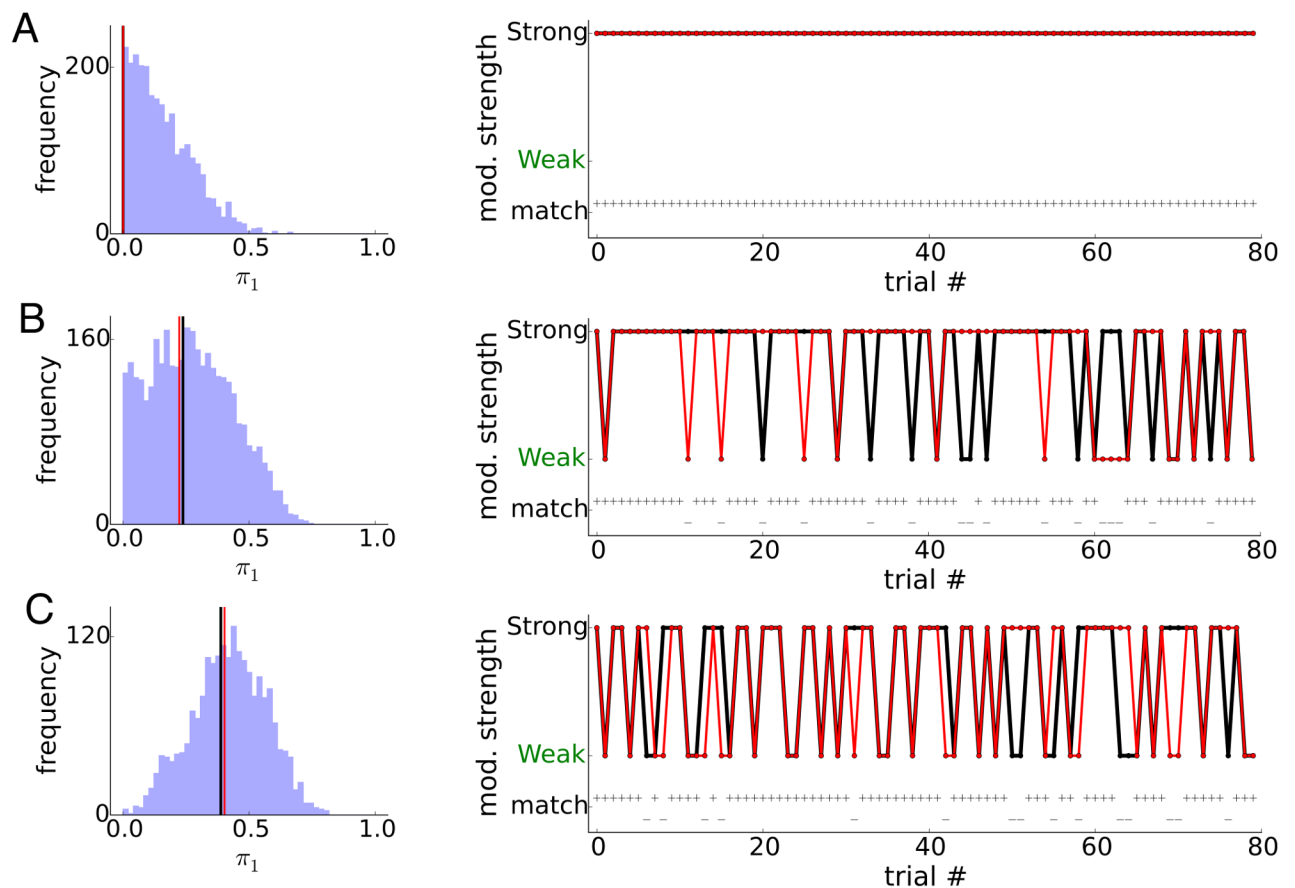


Fig 9. Mixture of modulated and non-modulated trials. A), B), C) 0%, 25% and 40% of the trials are weakly modulated. Left shows posterior distribution of π_1 and the black line is the ground truth value, red line is the mode of the marginal posterior. Right shows inferred modulation state, weak or strong, for each of the 80 trials. Black dots are ground truth, red dots are inferred state. Below them are indicators (+/-) of whether inferred matches ground truth.

<https://doi.org/10.1371/journal.pcbi.1005596.g009>

for $\langle Z_{m1} \rangle$ for categorization into weak or strong trials. We choose the threshold to be $M\pi_1^*$, where π_1^* is the posterior mode of π_1 . The trials whose $\langle Z_{m1} \rangle$ are larger than the threshold are deemed weakly modulated trials. As can be seen in Fig 9B and 9C, the LOST model categorizes a significant fraction, > 80%, of trials correctly, and also infers the mixture weights close to the true values. We note that in inferring the value of π_1 , we did not take the mean as we do for most other parameters due to its bounded nature and the bias this confers when π_1 is at or near 0, but rather take the mode of the posterior.

Data analysis

We analyze neurons from M1 of rat performing a self-paced lever push-hold-pull task [13, 41] that have been found to be significantly modulated to the theta rhythm in the LFP. These 2 neurons were recorded on separate electrodes, with “neuron 1” in Figs 10 and 11 recorded and spike sorted from a tetrode on a siliconprobe, and “neuron 2” recorded using juxtacellular (cell-attached) recording, where the spike trains did not need to be sorted. Both neurons were located in layer 1/2, and “neuron 2” identified morphologically as an interneuron, while “neuron 1” is likely to be an interneuron based on spike shape. We identified trials by selecting lever hold periods lasting more than 1 second followed by a large-amplitude pull, and analyzed the 1.2 second period encompassing the hold-pull period.

Slow fluctuations. Fig 10A shows that the spike trains appear to have long periods of low firing rate, interspersed with periods of higher firing within a trial, creating noticeable holes in the spike train. Applying the LOST model using an AR(1) latent state first picks up a latent state that follows these slow fluctuations to the firing rate, but with the faster oscillatory modulation absent. Fig 10B and 10C bottom show the power spectrum for the fluctuation I we infer for each neuron, and find that both are slow, in the 1-2Hz range. Interestingly, this slow fluctuation is weakly correlated between the neurons, $p < 6 \times 10^{-3}$, bootstrap test.

Shared theta oscillations. Next, we account for the faster fluctuation by including the slow fluctuation I into the CIF as described in Eq 47. After incorporating this known signal, we are able to infer the phases, Fig 11 for both neurons. The power spectral density of the inferred latent oscillation are similar, with a peak around 7-8Hz, Fig 11A and 11B bottom. Igarashi *et al* [13] found significant modulation to the LFP theta of these neurons using SP histograms, but did not find direct evidence of simultaneous modulation of the neurons from the spike-spike cross-correlation function, Fig 11D. In contrast, the oscillatory cross-correlation function of the independently inferred oscillations shown in Fig 11C, clearly shows evidence of simultaneous modulation of the neurons to a shared oscillation.

Weakly and strongly modulated trials to LFP theta in a single neuron. In the previous section, we treated all trials of neuron 1 as equally modulated and inferred a theta modulation in its spike train. Using a simple model to fit often results in finding an oscillation, as we did for neuron 1, but it may be that there is further structure that can be discovered by using a more complicated parametric model. Here, we analyze neuron 1 using the mixture model, and find a significant difference in the strength of oscillatory modulation between trials. Following the procedure described for the mixture model, we find that about 85% of the trials (H, pink) have a much stronger modulation, Fig 12A right, than the remaining trials (L, blue), Fig 12C top.

We emphasize this classification was made using only the spike trains. The PSTHs constructed separately from L trials and H trials show no significant difference Fig 12B left, so the L and H trials being detected do not reflect a difference in temporal trial-locked firing rate, or in the baseline firing rate. Since the classification into L and H reflects a dichotomous change in the structure of oscillatory modulation, we expect to find a similar change in the relationship

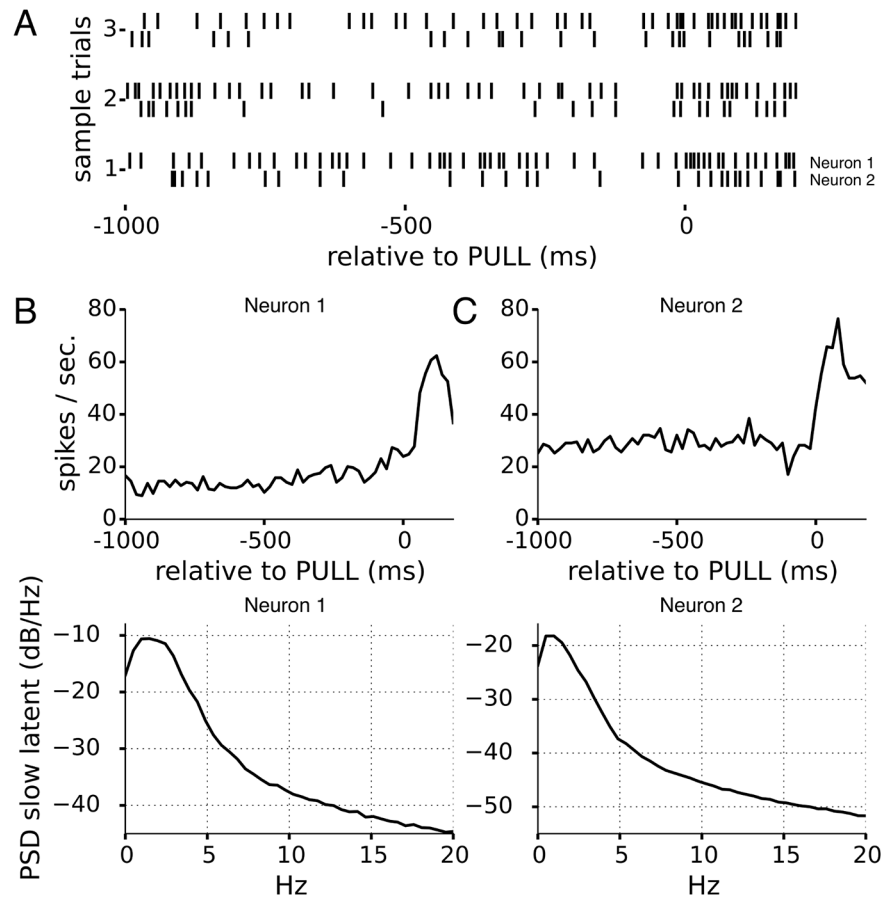


Fig 10. Two simultaneously-recorded motor cortical neurons: Slow fluctuations in firing rate for neuron 1 and 2. A) rasters from sample trials for the two neurons. There are noticeable holes in the rasters of the neuron 1, and neuron 2 also appears to have a similar slower fluctuation that is less noticeable. B, C) the PSTHs for the two neurons (top) show they both increase their activity around the time of PULL onset (0ms), and power spectral density of latent state show slow fluctuations on the order of 1-2 Hz. Correlation coefficient at 0 lag for the slow fluctuations inferred from neuron 1 and 2 are weakly correlated, with a correlation coefficient 0.15, $p < 1 \times 10^{-3}$, boot strap test.

<https://doi.org/10.1371/journal.pcbi.1005596.g010>

between the spike train and the LFP theta. It is possible that the L and H trials are indicative of absence or presence of the global LFP theta, that L trials show weak modulation because there is no modulating drive in those trials. The autocorrelation function of the high-pass filtered (above 1Hz) LFP constructed separately for the L trials and H trials, showed similar oscillatory structure, Fig 12B center. Further, the amplitude of the theta-band LFP in L trials and H trials also show no significant difference, Fig 12C, so global theta seems to be a reliable and perpetual presence. However, in Fig 12D, when we construct 2 separate LFP theta SP histograms, SP_L and SP_H , we find that SP_L is significantly more uniform than SP_H , with $R_L = 0.06 \pm 0.01$ and $R_H = 0.16 \pm 0.01$, boot strap test $p < 7 \times 10^{-4}$. Thus it appears that this neuron intermittently participates in the perpetual theta rhythmicity, resulting in distinct subsets of trials, some being weakly modulated and others being strongly modulated.

Discussion

We have defined the LOST model, together with accompanying posterior simulation technology, in order to detect the presence of oscillatory firing-rate modulation in a spike train, and

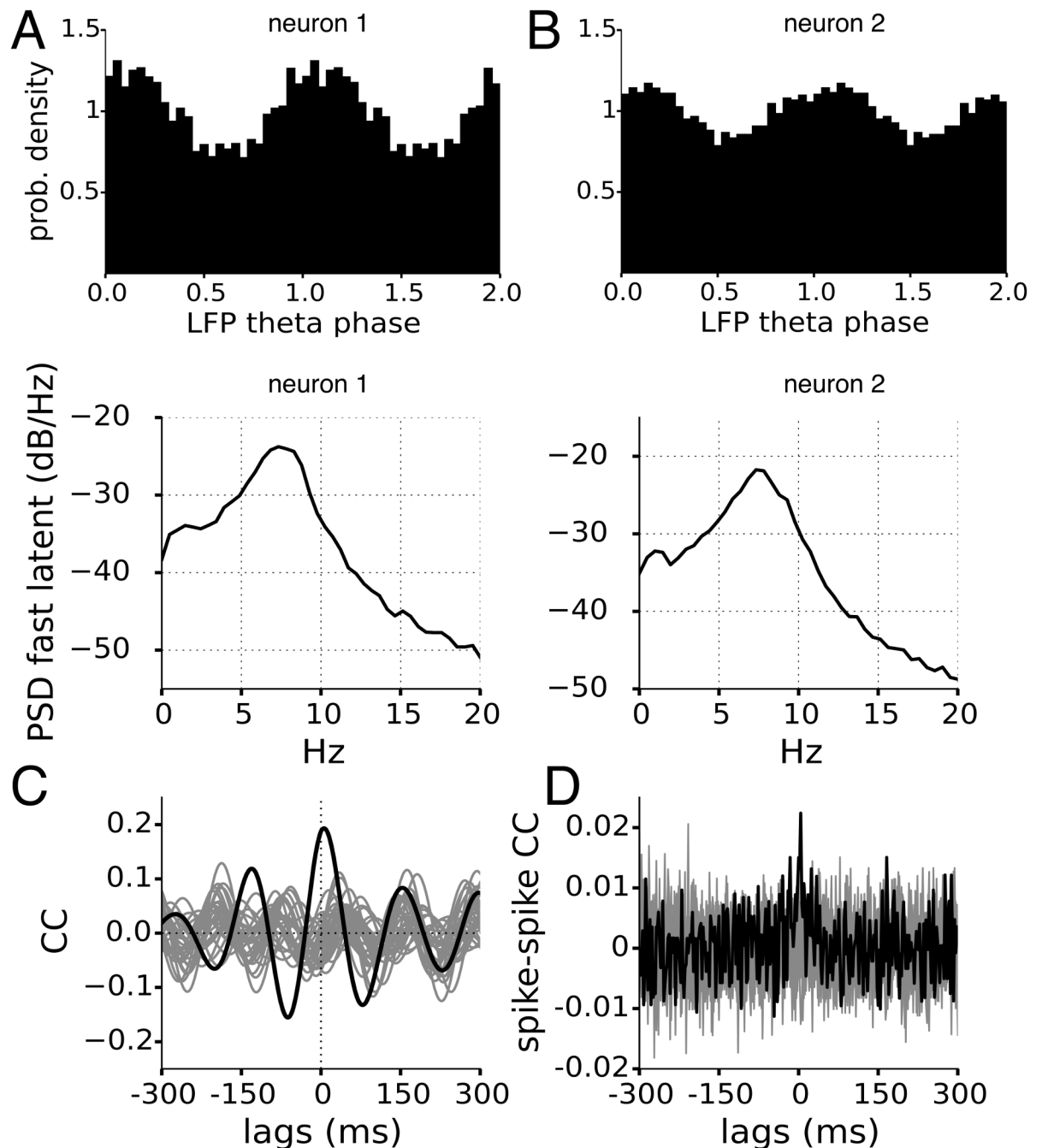


Fig 11. Two simultaneously-recorded motor cortical neurons: Simultaneous inference of theta oscillations using inferred slow fluctuation from Fig 10 as a known signal. A) SP histograms for neurons 1 and 2 with circular statistics $R = 0.13 \pm 0.01$ and 0.08 ± 0.01 . B) power spectral densities of inferred oscillation for the 2 neurons showing power in the theta range. C) The cross-correlation function of the inferred latent oscillations of the two neurons, grey traces from when trials shuffled. D) The spike-spike cross correlation function shows very little oscillatory structure.

<https://doi.org/10.1371/journal.pcbi.1005596.g011>

infer its phase of oscillation in the presence of a variety of non-stationarities in firing rate that are present in experimental data. Previous methods have assessed the oscillatory content in spike trains by comparing the spiking to a known oscillatory signal like a band-passed LFP [42, 43], by detecting oscillation directly from the spike train [12, 18, 44–46], or by point

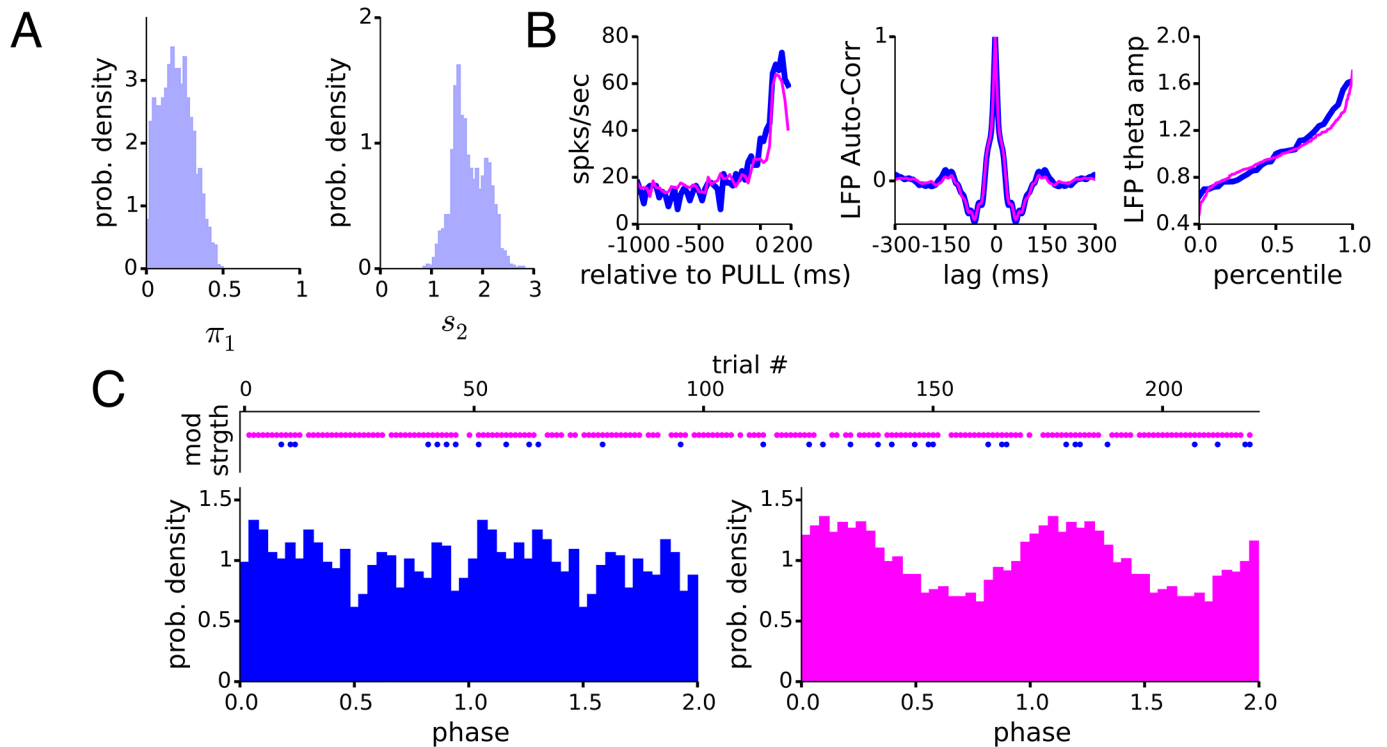


Fig 12. Motor cortical neuron (neuron 1, Figs 10 and 11) shows changing relationship to LFP theta oscillation in different trials. A) The posterior distributions of the mixture weight of the weakly modulated L trials, and relative modulation strength of the strongly modulated H trials. B) $PSTH_L$ (blue) and $PSTH_H$ (pink) overlap one another, showing no difference in average firing rate. C) LFP theta amplitudes calculated for all L and H trials on a per trial basis. Difference between cumulative distributions of per trial LFP theta amplitude not significant, $p < 0.78$, 2 sample KS-test. D) Top, the 33 (187) trials classified as weakly (strongly) modulated trials, blue (pink) dots, top. Bottom, the SP histogram calculated from L (H) trials, left (right). Inference of modulation strength using the spike train alone corresponds to the modulation strength to LFP theta. Circular statistics $R_L = 0.06 \pm 0.01$ and $R_H = 0.16 \pm 0.01$ are significantly different, bootstrap test $p < 7 \times 10^{-4}$.

<https://doi.org/10.1371/journal.pcbi.1005596.g012>

process regression using the oscillatory signal as a covariate [47]. The LOST model not only detects oscillatory modulation from the spike train itself, but does so in the presence of both spiking noise and oscillatory irregularity, and also allows extensions to inferring additional latent structure, such as non-stationarities in the modulational strength across trials. In addition, LOST is able to separately account for modulational signals of different frequency band, which has proven to be vital in the analysis of real spike trains. Structural priors on the latent state dynamics together with explicit consideration of spiking noise and oscillatory irregularity allows LOST to uncover oscillatory structure even when the firing rate is low, the modulation weak or when the oscillation is irregular, compared to methods that directly regressed the spiking probability on the raw spiking history itself without an intermediary latent state [12, 18, 19]. LOST uses the intermediary latent state to address the irregularity characteristic of neural oscillations. Another approach to modeling the spectral features of time series using Gaussian processes has recently been developed by Wilson *et al* [48], where different classes of covariance kernels, the SE and SM, respectively, model non-oscillatory and oscillatory structures, analogous to the real and imaginary roots of the characteristic polynomials. It would be interesting to compare the two approaches in future work.

The increased sensitivity and flexibility in specifying latent structures, should allow LOST to be used in investigating the role of oscillations in cognitive functions in the cortex. Investigators are increasingly interested in characterizing the change in neural responses to time-

varying stimulus or behavior [49, 50]. Theta and gamma oscillations in hippocampus change their coupling structure and prevalence during learning and memory acquisition [51, 52]. The inclusion of trial-specific structure independently of the LFP in the LOST model may also allow detection of increased recruitment of a given neuron into cell assemblies during periods when oscillations in the LFP are changing. Further, the variability of the timing and presence of oscillations in the LFP seen in many areas of the cortex and hippocampus [2, 13, 53], suggests oscillatory modulation in single neurons may likewise exhibit finer structure on a per-trial basis. Use of the LOST model in the analysis of such systems may reveal richer dynamics of recruitment into cell assemblies, and a better understanding of the role of single neurons in cognition.

Acknowledgments

We thank Yoshikazu Isomura at Tamagawa University Brain Science Institute for the rat data collected at Riken Brain Science Institute. We also thank James Scott at the University of Texas, Austin and Jesse Windle at Duke University, for helpful advice concerning the implementation of the Pólya-Gamma data augmentation method and programming assistance. And finally, we thank Uri Eden at Boston University, Pengcheng Zhou and Spencer Koerner at the Carnegie Mellon University for constructive discussion.

Author Contributions

Conceived and designed the experiments: REK KA.

Analyzed the data: KA.

Wrote the paper: KA REK.

References

1. Murthy VN, Fetz EE. Synchronization of neurons during local field potential oscillations in sensorimotor cortex of awake monkeys. *Journal of neurophysiology*. 1996 dec; 76(6):3968–82. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8985893>. PMID: 8985893
2. Donoghue JP, Sanes JN, Hatsopoulos NG, Gaál G, Omlor W, Patino L, et al. Neural Discharge and Local Field Potential Oscillations in Primate Motor Cortex During Voluntary Movements Neural Discharge and Local Field Potential Oscillations in Primate Motor Cortex During Voluntary Movements. *Journal of neurophysiology*. 1998; 79:159–173. PMID: 9425187
3. Engel AK, Fries P, Singer W. Dynamic predictions: oscillations and synchrony in top-down processing. *Nature reviews Neuroscience*. 2001; 2(10):704–716. <https://doi.org/10.1038/35094565> PMID: 11584308
4. Siapas AG, Lubenov EV, Wilson Ma. Prefrontal phase locking to hippocampal theta oscillations. *Neuron*. 2005 apr; 46(1):141–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15820700>. <https://doi.org/10.1016/j.neuron.2005.02.028> PMID: 15820700
5. Denker M, Roux S, Timme M, Riehle A, Grün S. Phase synchronization between LFP and spiking activity in motor cortex during movement preparation. *Neurocomputing*. 2007; 70(10-12):2096–2101. <https://doi.org/10.1016/j.neucom.2006.10.088>
6. Gregoriou GG, Gotts SJ, Zhou H, Desimone R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*. 2009; 324(5931):1207–1210. <https://doi.org/10.1126/science.1171402> PMID: 19478185
7. Fries P, Womelsdorf T, Oostenveld R, Desimone R. The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2008; 28(18):4823–4835. <https://doi.org/10.1523/JNEUROSCI.4499-07.2008>
8. Siegel M, Warden MR, Miller EK. Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*. 2009 dec; 106(50):21341–6. Available from: <http://>

www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2779828&tool=pmcentrez&rendertype=abstract. <https://doi.org/10.1073/pnas.0908193106>

9. Benchenane K, Peyrache A, Khamassi M, Tierney PL, Gioanni Y, Battaglia FP, et al. Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron*. 2010 jun; 66(6):921–36. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20620877>. <https://doi.org/10.1016/j.neuron.2010.05.013> PMID: 20620877
10. Ching S, Cimenser A, Purdon PL, Brown EN, Kopell NJ. Thalamocortical model for a propofol-induced α -rhythm associated with loss of consciousness. *Proceedings of the National Academy of Sciences*. 2010; 107(52):22665–22670. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1017069108>.
11. Chalk M, Herrero JL, Gieselmann Ma, Delicato LS, Gotthardt S, Thiele A. Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1. *Neuron*. 2010 apr; 66(1):114–25. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2923752&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1016/j.neuron.2010.03.013> PMID: 20399733
12. Sarma SV, Eden UT, Cheng ML, Williams ZM, Hu R, Eskandar E, et al. Using point process models to compare neural spiking activity in the subthalamic nucleus of Parkinson's patients and a healthy primate. *IEEE transactions on bio-medical engineering*. 2010 jun; 57(6):1297–305. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20172804>. <https://doi.org/10.1109/TBME.2009.2039213> PMID: 20172804
13. Igarashi J, Isomura Y, Arai K, Harukuni R, Fukai T. Oscillation Code for Neuronal Coordination during Motor Behavior. *Journal of Neuroscience*. 2013 nov; 33(47):18515–18530. Available from: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2126-13.2013>. PMID: 24259574
14. Mukamel EA, Pirondini E, Babadi B, Wong KFK, Pierce ET, Harrell PG, et al. A transition in brain state during propofol-induced unconsciousness. *Journal of Neuroscience*. 2014; 34(3):839–845. Available from: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5813-12.2014>. PMID: 24431442
15. Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*. 2005 oct; 9(10):474–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16150631>. <https://doi.org/10.1016/j.tics.2005.08.011> PMID: 16150631
16. Fell J, Axmacher N. The role of phase synchronization in memory processes. *Nature reviews Neuroscience*. 2011 feb; 12(2):105–18. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21248789>. <https://doi.org/10.1038/nrn2979> PMID: 21248789
17. Denker M, Roux S, Lindén H, Diesmann M, Riehle A, Grün S. The local field potential reflects surplus spike synchrony. *Cerebral Cortex*. 2011; 21(12):2681–2695. <https://doi.org/10.1093/cercor/bhr040> PMID: 21508303
18. Eden UT, Gale JT, Amirnovin R, Eskandar EN. Characterizing the spiking dynamics of subthalamic nucleus neurons in Parkinson's disease using generalized linear models. *Frontiers in integrative neuroscience*. 2012 jan; 6(June):28. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3379030&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.3389/fnint.2012.00028> PMID: 22723771
19. Deng X, Eskandar EN, Eden UT. A point process approach to identifying and tracking transitions in neural spiking dynamics in the subthalamic nucleus of Parkinson's patients. *Chaos*. 2013 dec; 23(4):046102. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3808419&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1063/1.4818546> PMID: 24387581
20. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of neurophysiology*. 2005; 93(September 2004):1074–1089. PMID: 15356183
21. Kass RE, Eden UT, Brown EN. *Analysis of Neural Data*. Springer; 2014.
22. Smith AC, Brown EN. Estimating a state-space model from point process observations. *Neural computation*. 2003 may; 15(5):965–91. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12803953>. <https://doi.org/10.1162/089976603765202622> PMID: 12803953
23. Kitagawa G. *Introduction to Time Series Modeling*. Chapman & Hall / CRC Monographs on Statistics & Applied Probability; 2010.
24. Huerta G, West M. Priors and component structures in autoregressive time series models. *Journal of the Royal Statistics Society Series B*. 1999; 61(4):881–899. <https://doi.org/10.1111/1467-9868.00208>
25. Polson NG, Scott JG, Windle J. Bayesian Inference for Logistic Models Using Pólya-Gamma Latent Variables. *Journal of the American Statistical Association*. 2013; 108(504):1339–1349. Available from: <http://www.tandfonline.com/doi/abs/10.1080/01621459.2013.829001>.
26. Allcroft DJ, Glasbey CA. A spectral estimator of ARMA parameters from thresholded data. *Statistics and Computing*. 2002; 12:369–376. <https://doi.org/10.1023/A:1020796314300>

27. Kass RE, Ventura V. A Spike-Train Probability Model. *Neural Computation*. 2001; 1720:1713–1720. <https://doi.org/10.1162/08997660152469314>
28. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. Chapman & Hall; 1995.
29. Tanner MA, Wong WH. The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*. 1987 jun; 82(398):528–540. Available from: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1987.10478458>. <https://doi.org/10.2307/2289463>
30. Frühwirth-Schnatter S. Data Augmentation and Dynamic Linear Models. *Journal of Time Series Analysis*. 1994; 15(2):183–202. <https://doi.org/10.1111/j.1467-9892.1994.tb00184.x>
31. Penttonen M, Kamondi A, Acsady L, Buzsaki G. Gamma frequency oscillation in the hippocampus of the rat: intracellular analysis in vivo. *European Journal of Neuroscience*. 1998 feb; 10(2):718–728. Available from: <http://doi.wiley.com/10.1046/j.1460-9568.1998.00096.x>. PMID: 9749733
32. Burgess N, O'Keefe J. Models of place and grid cell firing and theta rhythmicity. *Current opinion in neurobiology*. 2011 oct; 21(5):734–44. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3223517&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1016/j.conb.2011.07.002> PMID: 21820895
33. Polasek W, Jin S. Gibbs Sampling in AR Models with Random Walk Priors. In: *From Data to Knowledge: Theoretical and practical aspects of classification, data analysis, and knowledge organization*. Springer; 1996. p. 86–93.
34. Funkhouser HG. A Short Account of the History of Symmetric Functions of Roots of Equations. *The American Mathematical Monthly*. 1930; 37(7):357–365. <https://doi.org/10.2307/2299273>
35. Buesing L, Macke JH, Sahani M. Learning stable, regularised latent models of neural population dynamics. *Network: Computation in Neural Systems*. 2012; 23(1-2):24–47.
36. Sherman BYJ, Morrison WJ. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *The Annals of Mathematical Statistics*. 1950; 21(1):124–127. <https://doi.org/10.1214/aoms/1177729893>
37. Brown EN, Nguyen DP, Frank LM, Wilson Ma, Solo V. An analysis of neural receptive field plasticity by point process adaptive filtering. *Proceedings of the National Academy of Sciences*. 2001; 98(21):12261–12266. <https://doi.org/10.1073/pnas.201409398>
38. Okamoto H, Isomura Y, Takada M, Fukai T. Temporal integration by stochastic recurrent network dynamics with bimodal neurons. *Journal of neurophysiology*. 2007; 97(6):3859–3867. <https://doi.org/10.1152/jn.01100.2006> PMID: 17392417
39. Jones LM, Fontanini A, Sadacca BF, Miller P, Katz DB. Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proceedings of the National Academy of Sciences*. 2007; 104(47):18772–18777. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2141852&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1073/pnas.0705546104>
40. Ulrich KR, Carlson DE, Lian W, Borg JS, Dziras K, Carin L. Analysis of Brain States from Multi-Region LFP Time-Series. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, editors. *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc.; 2014. p. 2483–2491. Available from: <http://papers.nips.cc/paper/5624-analysis-of-brain-states-from-multi-region-lfp-time-series.pdf>.
41. Isomura Y, Harukuni R, Takekawa T, Aizawa H, Fukai T. Microcircuitry coordination of cortical motor information in self-initiation of voluntary movements. *Nature neuroscience*. 2009 dec; 12(12):1586–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19898469>. <https://doi.org/10.1038/nn.2431> PMID: 19898469
42. Fries P, Roelfsema PR, Engel AK, König P, Singer W. Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences*. 1997; 94(November):12699–12704. <https://doi.org/10.1073/pnas.94.23.12699>
43. Lepage KQ, Gregoriou GG, Kramer MA, Aoi M, Gotts SJ, Eden UT, et al. A procedure for testing across-condition rhythmic spike-field association change. *Journal of neuroscience methods*. 2013 feb; 213(1):43–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23164959>. <https://doi.org/10.1016/j.jneumeth.2012.10.010> PMID: 23164959
44. Mureşan RC, Jurjuţ OF, Moca VV, Singer W, Nikolić D. The Oscillation Score: An Efficient Method for Estimating Oscillation Strength in Neuronal Activity. *Journal of neurophysiology*. 2008;p. 1333–1353. PMID: 18160427
45. Climer JR, Newman EL, Hasselmo ME. Phase coding by grid cells in unconstrained environments: two-dimensional phase precession. *The European journal of neuroscience*. 2013 aug; 38(4):2526–41. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3912569&tool=pmcentrez&rendertype=abstract>. <https://doi.org/10.1111/ejn.12256> PMID: 23718553

46. Matzner A, Bar-Gad I. Quantifying Spike Train Oscillations: Biases, Distortions and Solutions. *PLOS Computational Biology*. 2015; 11(4):e1004252. Available from: <http://dx.plos.org/10.1371/journal.pcbi.1004252>. PMID: 25909328
47. Zhou P, Burton SD, Snyder AC, Smith MA, Urban NN, Kass RE. Establishing a Statistical Link between Network Oscillations and Neural Synchrony. *PLoS Computational Biology*. 2015; 11(10):1–25. <https://doi.org/10.1371/journal.pcbi.1004549>
48. Wilson AG, Adams RP. Gaussian Process Covariance Kernels for Pattern Discovery and Extrapolation. In: *Proceedings of the 30th International Conference on Machine Learning*. vol. 28; 2013. p. 15. Available from: <http://arxiv.org/abs/1302.4245>.
49. Czanner G, Eden UT, Wirth S, Yanike M, Suzuki Wa, Brown EN. Analysis of between-trial and within-trial neural spiking dynamics. *Journal of neurophysiology*. 2008; 99(5):2672–2693. <https://doi.org/10.1152/jn.00343.2007> PMID: 18216233
50. Mandelblat-Cerf Y, Paz R, Vaadia E. Trial-to-trial variability of single cells in motor cortices is dynamically modified during visuomotor adaptation. *The Journal of neuroscience*. 2009 dec; 29(48):15053–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19955356>. <https://doi.org/10.1523/JNEUROSCI.3011-09.2009> PMID: 19955356
51. Tort ABL, Komorowski RW, Manns JR, Kopell NJ, Eichenbaum H. Theta-gamma coupling increases during the learning of item-context associations. *Proceedings of the National Academy of Sciences*. 2009; 106(49):20942–20947. <https://doi.org/10.1073/pnas.0911331106>
52. van Vugt MK, Schulze-Bonhage A, Litt B, Brandt A, Kahana MJ. Hippocampal gamma oscillations increase with memory load. *The Journal of Neuroscience*. 2010; 30(7):2694–2699. <https://doi.org/10.1523/JNEUROSCI.0567-09.2010> PMID: 20164353
53. Yartsev MM, Witter MP, Ulanovsky N. Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature*. 2011 nov; 479(7371):103–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22051680>. <https://doi.org/10.1038/nature10583> PMID: 22051680