J Neurophysiol 105: 474–486, 2018.
First published November 10, 2017; doi:10.1152/jn.00642.2017.

Removal of Spurious Correlations Between Spikes and Local Field Potentials

Theodoros P. Zanos, Patrick J. Mineault, and Christopher C. Pack
Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

Submitted 21 July 2017; accepted in final form 3 November 2017

Zanos TP, Mineault PJ, Pack CC. Removal of spurious correlations between spikes and local field potentials. *J Neurophysiol* 105: 474–486, 2018. First published November 10, 2017; doi:10.1152/jn.00642.2017. Single neurons carry out important sensory and motor functions related to the larger networks in which they are embedded. Understanding the relationships between single-neuron spiking and network activity is therefore of great importance and the latter can be readily estimated from low-frequency brain signals known as local field potentials (LFPs). In this work we examine a number of issues related to the estimation of spike and LFP signals. We show that spike trains and individual spikes contain power at the frequencies that are typically thought to be exclusively related to LFPs, such that simple frequency-domain filtering cannot be effectively used to separate the two signals. Ground-truth simulations indicate that the commonly used method of estimating the LFP signal by low-pass filtering the raw voltage signal leads to artifactual correlations between spikes and LFPs and that these correlations exert a powerful influence on popular metrics of spike–LFP synchronization. Similar artifactual results were seen in data obtained from electrophysiological recordings in macaque visual cortex, when low-pass filtering was used to estimate LFP signals. In contrast LFP tuning curves in response to sensory stimuli do not appear to be affected by spike contamination, either in simulations or in real data. To address the issue of spike contamination, we devised a novel Bayesian spike removal algorithm and confirmed its effectiveness in simulations and by applying it to the electrophysiological data. The algorithm, based on a rigorous mathematical framework, outperforms other methods of spike removal on most metrics of spike–LFP correlations. Following application of this spike removal algorithm, many of our electrophysiological recordings continued to exhibit spike–LFP correlations, confirming previous reports that such relationships are a genuine aspect of neuronal activity. Overall, these results show that careful preprocessing is necessary to remove spikes from LFP signals, but that when effective spike removal is used, spike–LFP correlations can potentially yield novel insights about brain function.

INTRODUCTION

Much of our mechanistic understanding of brain function comes from extracellular recordings, which provide a direct measure of electrical activity near the tip of the recording electrode. The resulting voltage signal is composed of the spikes emitted by one or more neurons and the local field potentials (LFPs), which represent the total synaptic current in the neuronal circuit from which the recordings are obtained. Typically, the LFP is extracted by low-pass filtering the wideband voltage signal, whereas spikes are identified by high-pass filtering, thresholding, and subsequent sorting. The possibility of isolating both signals from the same electrode is of great interest because it allows the experimenter to relate the responses of individual neurons to those of the larger-scale circuits in which they are embedded.

Popular methods for studying spike–LFP relationships in the time domain include the spike-triggered average (STA) of the LFP, the spike-field coherence (SFC), spike–LFP phase-locking histograms, and the prediction of spike timing from LFP features. These types of analysis have proven useful for inferring the role of feedforward and feedback circuitry in functions as diverse as perception, attention, memory, and motor control (Andersen et al. 2014; Chalk et al. 2017; Destexhe et al. 1999; Fries et al. 2018; Jacobs et al. 2017; Pesaran et al. 2012; Saleh et al. 2017). Similar methods applied to the space domain permit comparisons of single-neuron activity to columnar systems in the cortex (Nauhaus et al. 2018; Xing et al. 2018). More generally, comparisons of spiking and LFP activity facilitate comparison with other signals, such as blood oxygenation level dependent (BOLD) and electroencephalographic (EEG) activity, that can be obtained noninvasively (Goense and Logothetis 2018).

Of course, before attempting to analyze the relationship between two types of signals, one must ensure that they can be estimated independently. LFP estimation often relies on the fact that the frequency content of LFPs and individual spike waveforms differ substantially, such that LFPs can be isolated by simply low-pass filtering the voltage signal. However, to the extent that the LFPs and spikes contain overlapping frequencies, this approach will fail to completely separate the two signals. The resulting *spectral contamination* is of particular concern when estimating the causal relationships between LFPs and spikes because contamination of one signal by the other will result in spurious correlations. Spectral contamination is most problematic when spikes and LFPs are estimated from the same electrode recordings, but it can also affect recordings obtained from separate electrodes, particularly when the corresponding spike density functions are highly correlated. Although various algorithms for efficient removal of spikes from LFPs have been proposed (David et al. 2017; Galindo-Leon and Liu 2017), a systematic appraisal of the consequences of this contamination in common analytical settings has not yet been attempted.

In this work we use simulations to estimate the extent of the spectral contamination of LFPs by individual spikes and spike trains, under conditions similar to those of typical extracellular recordings. We show that such contamination introduces artifactual spike–LFP relationships into many of the popular time-domain metrics mentioned earlier. In particular, strong spike–LFP coherence and spike-triggered LFPs can be obtained from random combinations of spikes and LFPs, if the former are not removed properly. We confirm these simulation results with real recordings from the primate visual system.

Address for reprint requests and other correspondence: T. Zanos, Montreal Neurological Institute, McGill University School of Medicine, Department of Neurology and Neurosurgery, 3801 University St., Room 888, Montreal, Quebec H3A 2B4, Canada (E-mail: theodoros.zanos@mcgill.ca).

0022-3077/11 Copyright © 2018 The American Physiological Society

Downloaded from jn.physiology.org on January 11, 2018

To address the issues associated with spike contamination we develop a novel Bayesian method that estimates veridical LFPs by removing action potential signatures from raw voltage signals. We show that application of this method to both simulations and real data has a substantial impact on measures of the relationships between spikes and LFPs. We also test two other spike removal methods and show that the Bayesian algorithm performs better in most cases. Overall, these results show that the way in which electrode signals are processed can significantly affect the conclusions drawn about the nature of spikes, LFPs, and interactions between these two signals.

METHODS

Electrophysiological recordings

EXPERIMENTAL SETUP. Two rhesus macaque monkeys took part in the experiments. Both underwent a sterile surgical procedure to implant a headpost and recording cylinder and, after recovery, monkeys were seated comfortably in a primate chair (Crist Instruments) and trained to fixate a small spot on a computer monitor in return for a liquid reward. Eye position was monitored at 200 Hz with an infrared camera (SR Research) and required to be within 3° of the fixation point for the reward to be dispensed. Recordings were obtained from well-isolated single units in areas MT/V5 and V6, both of which were identified based on anatomical magnetic resonance imaging scans and the physiological properties of the neurons (Fattori et al. 2009; Maunsell and Newsome 1987). All aspects of the experiments were approved by the Animal Care Committee of the Montreal Neurological Institute and were conducted in compliance with regulations established by the Canadian Council of Animal Care.

PROCEDURE AND VISUAL STIMULI. On each trial, the animal acquired fixation, after which the stimulus appeared and remained stationary for 200 ms. The stimulus then moved at a constant direction and speed for 500 ms. Stimuli were displayed at 85 Hz at a resolution of 1,920 × 1,200 pixels and the viewing area subtended 70 × 42° of visual angle at a distance of 42 cm. Stimuli consisted of sinusoidal gratings of optimal spatiotemporal frequency or random-dot patches displayed on a gray background (luminance of 70.3 cd/m^2). Stimulus size was optimized for each cell. Motion direction was sampled in 30° steps and each stimulus was repeated five times in blockwise random order.

Signal processing

Electrophysiological signals were recorded with a standard data acquisition system (Plexon Multichannel Acquisition Processor [MAP] System). Specifications of this system typically include a dedicated low-pass filter for LFP signals (four-pole high cut at 170 Hz) followed by digitization at 1 kHz. As we show in the following text, this type of filtering invariably introduces spurious correlations into measures of spike–LFP coherence. We therefore obtained custom hardware modifications to the MAP system that included wideband analog filters (two-pole high cut at 2.5 kHz) and a higher digitization rate (10 kHz). These modifications allowed us to detect spectral contamination of the LFPs by spikes; subsequent spike sorting and LFP analysis were performed through off-line digital filtering.

To remove line noise at 60 Hz, we assumed that the underlying signal had equal power in a small range of frequencies around 60 Hz and filtered the signal so that the power at 60 Hz was similar to that of surrounding frequencies. We transformed the wideband signal into the Fourier domain and minimized the squared error between the absolute value of the Fourier coefficients between 55 and 65 Hz and the function

$$a \exp(-b\left|f - f_0\right|^c) + d \quad (1)$$

Here f is frequency, f_0 is the peak frequency (nominally 60 Hz), a is the height of the 60 Hz peak, b controls its width, c adjusts the shape of the peak, and d is the baseline frequency content. This functional form provides a parsimonious parameterization of the frequency content around 60 Hz. We then corrected the 60 Hz artifact by filtering the signal in the Fourier domain with the function's inverse

$$\frac{d}{a \exp(-b\left|f - f_0\right|^c) + d} \quad (2)$$

The power spectrum of the corrected signal showed no apparent discontinuity around 60 Hz, in contrast to a simpler notch filter. A similar operation was performed around 180 Hz.

Simulations

To assess the likely contribution of spectral contamination to various measures of spike–LFP coherence, we analyzed a signal that, by construction, contained no temporal relationship between spikes and LFPs. This *composite* signal was obtained by taking the frequency content of the wideband signal obtained from our electrophysiological recordings, after removing the spikes using the spike removal algorithm presented in the following text, and shuffling the phases (Fig. 1*A*). To this signal we added back waveforms taken at random from the original electrophysiological signal, preserving the timing of each spike while scaling the waveforms to achieve a desired signal-to-noise ratio (SNR). To obtain reliable estimates of the underlying spike waveforms, we first selected typical waveforms that were within 3SDs of the mean waveform at all times. We then projected this ensemble of waveforms onto their first five principal components, obtaining a bank of denoised waveforms. To achieve a desired mean firing rate, Poisson distributed spurious spikes were added to the signal.

The LFP components of this signal thus had the same amplitude spectra as those of a real neural signal, but lacked any consistent relationship between spike times and LFP phases. Consequently, any apparent relationship between the two signals would necessarily be artifactual.

Data analysis

In an initial preprocessing step, we corrected the phase distortions introduced by analog filters by filtering the wideband signal backward in time with a digital Bessel filter that closely modeled the phase response of the analog filters used in the preamplifier (Nelson et al. 2008). We then filtered the wideband signal with digital filters similar to those used conventionally to estimate LFPs. The filters used were a fourth-order 170-Hz low-pass causal Butterworth filter (digital version of the filter used on the Plexon PreAmp separation for the LFP board) and a fourth-order two-pass Butterworth (same filter with zero phase delay). Filtering was implemented with Matlab software (The MathWorks).

For some analyses, we divided the LFP frequency spectrum into the following bands, similar to those used in previous studies (Khawaja et al. 2009; Ray et al. 2008): θ (4–8 Hz), α (8–12 Hz), β (16–24 Hz), low γ (25–55 Hz), and high γ (65–140 Hz). Since the θ, α, and β bands did not show significant spike contamination in any of our metrics, in both simulations and real data, for reasons of compactness we combined all three bands into a single "low-frequency" band (4–24 Hz).

Spike detection and sorting based on the high-pass filtered wideband signal were performed off-line using established methods (Quiroga et al. 2004). These results were compared for consistency with results from Plexon's Off-line Sorter software, which uses spike waveforms sampled at 40 kHz. Spike timing information was considered accurate if agreement between the results of the two spike sorting

Downloaded from jn.physiology.org on January 11, 2018

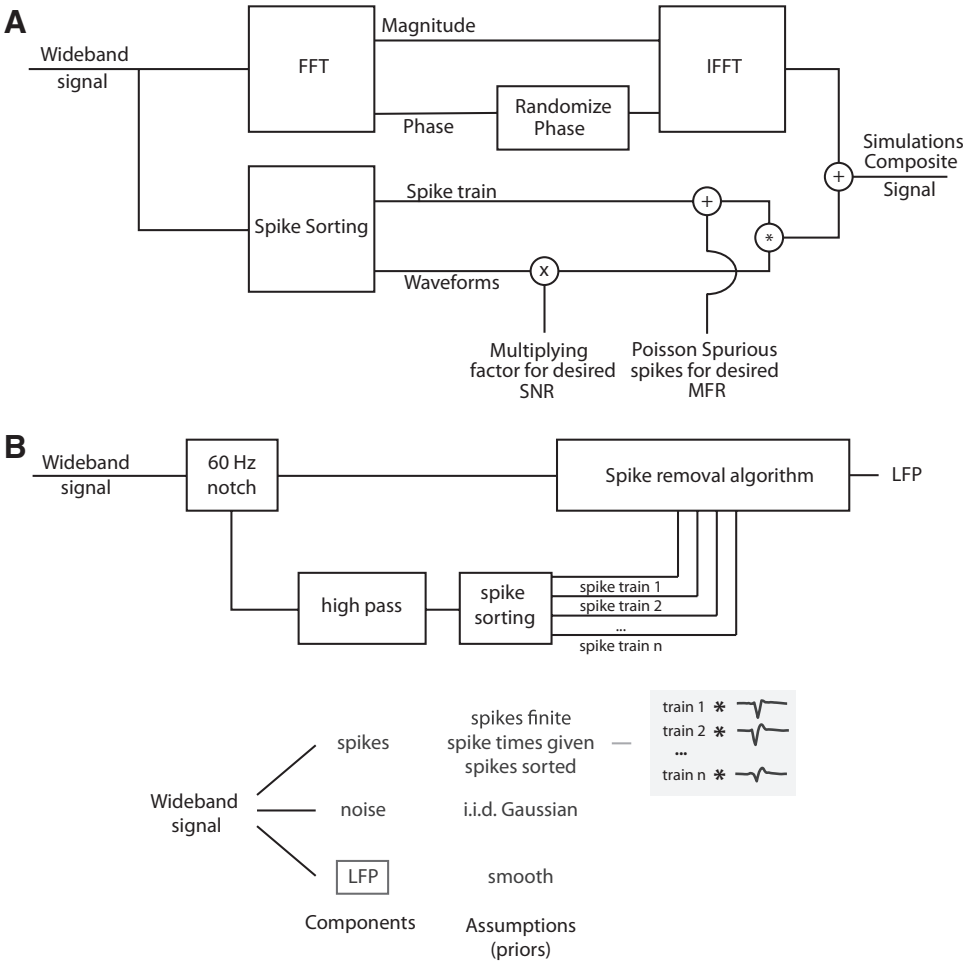


FIG. 1. Block diagrams of the ground-truth simulations and the spike removal algorithm. *A*: block diagram of the procedure followed in our simulations to create a composite signal. This signal is generated by randomizing the phase of raw signals recorded electrophysiologically and adding spike waveforms, scaled according to a desired signal-to-noise ratio (SNR). This procedure dissociates the local field potential (LFP) components of the signal from the spike times. *B*: block diagram of the spike removal algorithm. The algorithm assumes that the wideband voltage signal is generated by the sum of a low-frequency LFP component, spikes of limited duration occurring at known times, an offset, and white noise. These assumptions are embedded in a probabilistic model and the most likely LFP is then determined through Bayesian inference.

methods was above a criterion level of 95%; otherwise, the sorting parameters were altered until the agreement exceeded the criterion. In three cases where this level could not be reached, the data were rejected from further analysis.

SNRs were computed for each isolated unit as the peak-to-trough voltage of the mean spike waveform, divided by the square root of the mean power of the complete raw signal. We obtained the STA of the LFP by averaging local field potentials centered around the time of spikes (Dayan and Abbot 2001). We subsequently filtered the LFP with complex Morlet wavelets and used the magnitude of the coefficients as an estimate of the instantaneous power of the LFP at given frequencies (Goupillaud et al. 1984). We averaged the magnitude of the coefficients around the time of spikes, normalized by the average magnitude of the coefficients of the whole LFP signal, to obtain a normalized wavelet STA, which can reveal transient increases or decreases in the magnitude of oscillations at given frequencies around the time of spikes.

To test whether spikes preferred certain phases of the LFP, the instantaneous phase and amplitude of the LFP at 35 different frequencies (log spaced from 4 to 150 Hz) were first calculated with the Hilbert transform of different filtered versions of an LFP recording. A single phase value was then assigned at each frequency for every spike in a recording (Jacobs et al. 2007). By convention, a phase value of 180° corresponds to the positive (hyperpolarizing) peak of the cycle. Finally, histograms of the distributions of spike phase for different LFP frequencies across all spikes were compiled; 50 phase bins, ranging from 0 to 360° were used. The uniformity of the distribution of spike phase values for our frequency ranges was assessed for each case, using the Rayleigh statistical test, corrected for the number of frequency ranges (Rutishauser et al. 2010). A cell was considered phase-locked at a specific frequency range if the null hypothesis of uniformity of the phase distribution could be rejected at a corrected $P < 0.01$.

Spike-field coherence (SFC) was calculated as the power spectrum of the LFP STA divided pointwise by the sum of the power spectra of all LFP segments used to compute the STA (Fries et al. 2001). This measure can reveal phase-locking at high frequencies that is otherwise obscured in the STA of the LFP by low-frequency noise. A value of 1 at a given frequency of the SFC means that all spikes appear at the same phase for this frequency, whereas a value of 0 indicates no consistent relationship between the phase of the LFP at a given frequency and spikes.

Tuning curves were quantified for single units (spike count) and LFPs (normalized LFP band power), in the three frequency bands described earlier. We used Morlet wavelets to measure the instantaneous power at different frequencies of the LFP across time, as in the wavelet STA analysis. LFP tuning curves were derived by normalizing the mean magnitude of the Morlet coefficients over a time period that spanned 100–500 ms after the onset of stimulus motion against that in a baseline period consisting of 150 ms of spontaneous activity before the stimulus presentation. Average tuning vectors for each band were also computed for both single units and LFPs.

Spike removal algorithm

The purpose of the algorithm is to estimate the local field potential based on a measured wideband voltage trace of length n. We assume that this wideband signal $\mathbf{y}$ is the superposition of a low-frequency local field potential $\mathbf{w}$, high-frequency spike components $\boldsymbol{\eta}^k$, an offset μ, and white noise $\boldsymbol{\varepsilon}$ (Fig. 1*B*).

Downloaded from jn.physiology.org on January 11, 2018

$$\mathbf{y} = \mathbf{w} + \sum_{k=1}^{m} \boldsymbol{\eta}^k + \mu + \boldsymbol{\varepsilon} \quad (3)$$

Here m is the number of sorted neurons emitting spikes. The high-frequency component of the kth neuron $\boldsymbol{\eta}^k$ is created by the convolution of the neuron's spike train $\mathbf{s}^k$, assumed known (see Fig. 1*B*), and the neuron's spike waveform $\boldsymbol{\phi}^k$

$$\boldsymbol{\eta}^k = \boldsymbol{\phi}^k * \mathbf{s}^k \quad (4)$$

Importantly, waveforms $\boldsymbol{\phi}^k$ are not set to the mean waveforms determined by spike sorting or by spike-triggered averaging the wideband signal, as in previously proposed methods of eliminating spike remnants (Pesaran et al. 2002). Rather, the waveforms $\boldsymbol{\phi}^k$ are considered as free parameters and are estimated optimally as a part of the algorithm. The offset μ is introduced to compensate for the fact the mean of $\boldsymbol{\eta}^k$ is likely to be negative.

The generative model (*Eq. 3*) contains more unknowns ($\mathbf{w}$, $\boldsymbol{\phi}^k$, and μ) than knowns. Thus assumptions must be introduced to make the estimation of the local field potential $\mathbf{w}$ well-posed. We first assume that the local field potential is smooth or, equivalently, that most of its power lies in the low frequencies

$$p(\mathbf{w}) = N(0, \gamma^2\boldsymbol{\Gamma}) \quad (5)$$

Here $N(\mathbf{a}, \boldsymbol{\Sigma})$ represents a multivariate Gaussian with mean $\mathbf{a}$ and covariance $\boldsymbol{\Sigma}$. Capital gamma ($\boldsymbol{\Gamma}$) is a matrix that embodies our assumption of smoothness. Multiplication of a vector $\mathbf{x}$ by this matrix ($\boldsymbol{\Gamma}\mathbf{x}$) produces a low-pass filtered version of the target $\mathbf{x}$. The matrix was chosen to be circulant, a property necessary to make estimation of the local field potential tractable; definitions and properties of circulant matrices are discussed at length in the Supplemental APPENDIX.[1] Gamma (γ) controls the strength of the prior. Our second assumption is that $\boldsymbol{\varepsilon}$ is generated by a white noise process, $p(\boldsymbol{\varepsilon}) = N(0, \sigma^2\mathbf{I})$. Finally, we assume that spike waveforms $\boldsymbol{\phi}^k$ lie in a subspace $\mathbf{B}$

$$\boldsymbol{\phi}^k = \mathbf{B}\boldsymbol{\varphi}^k \quad (6)$$

We chose $\mathbf{B}$ so that spike waveforms were described by one parameter per time sample in the interval from −1 to +2 ms relative to the peak depolarization. This duration of 3 ms was chosen based on the traditional length of snippets used in spike sorting, generally ranging from 1.5 to 3 ms.

To estimate the local field potential $\mathbf{w}$, the spike waveforms $\boldsymbol{\phi}^k$, and the offset μ, we used Bayesian inference to obtain maximum a posteriori (MAP) model parameters. By Bayes' theorem, the log-posterior of the model is

$$p(\mathbf{w}, \boldsymbol{\varphi}^k, \mu|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{w}, \boldsymbol{\varphi}^k, \mu)p(\mathbf{w})$$
$$= k \exp\left[-\frac{1}{2\sigma^2}\sum_i \left(\mathbf{y} - \mathbf{w} - \sum_{k=1}^{m} \boldsymbol{\eta}^k - \mu \right)_i^2 - \frac{1}{2\gamma^2}\mathbf{w}'\boldsymbol{\Gamma}^{-1}\mathbf{w} \right] \quad (7)$$

Here k is a constant factor. By taking the log of this expression and setting partial derivatives with respect to the parameters to 0, we obtain the MAP estimates of the parameters

$$\bar{\mathbf{w}} = (\gamma^2\boldsymbol{\Gamma} + \sigma^2\mathbf{I})^{-1}\gamma^2\boldsymbol{\Gamma}\left[\mathbf{y} - \sum_k \mathbf{s}^k * (\mathbf{B}\bar{\boldsymbol{\varphi}}^k) - \bar{\mu} \right]$$
$$\bar{\boldsymbol{\varphi}}^k = (\mathbf{s}^k * \mathbf{B})^+ \left[\mathbf{y} - \bar{\mathbf{w}} - \sum_{j \neq k} \mathbf{s}^j * (\mathbf{B}\bar{\boldsymbol{\varphi}}^j) - \bar{\mu} \right]$$
$$\bar{\mu} = \frac{1}{n}\sum_i \left[\mathbf{y} - \bar{\mathbf{w}} - \sum_k \mathbf{s}^k * (\mathbf{B}\bar{\boldsymbol{\varphi}}^k) \right] \quad (8)$$

Here $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ denotes the Penrose–Moore pseudoinverse. These equations can be understood as follows. First, the optimal offset of the model $\bar{\mu}$ is simply the mean of the wideband signal with the LFP and spike contributions removed. Second, the optimal local field potential $\bar{\mathbf{w}}$ is related to the spike-free, mean-corrected wideband signal $\mathbf{z} = \mathbf{y} - \sum_k \mathbf{s}^k * (\mathbf{B}\bar{\boldsymbol{\varphi}}^k) - \bar{\mu}$ by matrix multiplication with $(\gamma^2\boldsymbol{\Gamma} + \sigma^2\mathbf{I})^{-1}\gamma^2\boldsymbol{\Gamma}$. Multiplying a vector by this matrix low-pass filters the vector and thus the model tells us that the optimal LFP is obtained by low-pass filtering the wideband signal after removing spike waveforms around the time of every spike.

The key equation is the second in the series, which governs the spike waveforms. By the properties of the pseudoinverse, $(\mathbf{s}^k * \mathbf{B})^+(\mathbf{s}^k * \mathbf{B})\mathbf{a} = \mathbf{a}$. Now to recover a waveform $\mathbf{a}$ from a train of waveforms $(\mathbf{s}^k * \mathbf{B})\mathbf{a}$, one simply needs to take the spike-triggered average of this signal normalized by the autocorrelation of the spike train. Thus the optimal spike waveform for the kth neuron is given by the decorrelated STA of the wideband signal minus the LFP, the spike contributions from other neurons, and the offset. However, since the optimal LFP is mostly determined by low-pass filtering the wideband signal, the overlap between spikes from different neurons is often negligible and spike trains are uncorrelated at short time lags due to the absolute refractory period, it follows that to a first approximation the optimal waveform for a given neuron is equal to the STA of the high-pass filtered wideband signal.

In practice, the optimal spike waveform determined by the model can vary substantially from the STA of the high-pass filtered wideband signal because of the various corrections it applies. A fast and effective model-based spike removal algorithm can be derived by decoupling the *Eq. 8* system and applying the conjugate gradient descent to solve the resulting equations, as shown in the APPENDIX. An implementation of this algorithm in Matlab is available from our website (http://apps.mni.mcgill.ca/research/cpack/lfpcode.zip). For the model to be fully specified, the strength of the prior relative to the noise γ^2/σ^2 must be estimated because this determines what the model considers as signal and what it discounts as noise. In the APPENDIX, we show how marginal likelihood (or evidence) optimization is used to estimate prior and noise levels. We also detail in the APPENDIX how to choose an appropriate form for the prior matrix $\boldsymbol{\Gamma}$ and how to optimize performance for long recordings.

Other spike removal algorithms

Two methods to remove spike waveforms from wideband signals were also tested in both simulated and real data. The first method, known as *mean spike removal*, simply subtracts the mean spike waveform from the wideband signal at the corresponding spike times. Here we used a 3-ms window (1 ms before the spike and 2 ms after) to recover the mean spike waveform. This was chosen for comparison with our Bayesian algorithm, which used a window of the same size. A second method uses linear *interpolation* to replace the wideband signal around the time of each spike, using the same time window as that in the other methods (3 ms).

RESULTS

Spectral composition of spike waveforms

Spikes are by definition fast events that are dominated by temporal frequencies far higher than those that make up the LFP signals. Thus it is natural to assume that low-pass filtering the raw voltage signal will remove the spikes and that what remains will be a pure estimate of the LFPs. This assumption can be examined by computing the power spectrum of a typical spike waveform and Fig. 2*A* shows one such waveform recorded extracellularly in our laboratory. As is typical in such recordings, the duration of the action potential is on the order of 1 ms and the corresponding waveform contains both a peak and a trough. Consequently, the power spectrum (Fig. 2*B*, *top*) is dominated by high frequencies. However, closer inspection of the power spectrum shows that there is a small amount of power at low frequencies in this fast waveform (as shown in

[1] The online version of this article contains supplemental data.

Downloaded from jn.physiology.org on January 11, 2018

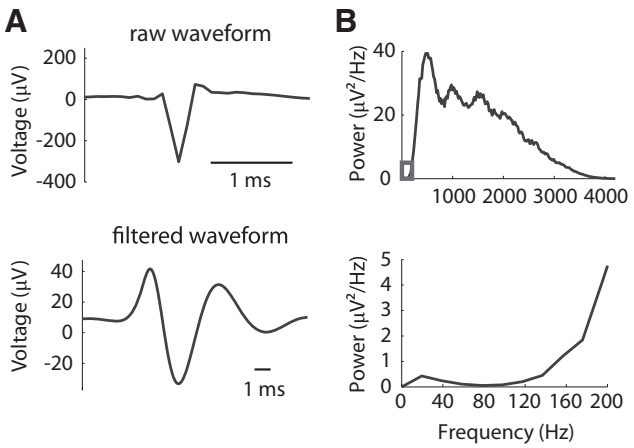


FIG. 2. Spike waveform spectral characteristics and low-pass filtering. *A*: waveform of an extracellularly recorded spike (*top panel*). The *bottom panel* shows a version of the same waveform following application of a low-pass filter with a cutoff frequency of 150 Hz. *B*: power spectrum of the spike waveform (*top panel*) and a magnified version of the spectrum for the frequencies typically associated with LFPs (*bottom panel*).

Fig. 2*B*, *bottom*), particularly in some of the frequencies (100–200 Hz) that are often examined in LFP studies (Gold et al. 2006).

The *bottom panel* of Fig. 2*A* shows the result of applying a low-pass filter to the spike waveform shown in the *top panel*. Although this filtering reduces the amplitude of the waveform, some residual signal remains and this residual retains the basic shape of the original waveform. Moreover, the residual signal is stretched in time over an interval determined by the cutoff frequency of the low-pass filter (here on the order of 10 ms).

Beyond the low-frequency content of individual spike waveforms, the temporal structure of spike trains can contain additional power at low frequencies (Bair et al. 1994). Consequently, if spikes are not properly removed, a succession of spikes can contain more power at low frequencies than an individual waveform would indicate. These considerations suggest that the commonly used technique of low-pass filtering the raw voltage signal is not entirely sufficient to isolate the LFP signals. To estimate the magnitude of this problem, we developed a set of ground-truth simulations that allowed us to quantify the effects of spike contamination in a context where the veridical LFP signal is known.

Spectral contamination in simulated data

The basis of the simulations was a phase-randomized wideband signal with the same frequency content as that of a real wideband signal (see METHODS and Fig. 1*A* for details). By combining this signal with spike waveforms taken from electrophysiological recordings (see following text), we obtained a *composite* signal similar to that typically recorded with electrophysiological methods but in which the LFPs and spike times were completely dissociated. This allowed us to perform ground-truth simulations in which any temporal relationship between spikes and LFPs was necessarily an artifact of the method of analysis. The duration of the phase-randomized signal in all our simulations was 3 min.

Figure 3 shows a short segment of the signals used in our ground-truth simulations. The *top panel* shows a phase-randomized wideband signal before (red line) and after the addition of spikes with a mean amplitude of 2 dB (blue line, offset downward to facilitate reading). The *middle panel* shows these same signals after low-pass filtering. Here the difference between the signals is much less visible, which suggests that in some analysis scenarios low-pass filtering may be sufficient to remove spike traces. The *bottom panel*, however, reveals that the deviations between the two signals (green line, scaled by a factor 3 to facilitate reading) are centered around the time of spikes (gray lines) and are highly stereotyped. These small spike remnants can bias analyses of LFPs, as we show in the next sections.

Spike-triggered average (STA) of the LFP

As shown in Fig. 2, a typical spike waveform contains power at low frequencies, which can potentially complicate the isolation of the LFP component of the raw voltage signal. In particular, although the magnitude of the low-frequency component is relatively small (Fig. 3, *bottom*), its importance could be amplified by analytical techniques aimed at finding the components of the LFP that are correlated with spikes. One common example of such an analysis is the STA of the LFP.

The STA of the LFP is obtained by averaging the LFP signals recorded around the time of each spike. In principle this method is useful for inferring causal relationships between the two signals and it is commonly used to study local neuronal synchronization related to the discharge of a specific neuron and to compute the average spike-related modulation in specific LFP bands. The shape of the STA reveals the nature of such correlations, whereas a flat STA suggests that the two signals are unrelated.

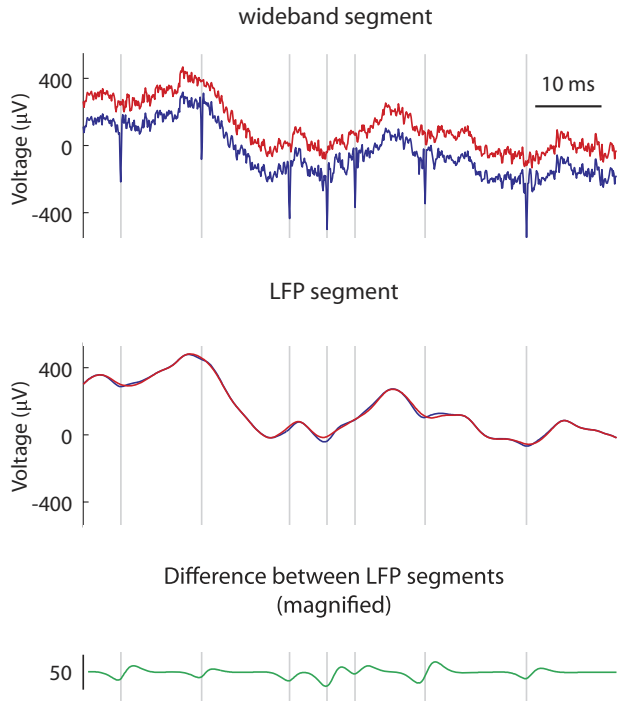


FIG. 3. Composite wideband and LFP signal segments with and without spike waveforms. Segment of a spike-free, phase-randomized wideband signal (*top panel*) before (red line) and after (blue line) the addition of spike waveforms. Filtered versions of these signals (*middle panel*) and their difference (*bottom panel*) reveal the effect of the addition of spike waveforms on the LFP component of the composite signal.

Downloaded from jn.physiology.org on January 11, 2018

We first examined the STA of a composite signal in which the SNR of the spike relative to the LFP was set to 2 dB and the mean spike rate was 9 spikes/s. Prior to calculation of the STA the raw composite signal was low-pass filtered to remove frequencies ≥170 Hz. However, inspection of the resulting STA in Fig. 4*A* (*top*) reveals clear residual structure, indicating that low-pass filtering is insufficient to separate spikes from LFPs. The artifact has a shape similar to that of the mean waveform, but has a longer duration, as expected from Fig. 2*A* (*bottom*). The exact shape of the artifact is determined by the properties of the low-pass filter.

For comparison we also computed the STA associated with a signal that was filtered with a single-pass Butterworth filter (Fig. 4*A*, *middle*), which is commonly implemented in analog circuits that are part of LFP recording hardware. In addition to the spurious STA, this filter also introduces frequency-dependent phase shifts, as reported previously (Nelson et al. 2008). Thus the commonly used approach of low-pass filtering the raw voltage signal appears to alter both the magnitude and timing of spike–LFP correlations.

The effects of this contamination are decomposed into different frequency bands in Fig. 4*B*. Here the STAs affected by spike contamination are shown in blue, whereas STAs obtained after the application of the spike removal algorithm introduced in the next section are shown for reference in red. The residual influence of spikes is present in all frequency bands and is particularly strong at low- and high-gamma frequencies.

Although the STA can detect linear relationships between spikes and LFPs, it is insensitive to higher-order correlations between these signals. The wavelet STA, however, can reveal transient changes in the relative instantaneous power in different frequency bands around the time of spikes, the signature of a more complex relationship between LFPs and spikes (Burns et al. 2010; also see METHODS). Figure 4*C* (*top*) shows that the wavelet STA is also sensitive to contamination by spikes. Here contamination is restricted to LFP frequencies ≥100 Hz at 0 ms time lag (deep red color). Note that the stripe pattern visible at low frequencies is due to chance fluctuations in the power at these frequencies and is not significant.

Bayesian spike removal algorithm

To address the contamination issues illustrated in Fig. 4, we developed an algorithm that removes spikes from raw voltage traces. The algorithm assumes that the wideband voltage signal is generated by the sum of a low-frequency LFP component, spikes of limited duration occurring at known times, an offset, and white noise (see Fig. 1*B*, *bottom diagram*, and METHODS). These assumptions are embedded in a probabilistic model and the most likely LFP is then determined through Bayesian inference. The optimal parameters of the model have straightforward interpretations (see METHODS and the APPENDIX): spike waveforms are essentially STAs of the high-pass filtered wideband signal, whereas the LFP is given by the low-pass filtered wideband signal after the subtraction of the contribution of spikes.

Figure 4*A* shows that the magnitude of the STA of the LFP when the composite signal is preprocessed through different methods. As mentioned earlier, a spurious STA is obtained following standard low-pass filtering (*top two panels*), whereas our spike removal algorithm greatly reduces this artifact (*bottom*). Figure 4*B* shows the STAs at different frequencies, with the LFP derived through standard low-pass filtering indicated in blue. The red lines show the STAs at different frequencies following application of our spike removal algorithm. The magnitude of the spurious correlation is now much smaller, particularly at high frequencies. Because the spike removal method removes the mean spike only at spike times, and thus will not remove all spike remnants when spike shapes for a given neuron are not completely stereotyped, some residual second-order correlations between spikes and LFPs could remain. Nevertheless, the method is effective at attenuating

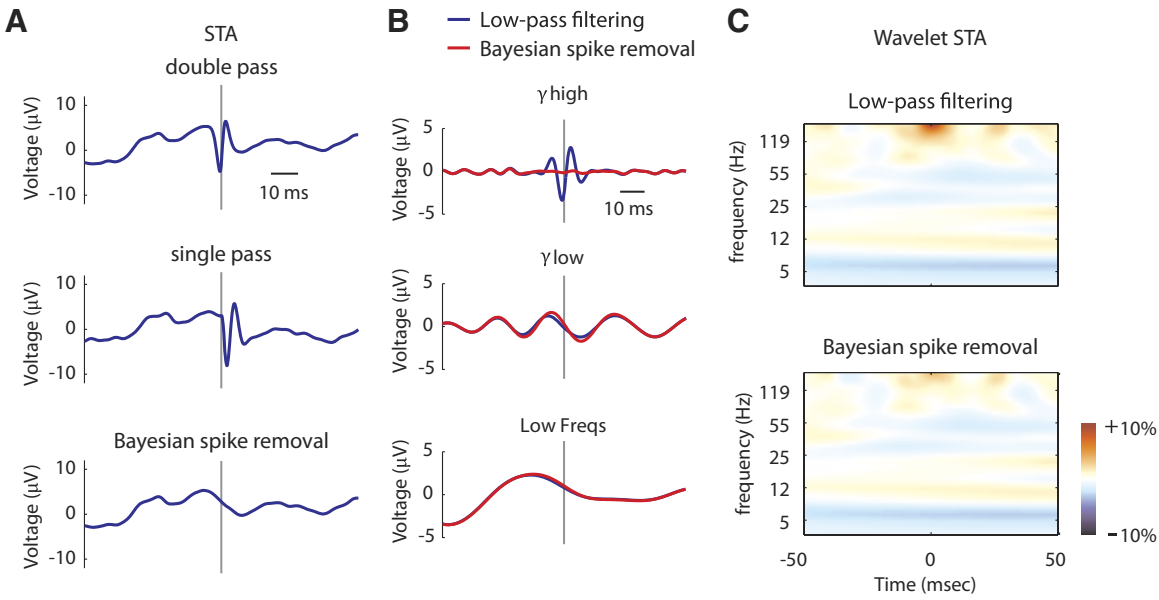


FIG. 4. STAs for the 2-dB simulated case. *A*: spike-triggered average (STA) of the LFP signal filtered with a Butterworth single-pass filter (*top panel*), a Butterworth double-pass (zero-phase) filter (*middle panel*), and a zero-phase filter after applying the spike removal algorithm (*bottom panel*). *B*: STAs of specific LFP frequency bands before (*blue trace*) and after (*red trace*) spike removal. *C*: normalized wavelet STA before (*top panel*) and after (*bottom panel*) spike removal.

Downloaded from jn.physiology.org on January 11, 2018

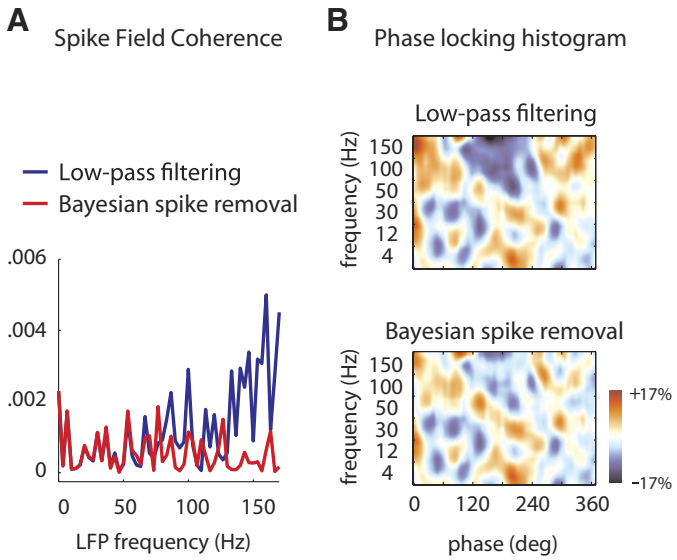


FIG. 5. Phase-locking analysis results for the 2-dB simulated case. *A*: spike-field coherence (SFC) of the LFP signal before (*blue trace*) and after (*red trace*) spike removal. *B*: phase-locking histogram before (*top panel*) and after (*bottom panel*) spike removal.

spurious correlations in the wavelet STA, as shown in Fig. 4*C*. Additional details on the derivation and implementation of the spike removal algorithm are given in the APPENDIX, which also covers more challenging conditions, such as long recordings or nonstationary spike shapes that result from electrode drift.

Phase synchronization

Another quantity commonly used to assess spike–LFP relationships is the synchronization between spikes and LFPs at different frequencies. This is typically manifested as a preference for spikes to occur at specific phases of the LFP oscillation cycle and such preferences may reveal synchronization of the local presynaptic activity. This synchronization can be an alternative way of encoding behaviorally relevant signals, particularly those related to visual attention (Fries et al. 2008), motor planning (Denker et al. 2007), motor output (Baker et al. 1999), and memory consolidation (Paz et al. 2008; Siapas et al. 2005). The frequency at which spikes are locked can also provide hints about the spatial and dynamic properties of the neural circuitry driving the neuron (Liu and Newsome 2006).

One method that is commonly used to assess the synchronization between spikes and LFPs is the spike-field coherence (SFC), obtained from the power spectrum of the STA. Figure 5*A* shows the SFC as a function of frequency for simulations in which the composite signal was preprocessed with low-pass filtering (blue) and our spike removal algorithm (red). When spikes are removed by low-pass filtering alone, prominent phase synchronization is seen at the higher frequencies and this artifactual relationship disappears following spike removal.

Figure 5*B* shows the phase-locking histogram, which measures the distribution of the instantaneous phase of the LFP at different frequencies at the time of spikes. Here it is clear that this measure is particularly sensitive to spectral contamination of the LFP. Note that the slant visible in the phase–frequency relationship is due to frequency-dependent phase shifts induced by the causal low-pass filter.

LFP tuning curves

A final quantity of interest is the tuning curve, which captures the magnitude of the spike or LFP response as a function of an experimentally manipulated sensory or motor variable. Comparison of tuning curves obtained from the two signals provides additional insight into local cortical processing, the architecture of such neuronal circuits (Liu and Newsome 2006), the spatial extent of the LFP (Katzner et al. 2009), and links of LFP features to synaptic inputs and neuronal outputs (Khawaja et al. 2009). We therefore examined the tuning of a composite signal in which untuned LFPs were combined with the waveforms of well-tuned spikes. For the latter both the waveforms and the tuning curves were taken from a single-unit recording of a neuron found in the middle temporal area (MT) of a macaque monkey. The tuning curve for the spiking responses of this neuron for the motion direction of a visual stimulus is shown in Fig. 6*A*.

Figure 6 shows the LFP tuning curve for the high gamma band, with low-pass filtering of the raw signal and after our spike removal algorithm was applied (Fig. 6*B*). Because the LFP signal is untuned by construction, successful spike removal should yield responses that respond roughly equally to all motion directions, whereas spike contamination will manifest itself as tuning for the downward-rightward motion shown in Fig. 6*A*. The results show that both LFP tuning curves are untuned for visual motion direction, suggesting that the spikes do not introduce spurious tuning into the LFPs, even at higher frequencies (high gamma band).

Summary of spike contamination in simulated data

Our simulation results reveal that spike contamination can introduce strong artifactual correlations between spikes and LFPs. In further stimulations we found that both spike amplitude and the firing rate of a neuron affect the magnitude of spike contamination in LFPs, with larger firing rates and spike amplitudes leading to greater contamination. Supplemental Figs. S1 and S2 and Table 1 provide summaries of the magnitude of spike contamination for different spike amplitudes and firing rates. As shown in Supplemental Fig. S1, the normalized amplitude of the STA of the LFP rises linearly as spike amplitude increases. Similarly, higher values of SNR

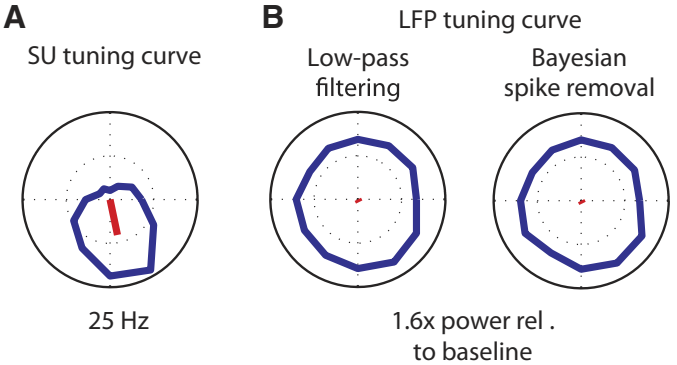


FIG. 6. Tuning curves for the 2-dB simulated case. *A*: direction tuning curves from single units recorded in macaque visual cortex. *B*: direction tuning curves from high-gamma LFPs before (*middle panel*) and after (*right panel*) spike removal. Lower-frequency bands exhibited similar tuning characteristics before and after spike removal and are omitted for space considerations.

Downloaded from jn.physiology.org on January 11, 2018

TABLE 1. *Results of Rayleigh tests for all simulations*

Case	*P* value (high gamma)	*P* value (low gamma)	*P* value (low frequencies)
SNR, dB			
−5	0.3931/0.8710	0.1184/0.1137	0.2748/0.2727
0	0.0087/0.9281	0.1232/0.1206	0.2749/0.2780
5	0/0.7800	0.1425/0.1472	0.2740/0.2900
10	0/0.0148	0.2197/0.3149	0.2758/0.1767
MFR, spikes/s			
5	0.0520/0.2772	0.1183/0.4079	0.3245/0.6200
10	0.0188/0.5650	0.4264/0.7205	0.3740/0.3300
20	0.0040/0.4500	0.2300/0.5600	0.3100/0.2900

P values computed with the Rayleigh test, before and after spike removal (value before/value after). A hypothesis of phase-locking stands for $P < 0.01$.

give higher artifactual values of the SFC and raise the probability of a neuron being considered phase-locked to specific frequency ranges (Table 1). A weaker effect is observed for high neuronal firing rates (Table 1).

Other spike removal approaches

As an alternative to simple low-pass filtering, one might attempt to estimate LFPs by removing individual spikes from the wideband signal. This has traditionally been accomplished using one of two methods. The first method involves subtracting each neuron's mean action potential waveform from the raw signal at the corresponding spike times. A second approach involves removing each action potential and interpolating the wideband signal across the resulting gap. Figure 7 compares the performance of these approaches with the Bayesian algorithm, for three of the metrics discussed earlier (LFP STAs, wavelet STAs, and phase-locking histograms). Here the simulated data are the same as those used in the previous figures.

The *top panel* of Fig. 7*A* shows that the mean spike removal method did not completely remove the residual spike correlation from the STA [a difference of 0.48 dB or 16% more residual compared with the Bayesian spike removal (green line)]. By contrast, the interpolation method (Fig. 7*A*, *middle panel*) seems to perform overly aggressive spike removal, in the process introducing an artificial positive deflection of the LFP STA around the moment of the spike (a difference of 1.68 dB or 44% more residual compared with the Bayesian spike removal). For the wavelet STA (Fig. 7*B*), the mean spike removal method (*top panel*) succeeds in attenuating the spurious correlations, as does the interpolation method (*middle panel*). Indeed the latter method slightly outperforms our Bayesian algorithm (*bottom panel*) for this particular measure (SD of the wavelet coefficients for the interpolation method is 0.0013, for the mean spike removal and the Bayesian algorithm 0.0015 and for the low-pass filter 0.0036). However, as shown in Fig. 7*C* (*middle panel*), interpolation can introduce significant spurious correlations into the phase locking histogram. For this measure mean spike removal (Fig. 7*C*, *top panel*) also fails to remove the spurious phase locking correlations, while the Bayesian algorithm shows little residual structure (*bottom panel*). These results are confirmed by the *P* values of Rayleigh statistical tests (0.008 for the mean spike removal, 0.004 for the interpolation method and 0.74 for the Bayesian algorithm, with the hypothesis of phase-locking for a value of $P < 0.01$).

Thus we find that removal of individual spikes is generally preferable to low-pass filtering the wideband signal. Of the methods for removing individual spikes, removal of the mean

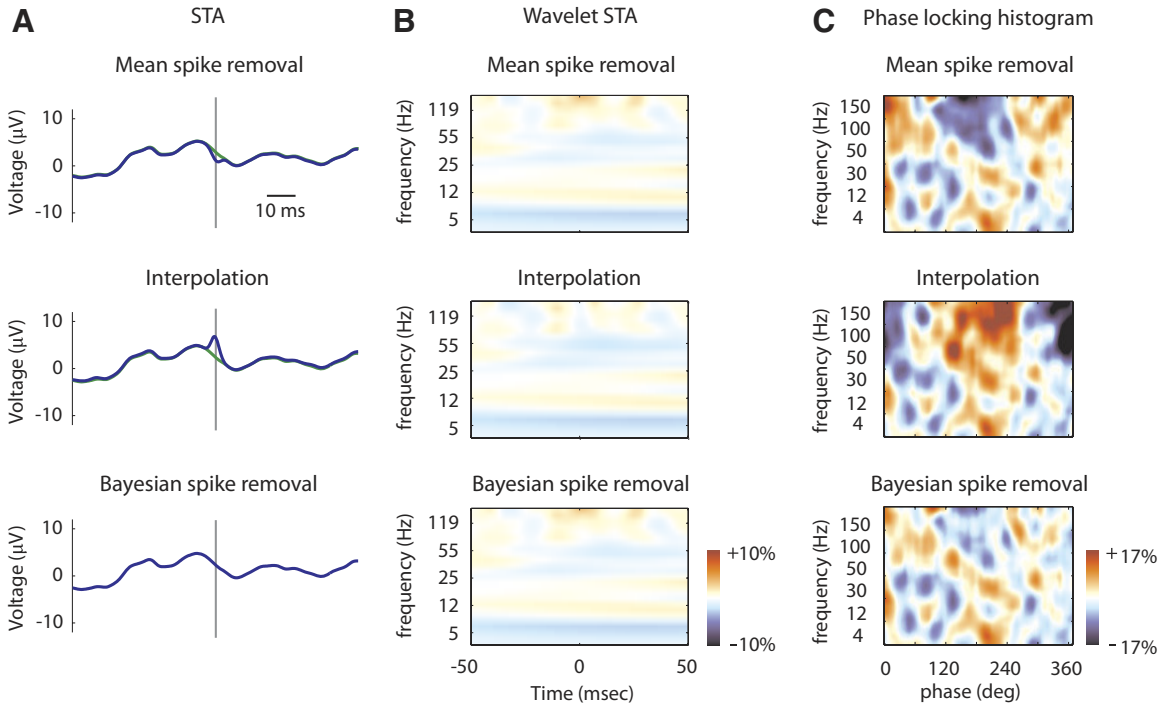


FIG. 7. Comparison of spike removal methods for the 2-dB simulated case. *A*: STAs of the LFP signal. *B*: normalized wavelet STAs. *C*: phase-locking histograms for simulated data in which no real spike–LFP correlations exist. Results are shown for the mean spike removal method (*top panels*), the interpolation method (*middle panels*), and the Bayesian algorithm (*bottom panels*). For the LFP STAs, the results for the Bayesian algorithm result are overlaid for comparison (green lines).

Downloaded from jn.physiology.org on January 11, 2018

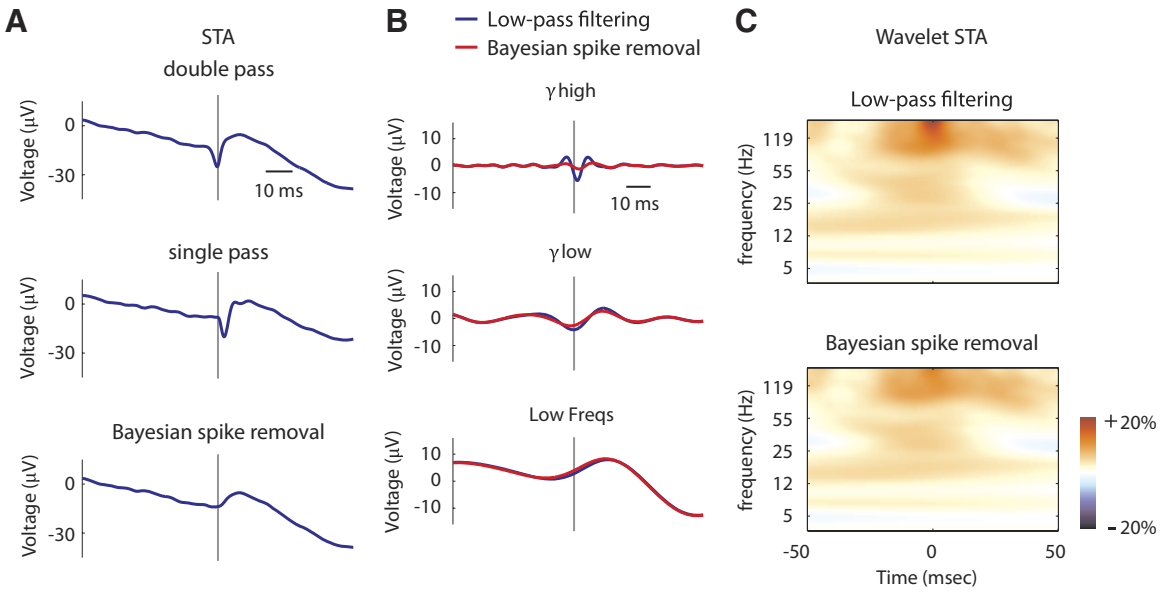


FIG. 8. STAs for an illustrative middle temporal area (MT) neuron. *A*: STA of the LFP signal filtered with a Butterworth single-pass filter (*top panel*), a Butterworth double-pass (zero-phase) filter (*middle panel*), and a zero-phase filter after applying the spike removal algorithm (*bottom panel*). *B*: STAs of specific LFP frequency bands before (*blue trace*) and after (*red trace*) spike removal. *C*: normalized wavelet STA before (*top panel*) and after (*bottom panel*) spike removal.

spike waveform most often leaves residual, artifactual spike–LFP correlations. Interpolation methods perform better by aggressively removing all signals related to each spike, but in the process they can introduce other types of spurious correlations. In most cases our Bayesian algorithm outperformed both previous approaches. Furthermore, as we show in the DISCUSSION and the APPENDIX, the Bayesian algorithm generalizes the advantages of these previous approaches, while casting the problem in a rigorous mathematical framework within which key assumptions can be made explicit.

Real recordings

SINGLE NEURONS. The previous sections showed that low-pass filtering cannot completely remove spike waveforms from composite signals constructed to have no correlation between spikes and LFPs. The resulting spectral contamination of the LFP can introduce significant bias into such measures as the STA, the wavelet STA, the SFC, and phase-locking histograms, especially for higher-frequency bands. By comparison, we found that the tuning of the LFP to stimulus features as measured by tuning curves is not significantly affected by spikes, even when spike responses are strongly tuned. We introduced a new spike removal algorithm that performs better than low-pass filtering.

To examine the consequences of these findings for signals obtained in vivo, we recorded single-unit activity and LFPs from two alert macaque monkeys. A total of 26 isolated neurons in areas V6 (19) and MT (7) neurons were analyzed. Both areas are specialized for the processing of visual motion (Fattori et al. 2009; Maunsell and Newsome 1987). In each case, we determined the selectivity of the single-unit activity for visual motion by drifting a sinusoidal grating or a random-dot stimulus in 12 different directions spaced evenly around the circle. Cells that failed to respond to the stimulus or that lacked direction selectivity (defined in METHODS) were excluded from the analysis.

As shown for a typical cell in Figs. 8 and 9, the trends in the real data qualitatively match those in our simulations. In several cases, including the one illustrated in Fig. 8, the LFP was correlated with the timing of spikes and this was apparent in both STA analysis and phase synchronization. These correlations were not removed by the spike removal algorithm. Thus our results are consistent with previous findings (Gregoriou et al. 2009; Whittingstall and Logothetis 2009) that indicate a relationship between spike timing and LFP phase. However, both the range of frequencies over which this relationship is found and the magnitude of the correlations are affected by the method of preprocessing the raw signals. For example, inspection of the spike-triggered coherence before spike removal

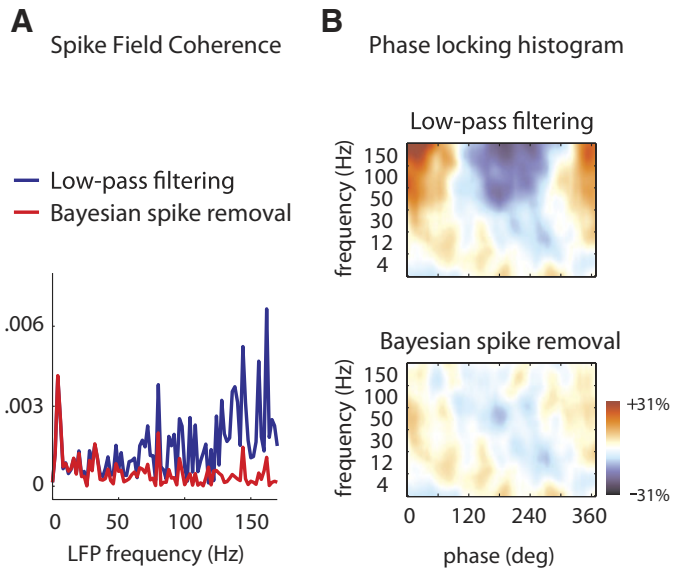


FIG. 9. Phase-locking analysis results for an illustrative MT neuron. *A*: SFC of the LFP signal before (*blue trace*) and after (*red trace*) spike removal. *B*: phase-locking histogram before (*top panel*) and after (*bottom panel*) spike removal.

Downloaded from jn.physiology.org on January 11, 2018

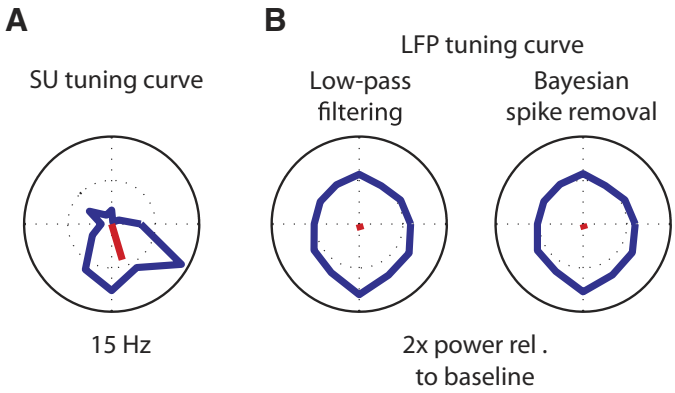


FIG. 10. Tuning curves for an illustrative MT neuron. *A*: direction tuning curves from a single unit. *B*: direction tuning curves for high-gamma LFPs before (*middle panel*) and after (*right panel*) spike removal. The neuron used for these tuning curves had a 1.6 dB SNR spike waveform.

(Fig. 9*A*, blue line) would lead one to conclude that high frequencies of the LFP are robustly correlated with spikes, whereas application of the spike removal algorithm shows that in fact only low frequencies of the LFP are modulated with spikes (Fig. 9*A*, red line).

Figure 10*B* shows LFP tuning curves before and after removal of the spikes of a robustly tuned neuron (Fig. 10*A*). As expected from our simulations, LFP tuning curves were essentially unaffected by spike removal. The next section quantifies the extent of this contamination for our population of cells.

POPULATION RESULTS. We assessed the level of spectral contamination for different coherence metrics for all our neuronal recordings and these results are shown in Fig. 11. Figure 11*A* shows that the normalized amplitude of the LFP STA (SNR) for various frequency bands, measured around the time of spikes (10 ms before to 10 ms after peak depolarization), drops significantly after the removal of spikes [average of 5 dB (68%) drop for high gamma and 1 dB (20%) decrease for low gamma]. The wideband signal, which contains both high and low frequencies, is similarly affected [average of 7 dB (80%) decrease]. In contrast, the STA at low frequencies is only weakly affected by the spike removal algorithm. A similar pattern was observed for the SFC, as shown in Fig. 11*B*. There is on average a 66% decrease of SFC values for the high gamma and a 30% decrease for the low gamma frequencies.

The effect of spike removal on phase-locking histograms follows a similar pattern across frequencies, as shown in Table 2. The phase-locking metric is dramatically affected by spike remnants, with the proportion of cells classified as phase-locked to high gamma frequencies dropping by half after spike removal.

In contrast to these results, Fig. 11*C* shows the correlation between the LFP and spike tuning curves, both before and after spike removal. For both high and low frequencies, the difference in the correlation coefficient of the two tuning curves is very small (average reduction of 0.06 and 0.03, respectively), indicating that spike contamination has little effect on LFP tuning curves. This conclusion is further supported by the fact that the points are clustered along the unity line, despite wide variation in the extent to which the tuning of the two signals is correlated. These results are consistent with our simulations (Supplemental Figs. S1 and S2).

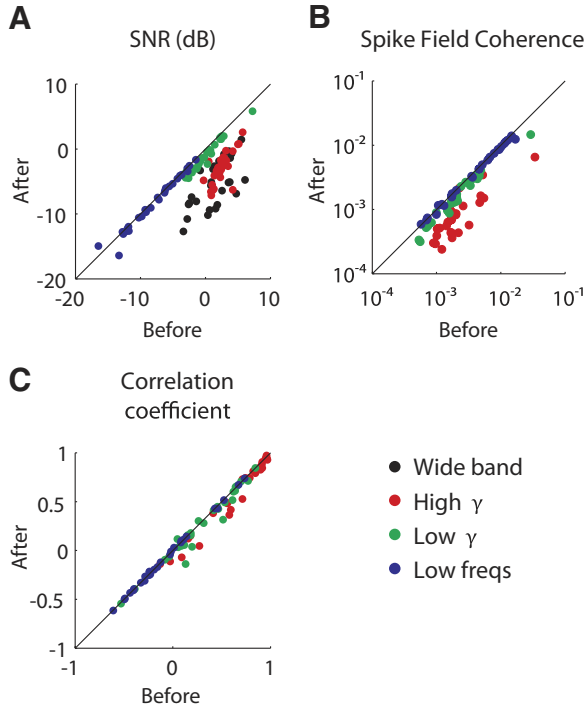


FIG. 11. Summary of analysis results for electrophysiological data. *A*: SNRs. *B*: mean SFC. *C*: correlation coefficients of the direction tuning curves for all real data cases for the wideband signal and 3 different bands of the LFP, before and after the application of the spike removal algorithm. Wideband (black) dots are not included in the SFC and correlation coefficient scatterplot since these measures were computed only for bands of the LFP.

Other approaches of spike removal on real data

For comparison with previous approaches, we also estimated the effects of mean spike removal and interpolation methods on our real data and the results are shown for an example recording in Fig. 12 (same example as that in Figs. 8 and 9). As in the simulated data, mean spike removal leads to larger LFP STAs and more structure in the phase-locking histogram (Fig. 12, *A* and *C*). In this specific example, linear interpolation performs very similarly to the Bayesian spike removal algorithm. However, across the population, both mean spike removal and interpolation yield larger estimates of spike–LFP relationships than our Bayesian algorithm (summary in Table 2). Although ground-truth is not available in the real data, our simulations from the preceding sections suggest that the greater synchronization found with other algorithms is likely to be artifactual.

DISCUSSION

In recent years there has been an increase in the sophistication of analytical methods for interpreting neuronal activity. Although these advances are important for understanding brain circuitry,

TABLE 2. *Percentage of neurons phase-locked to LFP frequencies for various spike removal algorithms*

Percentage of Phase-Locked Neurons	High *y*	Low *y*	Low Frequencies
Low-pass filtered LFP	100.0%	96.4%	67.9%
Mean spike removal	73.1%	84.6%	64.3%
Interpolation method	80.7%	88.5%	65.4%
Bayesian spike removal	50.0%	79.3%	64.3%

Downloaded from jn.physiology.org on January 11, 2018

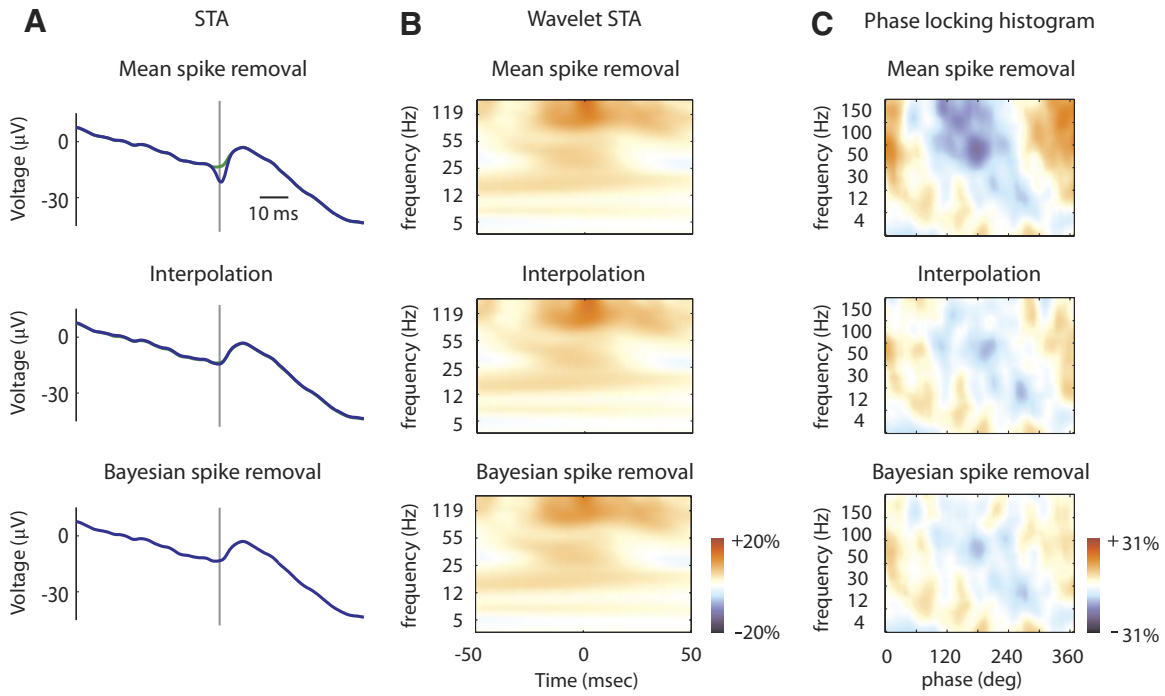


FIG. 12. Comparison of spike removal methods for an illustrative MT neuron. *A*: STAs of the LFP signal. *B*: normalized wavelet STAs. *C*: phase-locking histograms for real data recorded from macaque visual cortex. Results are shown for the mean spike removal method (*top panels*), the interpolation method (*middle panels*), and the Bayesian algorithm (*bottom panels*). For the LFP STAs, the results for the Bayesian algorithm are overlaid for comparison (green lines).

they are also easily influenced by particular assumptions about the way in which the raw signals should be processed. In this study, we have explored one issue related to such signal processing by quantifying the contamination of the LFP from action potential waveforms in both simulated and real data.

Our results show that estimates of the relationship between LFPs and spikes can be considerably affected by spike contamination. This results simply from the spectral characteristics of spike trains and individual spike waveforms: both have power at the frequencies used to study LFPs (Fig. 2*B*). Consequently, spikes and LFPs cannot be isolated from a raw voltage signal with simple frequency-domain filtering. To estimate the extent of this problem we performed simulations in which spike waveforms were added to phase-randomized LFP signals, such that the spectral profile of the LFP was maintained but there was no actual LFP–spike relationship. Nevertheless, when LFPs were estimated by low-pass filtering, our analysis showed artifactual spike–LFP STAs (Fig. 4) and spurious phase-locking of spikes to LFPs (Fig. 5). This effect was particularly strong at high frequencies. In contrast, tuning curves in response to visual stimuli were not strongly affected by spike contamination (Fig. 6), even at high frequencies ($\leq$140 Hz), unless the spike-to-LFP power ratio was unnaturally high ($>$5 dB). Similar results were obtained using real recordings with direction-tuned neurons from the primate visual system (Figs. 8–11). To address these issues we developed a novel Bayesian spike removal algorithm and demonstrated that it is capable of removing artifactual spike–LFP correlations.

Comparison with previous studies

Although the existence and potential importance of spectral contamination of LFPs by spikes has been recognized for some time (Pesaran et al. 2002), the issue has been addressed only rarely in recent studies. One common approach in previous work is to subtract the mean spike waveform from the raw signal at each spike time (Pesaran et al. 2002; Zanos et al. 2006).

The generative model underlying our spike removal algorithm was constructed to generalize this spike subtraction method. In fact, we show in the APPENDIX that our spike removal method reduces to the spike subtraction method when spike waveforms do not overlap, no assumptions are made on the smoothness of the LFP, and an offset term is not included in the model. The full spike removal algorithm thus works similarly to the spike subtraction method, but corrects for spike train autocorrelations and reduces low-frequency components in spike waveforms attributable to LFPs. Each of these adjustments corrects for artifacts introduced by spike subtraction, thus enabling more effective spike removal, as we show in detail in Figs. 7 and 12 and the APPENDIX.

Another proposed method of spike removal is to replace time samples around each spike with samples interpolated from the LFP before and after the spike (Galindo-Leon and Liu 2010; Jacobs et al. 2007; Okun et al. 2010). This approach, although again being superior to low-pass filtering, introduces the difficulty of selecting the method of interpolation, including the function used to interpolate and the length of its support on either side of spikes, without a firm theoretical foundation to justify these choices. Linear interpolation, which is commonly used on wideband signals for this purpose, not only does not completely remove spurious correlations but also introduces new ones, as shown in detail in Fig. 7. The difficulty of selecting a proper method could be alleviated by casting the interpolation as a Bayesian inference problem similar to the one considered here. Within this framework, the functional form of the interpolation would be adjusted by the statistics of

Downloaded from jn.physiology.org on January 11, 2018

the local field potential. A more fundamental issue is that when recording in areas with high transient firing rates, interpolation must sometimes be done over intervals several times longer than a single spike waveform. This is likely to introduce additional artifacts into the estimated LFP.

A recent method (David et al. 2010) uses a more sophisticated linear filtering approach to remove all spike–LFP correlations from the raw signal. In the APPENDIX, we show that our proposed spike removal algorithm reduces to this method in the limit of very long spikes. When very long spikes are assumed, however, the method essentially removes all linear relationships between spikes and LFPs, including legitimate correlations, and thus cannot provide an estimate of the degree to which spikes and LFPs are actually correlated. In contrast, our method focuses on removing the purely artifactual parts of the LFP signal that are introduced by spike waveforms of short duration ($\leq$3 ms), while preserving legitimate correlations that occur at longer timescales. Although the method of David et al. (2010) can be modified to accommodate the assumption of finite spikes, we show in the APPENDIX that this leads to residual artifacts around the time of spikes due to the way that the method handles edges.

Practical considerations for effective spike removal

Our spike removal algorithm assumes that the wideband signal derived from extracellular recordings is the sum of a low-frequency LFP signal, stereotyped spikes of finite length occurring at known times, noise, and an offset (Fig. 1*B*). For the spike removal process to be most effective, it is important that careful preprocessing be done so that the signal closely matches the assumptions of the generative model.

The form of the prior assumed on the LFP signal should be adjusted so that it closely matches the empirical autocorrelation properties of the observed signal; a practical procedure to choose the form of the prior is detailed in the APPENDIX. Line noise should be removed before spike removal because this effect is not dealt with by the generative model (see METHODS). Furthermore, analog filters in the recording hardware can cause frequency-dependent phase shifts in the wideband signal (Nelson et al. 2008), which may artificially lengthen the duration of the observed spike waveforms. This issue can be avoided by compensating for phase distortions by digital filtering (Nelson et al. 2008).

Our spike removal algorithm will remove artifactual linear correlations between spikes and LFPs regardless of whether the spikes in the wideband signal are completely stereotyped. In the latter case, however, residual high-frequency artifacts of random phase will be visible around the time of spikes. These artifacts may be picked up by second-order methods such as the wavelet STA and Volterra methods (Zanos et al. 2006, 2008). It is therefore crucial to perform careful spike sorting before removing spikes. It is also important that spike timings be accurate because spike timings that are inaccurate by as little as one or two samples will cause observed waveforms to shift appreciably (Sahani 1999). Spike timings that are obtained by thresholding a high-passed signal can be inaccurate; more accurate timings can be obtained by aligning spikes to their peak depolarization (Quiroga et al. 2004; Sahani 1999). Additionally, if spike waveforms shift over time, for example because of electrode drift, the algorithm should be applied on short segments of the data to obtain the best results; we show in the APPENDIX how to perform spike removal in chunks.

Many of these prescriptions are equally applicable to other methods of spike removal and, of course, they do not guarantee complete removal of spike artifacts from the wideband signal. Nonetheless, we found empirically that the spike removal algorithm diminished artifactual correlations to a great extent. When an interesting finding relies crucially on the quality of spike removal, it may be worthwhile to back up claims with ground-truth simulations, similar to those of Fig. 1*A*.

Hardware considerations

Several manufacturers offer products designed to record local field potentials (often designated *LFP boards*). These boards typically low-pass filter the wideband signal from a headstage, with a cutoff usually ranging from 150 to 300 Hz, and sample the resulting signal at a typical rate of 1 kHz. Our results indicate that the signals derived from these boards are at all frequencies appropriate for analyses that are insensitive to the relative timing of spikes and LFPs, for example to derive tuning curves (Figs. 6 and 10). However, our results also imply that analyses that are sensitive to temporal relationships between spikes and LFP, including the STA, spike-triggered coherence, phase-locking histograms, and predictions of spikes from LFPs, can be corrupted by the presence of spike remnants in LFPs derived from these boards. This contamination is especially strong at gamma frequencies $>$60 Hz.

It may be possible to work around this contamination, either by limiting analyses to frequencies $<$60 Hz, removing all linear correlation between spikes and LFPs using the methods of David et al. (2010), or by recording spikes and LFPs on different electrodes. However, a safer approach is to acquire the complete wideband signal during electrophysiological recording and to perform careful postprocessing using methods similar to those outlined here. For data that have already been recorded with LFP boards, it may be possible to recover the spike-free LFP by extending our algorithm to include a model of the processing performed by the boards, but this has not yet been attempted.

Conclusions

Numerous studies examining the relationship between spikes and LFPs have reported convincing and significant results. Important examples include studies using spikes and LFPs recorded from different electrodes (Gregoriou et al. 2009; Whittingstall and Logothetis 2009), spikes and cortical EEGs (Whittingstall and Logothetis 2009), or spikes, LFPs, and BOLD signals (Goense and Logothetis 2008) and results like those shown in Figs. 8 and 9 are consistent with some of these studies. However, our results indicate that commonly used methods of preprocessing may lead to overestimation of the association between spikes and LFPs. This is particularly the case when low-pass filtering is used as the only preprocessing step of the LFP, as has often been done in past work (e.g., Chalk et al. 2010; Litaudon et al. 2008; Rasch et al. 2008). The existence and magnitude of spurious correlations in such cases will depend on the nature of the data and the types of analyses applied. Particularly when higher frequencies of the LFP are analyzed with sophisticated nonlinear methods, the effect of even modest contaminations on the results can be significantly magnified (Zanos et al. 2006). On the other hand, successfully removing the spike waveforms from the raw

Downloaded from jn.physiology.org on January 11, 2018

recordings can reveal significant spike–LFP relationships that might otherwise have been masked (as in Fig. 9*B*).

We have introduced a spike removal algorithm that is designed to remove specific spike waveform(s) from the raw voltage signal, without affecting real spike–LFP correlations that are of great interest to the study of neural circuitry. In simulations, where ground-truth is available, our algorithm removed virtually all spike artifacts and the results of all subsequent analyses on the "spike-free" LFP did not show significant correlations between spike timing and any frequency band of the LFP. Similarly, when applied to real data the algorithm often eliminated artifactual spike–LFP relationships that would have emerged from LFPs obtained by low-pass filtering the raw voltage signal (Fig. 9). However, in other cases (Figs. 8 and 9) strong spike–LFP correlations remained after application of the spike removal algorithm and such synchronization will be an important topic for further study. Indeed, with accurate spike removal one can in principle study spike–LFP relationships at higher LFP frequencies (>150 Hz), despite their overlap with the spectral content of the spike waveforms.

ACKNOWLEDGMENTS

We thank Dr. Stavros Zanos for indicating the issue of spike contamination in previous local field potential studies and for providing analysis ideas, R. Cipriani for assistance in data collection, J. Coursol and C. Hunt for assistance with animal care and training, and N. Jabakhanji for help with computer programming.

GRANTS

This work was supported by a Canadian Institutes of Health Research Grant MOP-79352 to C. C. Pack and a Montreal Neurological Institute Centre for Commercialization and Research fellowship to T. P. Zanos.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

REFERENCES

Andersen R, Musallam S, Pesaran B. Selecting the signals for a brain–machine interface. *Curr Opin Neurobiol* 14: 720–726, 2004.

Bair W, Koch C, Newsome W, Britten K. Power spectrum analysis of bursting cells in area MT in the behaving monkey. *J Neurosci* 14: 2870–2892, 1994.

Baker SN, Kilner JM, Pinches EM, Lemon RN. The role of synchrony and oscillations in the motor output. *Exp Brain Res* 128: 109–117, 1999.

Burns SP, Xing D, Shelley MJ, Shapley RM. Searching for autocoherence in the cortical network with a time-frequency analysis of the local field potential. *J Neurosci* 30: 4033–4047, 2010.

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* 66: 114–125, 2010.

David SV, Malaval N, Shamma SA. Decoupling action potential bias from cortical local field potentials. *Comput Intell Neurosci* 2010: 393019, 2010.

Dayan P, Abbot LF. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press, 2001.

Denker M, Roux S, Timme M, Riehle A, Grün S. Phase synchronization between LFP and spiking activity in motor cortex during movement preparation. *Neurocomputing* 70: 2096–2101, 2007.

Destexhe A, Contreras D, Steriade M. Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. *J Neurosci* 19: 4595–4608, 1999.

Fattori P, Pitzalis S, Galletti C. The cortical visual area V6 in macaque and human brains. *J Physiol* (Paris) 103: 88–97, 2009.

Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291: 1560–1563, 2001.

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. *J Neurosci* 28: 4823–4835, 2008.

Galindo-Leon EE, Liu RC. Predicting stimulus-locked single unit spiking from cortical local field potentials. *J Comput Neurosci* 29: 581–597, 2010.

Goense J, Logothetis N. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr Biol* 18: 631–640, 2018.

Gold C, Henze DA, Koch C, Buzsáki G. On the origin of the extracellular action potential waveform: a modeling study. *J Neurophysiol* 95: 3113–3128, 2016.

Goupillaud P, Grossmann A, Morlet J. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration* 23: 85–102, 1984.

Gregoriou GG, Gotts SJ, Zhou H, Desimone R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324: 1207–1210, 2009.

Jacobs J, Kahana MJ, Ekstrom AD, Fried I. Brain oscillations control timing of single-neuron activity in humans. *J Neurosci* 27: 3839–3844, 2007.

Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M. Local origin of field potentials in visual cortex. *Neuron* 61: 35–41, 2009.

Khawaja FA, Tsui JMG, Pack CC. Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. *J Neurosci* 29: 13702–13709, 2009.

Litaudon P, Garcia S, Buonviso N. Strong coupling between pyramidal cell activity and network oscillations in the olfactory cortex. *Neuroscience* 156: 781–787, 2008.

Liu J, Newsome WT. Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *J Neurosci* 26: 7779–7790, 2016.

Maunsell JHR, Newsome WT. Visual processing in monkey extrastriate cortex. *Annu Rev Neurosci* 10: 363–401, 1987.

Nauhaus I, Busse L, Carandini M, Ringach D. Stimulus contrast modulates functional connectivity in visual cortex. *Nat Neurosci* 12: 70–76, 2008.

Nelson M, Pouget P, Nilsen E, Patten C, Schall J. Review of signal distortion through metal microelectrode recording circuits and filters. *J Neurosci Methods* 169: 141–157, 2008.

Okun M, Naim A, Lampl I. The subthreshold relation between cortical local field potential and neuronal firing unveiled by intracellular recordings in awake rats. *J Neurosci* 30: 4440–4448, 2010.

Paz R, Bauer EP, Paré D. Theta synchronizes the activity of medial prefrontal neurons during learning. *Learn Mem* 15: 524–531, 2008.

Pesaran B, Pezaris J, Sahani M, Mitra P, Andersen R. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci* 5: 805–811, 2012.

Quiroga RQ, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput* 16: 1661–1687, 2014.

Rasch MJ, Gretton A, Murayama Y, Maass W, Logothetis NK. Inferring spike trains from local field potentials. *J Neurophysiol* 99: 1461–1476, 2008.

Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS. Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci* 28: 11526–11536, 2008.

Rutishauser U, Ross IB, Mamelak AN, Schuman EM. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* 464: 903–907, 2010.

Sahani M. *Latent Variable Models for Neural Data Analysis*. Pasadena, CA: California Institute of Technology, 1999.

Saleh M, Reimer J, Penn R, Ojakangas CL, Hatsopoulos NG. Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron* 65: 461–471, 2010.

Siapas AG, Lubenov EV, Wilson MA. Prefrontal phase locking to hippocampal theta oscillations. *Neuron* 46: 141–151, 2015.

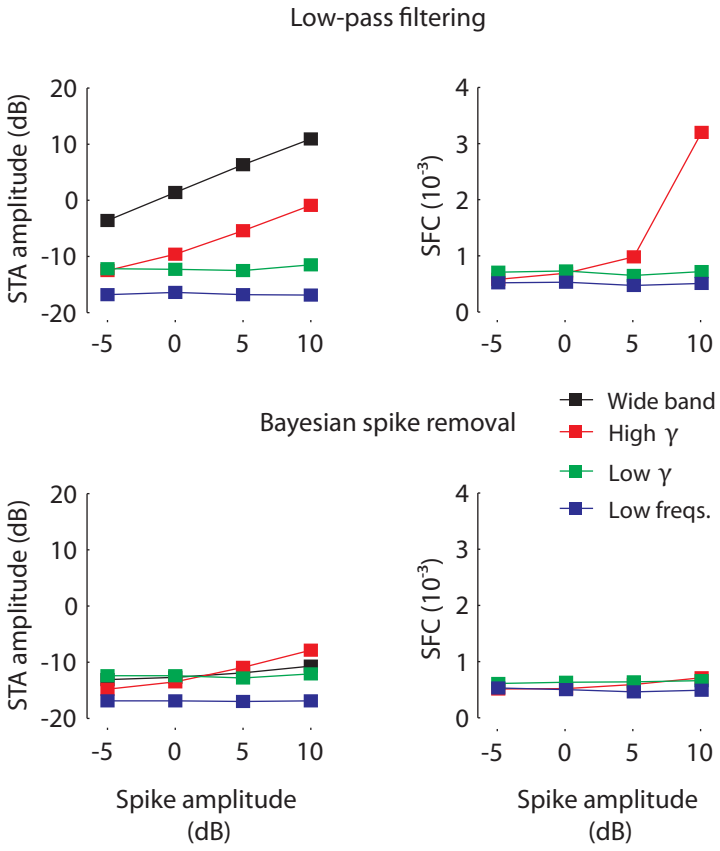
Whittingstall K, Logothetis NK. Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron* 64: 281–289, 2009.

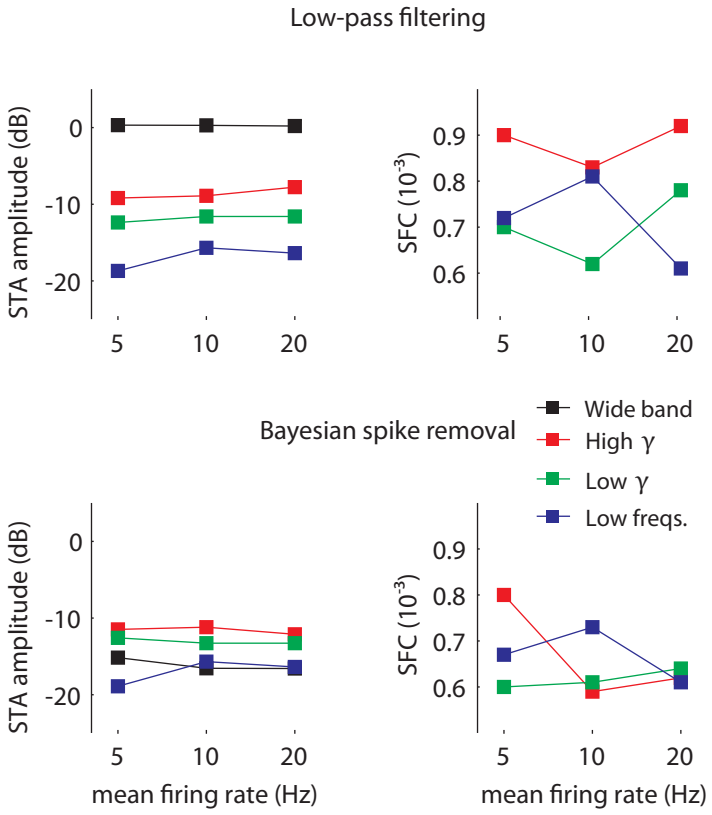
Xing D, Yeh C-I, Shapley RM. Spatial spread of the local field potential and its laminar variation in visual cortex. *J Neurosci* 29: 11540–11549, 2009.

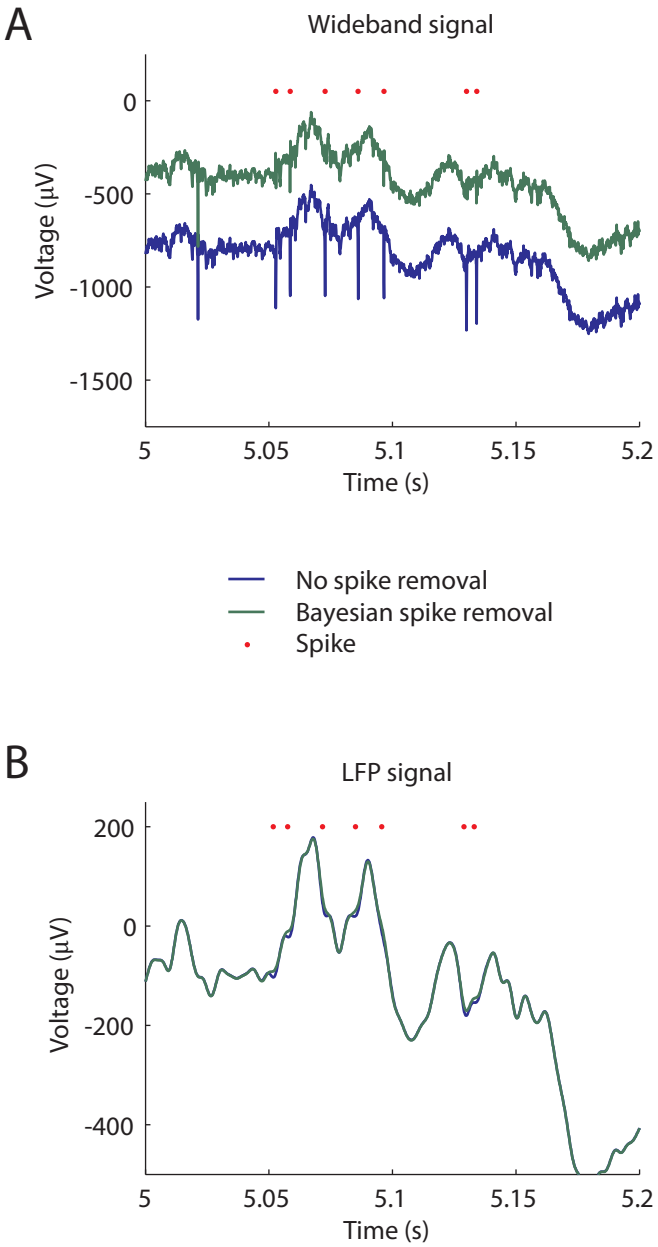
Zanos TP, Courellis SH, Berger TW, Hampson RE, Deadwyler SA, Marmarelis VZ. Nonlinear modeling of causal interrelationships in neuronal ensembles. *IEEE Trans Neural Syst Rehabil Eng* 16: 336–352, 2008.

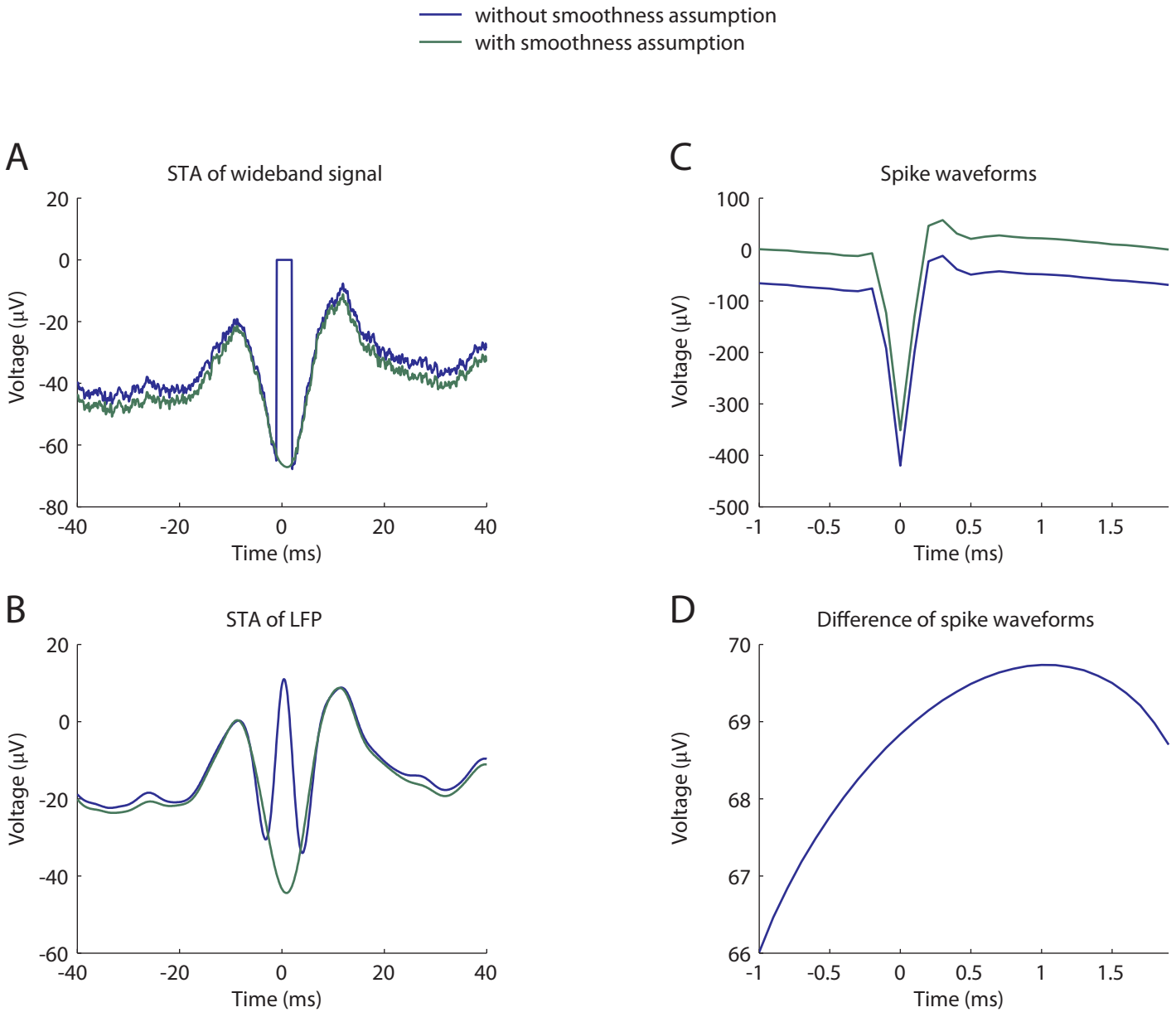
Zanos TP, Zanos SP, Courellis SH, Marmarelis VZ, Ojemann GA. Nonlinear dynamic modeling of the relationship between local field potentials and neural discharge in human temporal cortex. *Soc Neurosci Abstr* 748.18, 2006.

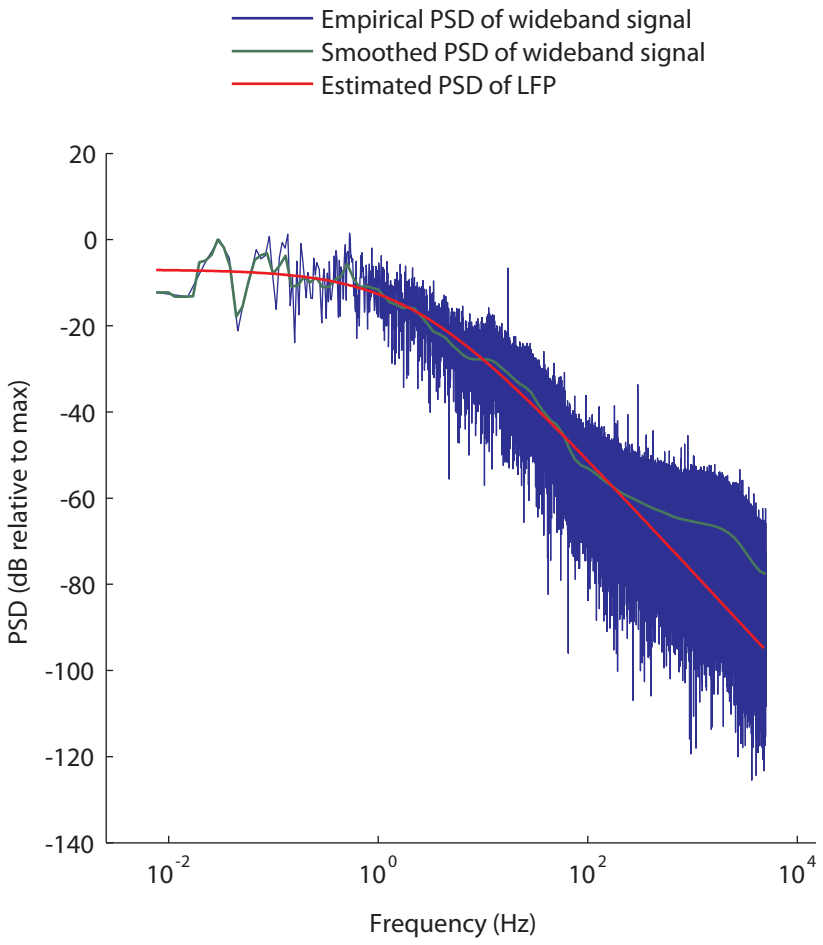
Downloaded from jn.physiology.org on January 11, 2018

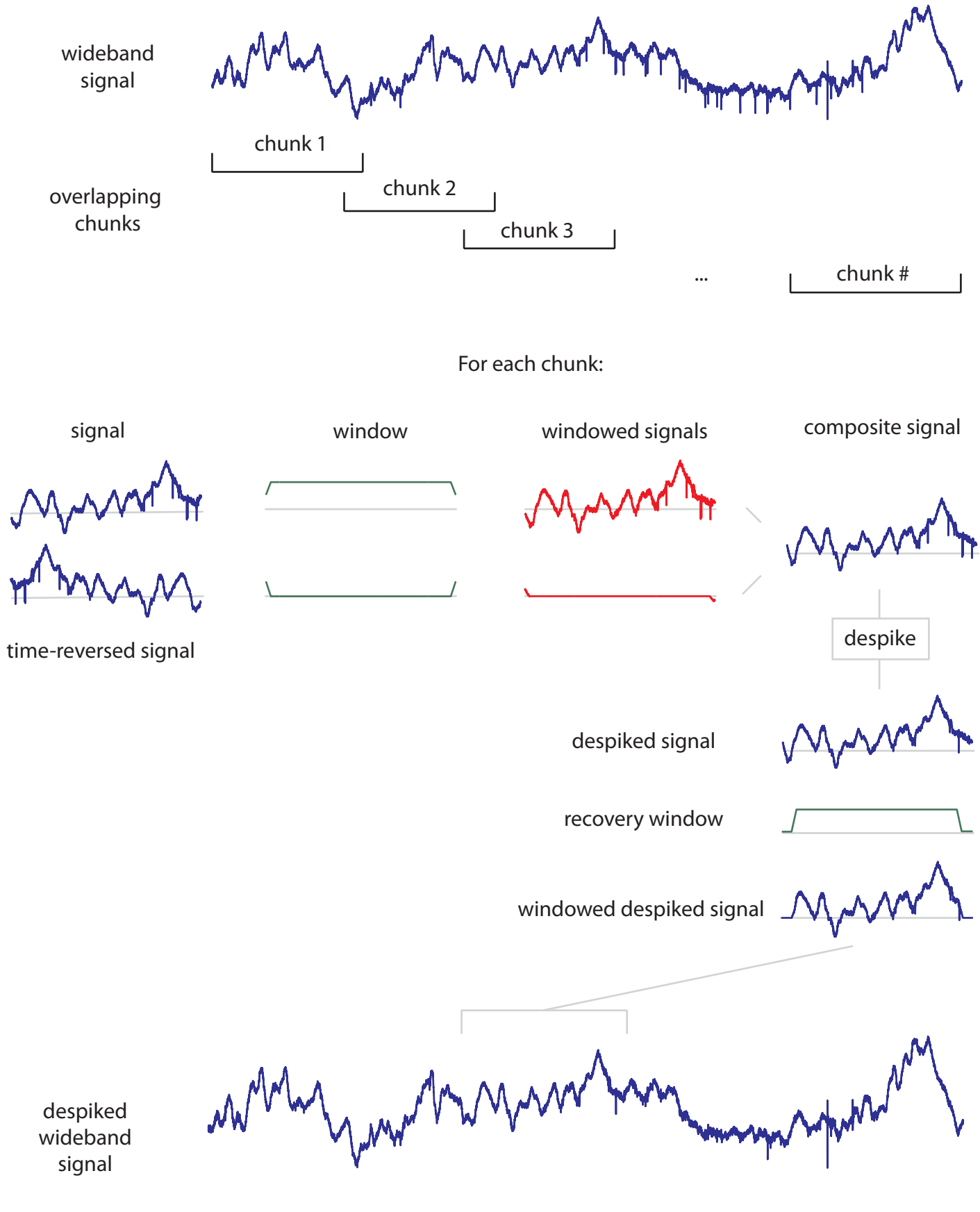

Low-pass filtering
STA amplitude (dB)
20
10
0
-10
-20
-5
0
5
10
SFC (10^{-3})
4
3
2
1
0
Wide band
High γ
Low γ
Low freqs.
Bayesian spike removal
Spike amplitude (dB)
Spike amplitude (dB)


Low-pass filtering
STA amplitude (dB)
0
-10
-20
5
10
20
SFC (10^{-3})
0.9
0.8
0.7
0.6
Bayesian spike removal
Wide band
High γ
Low γ
Low freqs.
mean firing rate (Hz)
mean firing rate (Hz)


A
Wideband signal
0
-500
-1000
-1500
Voltage (μV)
5
5.05
5.1
5.15
5.2
Time (s)
No spike removal
Bayesian spike removal
Spike
B
LFP signal
200
0
-200
-400
Voltage (μV)
5
5.05
5.1
5.15
5.2
Time (s)


without smoothness assumption
with smoothness assumption
A
STA of wideband signal
Voltage (μV)
Time (ms)
B
STA of LFP
Voltage (μV)
Time (ms)
C
Spike waveforms
Voltage (μV)
Time (ms)
D
Difference of spike waveforms
Voltage (μV)
Time (ms)


Empirical PSD of wideband signal
Smoothed PSD of wideband signal
Estimated PSD of LFP
20
0
-20
-40
-60
-80
-100
-120
-140
PSD (dB relative to max)
10^-2
10^0
10^2
10^4
Frequency (Hz)


wideband signal
chunk 1
overlapping chunks
chunk 2
chunk 3
...
chunk #
For each chunk:
signal
window
windowed signals
composite signal
time-reversed signal
despike
despiked signal
recovery window
windowed despiked signal
despiked wideband signal

Contents

Supplementary information 1
Model and parameter estimation 1
Relationship to spike removal by subtraction of the STA 5
Hyperparameter estimation 7
Empirical estimate of g 9
Choise of basis 10
Chunking 10
Relationship to LFP-spike correlation removal 12
Appendix - Properties of circulant matrices 14
List of supplementary figures 16
Supplementary Figure 1: Summary of results for all spike-amplitude simulations 16
Supplementary Figure 2: Summary of results for all firing-rate simulations 16
Supplementary Figure 3: Effect of spike removal in an example recording 16
Supplementary Figure 4: Necessity of smoothness assumption in a simulation 16
Supplementary Figure 5: Choice of $\mathbf{g}$ 16
Supplementary Figure 6: Chunking procedure 16
Supplementary References 17

Supplementary information

This section provides the mathematical details of our spike removal algorithm, its software implementation, and a comparison with other algorithms. All Matlab code is available as online supplementary material and from our website (http://apps.mni.mcgill.ca/research/cpack/lfpcode.zip).

Model and parameter estimation

The model is as specified in the main Methods section; we repeat it here for easy reference, with slight changes in notation. Our goal is to estimate the local field potential (LFP) based on a measured wideband voltage trace of length n. We assume that this wideband signal $\mathbf{y}$ is the superposition of a low-frequency local field potential $\mathbf{w}$, high-frequency spike components $\boldsymbol{\eta}^k$, an offset μ and white noise $\boldsymbol{\epsilon}$ (Figure 1B):

$$\mathbf{y} = \mathbf{w} + \sum_{k=1}^{m} \boldsymbol{\eta}^k + \mu + \boldsymbol{\epsilon} \tag{1}$$

Here m is the number of sorted neurons emitting spikes. The high-frequency component of the k^{th} neuron, $\boldsymbol{\eta}^k$, is created by the convolution of the neuron's spike train $\mathbf{s}^k$, assumed known (see Figure 1B), and the neuron's spike waveform $\mathbf{B}\boldsymbol{\varphi}^k$:

$$\boldsymbol{\eta}^k = C'(\mathbf{s}^k)\mathbf{B}\boldsymbol{\varphi}^k \tag{2}$$

Here the $C'(\mathbf{a})$ returns a circulant matrix whose first row is $\mathbf{a}$. The product $C'(\mathbf{a})\mathbf{b}$ returns the circular convolution of $\mathbf{a}$ and $\mathbf{b}$:

$$C'(\mathbf{a})\mathbf{b} = (a \circledast b)_i = \sum_j a_{[i-j]} b_j \tag{3}$$

Here $a_{[k]} = a_{(k \text{ modulo } n)+1}$. Circulant matrices have a number of properties that are crucial for the tractability of the model parameters; their properties are covered in detail in the last section. We call $\boldsymbol{\eta}^k$ the *waveform train* of the k^{th} neuron.

$\mathbf{B}$ is a matrix of basis functions which map the spike parameters $\boldsymbol{\varphi}^k$ onto a spike waveform. Typically, the number of parameters that describe the spike waveforms is much smaller than the length of the signal, which implies that $\mathbf{B}$ is much taller than it is wide.

Assumptions are as follows:

$$\begin{aligned} p(\mathbf{w}) &= N(0, \gamma^2 \boldsymbol{\Gamma}) \\ p(\boldsymbol{\epsilon}) &= N(0, \sigma^2 \mathbf{I}) \end{aligned} \tag{4}$$

Here $N(\mathbf{a}, \boldsymbol{\Sigma})$ represents a multivariate Gaussian with mean $\mathbf{a}$ and covariance $\boldsymbol{\Sigma}$. $\boldsymbol{\Gamma} = C'(F^{-1}(\mathbf{g}))$ is a matrix that embodies an assumption of smoothness, and $\boldsymbol{\Gamma}\mathbf{x}$ produces a low-pass filtered version of $\mathbf{x}$.

By Bayes' theorem, we have that the posterior probability of the parameters is, up to an additive constant:

$$\begin{aligned} -\log p(\mathbf{w}, \boldsymbol{\varphi}^k, \boldsymbol{\mu} | \mathbf{y}) \\ = \frac{1}{2\sigma^2}\left(\mathbf{y} - \mathbf{w} - \sum_k C'(\mathbf{s}^k)\mathbf{B}\boldsymbol{\varphi}^k - \mu\right)'\left(\mathbf{y} - \mathbf{w} - \sum_k C'(\mathbf{s}^k)\mathbf{B}\boldsymbol{\varphi}^k - \mu\right) \\ + \frac{1}{2\gamma^2}\mathbf{w}'\Gamma^{-1}\mathbf{w} \end{aligned} \tag{5}$$

We can solve for the MAP estimate of the parameters by taking partial derivatives of the posterior and setting derivatives to zero. The MAP estimates are given by:

$$\begin{aligned}\bar{\mathbf{w}} &= (\gamma^2\boldsymbol{\Gamma} + \sigma^2\mathbf{I})^{-1}\gamma^2\boldsymbol{\Gamma}\left(\mathbf{y} - \sum_k C'(\mathbf{s}^k)\mathbf{B}\bar{\boldsymbol{\varphi}}^k - \bar{\mu}\right)\\ \bar{\boldsymbol{\varphi}}^k &= \left(C'(\mathbf{s}^k)\mathbf{B}\right)^+\left(\mathbf{y} - \bar{\mathbf{w}} - \sum_{j\neq k} C'(\mathbf{s}^j)\mathbf{B}\bar{\boldsymbol{\varphi}}^j - \bar{\mu}\right)\\ \bar{\mu} &= \frac{1}{n}\hat{\mathbf{1}}'\left(\mathbf{y} - \bar{\mathbf{w}} - \sum_k C'(\mathbf{s}^k)\mathbf{B}\bar{\boldsymbol{\varphi}}^k\right)\end{aligned} \tag{6}$$

Here $\mathbf{A}^+$ is the pseudoinverse $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$, and $\hat{\mathbf{1}}$ is a vector of ones. It is possible to solve his system by taking initial guesses of the model parameters, recomputing the parameters using these equations in sequence and iterating. Convergence of this iterative solution is however too slow for this to be a practical solution because of the strong couplings[1] between the LFP, the waveforms and the offset.

Instead we isolate each parameter set individually, starting with the waveforms $\bar{\boldsymbol{\varphi}}^k$. To simplify the equations, we isolate $\bar{\boldsymbol{\varphi}}^k$ for a fixed value of k on the assumption that the other waveforms $\bar{\boldsymbol{\varphi}}^{j\neq k}$ are known in advance. We will explicitly compensate for the errors induced by this simplification later. We thus define $\mathbf{v} = \mathbf{y} - \sum_{j\neq k} C'(\mathbf{s}^j)\mathbf{B}\bar{\boldsymbol{\varphi}}^j$ and drop the indices in the equations to lighten the notation:

$$\begin{aligned}\bar{\mathbf{w}} &= \mathbf{M}(\mathbf{v} - \mathbf{D}\bar{\boldsymbol{\varphi}} - \bar{\mu})\\ \bar{\boldsymbol{\varphi}} &= \mathbf{D}^+(\mathbf{v} - \bar{\mathbf{w}} - \bar{\mu})\\ \bar{\mu} &= \frac{1}{n}\hat{\mathbf{1}}'(\mathbf{v} - \bar{\mathbf{w}} - \mathbf{D}\bar{\boldsymbol{\varphi}})\end{aligned} \tag{7}$$

Here $\mathbf{D} \equiv C'(\mathbf{s})\mathbf{B}$ and $\mathbf{M} = (\gamma^2\boldsymbol{\Gamma} + \sigma^2\mathbf{I})^{-1}\gamma^2\boldsymbol{\Gamma}$, a low-pass filter. $\mathbf{D}'\mathbf{x}$ computes the spike-triggered sum of the signal $\mathbf{x}$ while $\mathbf{D}\mathbf{x}$ computes the waveform train associated with the spike waveform $\mathbf{x}$. By substituting $\bar{\mu}$ into the first two equations, then isolating $\bar{\mathbf{w}}$ in the first equation and finally substituting it in the second equation, we obtain the normal equation:

$$\mathbf{D}'\mathbf{J}(\mathbf{I} - \mathbf{M})\mathbf{D}\bar{\boldsymbol{\varphi}} = \mathbf{D}'\mathbf{J}(\mathbf{I} - \mathbf{M})\mathbf{v} \tag{8}$$

Here $\mathbf{J} = \mathbf{I} - \frac{1}{n}\hat{\mathbf{1}}\hat{\mathbf{1}}'$ is a centering matrix which when applied to a vector yields the same vector but with its mean set to 0. This equation can be interpreted as follows: the STA ($\mathbf{D}'$) of the centered ($\mathbf{J}$) high-pass

[1] By strong coupling between parameters we mean that the terms of the Hessian of the log-posterior involving parameters of different types (LFP/waveforms/offset) are not negligible compared to terms of the Hessian involving parameters of similar type.

filtered ($\mathbf{I}-\mathbf{M}$) waveform train ($\mathbf{D}\bar{\boldsymbol{\varphi}}$) is equal to the STA of the centered high-pass filtered wideband signal.

While this equation is satisfyingly compact and intuitive, multiplication by $\mathbf{M}$ must be done through multiplication in the Fourier domain, and hence a Fourier and an inverse Fourier transform need to be performed for every column of $\mathbf{D}$; this operation is thus a bottleneck. Since spikes are assumed to be of finite duration, however, it follows that we can write $\mathbf{B} = [\mathbf{B}_s; \mathbf{0}]$, where $\mathbf{0}$ is a matrix of zeros of size $(n - q)$ by p, $q \ll n$ is the length of a spike and $p \leq q$ is the number of basis functions. Note that this requires time shifting the spike train $\mathbf{s}$ so that an entry equal to 1 indicates the beginning of a spike rather than its peak. Substituting and simplifying (8), we find:

$$\mathbf{D}'\mathbf{J}(\mathbf{I}-\mathbf{M})\mathbf{D} = \mathbf{B}_s'(\mathbf{k} * \mathbf{B}_s) - \frac{1}{\mathrm{n}}\left(\mathbf{s}'\hat{\mathbf{1}}\right)^2 \left(F^{-1}\left(\frac{\sigma^2}{\sigma^2+\gamma^2\mathbf{g}}\right)' \hat{\mathbf{1}}\right)\left(\mathbf{B}_s'\hat{\mathbf{1}}\right)\left(\mathbf{B}_s'\hat{\mathbf{1}}\right)' \tag{9}$$

Here $*$ denotes linear convolution with zero-padding, applied column-wise, and $\mathbf{k}$, the convolution kernel, is derived from $\mathbf{a} = \mathrm{F}^{-1}\left(\frac{\sigma^2}{\sigma^2+\gamma^2\mathbf{g}}|\mathrm{F}(\mathbf{s})|^2\right)$ by circular time shifting, with $k_i = a_{[i-q]}$. Since q is small, the convolution in (9) is inexpensive and the equation can be solved by preconditioned conjugate gradients at negligible cost. After solving equation (9) sequentially for every neuron, we plug these estimates in the system of equations (6) and solve to find:

$$\bar{\mu} = \frac{1}{n}\hat{\mathbf{1}}'\left(\mathbf{y} - \sum_k C'\left(\mathbf{s}^k\right)\mathbf{B}\bar{\boldsymbol{\varphi}}^k\right) \tag{10}$$

Thus $\bar{\mu}$ is equal to the mean of the wideband signal minus the mean of the spike contribution. Finally, the LFP $\bar{\mathbf{w}}$ is given by the first equation in the system (6); it is the low-pass filtered wideband signal minus the spike contribution. In practice, the experimenter will probably want to use his or her own filter or filterbank on the despiked wideband signal to obtain the LFP. Hence $\bar{\mathbf{w}}$ **is** never actually computed by the despiking algorithm; instead the algorithm works with and returns the despiked wideband signal, defined as:

$$\bar{\mathbf{z}} = \mathbf{y} - \sum_k C'\left(\mathbf{s}^k\right)\mathbf{B}\bar{\boldsymbol{\varphi}}^k - \bar{\mu} \tag{11}$$

Because we ignored cross terms when solving for the spike waveforms $\bar{\boldsymbol{\varphi}}^k$, and also because of possible numerical instability, solutions must be checked for convergence. It can be verified using equation (6) that for every k, the following auxiliary equation holds:

$$\mathbf{B}_s[\mathbf{B}_s; \mathbf{0}]' C(\mathbf{s}^k)\mathbf{J}(\mathbf{I}-\mathbf{M})\left(\mathbf{y} - \sum_j C'\left(\mathbf{s}^k\right)\mathbf{B}\bar{\boldsymbol{\varphi}}^j - \mu\right) = \mathbf{0} \tag{12}$$

This states that the STA of a high-pass filtered, centered despiked wideband signal projected onto the basis in which we express spike waveforms is 0. Convergence is attained when the largest deviation from 0 observed is smaller than some fraction of the standard deviation of $\mathbf{y}$. When convergence is not attained, the process is repeated. In practice, it rarely takes more than 2 iterations to reach convergence, even when there are several neurons emitting spikes.

Supplementary Figure S3A shows an example of a wideband signal before and after spike removal. The two signals have been offset vertically to facilitate comparison. By construction, the proposed method removes only the mean spike waveform around the time of every spike, and hence does not remove all traces of spikes when the spike waveform is variable. It also relies on the information given by the experimenter about the timing of spikes, and does not remove spikes which have not been detected. These limitations can be overcome in large part with good spike detection, alignment, and sorting. In these cases the method performs admirably, as seen in this figure, and results in an appreciable change in the wideband signal. The changes are much less conspicuous when looking at the LFP, shown in Figure S3B. The difference between the two signals consists of a barely visible artifact around the time of every spike. Although the artifact is very small for each spike, it is highly stereotyped: it always occurs at the same time relative to spikes, and always has the same shape, sign and relative phase. Thus, any technique that looks at temporal relationships between spikes and LFPs will amplify the artifact, masking legitimate spike-LFP relationships, as we demonstrate in the main text.

Relationship to spike removal by subtraction of the STA

The algorithm is conceptually similar to previously proposed methods of spike removal, in particular methods that subtract a mean spike determined by spike-triggered averaging the signal at the time of each spike (Pesaran et al. 2002; Zanos et al. 2006). Here we make this connection more explicit and highlight the differences between our proposed method and spike removal by STA subtraction.

Consider the case where the smallest interspike interval observed is equal to or longer than the duration of a spike, and the basis in which the waveform is expressed is given by $\mathbf{B} = [\mathbf{I}_q; \mathbf{0}]$; here $\mathbf{I}_q$ is the identity matrix of size q by q where q is the length of a spike. Furthermore, assume that no information is known about the LFP, such that $\gamma^2 \to 0$. Finally, assume that there is only one neuron spiking. Under these simplifying assumptions, the left hand side of equation (9) becomes:

$$\begin{aligned}\mathbf{D}'\mathbf{J}(\mathbf{I}-\mathbf{M})\mathbf{D}\bar{\boldsymbol{\varphi}} &= \mathbf{B}_s'(\mathbf{k} * \mathbf{B}_s\bar{\boldsymbol{\varphi}}) - \frac{1}{\mathrm{n}}\left(\mathbf{s}'\hat{\mathbf{1}}\right)^2\left(F^{-1}\left(\frac{\sigma^2}{\sigma^2+\gamma^2\mathbf{g}}\right)'\hat{\mathbf{1}}\right)\left(\mathbf{B}_s'\hat{\mathbf{1}}\right)\left(\mathbf{B}_s'\hat{\mathbf{1}}\right)'\bar{\boldsymbol{\varphi}} \\ &= r^2\left(\mathbf{I}-\frac{1}{n}\hat{\mathbf{1}}_{\mathbf{q}}{}'\right)\bar{\boldsymbol{\varphi}}\end{aligned} \tag{13}$$

Here r is the total number of spikes. The right-hand side becomes:

$$\begin{aligned}\mathbf{D}'\mathbf{J}(\mathbf{I}-\mathbf{M})\mathbf{y} &= r\left[\mathbf{I}_q, \mathbf{0}\right]\left(\mathbf{I}-\frac{1}{n}\hat{\mathbf{1}}'\right)\mathbf{y} \\ &= r^2\ \mathrm{STA}\left(\mathbf{y}-\mathrm{mean}(\mathbf{y})\right)\end{aligned} \tag{14}$$

Solving for $\bar{\boldsymbol{\varphi}}$ by inverting $\left(\mathbf{I} - \frac{1}{n}\widehat{\mathbf{1}}_{\mathbf{q}}{}'\right)$ using the Woodbury matrix identity, we find:

$$\begin{aligned}\bar{\boldsymbol{\varphi}} &= \mathrm{STA}\big(\mathbf{y} - \mathrm{mean}(\mathbf{y})\big) + \frac{rq}{n - rq}\mathrm{mean}\left(\mathrm{STA}\big(\mathbf{y} - \mathrm{mean}(\mathbf{y})\big)\right) \\ &= \mathrm{STA}\big(\mathbf{y} - \mathrm{mean}(\mathbf{y})\big) + c\end{aligned} \tag{15}$$

Under these assumptions, the spike waveform $\bar{\boldsymbol{\varphi}}$ is equal to the spike-triggered average of the centered wideband signal plus an offset c. Hence our proposed algorithm reduces to earlier proposals of removing spikes by subtracting the mean spike around the time of each spike in the limit of non-overlapping spikes and no assumptions on the smoothness of the LFP, save for the appearance of an offset c.

The appearance of this offset is puzzling, especially its sign which is the opposite of what intuition might dictate. This can be explained by the fact that the mean of the wideband signal will rise after spike removal, assuming spikes are negative. But the STA around the time of the spike will be zero after spike removal, which is inconsistent with this positive mean, and this will cause an artifactual downward dip in the STA around the time of a spike. The offset corrects this artifact.

Now that the relationship between our method and straightforward spike removal is clear, we can ask to what extent the unique features of our method are useful. The most important unique feature of our method is the assumption that the LFP is smooth. This assumption is crucial to obtain a correct estimate of the STA of the despiked signal when there is a real relationship between spikes and the LFP. This is illustrated in Figure S4 where we have created a surrogate signal composed of spikes and an LFP composed of low-frequency noise and a Gabor of smaller magnitude and longer time scale than a spike at every spike. Thus, there exists a true relationship between the spikes and the LFP. The STA of the despiked wideband signal when no assumptions are made has a prominent artifact around the time of a spike: it is flat and equal to 0. When an assumption of smoothness is made, however, the method correctly interpolates the signal around the time of the spike (Figure S4A). Figure S4B shows that the STA of the LFP obtained by smoothing the despiked signal without assumptions has an artifactual upward deflection around the time of the spike, while the STA under a smoothness assumption is correctly estimated. The proposed method accomplishes this better behavior around spikes by adjusting the spike waveform, as shown in Figure S4C. The difference between the two waveforms, illustrated in Figure S4D, appears to consist of a DC, linear, and quadratic trend. This shows that our method adjusts spikes waveforms by attenuating low-frequency trends which it attributes to the LFP.

It is important to understand that while the proposed method does noticeably better than previously proposed methods, it has its limitations. Our method works best when spikes are assumed to be short. When longer spike times are assumed (say, larger than 4 ms), the method works less well because in essence it must interpolate over a gap of several dozens of samples to determine what is part of the LFP and what is part of the spike. Paradoxically, therefore, the spike removal algorithm will be less aggressive in its spike removal when spikes are assumed to be too large. We therefore recommend

assuming spike lengths that are 3 ms or less; see the Choice of Basis section for another discussion of the choice of spike length.

Hyperparameter estimation

Recall that our model assumptions are that:

$$p(\mathbf{w}) = N(0, \gamma^2\mathbf{\Gamma});\ \mathbf{\Gamma} = C'(F^{-1}(\mathbf{g}))$$
$$p(\boldsymbol{\epsilon}) = N(0, \sigma^2\mathbf{I}) \tag{16}$$

Up to now, we have assumed that σ and γ are known. The strength of the prior relative to the noise is important as it determines what the model considers as signal and what it discounts as noise. These hyperparameters, which control the regularization of the model, cannot be determined by MAP, unlike regular parameters (Wu et al. 2006). Here we determine these parameters by optimizing the marginal likelihood of the model, a metric which takes into account both the quality of the model fit to the data and the number of degrees of freedom in the model to determine the optimal degree of regularization. This method is also known as evidence optimization (Bishop 2007; see chapter 3 for a detailed introduction to this subject). We begin by ignoring the uncertainty in the model due to the parameters of the spike. $\mathbf{z}$ is defined as before as the despiked wideband signal $\mathbf{z} = \mathbf{y} - \sum_k C'(\mathbf{s}^k)\mathbf{B}\bar{\boldsymbol{\varphi}}^k - \bar{\mu}$. $\bar{\boldsymbol{\varphi}}^k$ and $\bar{\mu}$ are assumed to have both been estimated according to the methods of the Model and Parameter Estimation section. The marginal likelihood of the model is defined as:

$$p(\mathbf{z}|\sigma, \gamma, \mathbf{g}) = \int p(\mathbf{z}|\mathbf{w}, \sigma)p(\mathbf{w}|\gamma, \mathbf{g})d\mathbf{w} \tag{17}$$

The marginal likelihood is thus the likelihood of the data (and therefore the model) with the uncertainty in the model parameters $\mathbf{w}$ marginalized out by integration. Unlike in MAP estimation, normalization constants that ensure that probabilities integrate to 1 are of crucial importance and thus are not ignored. The marginal likelihood is then:

$$p(\mathbf{z}|\sigma, \gamma, \mathbf{g}) = \frac{1}{\sqrt{2\pi|\sigma^2|}}\frac{1}{\sqrt{2\pi|\gamma^2\mathbf{\Gamma}|}}\int \exp\left(-\frac{1}{2\sigma^2}(\mathbf{z}-\mathbf{w})'(\mathbf{z}-\mathbf{w}) - \frac{1}{2\gamma^2}\mathbf{w}'\mathbf{\Gamma}^{-1}\mathbf{w}\right)d\mathbf{w} \tag{18}$$

Here $|\mathbf{M}|$ is the determinant of the matrix $\mathbf{M}$. The integral is performed by completing the square inside the exponential. Taking the negative log of the integral and grouping terms which are independent of the hyperparameters into a constant term k (compare eq. 3.86 in Bishop 2007), we find:

$$\begin{aligned} -\log p(\mathbf{z}|\sigma, \gamma, \mathbf{g}) \\ &= k + n\log\sigma + n\log\gamma + \frac{1}{2}\log|\mathbf{I}| + \frac{1}{2}\log|\mathbf{\Gamma}| + \frac{1}{2}\log|\sigma^{-2}\mathbf{I}^{-1} + \gamma^{-2}\mathbf{\Gamma}^{-1}| \\ &+ \frac{1}{2\sigma^2}(\mathbf{z}-\bar{\mathbf{w}})'(\mathbf{z}-\bar{\mathbf{w}}) + \frac{1}{2\gamma^2}\bar{\mathbf{w}}'\mathbf{\Gamma}^{-1}\bar{\mathbf{w}} \\ &= k + \frac{1}{2}\log|\sigma^2\mathbf{I} + \gamma^2\mathbf{\Gamma}| + \frac{1}{2\sigma^2}(\mathbf{z}-\bar{\mathbf{w}})'(\mathbf{z}-\bar{\mathbf{w}}) + \frac{1}{2\gamma^2}\bar{\mathbf{w}}'\mathbf{\Gamma}^{-1}\bar{\mathbf{w}} \end{aligned} \tag{19}$$

The leading term measures the model complexity, while the trailing terms measure the misfit of the model to the data; the optimal hyperparameters strike the best balance between model fit and complexity by minimizing their sum. To find optimal hyperparameters, the negative log marginal likelihood is minimized numerically.

Here circulant matrices are particularly useful in two ways. First, the normally problematic log determinant appearing in the marginal likelihood has the special form $\log|\sigma^2\mathbf{I} + \gamma^2\mathbf{\Gamma}| = \sum_i \log(\sigma^2 + \gamma^2 g_i)$ (see Circulant Matrices section for derivation), and is thus inexpensive to compute. Secondly, in the usual approach to evidence optimization (Bishop 2007), we find a MAP solution based on fixed hyperparameters and then determine optimal hyperparameters based on a fixed MAP solution, iterating until convergence. This iterative approach is taken because computing a MAP solution is usually expensive.

In contrast, the error term $\frac{1}{2\sigma^2}(\mathbf{z}-\bar{\mathbf{w}})'(\mathbf{z}-\bar{\mathbf{w}}) + \frac{1}{2\gamma^2}\bar{\mathbf{w}}'\mathbf{\Gamma}^{-1}\bar{\mathbf{w}}$ can be computed entirely in the Fourier domain as the discrete Fourier transform is an orthogonal transform, and thus preserves inner products up to a constant: $\mathbf{a}'\mathbf{b} = \frac{1}{n}F(\mathbf{a})'F(\mathbf{b})$. Remarkably, during evidence optimization, we do not need to perform any forward or inverse Fourier transforms. We found that this non-iterative approach, enabled by the choice of circulant matrices, was more computationally efficient than the usual iterative solutions by almost an order of magnitude.

We do, however, neglect the derivatives of $\mathbf{z}$ with respect to the hyperparameters (recall that $\bar{\boldsymbol{\varphi}}^k$ is a function of the hyperparameters), and we need to compensate for this fact. The complete algorithm is as follows:

```
Find optimal hyperparameters based on z = y
While convergence of hyperparameters and evidence is not reached
        While convergence of auxiliary equation is not reached
                Estimate each φ̄^k, μ̄
        End while
        Set z = y − Σ_k C(s^k)Bφ̄^k − μ̄
        Recompute hyperparameters based on z
End while
Return z
```

Convergence of hyperparameters is then usually reached in less than 5 iterations. This algorithm is implemented in the Matlab function `despikeLFP`. Our implementation is highly optimized and fast enough to be of practical use in day-to-day research. For instance, a wideband signal lasting about 3 minutes can be despiked in about 20 seconds on a medium-powered computer running 32-bit Matlab or in about 7 seconds on a high-powered computer running 64-bit Matlab on the Intel Core i7 platform. Computational times rise as $n \log n$, where n is the length of the wideband signal, because of the use of FFTs by the algorithm. The method can scale to arbitrarily long recordings by performing the despiking on short segments of data, an approach we detail in the Chunking section.

Empirical estimate of g

We have now shown how to optimize σ and γ using evidence optimization. There remains a vector of free hyperparameters $\mathbf{g}$ which controls our assumptions on the frequency content of the LFP. Choosing this vector properly is crucial, since in essence the assumed frequency content of the LFP is the only means through which the model can discriminate which portion of the STA around the time of a spike is artifactual and which portion is due to legitimate spike-LFP correlations (see Relationship to spike removal by STA subtraction section and Figure S4 for more on this).

If $p(\mathbf{w}) = N(\mathbf{w}|0, \gamma^2\mathbf{\Gamma})$, then the covariance of $\mathbf{w}$ is $\gamma^2\mathbf{\Gamma}_{\mathrm{ij}} = E_{\mathbf{w}}(w_i w_j)$, where $E_{\mathbf{w}}$ denotes the expected value over all $\mathbf{w}$. But because we constrain the prior matrix to be circulant, the covariance of $\mathbf{w}$ is completely described by its autocovariance, $\gamma^2\mathbf{\Gamma}_{\mathrm{ij}} = E_{\mathbf{w}}\left(E_k\left(w_k w_{[k+i-j]}\right)\right)$. The Fourier transform of this autocovariance is the expected power spectral density (PSD) of $\mathbf{w}$, and thus we have:

$$E_{\mathbf{w}}(|F(\mathbf{w})|^2) \propto \mathbf{g} \tag{20}$$

Thus, $\mathbf{g}$ should be matched to the expected PSD of the LFP. Here we have two difficulties. First, we never actually observe the LFP, only the wideband signal. Second, we typically observe only a handful of such wideband signals, thus even if $E_{\mathbf{w}}(|F(\mathbf{w})|^2) = E_{\mathbf{y}}(|F(\mathbf{y})|^2)$ because there are no spikes or noise in our recording, the mean empirical PSD of a handful of wideband signals is very noisy.

We resolved these issues by using our knowledge of the properties of the LFP and the wideband signal. We know that in a certain range of frequencies where the LFP has most of its power, say 1-150 Hz, it account for most of the power in the PSD of the wideband signal and therefore $E_{\mathbf{w}}(|F(\mathbf{w})|^2) \approx E_{\mathbf{y}}(|F(\mathbf{y})|^2)$ in this frequency range. We therefore fit a function to the PSD of the wideband signal in the range of 1 to 150 Hz and extrapolated this function at lower and higher frequencies to obtain $\mathbf{g}$. Extrapolation with highly nonlinear functions is unadvisable, so we used functions which were constant at the lowest frequencies and linear in log-log space at higher frequencies, consistent with previous reports (Bedard and Destexhe 2009).

We found that the function $-\exp(1 + \log x)$, which is constant for small x and decreases linearly for large x to provide an excellent fit to the PSD of the wideband signal within the range of 1 to 150 Hz. A procedure for fitting this function to a PSD is implemented in the Matlab function `fitLFPpowerSpectrum`. An example of such a fit is shown in Figure S5. Note that we purposefully set $\mathbf{g}$ to be lower than the empirical PSD of the wideband signal at the highest frequencies, as we know that most of the power at these frequencies is actually due to spikes. Some recordings may require a different function to be fit to the PSD of the wideband signal, for example when there is a peak in the PSD in a range of frequencies. Such a peak could happen for a variety of reasons, for example because of a low-pass filter in the recording system whose cutoff overlaps the PSD of the LFP or because of intrinsic properties of the recorded brain region. In that case, a sum of the function $-\exp(1 + \log x)$ and a logistic function could be used to fit the PSD of the LFP.

We obtained excellent results with this method of choosing $\mathbf{g}$ (see Figure S4 for an example). Other functional forms that closely followed the envelope of the PSD of the wideband signal also performed well. Choosing a $\mathbf{g}$ that was not adjusted to the statistics of the signal yielded less satisfactory results. For example, a choice of $\mathbf{g} = 1$ for frequencies smaller than a cutoff of 200 Hz and a vanishingly small value elsewhere performed poorly. We therefore highly recommend that $\mathbf{g}$ be selected on the basis of empirical PSD of the wideband signal.

Choise of basis

A final implicit set of hyperparameters is the basis $\mathbf{B}$. Our algorithm assumes that this basis has the form $[\mathbf{B}_s; \mathbf{0}]$, which as we showed earlier is an appropriate form when we assume that spikes are finite. The height of $\mathbf{B}_s$ corresponds to the duration of spikes measured in samples. As spike-sorting algorithms traditionally use snippets ranging in duration from about 1.5 to 3 ms, it is safe to assume that 3 ms is an upper bound for the duration of spike waveforms. Note that this duration does not correspond to the physical duration of a spike, which is shorter, but rather to the duration of the measured spike waveform, which is affected by the filters of the recording system. We chose the spike length to be equal to 3 ms (30 samples), and aligned spikes so that their peak was located at the 11th sample.

Thus, the basis $\mathbf{B}_s$ was taken to be the identity matrix of size 30x30. As we recorded in areas where neurons fire at high rates, and our sampling rate was relatively low, the spike waveforms $\boldsymbol{\varphi}^k$ were well constrained in this basis. When recording at higher sampling rates, or in areas with low firing rates, however, spike waveforms may be poorly constrained. In this case, $\mathbf{B}_s$ can be chosen to be undercomplete, thus parametrizing spike waveforms in a low-dimensional subspace. For example, we could express the waveforms in a spline basis with a higher density of knots around the time of the peak of spikes than elsewhere.

We must note, however, that this method has its limitations. Our algorithm is not well adapted to short recordings that contain a handful of spikes (say, less than 100), as it needs a sufficient amount of data to constrain the spike waveforms. When despiking trial data, therefore, one should perform the despiking on a continuous wideband signal, splitting the data into smaller chunks for trial analysis afterwards, if necessary.

Our implementation of the despiking algorithm automatically multiplies the chosen basis by a whitening matrix $\mathbf{W}$ obtained through a singular value decomposition, so that the basis internally by the algorithm is orthogonal, $\mathbf{B}_s'\mathbf{W}'\mathbf{W}\mathbf{B}_s = \mathbf{I}$. This tends to improve numerical conditioning appreciably. The implementation then expresses the spike waveforms $\boldsymbol{\varphi}^k$ in the original basis, so this is completely transparent to the end-user.

Chunking

When the signal is too long, it becomes inconvenient to perform the matrix operations required to estimate the local field potential. In addition, in long recordings electrode drift can cause spike waveforms to shift. We addressed these issues by splitting the signal into overlapping chunks, estimating the local field potential for each chunk, then stitching the results back to obtain the complete despiked signal.

The chunking scheme is illustrated graphically in Figure S6. Here we show a signal which is much shorter than one would use in reality for ease of visualization. The signal is split into overlapping chunks. Within each chunk, the signal is multiplied by an analysis window and added to a time reversed version of itself multiplied by 1 minus this window, thus creating a composite signal. The analysis window has a trapezoid shape:

$$f(i) = \begin{cases} \frac{1}{2} + \frac{1}{2}\frac{i}{N \cdot \text{overlap}} & i < N \cdot \text{overlap} \\ 1 & N \cdot \text{overlap} < i < N - N \cdot \text{overlap} \\ 1 - \frac{1}{2}\frac{i - (N - N \cdot \text{overlap})}{N \cdot \text{overlap}} & i > N - N \cdot \text{overlap} \end{cases} \quad (21)$$

Here N is the length of a segment and `overlap` is a variable that controls the degree of overlap between chunks. The edges are thus blended together to avoid discontinuities. The composite signal within each chunk is then despiked as in the previous sections. The despiked signals are put back together by multiplying each signal by a synthesis window and summing the windowed signals together. The synthesis window also has a trapezoid shape:

$$h(i) = \begin{cases} 0 & i < 2N \cdot \text{overlap} \\ \frac{(i - 2N \cdot \text{overlap})}{N \cdot \text{overlap}} & 2N \cdot \text{overlap} < i < 3N \cdot \text{overlap} \\ 1 & 3N \cdot \text{overlap} < i < N - 3N \cdot \text{overlap} \\ 1 - \frac{i - (N - 3N \cdot \text{overlap})}{N \cdot \text{overlap}} & N - 3N \cdot \text{overlap} < i < N - 2N \cdot \text{overlap} \\ 0 & i > N - 2N \cdot \text{overlap} \end{cases} \quad (22)$$

The support of the synthesis window is smaller than the size of the analysis window, thus discarding the edges within each chunk. For the initial segment the analysis and synthesis windows are of a different shape to avoid artifacts at the beginning of the recording signal. The analysis window is given by:

$$f_{\text{first}}(i) = \begin{cases} 1 & i < N - 4N \cdot \text{overlap} \\ 1 - \frac{1}{2}\frac{i - (N - 4N \cdot \text{overlap})}{N \cdot \text{overlap}} & N - 4N \cdot \text{overlap} < i < N - 2N \cdot \text{overlap} \\ 0 & i > N - 2N \cdot \text{overlap} \end{cases} \quad (23)$$

And the synthesis window is:

$$h_{\text{first}}(i) = \begin{cases} 1 & i < N - 6N \cdot \text{overlap} \\ \frac{i - (N - 6N \cdot \text{overlap})}{N \cdot \text{overlap}} & N - 6N \cdot \text{overlap} < i < N - 5N \cdot \text{overlap} \\ 0 & 0 \end{cases} \quad (24)$$

The analysis and synthesis windows for the last chunk are mirror inverses of those of the first chunk.

The number of chunks is determined by N, which is given by the user implicitly through the $\mathbf{g}$ parameter, and the variable `overlap`, which is given explicitly. Because the number of chunks must be an integer, however, the method automatically adjusts the overlap upwards so that chunks are equispaced and that recovery windows add up to 1 everywhere. The method is implemented in Matlab as the function `despikeLFPbyChunks`.

In simulations where ground truth was available, despiking by chunks gave essentially the same results as despiking an entire signal provided that chunks were large enough to obtain reliable estimates of spike waveforms. The chunk size should be a power of 2 for the fastest speeds as FFT routines are typically optimized for such cases (Frigo and Johnson 2005). We recommend using chunks that are a few minutes long (say 2-5 minutes) so that several hundred spikes will be present in each chunk.

Relationship to LFP-spike correlation removal

Our algorithm is closely related to the algorithm of David et al. 2010 which removes correlations between LFPs and spikes. To make this relationship clearer, we rewrite equation (7) under the assumption of a single spiking neuron as:

$$\begin{aligned} \bar{\mathbf{w}} &= \mathbf{M}(\mathbf{y} - \mathbf{D}\bar{\boldsymbol{\varphi}} - \bar{\mu}) \\ \bar{\mathbf{w}} &= F^{-1}\big(\mathbf{m}F(\mathbf{y})\big) - F^{-1}\big(\mathbf{m}F(\mathbf{D}\bar{\boldsymbol{\varphi}})\big) - a_1 \\ \text{despiked LFP} &= \text{LFP} - \text{smoothed contribution from spikes} - \text{offset} \end{aligned} \tag{25}$$

Here $\mathbf{m} = \frac{\gamma^2\mathbf{g}}{\sigma^2+\gamma^2\mathbf{g}}$ is a low pass filter. This can be compared to equation 7 of David et al. 2010:

$$\begin{aligned} L(t) &= L_0(t) - L_{\text{pred}}(t) \\ \text{despiked LFP} &= \text{LFP} - \text{portion of the LFP predictable by spikes} \end{aligned} \tag{26}$$

This suggests that the portion of the LFP predictable by spikes $L_{\text{pred}}(t)$ defined in David et al. 2010 may be similar to our smoothed contribution from spikes $F^{-1}\big(\mathbf{m}F(\mathbf{D}\bar{\boldsymbol{\varphi}})\big)$. This connection is made clearer by joining equations (8), (9) and (25) together, and setting the basis $\mathbf{B} = [\mathbf{I}_q; \mathbf{0}]$:

$$\begin{aligned} \left(\mathbf{k} * \mathbf{I}_q - \frac{q^2\big(\mathbf{s}'\widehat{\mathbf{1}}\big)^2}{n}\big(F^{-1}(1-\mathbf{m})'\widehat{\mathbf{1}}\big)\widehat{\mathbf{1}}_q'\right)\bar{\boldsymbol{\varphi}} &= r\big[\mathbf{I}_q, \mathbf{0}\big]F^{-1}\big((1-\mathbf{m})F(\mathbf{y})\overline{F(\mathbf{s})}\big) - a_1 \\ \mathbf{K}\bar{\boldsymbol{\varphi}} &= r\big[\mathbf{I}_q, \mathbf{0}\big]F^{-1}\big((1-\mathbf{m})F(\mathbf{y})\overline{F(\mathbf{s})}\big) - a_2 \\ \bar{\boldsymbol{\varphi}} &= r\mathbf{K}^{-1}\big[\mathbf{I}_q, \mathbf{0}\big]F^{-1}\big((1-\mathbf{m})F(\mathbf{y})\overline{F(\mathbf{s})}\big) - a_3 \end{aligned} \tag{27}$$

a_i are offsets and $\mathbf{K}$ is a convolution matrix derived from $\mathbf{a} = F^{-1}\big((1-\boldsymbol{m})|\mathrm{F}(\mathbf{s})|^2\big)$ by time shifting, $K_{ij} = a_{[i-j+1]}$. Hence our smoothed contribution from spikes is:

$$F^{-1}\left(\mathbf{m}F(D\overline{\boldsymbol{\varphi}})\right) = pF^{-1}\left(\mathbf{m}F(\mathbf{s})F\left(\left[\mathbf{I}_q,\mathbf{0}\right]'\mathbf{K}^{-1}\left[\mathbf{I}_q,\mathbf{0}\right]F^{-1}\left((1-\mathbf{m})F(\mathbf{y})\overline{F(\mathbf{s})}\right)\right)\right) - a_4$$
$$= pF^{-1}\left(\mathbf{m}F(\mathbf{s})F\left(\begin{bmatrix}\mathbf{K}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}F^{-1}\left((1-\mathbf{m})F(\mathbf{y})\overline{F(\mathbf{s})}\right)\right)\right) - a_4 \quad (28)$$

Meanwhile, the portion of the LFP predictable by spikes as defined in David et al. 2010 can be written in our notation as:

$$L_{\mathrm{pred}}(t) = \kappa F^{-1}\left(F(\mathbf{s})F\left(\begin{bmatrix}\mathbf{J} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}F^{-1}\left(\mathbf{m}\frac{F(\mathbf{y})\overline{F(\mathbf{s})}}{|F(\mathbf{s})|^{\mathbf{2}}}\right)\right)\right) \quad (29)$$

Notice the similarity between these two equations. This is especially striking when taking into account that $\mathbf{K}^{-1}$ is related, but not equal to $\boldsymbol{\Lambda} = F^{-1}\left(1/\left((1-\mathbf{m})|F(\mathbf{s})|^2\right)\right)$. Now suppose that $\mathbf{K}^{-1}\left[\mathbf{I}_q,\mathbf{0}\right] \approx \left[\mathbf{I}_q,\mathbf{0}\right]C'(\boldsymbol{\Lambda})$. This would hold if the support of $\boldsymbol{\Lambda}$ is smaller than the length of spikes. Hence, taking q to be large, our smoothed spike contributions becomes:

$$\approx pF^{-1}\left(\mathbf{m}F(\mathbf{s})F\left(\begin{bmatrix}\mathbf{I_q} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}F^{-1}\left(\frac{F(\mathbf{y})\overline{F(\mathbf{s})}}{|F(\mathbf{s})|^{\mathbf{2}}}\right)\right)\right) - a_4 \quad (30)$$

If we take the further step of assuming that the support of $F^{-1}(\mathbf{m})$ is smaller than the length of spikes, the two equations become equal, modulo a different handling of offsets. Hence the methods are essentially equivalent in the limit of long spikes. With short spikes, however, the order of the filters and the projection matrix in each method is different, and hence they handle edges around the time of spikes differently. Specifically, in our method the smoothed spike contribution is obtained by first taking the STA of a high-passed filtered wideband signal, keeping the first q elements of this STA, then undoing the high-pass filtering and compensating for the spike-train autocorrelation to obtain a spike waveform, and finally convolving this waveform with the spike train and smoothing it. In contrast, the David et al. method takes a STA of the low-pass filtered wideband signal, compensates for the autocorrelation of the signal, then keeps the first q elements of this STA to obtain the smoothed waveform which is convolved with the spike train to obtain a smoothed spike contribution.

For small q (short spikes), keeping the q first elements of the STA of a low-passed filtered signal (the smoothed mean spike) will lead to artifacts at the edges of spikes. The exact way that the edges are handled in the David et al. method is further complicated in this method by the use of Hanning windows to perform the required Fourier transforms. The edge handling of the method of David et al. 2010 should not be considered a design flaw, as the authors were interested in the scenario where q is large,

in which case the edge handling is sufficient; it does, however, render their method inappropriate for removing spikes in the scenario where q is small (short spikes).

Appendix - Properties of circulant matrices

A circulant matrix is defined as:

$$C(\boldsymbol{x}) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & & x_{n-1} \\ x_{n-1} & x_n & x_1 & & x_{n-2} \\ \vdots & & & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix} \tag{31}$$

By this definition, circulant matrices are closed under addition and $C(\boldsymbol{a}) + C(\mathbf{b}) = C(\mathbf{a} + \mathbf{b})$. The transpose of a circulant matrix is circulant. The product of two circulant matrices is circulant. Note also that the identity matrix $\mathbf{I}$ is circulant. Multiplication of the transpose of a circulant matrix by a vector corresponds to a circular convolution:

$$C'(\mathbf{a})\mathbf{b} = (a \circledast b)_i = \sum_j a_{[i-j]} b_j \tag{32}$$

Here $a_{[k]} = a_{(k \text{ modulo } n)+1}$ with n being the length of $\mathbf{a}$. By the circular convolution theorem we have that:

$$C'(\mathbf{a})\mathbf{b} = F^{-1}\big(F(\boldsymbol{a})F(\boldsymbol{b})\big) \tag{33}$$

Here $F(\mathbf{a})$ is the discrete Fourier transform of $\mathbf{a}$. This implies the following properties:

$$\begin{aligned} C'(\mathbf{a})\mathbf{b} &= C'(\mathbf{b})\mathbf{a} \\ C(\mathbf{a})\mathbf{b} &= \mathrm{F}^{-1}\left(\overline{F(\boldsymbol{a})}F(\boldsymbol{b})\right) \\ \big(C'(\mathbf{a})\big)^{-1} &= C'\left(F^{-1}\left(\frac{1}{F(\mathbf{a})}\right)\right) \end{aligned} \tag{34}$$

The determinant of a circulant matrix can be found by noting that:

$$\begin{aligned} C'(\mathbf{a})\mathbf{b} &= \mathrm{F}^{-1}\big(F(\boldsymbol{a})F(\boldsymbol{b})\big) \\ &= \hat{\mathbf{F}}^{-1}\mathrm{diag}(\hat{\mathbf{F}}\mathbf{a})\hat{\mathbf{F}}\mathbf{b} \end{aligned} \tag{35}$$

Here $\hat{\mathbf{F}}$ is the discrete Fourier transform matrix (DFT matrix) which maps a vector onto its discrete Fourier transform. Hence:

$$\begin{aligned} |C'(\mathbf{a})| &= \left|\hat{\mathbf{F}}^{-1}\mathrm{diag}\left(\hat{\mathbf{F}}\mathbf{a}\right)\hat{\mathbf{F}}\right| \\ &= \left|\hat{\mathbf{F}}^{-1}\right|\left|\mathrm{diag}\left(\hat{\mathbf{F}}\mathbf{a}\right)\right|\left|\hat{\mathbf{F}}\right| \\ &= \left|\mathrm{diag}\left(\hat{\mathbf{F}}\mathbf{a}\right)\right|\left|\hat{\mathbf{F}}\hat{\mathbf{F}}^{-1}\right| \\ &= \prod_i \mathrm{abs}\left(\hat{\mathbf{F}}\mathbf{a}\right)_i \end{aligned} \tag{36}$$

This expression is valid for $\mathbf{a} \in \mathbb{R}^N$. Therefore the log-determinant of a circulant matrix is:

$$\log|C'(\mathbf{a})| = \sum_i \log|F(\mathbf{a})_i| \tag{37}$$

A symmetric circulant matrix corresponds to circular convolution by a symmetric kernel, and its corresponding Fourier coefficients are real. A symmetric circulant matrix whose corresponding Fourier coefficients are positive is positive definite and therefore is a valid covariance matrix.