

Title:

Discovering reoccurring patterns in electrophysiological recordings

Authors:

Bart Gips^{a*}, Ali Bahramisharif^{a,b}, Eric Lowet^{a,c}, Mark J. Roberts^{a,c}, Peter de Weerd^{a,c}, Ole Jensen^a, Jan van der Eerden^a

*corresponding author:

bart.gips@donders.ru.nl

Donders Institute for Brain, Cognition and Behaviour

Kapittelweg 29, 6525 EN Nijmegen

Affiliations:

^a Radboud University, Donders Institute for Brain, Cognition and Behaviour, 6525 EN Nijmegen, The Netherlands

^b Department of psychiatry, Academic Medical Center, PO Box 22660, 1100 DD Amsterdam, The Netherlands

^c Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD Maastricht, the Netherlands

Keywords:

Evoked response, Theta, gamma, Oscillation, Markov Chain Monte Carlo

Abstract

Fourier-based techniques are used abundantly in the analysis of electrophysiological data. However, these techniques are of limited value when the signal of interest is non-sinusoidal or non-periodic. We present a new data-driven method for discovering reoccurring temporal patterns in electrophysiological data. We call this method sliding window matching (SWM). SWM not only can be applied to identify reoccurring patterns of which the shape and spectral content is known, but more importantly it is effective for detecting reoccurring but unknown patterns even when they appear non-periodically. To demonstrate this, we first validated the method on simulated datasets, and we checked that for recordings with a sufficiently low noise level the SWM results were consistent with results from the widely used phase alignment (PA) method. After this, we used SWM on oscillations in local field potential (LFP) recordings from the rat hippocampus and monkey V1. The application of SWM yielded two interesting findings. We could show that rat hippocampal theta and monkey V1 gamma oscillations were both skewed (i.e. asymmetric in time), rather than being sinusoidal. Furthermore, gamma oscillations in monkey V1 were skewed differently in the superficial compared to the deeper cortical layers. Second, we used SWM to analyze responses evoked by stimuli or microsaccades even when the onset timing of stimulus or microsaccades was unknown. We conclude that the proposed method has wide applicability in the exploration of noisy time series data where the onset times of particular events are unknown by the experimenter such as in resting state and sleep recordings.

1 Introduction

Neural oscillations are an intrinsic property of a working brain. Their amplitude reflects whether the activity of neurons in a particular brain region synchronizes or desynchronizes, whereas their phase might provide information on the communication between brain regions (Fries, 2005; Pfurtscheller and Lopes da Silva, 1999; Varela et al., 2001). Fourier-based approaches are commonly used to extract phase and amplitude information of neurophysiological data. However, there are no *a priori* reasons for neuronal oscillations to be sinusoidal. In such cases higher harmonics of the base frequency are required to fully describe the repetitive pattern. Therefore, Fourier-based descriptions of data in terms of power and phase in a relatively narrow frequency band may not adequately characterize the data. The importance of non-sinusoidal waveforms is discussed in many other works. It has been suggested for example that the peaks and troughs of oscillations have different functions (Haegens et al., 2011; Jensen et al., 2014; Klimesch et al., 2007). This was supported by Mazaheri and Jensen (Mazaheri and Jensen, 2008), who found that the peaks are more strongly modulated than the troughs in human occipital alpha oscillations (~ 10 Hz). Moreover, Belluscio and colleagues (Belluscio et al., 2012), investigated the shape of rat hippocampal theta oscillations (~ 8 Hz). They showed that hippocampal theta oscillations are sawtooth-like rather than sinusoidal waveforms. This sawtooth-like shape may point to a global signal temporally ordering neuronal firing along the long down-going flank (Gips et al., 2016; Jensen et al., 2014; Lisman and Jensen, 2013). As the amplitude of the rat hippocampal theta oscillations is very high relative to the background noise, the skewness of the sawtooth shape could be demonstrated after applying a simple band-pass filter (1-80 Hz) and detecting local maxima and minima. This might not be an option when the signal-to-noise ratio is low, as is the case for example in the non-invasive MEG or EEG recordings. It would be useful to have a method that identifies reoccurring patterns (i.e. motifs)

even if they are non-sinusoidal, and for a broad range of signal-to-noise ratios. Here, we propose a sensitive and broadly applicable method for discovering reoccurring patterns in time series data which we call sliding window matching (SWM).

Methods often used to deal with non-stationary signals are time-frequency analyses such as short time Fourier transforms or some sort of wavelet transform (Le Van Quyen et al., 2001; Lyons, 2003; Sejdić et al., 2009). These methods, however, generally still work within the Fourier framework of sinusoidal basis functions, so they tend to share some of the disadvantages mentioned above.

Other methods have been proposed such as empirical mode decomposition (Flandrin et al., 2004; Huang, 2005; Huang et al., 1998) or those based on single spectrum analysis (Bonizzi et al., 2014, 2012; Ghil, 2002; Vautard and Ghil, 1989; Vautard et al., 1992). These try to deconstruct the signals without using sine and cosine basis functions. But these methods do tend to assume that the components are periodic, albeit that the frequency does not have to be fully constant over time for some of these methods. Brain signals, however, may often not be periodic, but instead appear and disappear on non-predictable intervals.

In this paper we will show that SWM is well-suited for detecting repeated and non-sinusoidal components that are not necessarily temporally periodic in recorded signals. Furthermore, we will demonstrate its capabilities for characterizing (periodic) oscillatory LFP recordings in rat hippocampus and monkey V1 and we will compare it to a more commonly used short Fourier transform method that uses wavelet convolution. In these demonstrations, SWM is effective in discovering not only motifs with known shape that occur at a given time relative to a known trigger event, but also in discovering motifs with unknown shape and those that occur at unpredictable moments in the time series.

2 Methods

We present a new algorithm that finds repeating patterns in a dataset that we call *motifs* (see Figure 1 for an illustration). All analyses were performed using MATLAB (MathWorks).

2.1 Sliding window matching

The SWM algorithm positions windows over the data in such a way that the contents in all windows are maximally similar. It does this by sliding the windows separately across the data using a Monte-Carlo approach until their contents are maximally similar (Figure 1, also see Algorithm 1). In doing this, the algorithm will find patterns (motifs) that are repeated in the data. It is important to note that the motif locations do not have to be periodic. I.e. the temporal spacing between the motifs in the data can be variable.

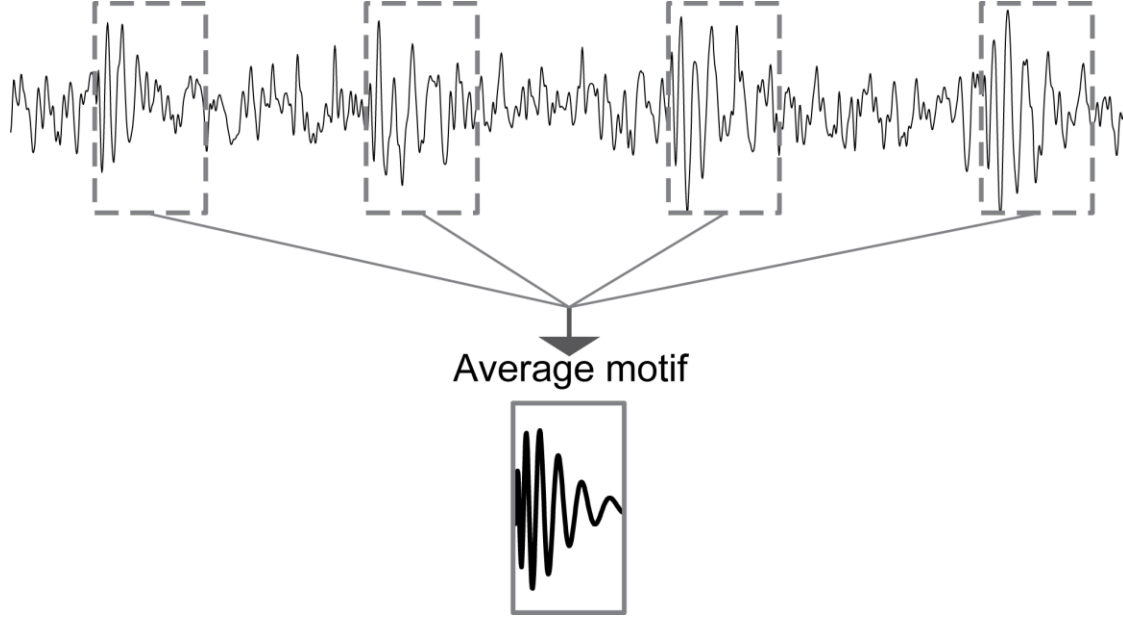


Figure 1. Conceptual illustration of the SWM algorithm and the concept of motifs. SWM tries to find reoccurring patterns (i.e. motifs) in the data (top trace) by sliding windows (dashed grey rectangles overlaid on the data trace) across the time dimension. The windows are moved randomly and the contents of the windows are compared to the mean across all the windows (average motif, bottom). When the SWM algorithm has converged, all windows contain a pattern that looks maximally like the average pattern of all windows. In this way, the motif and its occurrences in the data have effectively been detected.

2.1.1 The cost function

The SWM algorithm works by minimizing a cost function (Eq. 1) evaluating the similarity of all windows. When the minimum is found, a motif, i.e. the average signal in the windows, has been identified. N is the number of windows and L is the window length:

$$J = \frac{1}{2(N-1)} \sum_{i,j=1}^N d_{i,j}^2 \quad (1)$$

$$d_{i,j}^2 = \frac{1}{L} (\mathbf{z}_i - \mathbf{z}_j)^T \cdot (\mathbf{z}_i - \mathbf{z}_j) \quad (2)$$

$$\mathbf{z} = \frac{\mathbf{x} - \mu_x}{\sigma_x} \quad (3)$$

The more similar the windows are, the lower the variance across them and therefore the lower the cost J . The z-scoring (Eq. 3) is done for every window individually to counter signal drift and amplitude fluctuations. Additionally it ensures that the algorithm does not converge to periods of low-amplitude signals dominated by noise. Without this normalization, cost will be lowest in the windows only containing background noise, i.e. without any additional variance caused by any motif which has greater amplitude than the background noise. The value J measures the correlation among of the signals in the

windows, i.e. the higher the correlation coefficients between all windows, the lower the cost (see Supplemental material).

2.1.2 Minimizing the cost function

Our algorithm uses a Markov Chain Monte Carlo (MCMC) sampling technique (Hastings, 1970; Metropolis et al., 1953) to minimize J (see Algorithm 1). MCMC sampling techniques are often used to estimate probability density functions (Andrieu et al., 2003). It is important to note that we use MCMC sampling strictly for finding the global minimum of the cost function J and as such we do not use it in the context of Bayesian statistical models (Carlin and Chib, 1995). In most iteration steps a single window is moved to a random new position and the move is accepted if it decreases the cost J . If the cost increases with the new position, it is accepted only with a certain probability. This acceptance probability decreases with the change in cost (Algorithm 1, line 8: $p_{\text{accept}} = \exp(-\Delta J/T)$) where T denotes the ‘temperature’ factor. If T is high, window shifts that increase the cost (i.e. away from the nearest minimum) are more likely to be accepted. In some iterations the algorithm moves all windows at once in the same direction with the same step size (lines 17-25 in Algorithm 2 in Appendix A). The cost associated with the new window positions is evaluated and the move is again accepted with a probability based on the change in cost function. This makes it possible for the algorithm to jump to a different minimum with a low cost that is hard to reach by way of the other MCMC steps consisting of a change of a single window position.

To avoid that the minimization algorithm gets stuck in local minima, we used a Parallel Tempering (Frenkel and Smit, 2001) strategy to find the global minimum of the cost function J (see Algorithm 2 in Appendix A). This algorithm effectively minimizes at several temperatures in parallel. Therefore the method is computationally more demanding than straightforward isothermal minimization or simulated annealing (Kirkpatrick et al., 1983). Using information of these different temperatures, the algorithm can escape from local minima. The efficiency of the method to find the global minimum implies that the final solution is independent of the initial window locations.

Initialization:

- 1 Define window length (L) and minimal window spacing (G).
- 2 Initialize window positions X_1, X_2, \dots, X_N .
These window positions are separated by twice the minimal window spacing, i.e. $X_n - X_{n-1} = 2G$. So the total number of windows (N) depends on the length of the data time series, L and G .
Alternatively, for investigating evoked responses, a number of windows per trial can be given instead.
- 3 Calculate initial cost J_0 (Eq. 1) ;
- 4 **while** $i < \text{maxIterations}$ **do**
- 5 Pick a random window position X_n
- 6 Find a new allowed position X_n' for the window by moving it randomly along the time axis. (allowed positions: at least G separated from the preceding and following window);
- 7 Calculate the cost using X_n' , J' , and with this the change in cost function $\Delta J = J' - J_{i-1}$

```

8      Calculate acceptance probability  $p_{\text{accept}} = \exp(-\Delta J/T)$ 
9      if rand(0,1) <  $p_{\text{accept}}$  then
10         Accept change:
11          $J_i \leftarrow J'$ 
12          $X_n \leftarrow X'_n$ 
13      else
14         Do not accept change:
15          $J_i \leftarrow J_{i-1}$ 
16      end
17       $i \leftarrow i+1$ 
18  end

```

Algorithm 1. General algorithm for finding reoccurring shapes, i.e. motifs, in electrophysiological data using a MCMC technique.

2.1.3 Parameters and output

The output of the algorithm is the window position vector \mathbf{X}_q . When the algorithm converges, the output vector \mathbf{X}_q and the motif found may still depend on number of windows N , window length L and minimal window spacing G (see Figure 2A). For example when the data contains multiple oscillatory components of different frequencies, L of 100-150 ms may find the alpha-band oscillations (8-12 Hz) if present, whereas longer L will lead to waveforms in the theta (4-8 Hz) or delta (1.5-4 Hz) range. When studying high-frequency phenomena in electrophysiological data such as the shape of gamma oscillations (see Figure 7), it is recommended to use high-pass filtering below the frequency of interest before applying SWM, to make sure that the algorithm will not converge to a low-frequency motif with large signal-to-noise ratio.

Evoked responses in electrophysiological recordings generally have a higher signal-to-noise ratio than neural oscillations and are expected to have low variability across occurrences. Therefore SWM tends to converge to evoked responses if their timescale is similar to L , and the number of windows (N , the length of vector \mathbf{X}_q) is roughly equal to or lower than the number of evoked responses in the data. It should be realized that the motif may be placed at an arbitrary position within the window, especially when L is not exactly equal to the length of the evoked response (Figure 2B). The guard parameter G is the minimal distance between two window positions. If a new window location is less than G removed from a neighboring window (Figure 2A), the move is not accepted. This ensures that the algorithm will not stack all windows on the same piece of data, which would lead to a zero cost. G is generally best chosen to be of similar size to the motif one is trying to find, i.e. large enough that the windows will not have too much overlap, but small enough that they have enough space to be positioned over the data.

It is important to note that the precise values for N , L and G are not crucial for the motif (shape) as long they are chosen in the order of magnitude as discussed above (i.e. estimated number of occurrences for N and estimated length of the motif of interest for L and G)

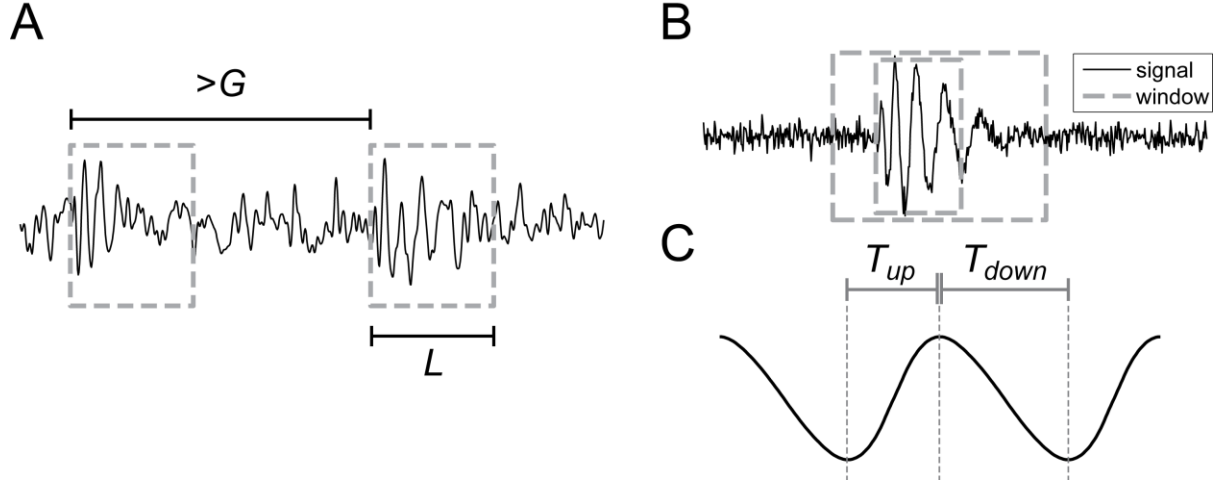


Figure 2. **A** L is the length of the windows, whereas G determines the minimal distance between two successive windows. In the case of having multiple components in the data, the choice of L affects the results. Since having noise in the data increases the cost (Eq. 1-3), SWM will converge to a motif that has high signal to noise ratio across the entire window and length of at least L . **B** Setting the window length to different values can cause the main motif features to be positioned differently within the window. The smaller window may tightly hug the motif, but the larger window may contain the motif in different positions within it. **C** T_{up} and T_{down} are the temporal duration of the up-going and down-going flank used to calculate the skewness index (Eq. 4). In the case of a perfect sawtooth waveform, either T_{up} or T_{down} approaches the limit of 0.

2.2 Analysis of the motifs

2.2.1 Skewness Index

We define the skewness index (SI) in equation 4 (see Figure 2C), where T_{up} is the time from trough to peak and T_{down} is the time from peak to trough.

$$SI = \frac{T_{up} - T_{down}}{T_{up} + T_{down}} \quad (4)$$

The SI of a temporally symmetrical waveform (e.g. a sinusoid) is zero, whereas the SI of a sawtooth fully skewed to the left or the right approaches the limits of -1 and 1 respectively.

2.2.2 Statistics

Since individual time-windows are too noisy for a reliable estimation of T_{up} and T_{down} , we used a bootstrap approach (Efron, 1979) to assess the accuracy of the measured skewness index (SI). Estimation of mean SI and standard error was done by generating new average motifs by sampling with replacement N windows from the N windows found by either the SWM or the PA algorithm. This was done 1000-fold to yield a probability distribution for the SI. Values for T_{up} and T_{down} are discrete due to the fixed sampling rate of the original signal; this therefore leads to discrete values of the SI. However, bootstrapping techniques require a continuous measure. To overcome this, the motifs are interpolated using cubic spline interpolation to a sampling rate of 100 kHz, giving smooth probability distributions of the SI. Standard errors of the mean SI are calculated from these probability distributions.

2.3 Validation

In order to validate our SWM method and to illustrate its potentials, we compared its results with those of the established phase alignment (PA) method (van Kerkoerle et al., 2014). With both methods we identified motifs in synthesized data as well as in empirical data from rat hippocampus and monkey V1.

2.3.1 Phase-alignment method

With PA we mean a method of cutting out and aligning short sections of a signal according to instantaneous phase in a predefined narrow frequency band.

We performed a complex Gabor wavelet (Eq. 5) convolution (Delprat et al., 1992) to calculate a time-resolved estimate of the phase and amplitude at the frequency of interest (f_0).

$$w(\tau) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\tau^2}{2\sigma^2}\right) \exp(-i2\pi f_0 \tau) \quad (5)$$

In the current work we applied a wavelet with bandwidth of $1/(2\pi\sigma) = 2$ Hz. Similar to the SWM algorithm, windows with a length of about $1/f_0$ can be defined, centered at the same instantaneous phase in the frequency band centered at f_0 . From the set of aligned windows, the ones with high instantaneous amplitude can be selected to extract the motifs with putatively highest signal-to-noise ratio. For our purpose of validating SWM based on PA, we extracted the windows centered on the peak of the oscillation waveform (phase = 0). Also we extracted the same number of windows in the PA and in the SWM analyses of each data set.

2.3.2 Simulated data

Two hundred time series (i.e. trials) of 1 s and a sampling rate of 1 kHz were generated containing pink (1/f) noise together with a 10 Hz periodic sawtooth signal (Figure 3A) that acted as the motif of interest. The signal to noise ratio (SNR) of the simulated data shown in Figure 3 and Figure 4 was calculated as the ratio between the time-domain variances of the isolated noise and the sawtooth signal.

The pink noise was generated by applying an inverse Fourier transform to a spectrum with amplitudes equal to $1/f$. In order for the amplitude to not diverge at low frequencies, the pink spectrum is cut off below 1 Hz. The amplitudes in the spectrum were then multiplied by a signal-to-noise factor to make the noise stronger or weaker relative to the sawtooth signal (see Figure 3B). The phase for every Fourier coefficient was randomly assigned a value between $-\pi$ and $+\pi$. In this manner the artificial signal had characteristics similar to experimental recordings: most notably the $1/f^2$ shape of the power spectral densities found in the human extracranial scalp recordings (Gilden et al., 1995; Gilden, 2001; Miller et al., 2009) and animal intracortical recordings (Baranauskas et al., 2012; Henrie and Shapley, 2005).

2.3.3 Rat hippocampus

The recordings used in this study were previously included in two other studies, and data collection methods were described in detail previously (Bieri et al., 2014; Zheng and Colgin, 2015). In brief: electrodes were implanted in the hippocampal area CA1 and the hippocampal fissure of three male Long-Evans rats to record the local field potential (LFP). The LFP recordings were made while the rats were freely moving around in their individual cages (~ 40 cm \times 40 cm \times 40 cm) built from clear acrylic and containing enrichment materials (e.g., plastic balls, cardboard tubes, and wooden blocks). Rats were

allowed to recover from surgery for at least 1 week prior to data collection. During the data collection period, rats were placed on a food deprivation regime that maintained them at ~90% of their free-feeding body weight. All experiments were conducted according to the United States National Institutes of Health Guide for the Care and Use of Laboratory Animals under a protocol approved by the University of Texas at Austin Institutional Animal Care and Use Committee. The data was acquired with a sampling rate of 2000 Hz. Around 10 minutes of continuous data were analyzed from each rat.

2.3.4 Monkey V1

In the experimental setup, a monkey sat in a dark room with its eyes directed towards a computer screen. After one second of fixating on a central dot on a grey screen, a square-wave grating of variable orientation appeared in the para-foveal vision of the monkey. The monkey had to detect a change in the color of the fixation spot and report it by making a saccadic eye movement to a target. The onset of the color change was variable and the final saccade was not included in the data that are analyzed in the present work. During this experiment the local field potential in the primary visual cortex of the monkey was recorded at different depths (16 contact points with 150 μm spacing). The data was recorded with a sampling rate of 1000 Hz. Eye position was recorded using an infrared camera system (Thomas Recoding) with 240 Hz sampling. In total, 402 trials were recorded. All procedures were in accordance with the European Communities Council Directive 1986 (86/609/EEC) and in EU directive 2010/63 EU as well as approved by the local ethics committee (Radboud University Dier Experimenten Commissie). Following initial training, the monkey was implanted with a titanium head holder (Crist Instrument) and a recording chamber (Crist Instrument) above V1/V2 under general anesthesia and sterile conditions. In a second surgery, after further training, a craniotomy was made above V1/V2.

3 Results

3.1 Characteristics of neuronal oscillations

SWM can be applied to identify the wave-shape of non-sinusoidal oscillations. We used SWM first on synthetic data and compared it to the phase alignment (PA) method. After this, the SWM algorithm was applied to experimental rat and monkey recordings to assess the skewness of rat hippocampal theta waves and monkey visual induced gamma.

3.1.1 Detecting the shape of oscillatory signals in simulated data

Simulated data with a periodic sawtooth function embedded in $1/f$ noise was constructed (Figure 3A and B, different SNR values are coded by color. Blue: no noise; purple: 0.1; red: 1.0, see section 2.3.2). Figure 3C shows the average signal as found by PA using a complex Gabor wavelet centered on $f_0 = 10$ Hz with a bandwidth of $1/(2\pi\sigma) = 2$ Hz (Eq. 5). Figure 3D shows the motif found by SWM using a window length of $L = 200$ ms. Both the PA and SWM approach identified the periodic sawtooth shape without noise (Figure 3C, D; blue trace), with PA even returning the exact sawtooth when $\text{SNR} \rightarrow \infty$. In PA a stronger bias towards sinusoidal motifs was found: the motifs found at lower SNR values (red and purple traces in Figure 3C-D) by using PA (Figure 3C) are much more sinusoidal than those found by SWM (Figure 3D). Quantitatively this is shown in Figure 4. We varied the SNR and applied both PA and SWM to

find the sawtooth motif and determine its skewness (see section 2.3 and Figure 2C). The addition of noise decreased the measured skewness index of the mean sawtooth shape in all cases. Note however that SWM yielded skewness estimates close to the noiseless value for SNRs as low as 0.2, whereas PA still gave a ~25% reduction at SNR = 1.

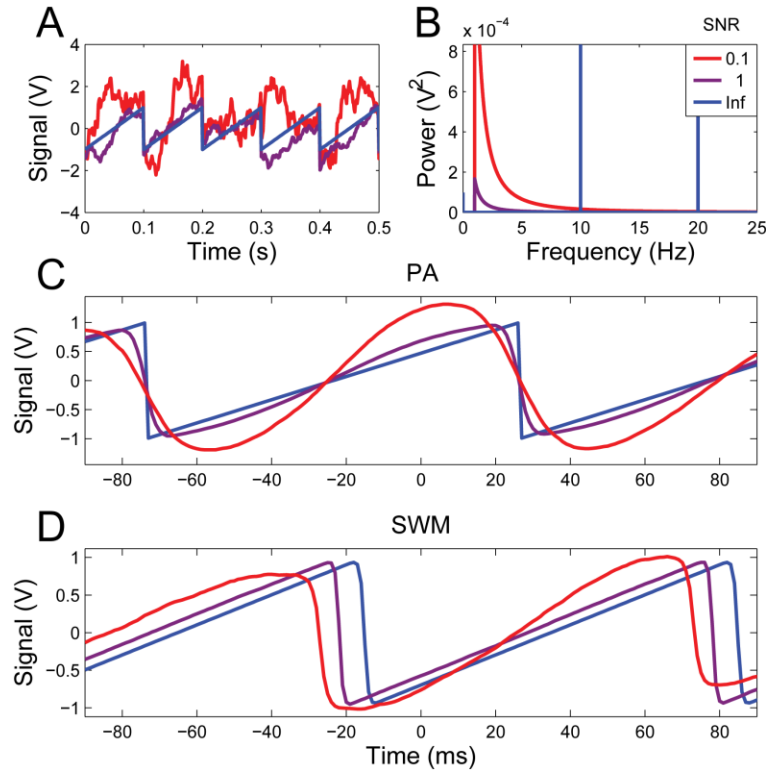


Figure 3. Results of applying phase alignment and the SWM algorithm to simulated data. **A.** Examples of three constructed signals: a regular 10 Hz sawtooth without (blue) and two with 1/f noise (purple and red). **B** Power spectral densities of the four signals. All three spectra have the same narrow peaks at 10 Hz and the first higher harmonic at 20Hz. **C, D.** The resulting motifs found by PA (C) and the SMW algorithm (D), traces indicate different signal to noise ratio levels as in B.

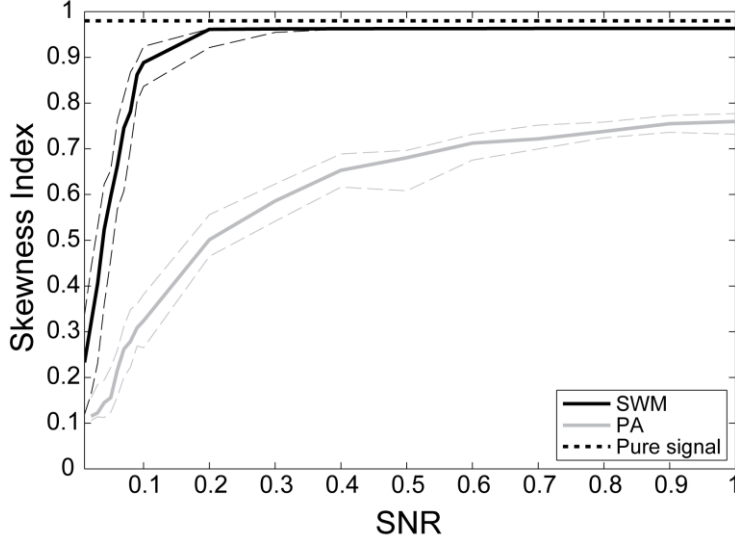


Figure 4. Comparison of the skewness indices found by using the SWM and PA methods for extracting the mean waveform for simulated data with different signal-to-noise ratio (SNR). The skewness index was 0.98 in the noiseless case (black dotted line). The grey and black lines show the median of 20 simulations for every SNR level using PA and SWM respectively to estimate skewness. Dashed lines show the corresponding 0.25 and 0.75 quartiles. The SWM method is superior to PA in estimating the skewness of the signal with lower SNR. Already at a SNR of 0.2 the method found a skewness index close to the noiseless case.

3.1.2 Rat hippocampal theta

We applied SWM to LFP recordings from the rat hippocampal region CA1 and hippocampal fissure while the rats were freely moving around in their individual cages using SWM (Figure 5A) and PA method (Figure 5B). Before applying SWM, the data were high-pass filtered at 5 Hz using a 4th order Butterworth IIR filter. The window length L was then set to 200 ms to capture ~ 1.5 theta cycles. The minimal window spacing G was set to 100 ms (i.e. having a maximal overlap of 100 ms) such that two subsequent windows could capture subsequent theta waveforms. Since the theta oscillations are quite prominent in LFP recordings in the hippocampal regions, we set the number of windows N in such a way to achieve $\sim 75\%$ data covered by the windows. Figure 5A and B show that a triangular-shaped periodic theta waveform can be found in both area CA1 and the hippocampal fissure. Figure 6C shows the corresponding skewness indices. This is in line with what Belluscio and colleagues (Belluscio et al., 2012) have demonstrated. They measured mean skewness indices of -0.3 and -0.15 in CA1 when the animals were running in a linear maze or in REM sleep respectively (Belluscio and colleagues used a different measure called the asymmetry index in their work, defined as the logarithm of the ratio between T_{up} and T_{down} , here we have converted their reported values to our SI definition for easy comparison). The right panel in Figure 5A illustrates that SWM does not explicitly control the position of the motif within the window. Hence, although the motif found in area CA1 in Rat 2 is in antiphase with the motifs found in CA1 in the other two rats, this does not reflect meaningful differences (see Figure 2C). Figure 5C indicates that both SWM and PA are sufficiently sensitive for estimating the skewness in the rat hippocampal theta. This was to be expected in view of the very high SNR for the theta waveforms (cf Figure 4).

The right panels in A and B show that SWM identifies a gamma component, locked to the theta waveform, which is missed by PA. The existence of such a component could indicate either phase-amplitude

coupling or phase-phase coupling. The latter is in agreement with the findings by Belluscio and colleagues (Belluscio et al., 2012). Our observation counters recent criticism (Aru et al., 2014) attributing apparent cross-frequency coupling to non-sinusoidal wave shapes: the motifs found in CA1 (Figure 5A) using SWM suggest a true cross-frequency coupling between theta and gamma. Note that we consider our interpretation as tentative; further quantitative analysis is required to confirm this finding.

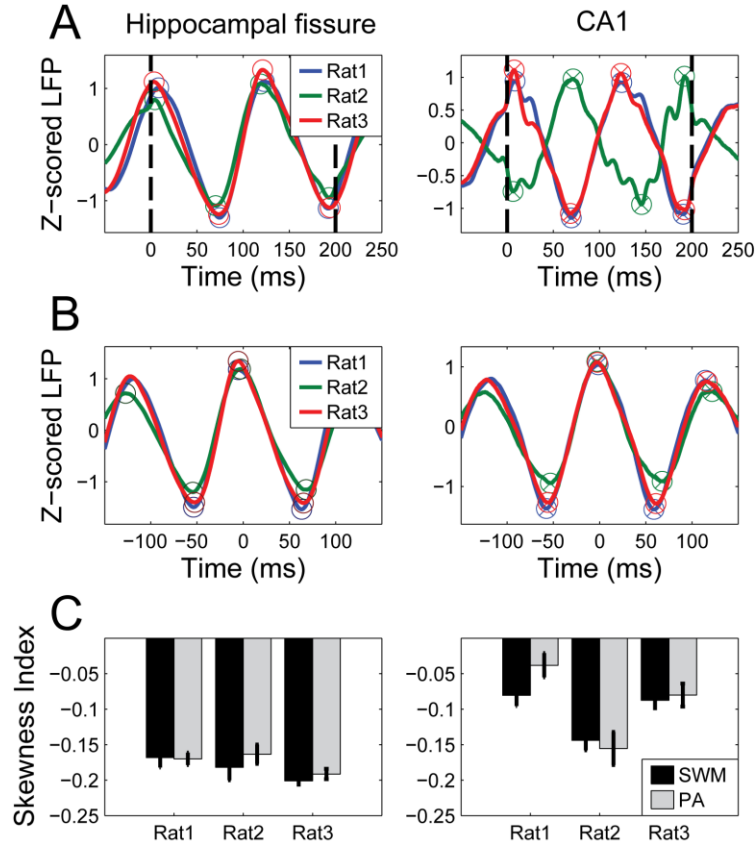


Figure 5. **A** Found theta shapes in rat LFP recordings in hippocampal fissure (left) and area CA1 (right) in 3 different rats. Vertical black dashed lines are the borders of the windows found by the SWM algorithm. The open circles denote the average position of the detected peaks and troughs for estimating T_{up} and T_{down} for calculating skewness using equation 4. **B** Same as A, but using PA rather than SWM. **C** Skewness of the waveforms in A and B. Black bars denote standard error of the mean as estimated by the bootstrap procedure (see section 2.2.2). Black bars indicate results from SWM (see A), grey bars correspond to the waveforms found using PA (see B).

3.1.3 Monkey V1 gamma

We applied the SWM algorithm to laminar LFP recordings from monkey V1 (see section 2.3.4) during presentation of a square-wave grating presented peripherally. Figure 6 shows the time-frequency representation of power around stimulus onset. A broadband evoked response can be seen, followed by an increase in gamma power as highlighted by the dashed white rectangle. The first 200 ms after stimulus onset were omitted from the analysis to avoid the evoked response and therefore focus on the induced gamma oscillations.

It has been proposed that two spatially distinct generators of gamma oscillations are active in V1 (Maier et al., 2010), correlation coefficients were calculated for the data filtered between 35 and 55 Hz for the pairs of the 16 contact points on the laminar probe. Figure 7A shows these correlation coefficients, together with the border of pia and cortex and the border between granular and infragranular layers (see Figure S 2), labeled L1 and L4 respectively. The figure clearly shows two clusters of correlated gamma activity, one in the superficial layers (L1-L4) and one in the infragranular layers.

The SWM algorithm was applied to the post-stimulus data to detect the shape of these gamma oscillations for the 16 contact points on the electrode shank separately. The data was high-pass filtered at 25 Hz to remove signals in the 5 – 25 Hz band evoked by micro-saccades (Ito et al., 2011; Lowet et al., 2016) (Figure 6 and Figure 9). Otherwise the SWM algorithm would converge to these evoked responses. After the motifs were identified by SWM, they were extracted from the raw, unfiltered data, to restore any possible low-frequency components present in the motifs. The skewness index was calculated for these 16 motifs and is shown in Figure 7B. These skewness indices are different for the gamma motifs found above the granular layer, compared to those present below it. The superficial motif derived from the signals from all L1-L4 contact points and the deep layer motif, derived from the contact points below L4 are shown in Figure 7C. The superficial skewness was significantly different from the deep layer skewness (0.2033 ± 0.0034 vs -0.0596 ± 0.0044 . Mean \pm standard error; $p < 0.001$ in a two-sample t-test). This result could be replicated using PA instead of SWM (Figure S 3B-D), but the skewness was less pronounced and not significantly different from zero in all contact points. This difference between PA and SWM can be understood by the lower sensitivity of PA, as discussed in section 3.1.1. In summary, SWM was able to detect a significant skewness in the induced gamma responses; moreover, the skewness was different for the superficial and deep gamma components.

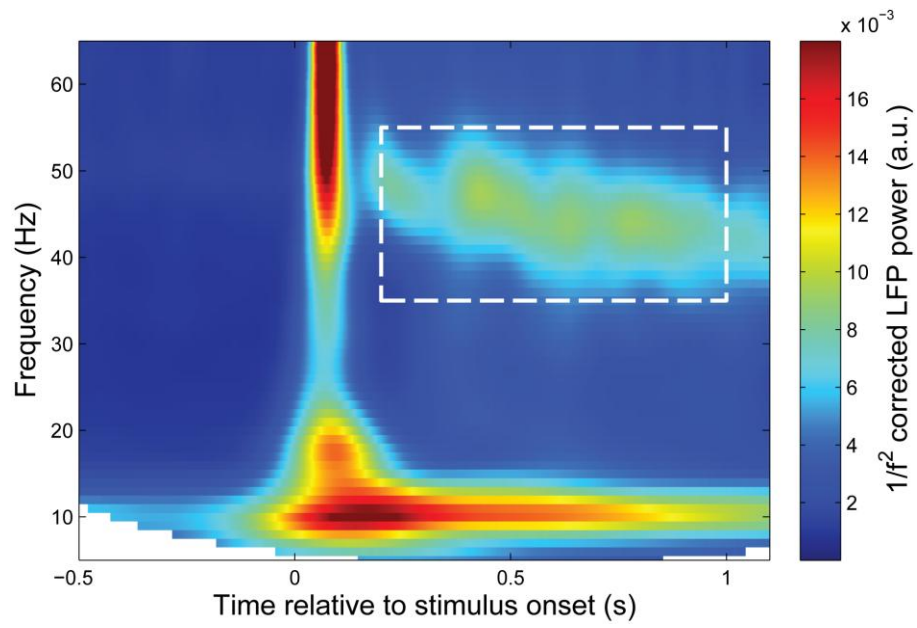


Figure 6. Time-frequency representations of monkey V1 LFP recordings. The power spectra are corrected by multiplying with f^2 to make both the alpha (10 Hz) and gamma band activity visible in the same figure. The dashed rectangle emphasizes the band-limited gamma component present in the post-stimulus period in the recordings.

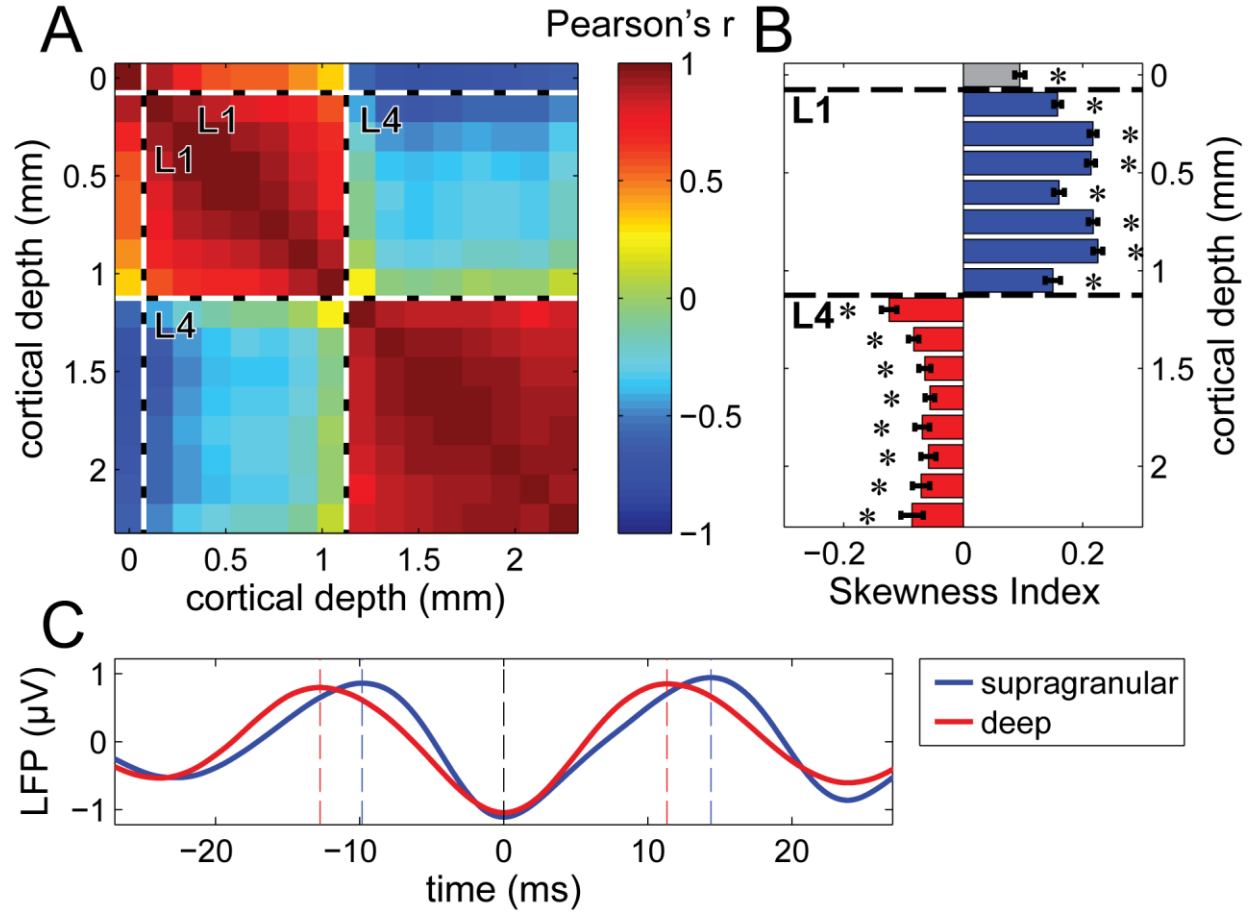


Figure 7. Properties of stimulus induced gamma-band activity in LFP recordings in monkey V1. **A** Pearson's correlation coefficient of gamma-band filtered data (35-55 Hz, corresponding to the induced activity in the dashed box in Figure 6). The supragranular contact points show a coherent gamma-band signal that was distinct from the signal measured by the deep contact points. Dashed lines labeled L1 and L4 indicate detected border (see Figure S 2) of dura/L1 and L4/L5 respectively. **B** Mean skewness of detected gamma waveforms at each contact point. Error bars denote standard error of the mean estimated by a 1000-fold bootstrap method (see section 2.2.2), Asterisks indicate significant differences from zero ($p < 0.001$, two-sample t-test). Blue denotes electrode contact points in the cortex situated above the L4-L5 boundary, red denotes contact points below this boundary. **C** The averaged gamma motifs found in the supragranular electrodes (blue in panel B) aligned to the central trough with the average motif found in the signal recorded deeper in the cortex (red in panel B). The supragranular gamma was significantly more positively skewed ($P < 0.001$, estimated with a 1000-fold bootstrap. See section 2.2.2) than the deep gamma with respective SI of 0.2033 ± 0.0034 versus -0.0596 ± 0.0044 (mean \pm standard error).

3.2 Detecting stimulus evoked potentials

Evoked responses are patterns which consistently occur locked to an event. We demonstrate how SWM is beneficial when detecting evoked potentials by finding stimulus- and saccade-evoked potentials using data from monkey V1 recordings. By having access to the actual event times of the evoked responses (i.e. the stimulus or saccade onset times) we can evaluate SWM's performance. In general, there is inherent ambiguity as to the exact locking of evoked potentials in V1. There could, for instance, be locking to the onset or the offset of the event, or there may be no consistent and precise locking. Since SWM is data-driven, we demonstrate that it can work around this ambiguity.

We applied SWM to the laminar LFP recordings in monkey V1 (see section 2.3.4). The onset of the stimulus evoked a typical response in the local field potential of the monkey's V1. To demonstrate the (temporal) accuracy of the SWM algorithm, we used it to analyze these evoked responses. We cut out 1 second of data in each of the 402 trials centered on the stimulus onset. We then performed SWM using 1 window per trial and a window length $L = 300$ ms. The dataset contained recordings at 16 different cortical depths, therefore, the windows were 2-dimensional arrays (16 cortical depths by 300 time points). The cost function was calculated across all cortical depths at once. This resulted in a 2-dimensional motif as shown in Figure 8B. The evoked response after aligning the data at the stimulus onset (captured by a photodiode attached to the screen) is shown in Figure 8A. In Figure 8C the difference between the evoked response and the motif found by SWM is shown. Only slight differences (compare scaling between C and A/B) could be discerned between the two responses. This can also be seen in Figure 8D where the window positions yielded by SWM have a distribution that is only around 4 ms wide (full width at half maximum). Note that in Figure 8B the motif found by SWM is aligned to the evoked response in A by shifting it temporally (i.e. time point $t=0$ on the x-axis in Figure 8B is not found by SWM, but by comparison to Figure 8A).

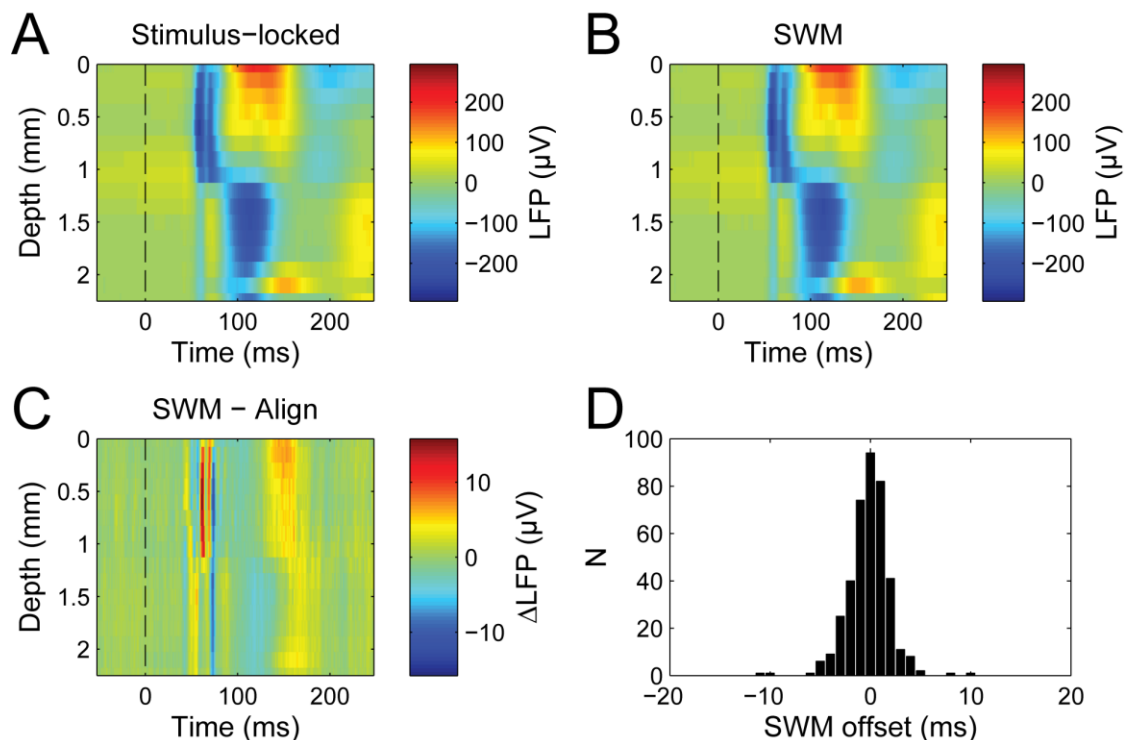


Figure 8. Detecting the stimulus evoked response using SWM **A** The average response evoked by the stimulus in the LFP across the different layers of V1. The trials were aligned to the stimulus onset before averaging. The dashed vertical line indicates the stimulus onset. **B** The motif (average window contents) as found by SWM. Note that it is virtually identical to the evoked response shown in **A**. The vertical dashed line indicates the mean location of the stimulus onset **C** The difference between the LFPs in **B** and **A**. **D** Histogram showing the temporal location of the windows relative to the stimulus onset (at 0 ms). This serves as an indication of the jitter across trials between stimulus onset and evoked response.

3.3 Detecting microsaccades

When an experimental paradigm demands a monkey to fixate, the monkey is able to keep its eyes relatively still, but will still make small eye movements around the fixation spot called microsaccades (Rolfs, 2009). Microsaccades also evoke a response in V1 (Bosman et al., 2009; Ito et al., 2011).

In the dataset used, eye movements were recorded allowing us to extract the responses in the LFP related to microsaccades; in this case eye movements in the order of 0.1 to 0.8 degrees occurring at 3-4 Hz. The temporal locations of microsaccade onsets were determined by thresholding the eye velocity signal (Engbert and Kliegl, 2003), using a velocity threshold of 5 times the standard deviation and a minimum saccade duration of 10 ms. We tested the quality of SWM by determining the microsaccade moments by applying SWM to the LFP data, without using the eye tracker recordings. SWM was applied to data recorded at least 200 ms after stimulus onset. The trial lengths were variable; the mean trial length was ~1100 ms. The window length was $L = 150$ ms (to match the size of the saccade triggered response, see Figure 9A) using on average 2 windows per trial. This probably underestimates the number of microsaccades present in the dataset. This was done to avoid windows not containing a saccade evoked response (e.g. when a trial contains 2 saccades, but 3 windows per trial are used, one window will by necessity contain noise, rather than the saccade related response). The average eye tracker-based saccade locked response and the SWM motif are shown in Figure 9A-B with the corresponding mean eye speed shown in C and D. The peak in Figure 9C and D, indicates that most SWM windows are correlated with microsaccades. However, the peak in the eye velocity profile found by SWM was lower and wider than the peak in Figure 9C. This can be caused by jitter between the onset of eye movement and the start of the evoked response. The jitter between microsaccades and SWM windows was clearly larger than between stimulus onset evoked potentials and SWM shown in Figure 8D. This can be partly explained by the low sampling rate of the eye tracking (240 Hz) compared to 1000 Hz of the LFP, see section 2.3.4. A second cause was that SWM converged to a configuration where not all window positions corresponded to a microsaccade. Since in the present case we did have sufficient eye movement information, these two effects could be disentangled. Based on eye velocity we conclude that 443 out of 615 converged windows contained an actual saccade; the others contained an unexplained similarly-looking alpha waveform, see Figure S 1.

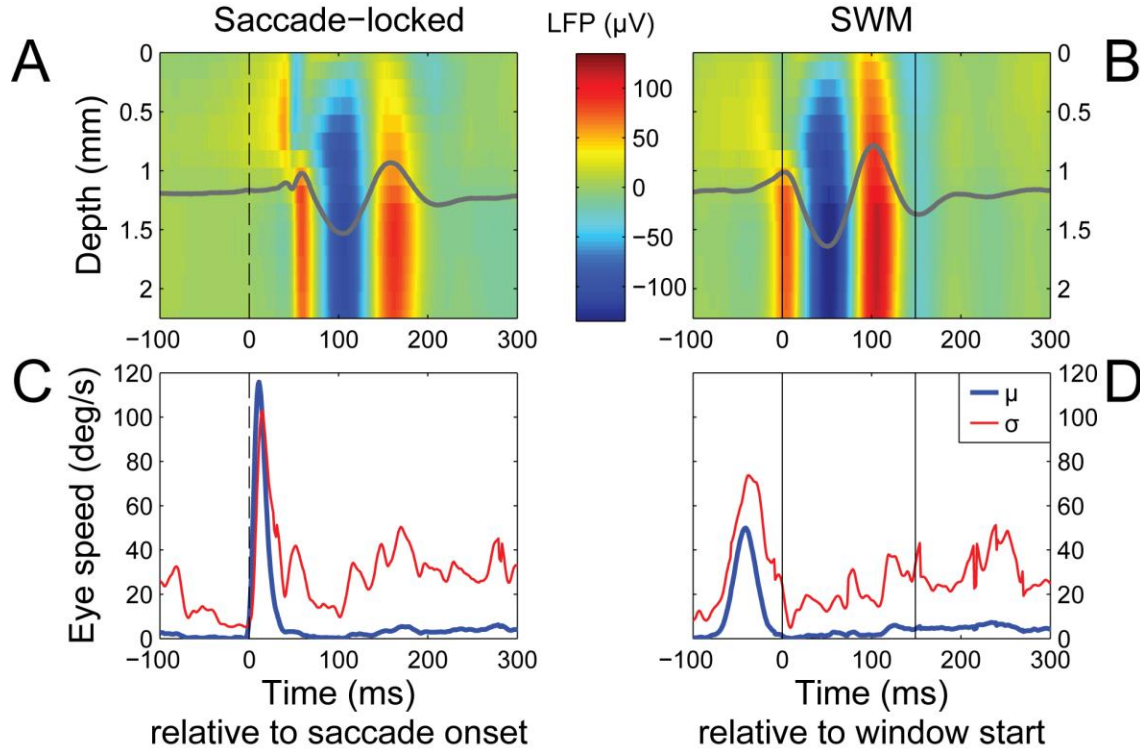


Figure 9. Detecting the microsaccade evoked response using SWM. **A.** The average response evoked by the microsaccade in the LFP across the different layers of V1. The trials were averaged aligned to the saccade onset. The dashed vertical line indicates the saccade onset. The overlaid dark trace shows the mean LFP across all 16 contact points. **B.** The motif (average window content) as found by SWM using 150 ms window length (L). The vertical black lines indicate the start and end of the window respectively. The time axis is relative to window start. Dark trace shows mean across the 16 contact points as in A. **C.** The eye velocity averaged across all the data segments locked to the onset of detected saccades (blue curve) and the standard deviation of the eye velocity across all detected saccades (red curve). Note that by definition this shows the average velocity profile of all microsaccades that were detected. **D.** Mean eye velocity (blue) and corresponding standard deviation (red) as in C, but now locked to the windows found by SWM. As in B, time axis is relative to window start and black vertical lines indicate start and end of the 150 ms windows.

4 Discussion

In this paper we have proposed a new method, called *sliding window matching* (SWM), for detecting and analyzing reoccurring patterns (motifs) in electrophysiological datasets. We demonstrated that we could find and characterize waveforms of sustained neuronal oscillations as well as those evoked by a stimulus without prior knowledge of the timing of occurrence of these motifs.

4.1 Application to oscillatory signals

SWM can be used to extract the shape of cortical oscillations when signal to noise ratio is low (Figure 4). We validated this in simulated data, showing that Fourier based approaches are biased towards finding sinusoidal motifs, whereas SWM correctly identifies more general waveforms. The SWM method was able to replicate the skewed waveforms in rat hippocampal theta as shown by Belluscio and colleagues (Belluscio et al., 2012). Secondly, it was able to directly show a gamma component in the local field potential phase-coupled to the theta waveform in area CA1 (Figure 5A). When applied to laminar

recordings in monkey V1, the gamma motif induced by a grating stimulus could be detected and, interestingly, gamma oscillations were also skewed. A possible cause of this skewness is a difference in neuronal synchrony during the up-going versus the down-going flank of the gamma activity. Gamma activity involves an interplay between excitatory and inhibitory neuronal populations (Bartos et al., 2007; Buzsáki and Wang, 2012; Tiesinga and Sejnowski, 2009). If one population fires more synchronously (i.e. more concentrated in time) than the other, then the corresponding flank (T_{up} or T_{down}) may last for a shorter time.

Furthermore, analysis of monkey V1 data replicated the finding of separate gamma oscillations in the deep and superficial layers (Maier et al., 2010); this was accompanied by the significant difference in skewness between deep and superficial gamma waveforms ($p < 0.001$). In principle this difference could result from a change in the polarity of the signal. Considering that the absolute correlation coefficient of the gamma-band activity between the clusters was lower than within (as is the gamma-band coherence, which should be of equal magnitude between and within the clusters if the superficial and deep gamma signals merely differed in their polarity, see Figure S 3A), it is unlikely that the two gamma oscillations are generated by the same source, with only the polarity flipped. Hence, it is more likely that the recorded activity originated from two distinct gamma sources having some partial synchrony (note the non-zero correlation/coherence between the superficial and deep gamma activity). One of these sources could be observed in the superficial layers of V1, the other in the deeper layers. To further investigate these two sources, the method proposed in the present paper, could be combined with a CSD analysis (Ainsworth et al., 2011; Bollimunta et al., 2008; Pettersen et al., 2006; Xing et al., 2012) of high-density laminar probe data to yield more definitive results.

4.2 Application to evoked responses

By applying SWM to LFP recordings from V1 in a macaque monkey, we demonstrated how SWM could determine the moment at which the response evoked by a visual stimulus occurred (Figure 8). The delay between stimulus onset and evoked response could conceivably be jittered on a trial-by-trial basis, possibly caused by variable hysteresis effects and delays in the neural network. In such cases, SWM could identify this jitter.

To illustrate the power of SWM, we applied SWM to locate micro-saccade evoked potentials where the locking of the evoked potential to the event is ambiguous (Figure 9). Compared to stimulus evoked responses, the detection of responses evoked by microsaccades is more difficult. The following points could play a role: 1) The responses evoked by the microsaccades were less similar across trials, compared to the stimulus evoked responses. This could be due to differences in microsaccade amplitudes and directions as well as effects from the activity preceding a microsaccade 2) The saccade evoked responses were smaller in amplitude than those evoked by the grating stimuli, leading to a reduced signal-to-noise ratio. 3) The time span between saccade onset and saccade evoked response was more variable than that between stimulus onset and stimulus evoked response. With the available eye tracking equipment, any physiological variability will be enhanced by the discrepancy between the sampling rate of the LFP recording (1 kHz) and that of the eye tracking equipment (240 Hz). 4) Other components (e.g. alpha-band neuronal oscillations) in the signal had a shape that is similar to the

saccade evoked response. This would cause some of the SWM windows to converge to parts of the data containing high amplitude alpha oscillations (172 out of 615 windows did not contain an actual saccade; instead they contained a similarly-looking alpha waveform, see Figure S 1). The fact that it seems quite difficult to disentangle the two based on shape alone, raises the question to what extent alpha oscillations and (micro-) saccades are related. To answer this question further research is needed.

4.3 Comparison to other techniques and recommendations

SWM is not the only analysis method that uses sliding windows. Several techniques using some form of template matching (Brunelli, 2009) have been used by others in electrophysiological signal analysis (O’Keefe and Recce, 1993) and artifact rejection (Li et al., 2006). However, it differs from these techniques in the sense that SWM does not require an *a priori* template (the motif), but discovers it from the data. In that sense it is similar to the Woody filter (Woody, 1967). The Woody filter, however, requires a priori knowledge of the approximate temporal locations of the motifs, whereas SWM will automatically detect these.

In the context of the analysis of non-sinusoidal, non-stationary signals, there exists a broad literature based on *empirical mode decomposition* (EMD) (Flandrin et al., 2004; Huang, 2005; Huang et al., 1998). EMD is able to separate the original signal into intrinsic mode functions. After obtaining these intrinsic mode functions, a Hilbert transform can be performed to derive an analytic representation of these modes to estimate the instantaneous phase. From then on, the motifs can be extracted by aligning parts of the data with the same phase similar to the phase alignment method mentioned in section 2.3.1. These empirical modes are free to change their frequency content over time, which provides greater flexibility and it requires less a priori knowledge (such as the central frequency and the bandwidth of the filter used in the PA method). On the other hand, this means that the correct mode needs to be selected from all intrinsic mode functions that have been extracted. Also, EMD is likely to lead to multiple modes containing similar spectral content, just at different moments in time. This means that the motifs of interest may be split across multiple modes. As a control, we applied this hybrid EMD-PA method to the signal with an SNR of 0.5 from Figure 4. We selected the mode that most looked like a band-limited 10 Hz signal by hand, and found that it underperformed compared to regular PA with a detected SI of 0.49 ± 0.08 .

Besides this, EMD is designed to find oscillatory modes, not repeating patterns. This fact makes it more difficult for EMD-based techniques to deal with irregularly popping up motifs that SWM is designed to detect. This makes EMD suited to extracting the periodic theta and gamma waves in rat hippocampus and monkey V1 in sections 3.1.2 and 3.1.3, but not suited for finding evoked responses such as in sections 3.2 and 3.3.

SWM relies on a parallel tempering MCMC sampling technique. This is an optimization method that is powerful in avoiding convergence to local minima and finding the global minimum. It is computationally more demanding than techniques that do not explicitly try to minimize a certain cost function, such as phase alignment or EMD. This effect is worse for larger datasets, since, unlike PA, convergence time for SWM does not scale linearly with sample size. The size of the energy landscape that SWM needs to sample to find the minimum in the cost function increases roughly exponentially with the size of the

dataset. For reference, detecting the stimulus evoked responses in the 402 trials each consisting of 16 electrodes having recorded 1000 data points required on the order of hours on a modern machine, whereas a method similar to PA (as used in Figure 3 and Figure 4) would only last around 10 seconds. Compared to SWM, PA can be a much faster technique that yields a unique solution. However the motifs need to have a high signal to noise ratio in a narrow frequency band. Secondly, because of the focus on a narrow frequency band, PA will return a mixture of motifs if the data contains two or more different motifs with overlap in the spectral content. In its present form, in this case, SWM will converge to the dominant motif, i.e. the motif with the lowest variability and highest SNR. This may partly explain the sharp transition of the gamma skewness from positive to negative when going from electrode contact point 8 to 9 in Figure 7B. In such a case, the algorithm could be extended to include clustering. Instead of comparing the contents of a window to all others to calculate the cost, the window is only compared to the other windows in the same cluster. Secondly, instead of only attempting to slide a window, the algorithm should also attempt to move a window to a different cluster and then estimate the change in cost function to calculate the acceptance probability. This could yield other, less dominant motifs in the other clusters.

When investigating oscillatory waveforms, it is arbitrary to which “phase” of the oscillation SWM converges. This is visible in Figure 5B: the motif found in area CA1 in Rat 2 is in antiphase with the motifs found in CA1 in the other two rats. In contrast, when using PA, the phase of the extracted motif is, by definition, controlled. In principle this is not a crucial deficiency of the algorithm and is mainly important for visualization purposes.

Despite the fact that SWM is data driven, it may generally help to use some a priori knowledge about the motifs one wants to detect. During a preprocessing step, irrelevant components of the data can sometimes be removed. For example, to detect the gamma motif in monkey V1 (Figure 7), it was helpful to filter out high-amplitude, low-frequency components that were not of interest. Note that if the window length is small relative to the period of a low-frequency oscillation, this high-pass filtering with this frequency threshold resembles linear detrending the data in each window. One could consider changing the cost function to explicitly incorporate linear detrending, making less preprocessing necessary. The same holds for parameters such as the approximate length of the motif; the parameter of window length L can determine to a large extent to which component the algorithm converges, so an approximate length of the motif of interest should be known a priori. Parameters that determine the rate of convergence of the Parallel Tempering technique (i.e. k and T , see Algorithm 2 in Appendix A) also need to be set when applying SWM. Good values for these parameters depend on the dataset. It is therefore recommended to perform a few test runs with different ranges of temperature factors (T , generally a set of logarithmically spaced temperatures) and switching probabilities (k) and investigate the trajectories of the cost function, to make sure that: 1) states are switched often enough between different temperatures. This can be achieved by increasing the number of parallel temperatures or decreasing the factor k . 2) The highest and lowest temperatures should both stay on a fairly constant cost value at the end of the run (indicating the maximum and the minimum cost value, respectively). See Figure S 4 for an illustration.

4.4 Possible applications in neuroscience

The different gamma motifs in monkey V1 in different laminae add some interesting insight regarding the cortical source and the mechanism behind gamma oscillations. In addition, the findings described open the possibility that SWM could contribute to functional dissociations among cortical layers. In the present work, analysis has been limited to LFP measurements, but it may be interesting to detect action potentials locked to the extracted waveforms. For example, in the gamma range, if pyramidal cells and interneurons lock their spiking to different flanks of the gamma cycle, this may be informative with regard to the mechanisms generating the gamma rhythm. As has been proposed before (Bartos et al., 2007; Tiesinga and Sejnowski, 2009), the gamma rhythm can be dictated either by the population of interneurons alone, or by an interplay between both the pyramidal cells and the interneurons.

In the present study, we demonstrated the reliability of SWM in cases in which the triggering information (stimulus onsets and microsaccades) is available, or in which the effects on the recorded activity is relatively well-known. However, SWM can be useful in discovering fully unknown waveform signatures in cases where also the triggering information is unknown. This can be the case when studying events in electrophysiological signals during sleep (e.g. sleep spindles (De Gennaro and Ferrara, 2003) and sharp wave-ripple complexes (Buzsáki, 1986)), or identification of interictal discharges in epilepsy patients (de Curtis et al., 2012), or even when studying higher cognitive processes, such as decision making. Secondly, SWM is not bound to electrophysiological data and besides LFP, MEG and EEG recordings, SWM can be used on any other dataset that is expected to contain reoccurring patterns (such as resting-state fMRI (Harrison et al., 2015)). Secondly, SWM may be employed for artifact detection. When artifacts in the recorded signal have a consistent shape, such as eye-blinks in EEG and MEG recordings, SWM can be employed to find them such that they can be discarded or regressed out.

5 Conclusion

We introduced SWM as a data-driven approach for detecting reoccurring patterns in (neurophysiological) data. Using simulated and empirical data, we have shown that SWM can reliably detect the shape of oscillatory waveforms and evoked responses. The application of SWM to neurophysiological data opens up new possibilities in understanding how brain signals are generated.

6 Acknowledgements

We are extremely grateful to Laura Lee Colgin at the University of Texas at Austin for providing us with the data from rat hippocampal recordings used in this work. The current work was supported by the Organization for Scientific Research (NWO) through the ALW Open Competition Grant (822-02-011).

7 References

Ainsworth, M., Lee, S., Cunningham, M.O., Roopun, a. K., Traub, R.D., Kopell, N.J., Whittington, M. a.,

2011. Dual Gamma Rhythm Generators Control Interlaminar Synchrony in Auditory Cortex. *J. Neurosci.* doi:10.1523/JNEUROSCI.2209-11.2011
- Andrieu, C., De Freitas, N., Doucet, A., Jordan, M.I., 2003. An introduction to MCMC for machine learning. *Mach. Learn.* 50, 5–43.
- Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., Singer, W., Vicente, R., 2014. Untangling cross-frequency coupling in neuroscience. *Curr. Opin. Neurobiol.* 31, 51–61. doi:10.1016/j.conb.2014.08.002
- Baranauskas, G., Maggiolini, E., Vato, A., Angotzi, G., Bonfanti, A., Zambra, G., Spinelli, A., Fadiga, L., 2012. Origins of 1/f² scaling in the power spectrum of intracortical local field potential. *J. Neurophysiol.* 107, 984–994. doi:10.1152/jn.00470.2011
- Bartos, M., Vida, I., Jonas, P., 2007. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* 8, 45–56. doi:10.1038/nrn2044
- Belluscio, M. a, Mizuseki, K., Schmidt, R., Kempter, R., Buzsáki, G., 2012. Cross-frequency phase-phase coupling between θ and γ oscillations in the hippocampus. *J. Neurosci.* 32, 423–35. doi:10.1523/JNEUROSCI.4122-11.2012
- Bieri, K.W., Bobbitt, K.N., Colgin, L.L., 2014. Slow and Fast Gamma Rhythms Coordinate Different Spatial Coding Modes in Hippocampal Place Cells. *Neuron* 82, 670–681. doi:10.1016/j.neuron.2014.03.013
- Bollimunta, A., Chen, Y., Schroeder, C.E., Ding, M., 2008. Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques. *J. Neurosci.* 28, 9976–88. doi:10.1523/JNEUROSCI.2699-08.2008
- Bonizzi, P., Karel, J., De Weerd, P., Lowet, E., Roberts, M., Westra, R., Meste, O., Peeters, R., 2012. Singular spectrum analysis improves analysis of local field potentials from macaque V1 in active fixation task, in: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 2945–2948. doi:10.1109/EMBC.2012.6346581
- Bonizzi, P., Karel, J.M.H., Meste, O., Peeters, R.L.M., 2014. Singular spectrum decomposition: A new method for time series decomposition. *Adv. Adapt. Data Anal.* 06, 1450011. doi:10.1142/S1793536914500113
- Bosman, C. a, Womelsdorf, T., Desimone, R., Fries, P., 2009. A microsaccadic rhythm modulates gamma-band synchronization and behavior. *J. Neurosci.* 29, 9471–80. doi:10.1523/JNEUROSCI.1193-09.2009
- Brunelli, R., 2009. Template Matching Techniques in Computer Vision: Theory and Practice. Wiley Publishing.
- Buzsáki, G., 1986. Hippocampal sharp waves: Their origin and significance. *Brain Res.* 398, 242–252. doi:10.1016/0006-8993(86)91483-6
- Buzsáki, G., Wang, X.-J., 2012. Mechanisms of gamma oscillations. *Annu. Rev. Neurosci.* 35, 203–25. doi:10.1146/annurev-neuro-062111-150444
- Carlin, B.P., Chib, S., 1995. Bayesian Model Choice via Markov Chain Monte Carlo Methods. *J. R. Stat. Soc. Ser. B* 57, pp. 473–484.
- de Curtis, M., Jefferys, J.G.R., Avoli, M., 2012. Interictal Epileptiform Discharges in Partial Epilepsy: Complex Neurobiological Mechanisms Based on Experimental and Clinical Evidence, in: Jasper's Basic Mechanisms of the Epilepsies.

- De Gennaro, L., Ferrara, M., 2003. Sleep spindles: An overview. *Sleep Med. Rev.* doi:10.1053/smr.v.2002.0252
- Delprat, N., Escudié, B., Guillemain, P., Kronland-Martinet, R., Tchamitchian, P., Torrèsani, B., 1992. Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies. *Inf. Theory, IEEE Trans.* 38, 644–664.
- Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife. *Ann. Stat.* 7, 1–26. doi:10.1214/aos/1176344552
- Engbert, R., Kliegl, R., 2003. Microsaccades uncover the orientation of covert attention. *Vision Res.* 43, 1035–1045. doi:10.1016/S0042-6989(03)00084-1
- Flandrin, P., Rilling, G., Goncalves, P., 2004. Empirical Mode Decomposition as a Filter Bank. *IEEE Signal Process. Lett.* 11, 112–114. doi:10.1109/LSP.2003.821662
- Frenkel, D., Smit, B., 2001. Understanding molecular simulation: from algorithms to applications. Academic Press.
- Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–80. doi:10.1016/j.tics.2005.08.011
- Ghil, M., 2002. Advanced spectral methods for climatic time series. *Rev. Geophys.* 40, 1003. doi:10.1029/2000RG000092
- Gilden, D., Thornton, T., Mallon, M., 1995. 1/f noise in human cognition. *Science* (80-.). 267, 1837–1839. doi:10.1126/science.7892611
- Gilden, D.L., 2001. Cognitive emissions of 1/f noise. *Psychol. Rev.* 108, 33–56. doi:10.1037/0033-295X.108.1.33
- Gips, B., van der Eerden, J.P.J.M., Jensen, O., 2016. A Biologically Plausible Mechanism for Neuronal Coding Organized by the Phase of Alpha Oscillations. *Eur. J. Neurosci.* doi:10.1111/ejn.13318
- Haegens, S., Nácher, V., Luna, R., Romo, R., Jensen, O., 2011. α -Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc. Natl. Acad. Sci. U. S. A.* 108, 19377–82. doi:10.1073/pnas.1117190108
- Harrison, S.J., Woolrich, M.W., Robinson, E.C., Glasser, M.F., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2015. Large-scale Probabilistic Functional Modes from resting state fMRI. *Neuroimage* 109, 217–231. doi:10.1016/j.neuroimage.2015.01.013
- Hastings, W.K., 1970. Monte carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97–109. doi:10.1093/biomet/57.1.97
- Henrie, J., Shapley, R., 2005. LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J. Neurophysiol.* 479–490. doi:10.1152/jn.00919.2004.
- Huang, N.E., 2005. Introduction to the Hilbert-Huang Transform and its related mathematical problems. *Hilbert-Huang Transform Its Appl.* 1–26. doi:doi:10.1142/9789812703347_0001
- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 454, 903–995. doi:10.1098/rspa.1998.0193
- Ito, J., Maldonado, P., Singer, W., Grün, S., 2011. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. *Cereb. Cortex* 21, 2482–97.

doi:10.1093/cercor/bhr020

- Jensen, O., Gips, B., Bergmann, T.O., Bonnefond, M., 2014. Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci.* 37, 357–69. doi:10.1016/j.tins.2014.04.001
- Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science* 220, 671–80. doi:10.1126/science.220.4598.671
- Klimesch, W., Sauseng, P., Hanslmayr, S., 2007. EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi:10.1016/j.brainresrev.2006.06.003
- Le Van Quyen, M., Foucher, J., Lachaux, J., Rodriguez, E., Lutz, A., Martinerie, J., Varela, F.J., 2001. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *J. Neurosci. Methods* 111, 83–98. doi:http://dx.doi.org/10.1016/S0165-0270(01)00372-7
- Li, Y., Ma, Z., Lu, W., Li, Y., 2006. Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach. *Physiol. Meas.* 27, 425.
- Lisman, J.E., Jensen, O., 2013. The θ - γ neural code. *Neuron* 77, 1002–16. doi:10.1016/j.neuron.2013.03.007
- Lowet, E., Roberts, M.J., Bosman, C.A., Fries, P., De Weerd, P., 2016. Areas V1 and V2 show microsaccade-related 3-4-Hz covariation in gamma power and frequency. *Eur. J. Neurosci.* 43, 1286–96. doi:10.1111/ejn.13126
- Lyons, R., 2003. dsp tips & tricks - the sliding DFT. *IEEE Signal Process. Mag.* 20, 74–80. doi:10.1109/MSP.2003.1184347
- Maier, A., Adams, G.K., Aura, C., Leopold, D. a, 2010. Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. *Front. Syst. Neurosci.* 4, 1–11. doi:10.3389/fnsys.2010.00031
- Mazaheri, A., Jensen, O., 2008. Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J. Neurosci.* 28, 7781–7787. doi:10.1523/JNEUROSCI.1631-08.2008
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* 21, 1087–1092. doi:doi:10.1063/1.1699114
- Miller, K.J., Sorensen, L.B., Ojemann, J.G., den Nijs, M., 2009. Power-law scaling in the brain surface electric potential. *PLoS Comput. Biol.* 5, e1000609. doi:10.1371/journal.pcbi.1000609
- O’Keefe, J., Recce, M.L., 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3, 317–30. doi:10.1002/hipo.450030307
- Pettersen, K.H., Devor, A., Ulbert, I., Dale, A.M., Einevoll, G.T., 2006. Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. *J. Neurosci. Methods* 154, 116–33. doi:10.1016/j.jneumeth.2005.12.005
- Pfurtscheller, G., Lopes da Silva, F.H., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–57. doi:10.1016/S1388-2457(99)00141-8
- Rolfs, M., 2009. Microsaccades: small steps on a long way. *Vision Res.* 49, 2415–41. doi:10.1016/j.visres.2009.08.010

- Sejdić, E., Djurović, I., Jiang, J., 2009. Time–frequency feature representation using energy concentration: An overview of recent advances. *Digit. Signal Process.* 19, 153–183. doi:10.1016/j.dsp.2007.12.004
- Tiesinga, P., Sejnowski, T.J., 2009. Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? *Neuron* 63, 727–32. doi:10.1016/j.neuron.2009.09.009
- van Kerkoerle, T., Self, M.W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., Roelfsema, P.R., Kerkoerle, T. Van, Self, M.W., Dagnino, B., 2014. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci.* 111, 14332–14341. doi:10.1073/pnas.1402773111/-/DCSupplemental
- Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. doi:10.1038/35067550
- Vautard, R., Ghil, M., 1989. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Phys. D Nonlinear Phenom.* 35, 395–424. doi:10.1016/0167-2789(89)90077-8
- Vautard, R., Yiou, P., Ghil, M., 1992. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Phys. D Nonlinear Phenom.* 58, 95–126. doi:10.1016/0167-2789(92)90103-T
- Woody, C.D., 1967. Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Med. Biol. Eng.* 5, 539–554. doi:10.1007/BF02474247
- Xing, D., Yeh, C.-I., Burns, S., Shapley, R.M., 2012. Laminar analysis of visually evoked activity in the primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 109, 13871–6. doi:10.1073/pnas.1201478109
- Zheng, C., Colgin, L.L., 2015. Beta and Gamma Rhythms Go with the Flow. *Neuron* 85, 236–237. doi:10.1016/j.neuron.2014.12.067

8 Appendices

8.1.1 Appendix A. Parallel tempering SWM algorithm

	Initialization:
1	Define window length (L) and minimal window spacing (G).
2	Define Q different temperatures T_1, T_2, \dots, T_Q with T_q decreasing from T_1 to T_Q .
3	Initialize window starting positions $\mathbf{x}_q = [X_1, X_2, \dots, X_N]$ for every temperature T_q as in Algorithm 1.
4	Calculate initial cost $J_{q,0}$ (Eq. 1) for every parallel temperature;
5	while $i_{out} < \text{maxIterations}$ do
6	for $i = 1:\text{maxSubIterations}$
7	Perform while-loop in algorithm 1 (lines 4–18) for every parallel temperature separately
8	end
9	Pick a pair of neighbouring temperatures at random: T_q and T_{q+1} .

```

10       $p_{\text{swap}} = \exp(k * (J_q - J_{q+1}) / (T_q - T_{q+1}))$ 
11      if       $\text{rand}(0,1) < p_{\text{swap}}$ 
12          Swap states between parallel temperatures:
13           $\mathbf{x}' = \mathbf{x}_q$ 
14           $\mathbf{x}_q \leftarrow \mathbf{x}_{q+1}$ 
15           $\mathbf{x}_{q+1} \leftarrow \mathbf{x}'$ 
16      end

17      for       $q = 1:Q$ 
18          Move all positions in  $\mathbf{x}_q$  randomly by the same amount/step
          (instead of only a single window as in (line 5-6 in
          Algorithm 1):  $\mathbf{x}' = \mathbf{x}_q + \text{step}$ 
19          Calculate the new cost  $J'$  using  $\mathbf{x}'$  and with this the change
          in cost function  $\Delta J = J_{q,i-1} - J'$ ;
20          Calculate acceptance probability  $p_{\text{accept}} = \exp(\Delta J / T)$ 
21          if  $\text{rand}(0,1) < p_{\text{accept}}$  then
22               $J_{q,i} \leftarrow J'$ 
23               $\mathbf{x}_q \leftarrow \mathbf{x}'$ 
24          end
25      end
26       $i_{\text{out}} \leftarrow i_{\text{out}} + 1$ 
27  end

```

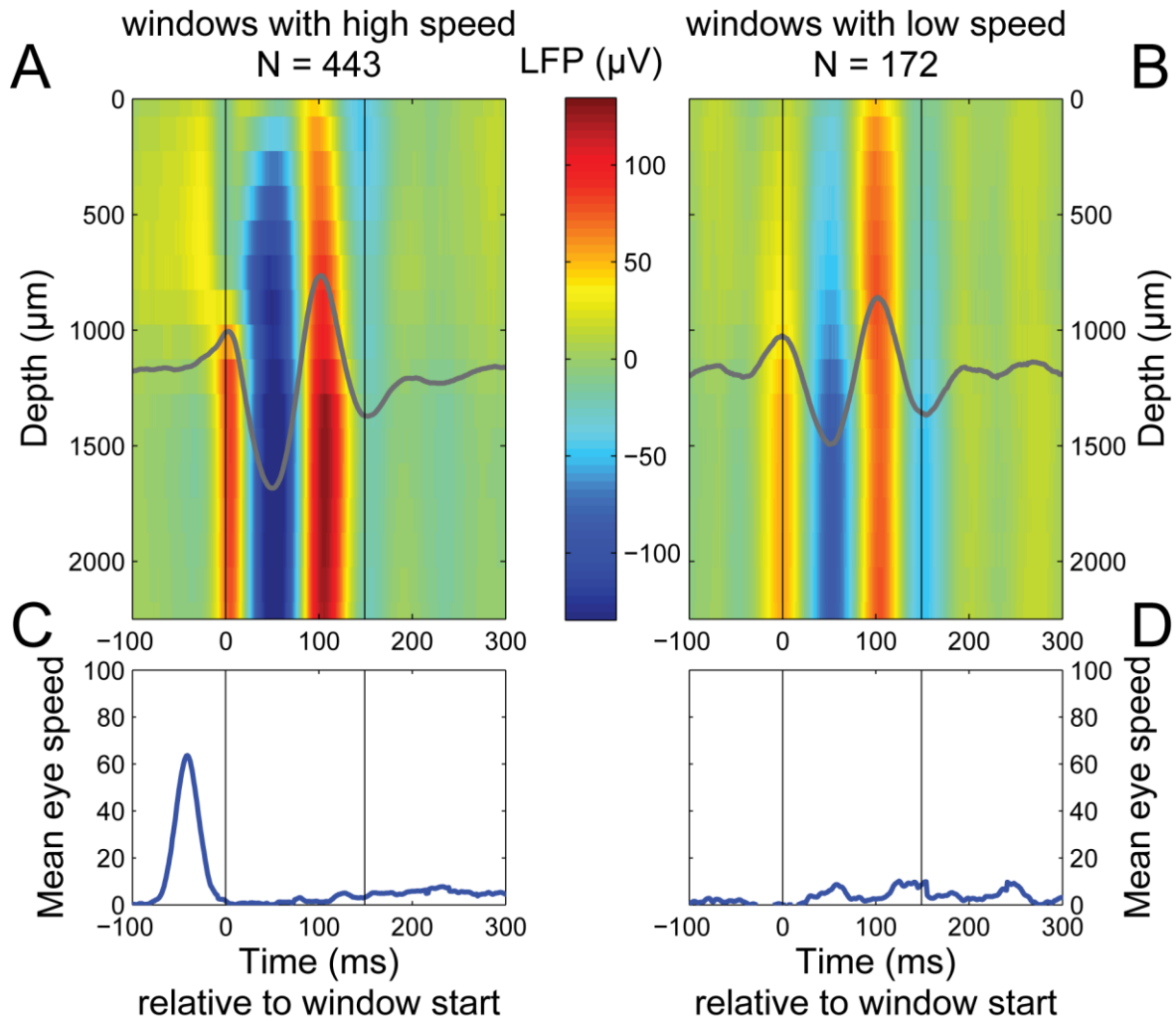
Algorithm 2. Combining Algorithm 1 with Parallel Tempering to better find the global minimum of cost function J .

9 Supplemental material

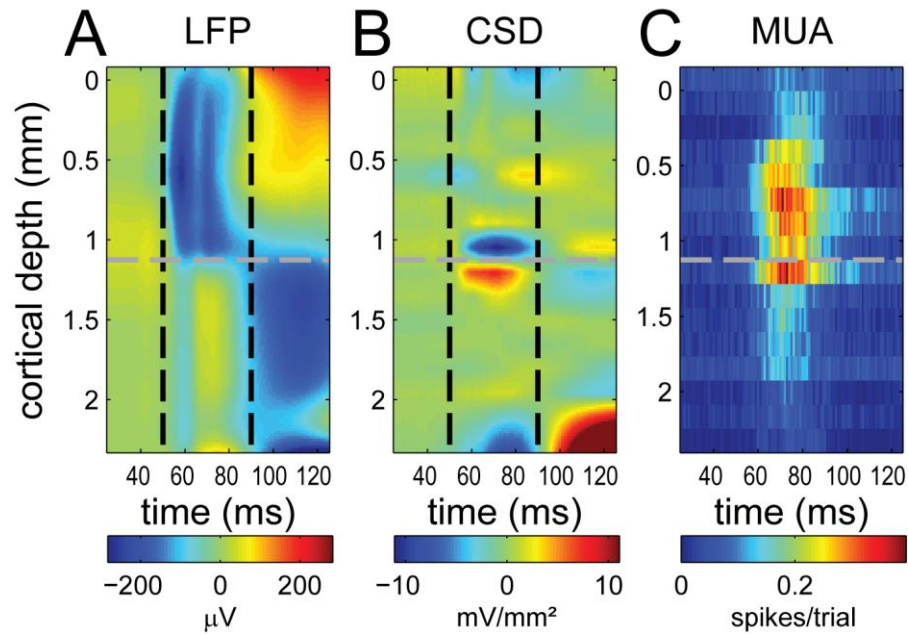
Calculation of J is equivalent to one minus the mean of the off-diagonal elements in the Pearson correlation matrix of all windows, multiplied by the number of windows (N). This means that the maximum of J is equal to N (all windows are completely uncorrelated, the correlation matrix is an identity matrix) and the minimum value for J is 0 when all windows are fully correlated (the correlation matrix is a unit matrix, the contents of all windows are perfectly correlated). See equation 6 for this equivalent definition (where Q_{ij} is the element on the row i and column j in the correlation matrix calculated between all windows, see equation 6).

$$J = N * \left(1 - \frac{2 * \sum_{i=1}^N \sum_{j=i+1}^N Q_{ij}}{N * (N - 1)} \right) \quad (6)$$

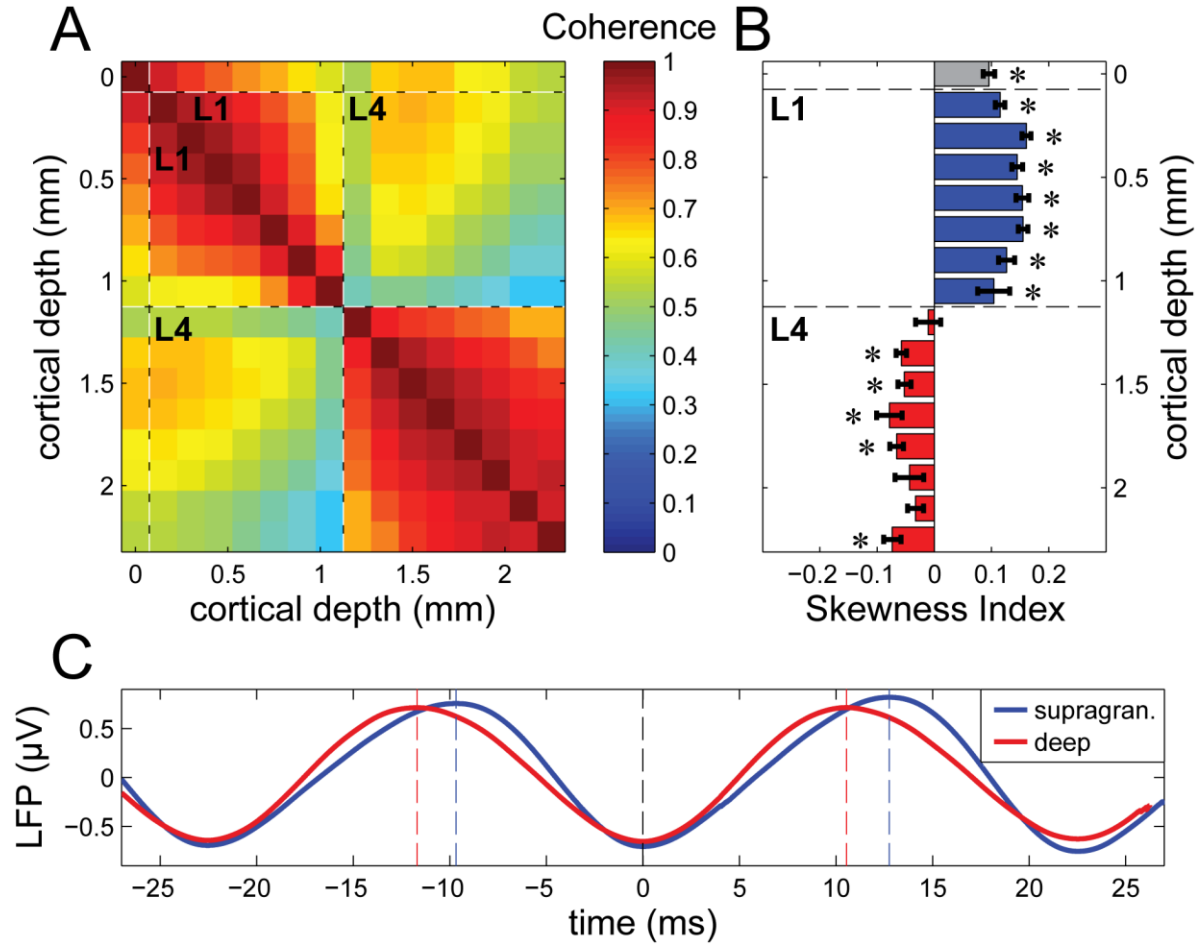
$$Q_{ij} = \frac{\mathbf{z}_i \cdot \mathbf{z}_j}{L - 1} \quad (7)$$



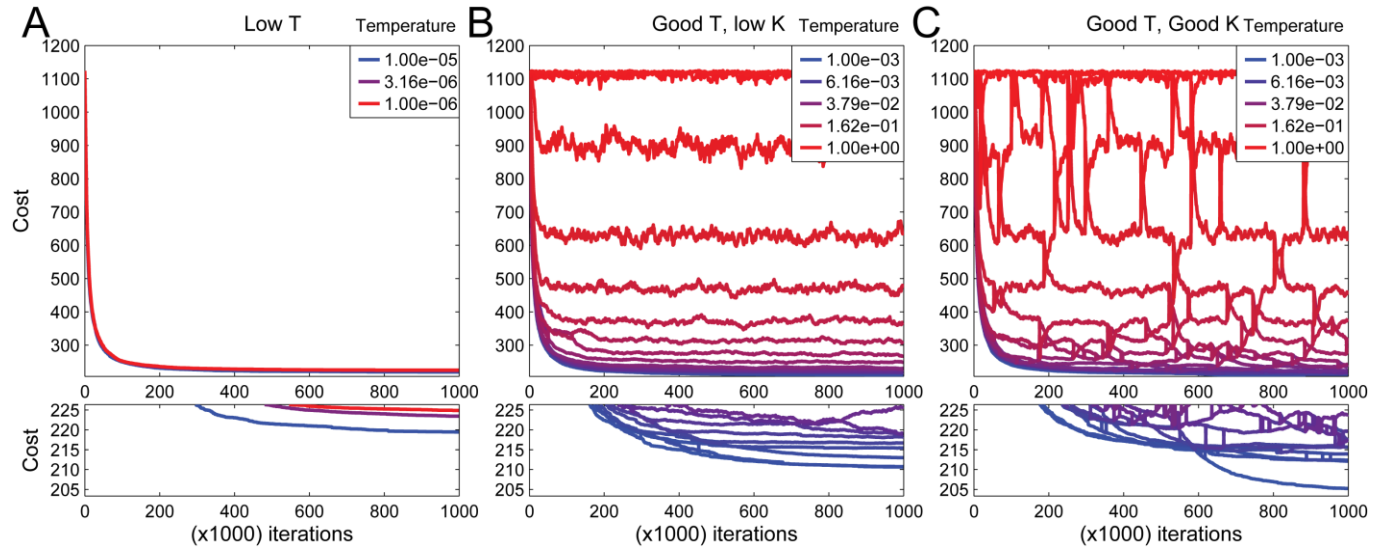
S 1. Splitting up SWM windows based on eye velocity.



S 2. Detection of layer 4/5 boundary and corresponding gamma skewness shift. The position of the layer 4/5 boundary was taken as the position of the first sink (blue) – source (red) reversal in the CSD (panel B) shortly after stimulus onset (50-90 ms after stimulus onset, see black dashed lines). This position is indicated by a gray dashed line across all panels. Depth 0 corresponds to the top of the cortex, the location where the first negative visually evoked response (A) and first spiking response was observed (C). **A** Local field potential as a function of cortical depth after stimulus onset averaged across trials. Interpolated along the laminar axis using cubic spline interpolation. **B** The negative value of the second derivative of A along the laminar axis (an estimate of the current source density, CSD. Blue color indicate sinks, red color indicate sources). **C** Heat map showing multi unit activity.



S 3. Properties of stimulus induced gamma-band activity in LFP recordings in monkey V1. **A** Spectral coherence across cortical depth in the gamma band (40-50Hz). **B-C** Analogous to **Figure 7B-C**, but obtained by using PA with central frequency f_0 of 45 Hz with a bandwidth of $1/(2\pi\sigma) = 2$ Hz (Eq. 5) instead of SWM.



S 4. Examples of cost function trajectories. Illustration how temperature parameters change the accuracy of SWM. Each panel (A-C) shows the cost function as a function of iteration number. Red colors indicate the cost of a high temperature state, whereas blue curves describe the cost trajectory of low temperature states (see legends). The top panels show the full trajectories, whereas the bottom panels show the cost value close to the minimum for easier comparison between the three different parameter sets. All three cases were performed on the same data set. **A** The amount of different temperatures is set low (3 different temperatures were used, see legend) and the individual temperatures are too low ($< 10^{-5}$) SWM will quickly converge to a minimum, but this will be a local one. **B** The amount of different temperatures (20) and their values is set well, spaced logarithmically between 1 and 10^{-3} . The highest temperature stay at the maximum cost value, whereas decreasing temperatures span the space between the maximum and the found minimum. However, the parameter that determines the probability of swapping temperatures between states (k in line 10 of Algorithm 2) is set too low causing this swapping to never happen. This makes the algorithm run independently for all 20 temperatures and in essence not use parallel tempering at all. **C** Both temperatures (same values as in B) and k (200) are set properly. Parallel tempering is in effect as can be witnessed by the jumps in the cost trajectories indicating swapping of states between all temperatures. The minimum found in this last case is therefore lowest of the three.