A multivariate, multitaper approach to detecting and estimating harmonic response in cortical optical imaging data

A.T. Sornborger [a,*], T. Yokoo [b]

[a] Department of Mathematics and Faculty of Engineering, University of Georgia, Athens, GA 30602, USA
[b] Division of Abdominal Imaging and Intervention, MGH Imaging, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02115, USA

ARTICLE INFO

ABSTRACT

The efficiency and accuracy of cortical maps from optical imaging experiments have been improved using periodic stimulation protocols. The resulting data analysis requires the detection and estimation of periodic information in a multivariate dataset. To date, these analyses have relied on discrete Fourier transform (DFT) sinusoid estimates. Multitaper methods have become common statistical tools in the analysis of univariate time series that can give improved estimates. Here, we extend univariate multitaper harmonic analysis methods to the multivariate, imaging context. Given the hypothesis that a coherent oscillation across many pixels exists within a specified bandwidth, we investigate the problem of its detection and estimation in noisy data by constructing Hotelling's generalized T^2-test. We then extend the investigation of this problem in two contexts, that of standard canonical variate analysis (CVA) and that of generalized indicator function analysis (GIFA) which is often more robust in extracting a signal in spatially correlated noise. We provide detailed information on the fidelities of the mean estimates found with our methods and comparison with DFT estimates. Our results indicate that GIFA provides particularly good estimates of harmonic signals in spatially correlated noise and is useful for detecting small amplitude harmonic signals in applications such as biological imaging measurements where spatially correlated noise is common. We demonstrate the power of our methods with an optical imaging dataset of the periodic response to a periodically rotating oriented drifting grating stimulus experiment in cat visual cortex.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In a pair of papers, Kalatsky and Stryker have demonstrated that by using a periodic experimental design one can efficiently reconstruct cortical maps with relatively short datasets (Kalatsky and Stryker, 2003; Kalatsky et al., 2005) using Fourier techniques. In similar contexts, statistical, multitaper Fourier techniques have been used to distinguish differential and average dynamics in periodically ordered optical imaging data (Sornborger et al., 2005; Pesaran et al., 2004). Any time-periodic signal may be represented as the sum of a set of sinusoids equally spaced in frequency, usually called harmonics. It is of considerable interest to have accurate and robust methods for their detection and estimation. Therefore, in this paper, we present new multitaper-based, multivariate methods for sensitively detecting and estimating harmonic content in noisy optical imaging data.

Multitaper methods for digital spectral and harmonic analysis have seen widespread application since their introduction (Thomson, 1982, 2000; Mitra and Pesaran, 1999; Tauxe, 1993; Thomson and Schild, 1997). Multitaper methods were introduced for bioimaging applications by Mitra and Pesaran (1999). We have used the univariate multitaper harmonic analysis method with considerable success to reconstruct the response to stimulus in stimulus-locked uni- and multivariate (imaging) data (Sornborger et al., 2003a, 2005; Pesaran et al., 2004). We have also applied these methods to reconstruct periodically propagating electrical potentials on the mouse heart imaged with voltage sensitive dyes (Sornborger et al., 2003b).

Multitaper spectral analysis has been considered recently in the multivariate context (Walden, 2000). The purpose of spectral analysis is primarily to understand variance in the data. Bivariate, multitaper coherency estimates have been used extensively in electrophysiological data (Mitra and Bokil, 2008; Pesaran et al., 2002). Multivariate, multitaper methods have been investigated for the detection of coherence in imaging data (Mitra and Pesaran, 1999; Mitra and Bokil, 2008). However, it is often also important to detect and estimate coherent oscillations in noisy data. This paper considers the generalization of multitaper harmonic analysis methods to

* Corresponding author. Tel.: +1 706 583 0854; fax: +1 706 542 5907.
  E-mail addresses: ats@math.uga.edu (A.T. Sornborger),
tyokoo@partners.org (T. Yokoo).

the multivariate context, i.e. the detection and estimation of coherent oscillations in noisy imaging data.

In order to motivate the usefulness of multivariate methods consider the commonly used method of calculating a region-of-interest (ROI) used by many experimentalists. Typically, a user manually defines a region within which a consistent response is believed to exist, then averages all pixels in the region to improve the signal-to-noise ratio (SNR). Multivariate statistical methods are designed to perform a similar averaging over correlated variables (pixel timecourses in imaging data) that avoids manual definition of an ROI. Multivariate techniques can improve signal estimation significantly relative to ROI methods because they make use of the covariance structure of the data to determine spatial regions that contain the signal of interest. Such regions can easily be overlooked when ROIs are determined manually. Particularly in low SNR data, the user will often define an ROI in places where the background is large but other areas where signal may be found are overlooked.

In this paper, we introduce a new, fully multivariate set of multitaper harmonic analysis methods beginning with the introduction of Hotelling's T^2-statistic for the detection of multivariate harmonics using multitaper methods. We then present two auxiliary multivariate statistical optimization techniques in the framework of multitaper analysis, that can aid signal estimation. The first of the two is the classical canonical variate analysis (CVA) method. The second method is based on a regularization method called generalized indicator function analysis (GIFA) (Yokoo et al., 2001). GIFA has been developed to circumvent the problem of matrix inversion and often produces higher fidelity estimates than CVA.

The main contribution of this paper is that our methods are exceptionally good at estimating multivariate harmonic signals in data with spatially correlated noise. Spatially correlated noise often arises in biological imaging data due to the fact that noise processes are typically physiologically related. A typical example in optical imaging measurements of the brain is the large contribution to the variance in the vasculature due to vasomotion, respiration and the cardiac pulse (Sornborger et al., 2003a). Spatially and temporally correlated noise is also seen in MEG/EEG data (de Munck et al., 2002).

In Section 2, we present our multivariate harmonic detection and estimation methods. In Section 3, we present a detailed exploration of the dependence of the fidelity of the estimated harmonic content on the number of dimensions of the multivariate data, the time–bandwidth product, the amplitude of the signal and the amount of spatial correlation in the noise. For all of our simulated data, we use an AR(2) random noise process that is commonly used as a test case in the literature (Ulrych and Bishop, 1975; Box and Jenkins, 1976). Finally, we apply our methods to the detection and estimation of the response to periodic stimulation in optical imaging data of cat primary visual cortex.

2. Methods

2.1. The probabilistic model

We consider a vector of P, real-valued, stationary random processes, $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_P(t)]^\dagger$ with $t = [1, \ldots, T]$, each of whose mean is varying in time. In imaging data, P would denote the number of pixels in an image and each vector would represent an image. More generally, P could denote the number of electrodes in multielectrode electrophysiological data, or the number of voxels in an fMRI dataset. Throughout this paper, † indicates the hermitian transpose and * the complex conjugate. Our data consist of a realization of the vector process, $\mathbf{x}(t)$, sampled discretely in time. We wish to determine the existence of a sinusoidal signal at a frequency ω in our multivariate data and estimate its complex amplitude. In

many experimental situations, this means ω is known. For instance, a movie of concatenated measurements of the response to a stimulus, measurement of the first-order response to a forced oscillation or the beating of a heart for a short duration (20 or so periods).

We model our vector process as

$$
\begin{aligned}
\mathbf{X}(t) &= E\{\mathbf{X}(t)\} + \boldsymbol{\eta}(t) \\
&= \boldsymbol{\mu}e^{i\omega t} + \boldsymbol{\mu}^* e^{-i\omega t} + \boldsymbol{\eta}(t)
\end{aligned}
\tag{1}
$$

Here, $\boldsymbol{\mu}$ is a fixed vector of length P, $\boldsymbol{\eta}$ is a random vector of length P and ω is the frequency at which we wish to detect a signal. We note here that, although for the mathematical development of our method we focus on a single frequency ω, in practice, we vary ω across a fine mesh of frequencies, searching non-parametrically for harmonic content.

An equivalent Fourier representation of (1) evaluated at frequency ω is

$$
FT\{\mathbf{X}(t)\}|_\omega = \boldsymbol{\mu} + \boldsymbol{\epsilon}.
\tag{2}
$$

where $FT\{\cdot\}$ is the standard Fourier transform operator, $\epsilon = FT\{\boldsymbol{\eta}(t)\}|_\omega$ is a P-dimensional random vector and $E\{\epsilon\} = 0$. We model the noise $\epsilon = [\epsilon_1, \ldots, \epsilon_P]^\dagger$ as a P-dimensional complex normal process, $\boldsymbol{\epsilon} \sim N_P^c(\mathbf{0}, \boldsymbol{\Sigma})$, where the $P \times P$ covariance matrix $Cov\{\epsilon\} = E\{\epsilon\epsilon^\dagger\} \equiv \boldsymbol{\Sigma}$. With this model, $\boldsymbol{\mu} + \boldsymbol{\epsilon} \sim N_P^c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Therefore, in the multitaper context, our goal is the detection and the estimation of the complex amplitude $\boldsymbol{\mu}$, given M tapered eigenestimates of the signal $\mathbf{X}(t)$ at frequency ω.

Further modeling of the signal, such as its phase is not necessarily helpful in optical imaging data if the delay of the hemodynamic response varies between individuals. For this reason, instead of modeling the phase, we use our methods to calculate an estimate of both the amplitude and phase of the harmonic signal from the data.

2.2. Multivariate, multitaper mean and covariance estimation

Multitaper harmonic analysis methods (Thomson, 1982; Percival and Walden, 1993) make use of a set of data tapers $h_m(t)$, $m = [1, \ldots, M]$ to window the data before applying a Fourier transform. The tapers are called Slepian sequences; the important property of the tapers is that they are solutions to a frequency concentration problem and are ordered in terms of decreasing frequency concentration and are maximally-concentrated as long as $M < 2TW - 2$. In our analyses, we will always take $M = 2TW - 3$. Because of the frequency concentration of the tapers, tapered spectral estimates at a frequency f exhibit minimal leakage from frequencies outside $[f - W, f + W]$. Multitaper harmonic estimates make use of a regression estimate over the tapers in order to obtain statistically consistent mean harmonic amplitude and phase estimates at a given frequency. For a complete description of univariate harmonic estimation, see (Thomson, 1982; Percival and Walden, 1993). The univariate multitaper approach may be generalized immediately with the introduction of the multivariate data model presented above. Our notation follows Percival and Walden (1993). Univariate spectral eigenestimates, J_m, generalize to vector eigenestimates

$$
\mathbf{J}_m = \boldsymbol{\mu}H_m + \boldsymbol{\epsilon}_m
\tag{3}
$$

where

$$
\mathbf{J}_m = \sum_t \mathbf{X}(t) h_m(t) e^{-i\omega t},
\tag{4}
$$

and

$$
\boldsymbol{\epsilon}_m = \sum_t \boldsymbol{\eta}(t) h_m(t) e^{-i\omega t}.
\tag{5}
$$

where h_m is the mth taper (Slepian sequence) and $H_m = \sum_t h_m(t)$.

If we can assume that the spectrum of the noise is continuous, there exists a positive constant W such that the spectrum is approximately white within the interval $(\omega - W, \omega + W)$. In this case, the ϵ_m terms can be shown (Percival and Walden, 1993) to be pairwise uncorrelated (independent under the assumption of normally distributed noise) and the covariance

$$E\{\boldsymbol{\epsilon}_m \boldsymbol{\epsilon}_m^\dagger\} = E\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\dagger\} \equiv \boldsymbol{\Sigma} \tag{6}$$

for all m.

The least squares mean estimate of the signal $\boldsymbol{\mu}$ is given by

$$\hat{\boldsymbol{\mu}} = \sum_{m=1}^{M} \frac{\mathbf{J}_m H_m}{\mathcal{H}}, \tag{7}$$

where $\mathcal{H} = \sum_{l=1}^{M} H_l^2$.

Since the tapered residuals ϵ_m are distributed as $N_P^c(\mathbf{0}, \boldsymbol{\Sigma})$ for any m and $\{\mathbf{J}_1, \ldots, \mathbf{J}_M\}$ are independent samples of $\boldsymbol{\epsilon}$ under the null hypothesis, $\{H_0 : \boldsymbol{\mu} = \mathbf{0}\}$, an unbiased estimate of $\boldsymbol{\Sigma}$ may be calculated as

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{M-1} \sum_{m=1}^{M} (\mathbf{J}_m - \hat{\boldsymbol{\mu}} H_m)(\mathbf{J}_m - \hat{\boldsymbol{\mu}} H_m)^\dagger. \tag{8}$$

For later convenience, we now define $(M-1)\hat{\boldsymbol{\Sigma}} \equiv \hat{\mathbf{K}} \sim \sum_m \boldsymbol{\epsilon}_m \boldsymbol{\epsilon}_m^\dagger$, the sum of squares and cross-product matrix of the residuals.

2.3. Testing for statistical significance

Given a realization, $\mathbf{x}(t)$, and its harmonic eigenestimates, $\{\mathbf{j}_1, \ldots, \mathbf{j}_M\}$, of $\{\mathbf{J}_1, \ldots, \mathbf{J}_M\}$, signal detection of the Fourier amplitude $\boldsymbol{\mu}$ at ω can be performed as a statistical test on the mean estimate $\hat{\boldsymbol{\mu}}$ against the null (zero-mean) hypothesis, $\{H_0 : \boldsymbol{\mu} = \mathbf{0}\}$, using the multivariate t-test, known as Hotelling's T^2-test. By assumption, the noise is multi-normally distributed, $\boldsymbol{\epsilon} \sim N_P^c(\mathbf{0}, \boldsymbol{\Sigma})$. Therefore, under the null hypothesis H_0, the mean estimate $\hat{\boldsymbol{\mu}}$ is distributed as

$$\hat{\boldsymbol{\mu}} \sim N_P^c \left(\mathbf{0}, \frac{\boldsymbol{\Sigma}}{\mathcal{H}} \right), \tag{9}$$

Using the above mean and variance estimates, Hotelling's T^2-statistic may be constructed (Giri, 1965) as

$$T^2 = \mathcal{H} \hat{\boldsymbol{\mu}}^\dagger \hat{\mathbf{K}}^{-1} \hat{\boldsymbol{\mu}}, \tag{10}$$

whose null distribution can be used to test H_0 in the usual way.

When $P < M - 1$, that is, the number of tapered eigenestimates is equal to or larger than the number of variables in $\mathbf{X}(t)$, the $P \times P$ sample covariance $\hat{\mathbf{K}}$ is full rank (i.e. invertible) and $T^2 \nu_\epsilon / \nu_s$ is distributed as a central F-variate, F_{ν_s, ν_ϵ}, where $\nu_\epsilon = 2(M - P)$ and $\nu_s = 2P$. If the T^2-statistic from the samples, $\{\mathbf{j}_1, \ldots, \mathbf{j}_M\}$, exceeds the upper $100(1 - \alpha)\%$ point of the distribution of F_{ν_s, ν_ϵ}, where α is a user-defined Type I error probability, we conclude that the mean estimate $\hat{\boldsymbol{\mu}}$ is significant, and we conclude that the signal exists at frequency ω with confidence $1 - \alpha$.

In certain applications, such as imaging, it may be unrealistic to pick a bandwidth that will allow as many frequency concentrated tapers as the number of components of $\boldsymbol{\mu}$. For example, consider movies with $P = 10,000$ pixels per frame, but where the spectrum is only effectively white within a bandwidth $W = 7/T$, giving $M = 2TW - 3 = 11$ tapered eigenestimates $\mathbf{J}_m$. When $P \geq M - 1$, the $P \times P$ sample covariance matrix is singular, and cannot be inverted. This apparent difficulty can be circumvented by noting that the data spans a smaller subspace of $\mathbb{C}^P$ and subsequent analysis can be performed within this subspace. The standard method (Friston et al.,

1995) to do this is to construct a pseudoinverse by first forming a mean-corrected eigenestimate matrix with rank $M - 1$

$$\mathbf{C} = [\mathbf{j}_1 - \hat{\boldsymbol{\mu}} H_1, \ldots, \mathbf{j}_M - \hat{\boldsymbol{\mu}} H_M]. \tag{11}$$

This matrix is a P by M matrix where the mean (times a weighting factor H_i) has been subtracted from the eigenestimates and therefore each column represents an estimate of the noise. Then, we perform a singular value decomposition

$$\mathbf{C} = \boldsymbol{\Psi}\mathbf{D}\mathbf{A}^\dagger. \tag{12}$$

where $\boldsymbol{\Psi}$ is a $P \times (M - 1)$ matrix of orthonormal left eigenvectors, $\mathbf{D}$ is a diagonal matrix of $M - 1$ positive eigenvalues, and $\mathbf{A}$ is an $(M - 1) \times M$ matrix of orthonormal right eigenvectors. This decomposition allows us to represent the P-dimensional data in an $M - 1$ dimensional space, i.e. the number of variables has been reduced to $M - 1$. The covariance matrix is now non-singular and Hotelling's T^2-statistic is

$$T^2 = \mathcal{H} \hat{\boldsymbol{\mu}}^\dagger \boldsymbol{\Psi}\mathbf{D}^{-2}\boldsymbol{\Psi}^\dagger \hat{\boldsymbol{\mu}} \tag{13}$$

and is distributed as

$$T^2/(M - 1) \sim F_{2(M-1),2}. \tag{14}$$

It is crucial to note that the SVD (12) is lossless, i.e. no eigenvectors were truncated from the space of the matrix $\mathbf{C}$. Here, the SVD is used only for the calculation of the squared pseudoinverse, $\mathbf{D}^{-2}$, in the space of the eigenestimates.

In summary, when the number of tapered estimates is small relative to the number of pixels in the dataset, as is typical in imaging data, statistical significance must be estimated in the space spanned by the vectors making up the mean-corrected eigenestimate matrix (i.e. vectors representing noise within the bandwidth W).

2.4. Auxiliary methods

In the above discussion, we set forward a method for detecting harmonic signal in noisy, multivariate data. Once the statistical significance of $\hat{\boldsymbol{\mu}}$ is established, we would like to look for the components of $\hat{\boldsymbol{\mu}}$ that gave rise to the large T^2 value. To do this, we consider an unknown unit vector $\boldsymbol{\phi}$ in $\mathbb{C}^P$, and its projection onto $\hat{\boldsymbol{\mu}}$. For multinormally distributed noise, ϵ, and arbitrary vector $\boldsymbol{\phi}$, the distribution of $\boldsymbol{\phi}^\dagger \hat{\boldsymbol{\mu}}$ is known to be univariate normal with parameters,

$$\boldsymbol{\phi}^\dagger \hat{\boldsymbol{\mu}} \sim N(\boldsymbol{\phi}^\dagger \boldsymbol{\mu}, \boldsymbol{\phi}^\dagger \boldsymbol{\Sigma} \boldsymbol{\phi}/\mathcal{H}) \tag{15}$$

(see (Anderson, 1984), Chapter 2, Theorem 2.4.1). Since $\hat{\boldsymbol{\mu}}$ is an estimate, it contains both signal and noise components. It is our goal to find a vector $\boldsymbol{\phi}$ that best represents the signal content in $\hat{\boldsymbol{\mu}}$. It is important to stress here that in this context $\boldsymbol{\phi}$ is not necessarily an estimate of the mean vector $\boldsymbol{\mu}$.

A natural measure of the signal power that is parallel to $\boldsymbol{\phi}$ is

$$S(\boldsymbol{\phi}) = \boldsymbol{\phi}^\dagger \hat{\mathbf{K}}_s \boldsymbol{\phi} \tag{16}$$

where

$$\hat{\mathbf{K}}_s = \mathcal{H} \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^\dagger, \tag{17}$$

Similarly, a natural measure of the power in the noise, proportional to the variance of the eigenestimates $[\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_M]$ projected along $\boldsymbol{\phi}$, is

$$N(\boldsymbol{\phi}) = \boldsymbol{\phi}^\dagger \hat{\mathbf{K}} \boldsymbol{\phi} = \boldsymbol{\phi}^\dagger \hat{\boldsymbol{\Sigma}} \boldsymbol{\phi}(M - 1). \tag{18}$$

Both $S(\boldsymbol{\phi})$ and $N(\boldsymbol{\phi})$ are positive semi-definite, Hermitian operators with real, non-negative eigenvalues. Furthermore, under the null hypothesis $\hat{\mathbf{K}}$ and $\hat{\mathbf{K}}_s$ are Wishart distributed (Giri, 1965), with parameters $\boldsymbol{\Sigma}$ and $\nu_\epsilon = 2(M - P)$ and $\nu_s = 2P$ degrees of freedom, respectively.

Given the above expressions for signal $S(\boldsymbol{\phi})$ and noise $N(\boldsymbol{\phi})$, there are a number of approaches that we may take in order to find a projection vector $\boldsymbol{\phi}$ with favorable statistical properties. In the following two sections, we present methods based on two approaches, the classical method of canonical variate analysis (CVA) and a regularization method called generalized indicator function analysis (GIFA) (Yokoo et al., 2001) that is often more robust relative to CVA.

2.4.1. Canonical variate analysis

The CVA approach is to maximize the univariate t^2-statistic of the projection,

$$t^2(\boldsymbol{\phi}) = \frac{S(\boldsymbol{\phi})}{N(\boldsymbol{\phi})}. \tag{19}$$

as a function of the projection vector, $\boldsymbol{\phi}$. The ratio t^2 is a Rayleigh quotient whose maximization over $\boldsymbol{\phi} \in \mathbb{C}^P$ results in the generalized eigenvalue problem

$$(\hat{\mathbf{K}}_s - \rho\hat{\mathbf{K}})\boldsymbol{\phi} = 0 \tag{20}$$

For the present application, $\hat{\mathbf{K}}_s$ is a rank one matrix and the rank of $\hat{\mathbf{K}}_s/\hat{\mathbf{K}}$ is consequently also one. Since there is only one non-zero characteristic root

$$tr\{\mathcal{H}\hat{\mathbf{K}}^{-1}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^{\dagger}\} = \rho \tag{21}$$

and therefore

$$\mathcal{H}\hat{\boldsymbol{\mu}}^{\dagger}\hat{\mathbf{K}}^{-1}\hat{\boldsymbol{\mu}} = \rho, \tag{22}$$

which shows that the maximum of $t^2(\boldsymbol{\phi})$ over $\boldsymbol{\phi}$ is the T^2-statistic. Thus, the type 1 error probability, α, and the threshold, τ_α^2, of the T^2-test are related by

$$\begin{aligned} \alpha &= \Pr\{\rho_{max} > \tau_\alpha^2 | H_0\} \\ &= \Pr\{T^2 > \tau_\alpha^2 | H_0\} \end{aligned} \tag{23}$$

where $\rho_{max} = \max(t^2(\boldsymbol{\phi}))$, i.e. the maximum characteristic root.

We therefore identify the eigenvector $\boldsymbol{\phi}_{max}$ with the largest eigenvalue ρ_{max} as the component of $\hat{\boldsymbol{\mu}}$ that, when projected on the data $\{\mathbf{j}_m\}$, has the best signal-to-noise ratio. Somewhat imprecisely, we denote $\hat{\boldsymbol{\mu}}_{CVA} \equiv (\boldsymbol{\phi}_{max}, \hat{\boldsymbol{\mu}})_x^* \boldsymbol{\phi}_{max}$, the "estimated" multivariate signal component.[1] If ρ_{max} is greater than the $100(1-\alpha)\%$ percentile determined by the inverse of the cumulative F-distribution of the T^2-statistic, $\tau_\alpha^2 = (\nu_s/\nu_\epsilon)F_{cdf}^{-1}(1-\alpha, \nu_s, \nu_\epsilon)$, the alternative hypothesis, $H_0 : \boldsymbol{\mu} \neq \mathbf{0}$, is preferred and we retain $\hat{\boldsymbol{\mu}}_{CVA}$ as a statistically significant harmonic component of $\hat{\boldsymbol{\mu}}$. Otherwise, we claim not to have detected any harmonic content.

When $P > M - 1$, the P-dimensional generalized eigenvalue problem can be reduced without loss to an $(M-1)$-dimensional problem and solved. Details for this case are given in the Appendix A.1.

2.4.2. Generalized indicator function analysis

The GIFA approach is to quantify the success of $\boldsymbol{\phi}$ by defining an alternative performance measure called the signal-noise difference (Yokoo et al., 2001).

$$G(\boldsymbol{\phi}; \tau_\alpha^2) = S(\boldsymbol{\phi}) - \tau_\alpha^2 N(\boldsymbol{\phi}), \tag{24}$$

where τ_α^2 is a user-defined, threshold value of Hotelling's T^2-test at level α. This weighting of the noise operator is motivated by the fact (Ennis and Johnson, 1993) that the probability

$$\begin{aligned} \alpha &= \Pr\{T^2 > \tau_\alpha^2 | H_0\} \\ &= \Pr\{\boldsymbol{\phi}^{\dagger}(\hat{\mathbf{K}}_s - \tau_\alpha^2\hat{\mathbf{K}})\boldsymbol{\phi} > 0 | H_0\}, \end{aligned} \tag{25}$$

implying that the GIFA test criterion is identical to the T^2 test.

If we maximize $G(\boldsymbol{\phi}; \tau_\alpha^2)$ over all $\boldsymbol{\phi} \in \mathbb{C}^P$ with the constraint that $\boldsymbol{\phi}$ be a unit vector, $\boldsymbol{\phi}^{\dagger}\boldsymbol{\phi} = 1$, the standard Lagrangean procedure leads to a regular, symmetric eigenvalue problem

$$(\hat{\mathbf{K}}_s - \tau_\alpha^2\hat{\mathbf{K}})\boldsymbol{\phi} = \gamma\boldsymbol{\phi}, \tag{26}$$

where γ is the eigenvalue and $\boldsymbol{\phi}$ is the eigenvector. There is at most one positive solution eigenpair when Rank$(\hat{\mathbf{K}}_s) = 1$. Observe that a solution $\boldsymbol{\phi}$ of this eigenvalue problem satisfies the property

$$S(\boldsymbol{\phi}) - \tau_\alpha^2 N(\boldsymbol{\phi}) = \gamma. \tag{27}$$

Rearranging and using the non-negativity of $N(\boldsymbol{\phi})$, we have

$$\frac{S(\boldsymbol{\phi})}{N(\boldsymbol{\phi})} = \tau_\alpha^2 + \frac{\gamma}{N(\boldsymbol{\phi})} > \tau_\alpha^2 \tag{28}$$

if $\gamma > 0$. Therefore, rejection of H_0 by Hotelling's T^2-test, i.e. $T^2 < \tau_\alpha^2$, is equivalent to the existence of a positive eigenvalue of (26). Then the eigenvector associated with $\gamma > 0$ can be identified as the signal component that is responsible for the large T^2 value. This method has an advantage over CVA when the covariance matrix is singular, since solving (26) does not involve the inversion of a matrix.

For this method, $\hat{\boldsymbol{\mu}}_{GIFA} \equiv (\boldsymbol{\phi}_{\gamma>0}, \hat{\boldsymbol{\mu}})_x^* \boldsymbol{\phi}_{\gamma>0}$ gives the estimated multivariate signal component.

When $P > M - 1$, the P-dimensional generalized eigenvalue problem can be reduced to an $(M-1)$-dimensional problem and solved. Details for this case are given in Appendix.

3. Results

In order to quantify the success of our approach, we applied our methods to simulated data. We embedded a known signal in simulated noise, and varied the signal amplitude, the number of variables (e.g. pixels) in the data and the number of tapers used, i.e. the desired frequency resolution. Because biological signals, the signals of most interest to us, often contain spatially correlated noise, we compared the results of our methods for spatially uncorrelated noise with spatially correlated noise where both spatial signals were in a background with a colored (temporally correlated) frequency spectrum. Finally, we tested our methods on experimental measurements of the intrinsic optical signal in cat visual cortex.

3.1. Spatially uncorrelated noise

In order to study a harmonic signal in spatially uncorrelated noise, we generated $T = 500$ images of noise $\eta(t)$ from the temporally correlated AR(2) process

$$\eta(t) = \frac{3}{4}\eta(t-1) - \frac{1}{2}\eta(t-2) + \mathbf{n}(t), \tag{29}$$

where $\mathbf{n}(t)$ is a multinormally distributed random vector independent of t with covariance matrix $\mathbf{I}$. To this was added a separable signal $\mathbf{s}(t)$, that consisted of a one dimensional 'image' with P pixels taking values in $\mathbf{s}(t) = c[1, 0, -1, 0, 1, 0, -1, \ldots]$ whose amplitude, $c = c(t) = \mu \sin 2\pi kt/T$ was sinusoidally modulated in time. Combining the signal and the noise, our data consisted of the sum

$$\mathbf{x}(t) = \mathbf{s}(t) + \eta(t). \tag{30}$$

The simulated data, therefore, had a constant average SNR $(= \sqrt{||\boldsymbol{\mu}||^2/\text{tr}\,\boldsymbol{\Sigma}})$ per pixel. Fig. 1 shows typical simulated signal and noise pixel timecourses that we tested our methods against.

In Fig. 2, we show signal reconstruction fidelities for our methods. Fidelities, $B(P, TW, \mu, \alpha) = (\hat{\boldsymbol{\mu}}, \boldsymbol{\mu})/|\boldsymbol{\mu}||\hat{\boldsymbol{\mu}}|$, indicate how well-correlated the recovered means from our methods are with the actual known signal. In this figure, all fidelities were calculated for a sinusoid at wavenumber, $k = 75.5$. This value of k put the sinusoid at the wavenumber of maximum noise amplitude for the AR(2)

---

[1] Note that $\lim_{M\to\infty}\phi = \Sigma^{-1}\mu \neq \mu$. That is, the limit of the CVA eigenvector is not μ but $\Sigma^{-1}\mu$. Therefore, the eigenvector may never converge to the true mean, even for an arbitrarily large dataset.
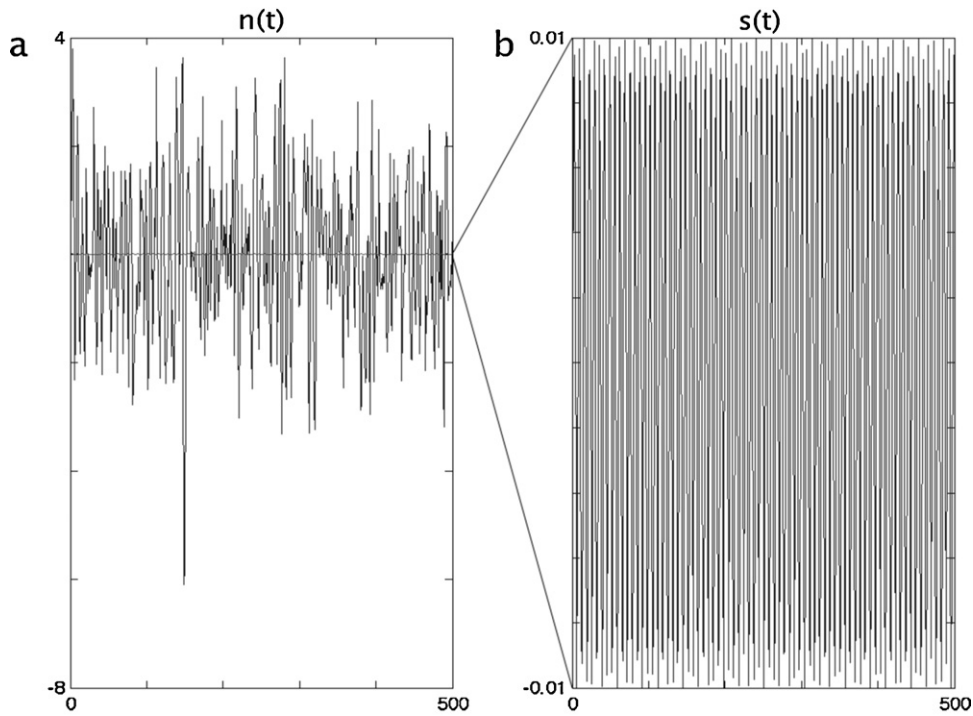
Fig. 1. Simulated signal and noise. (a) A typical AR(2) noise pixel timecourse, $n(t)$, generated as indicated in the text plotted along with a sinusoidal signal, $s(t)$, with amplitude, $\mu = 0.01$. Note: in (a), the signal is barely visible because of its low amplitude relative to the noise. (b) The sinusoidal signal amplified so that it may be easily seen.

process. Each point in the graphs represents the average of 100 fidelities. We calculated fidelities for the multivariate estimates, $\hat{\mu}_{GIFA}$ and $\hat{\mu}_{CVA}$, as well as the univariate estimates, $\hat{\mu}_{raw}$, which we will use to refer to the estimate given in Eq. (7) and $\hat{\mu}_{DFT}$, which we will use to refer to an estimate from the standard method of calculating the vector of DFT coefficients directly from the data with no

tapering or averaging. We also calculated the fidelity for the regularization method referred to in Appendix A.2, in which an SVD is calculated for the dataset, the number of singular eigenvectors is truncated, then a harmonic estimate is performed. We denote this SVD regularization method, $\hat{\mu}_{reg}$. For the regularized estimates, we truncated the original dataset to 200 singular eigenvectors (from a
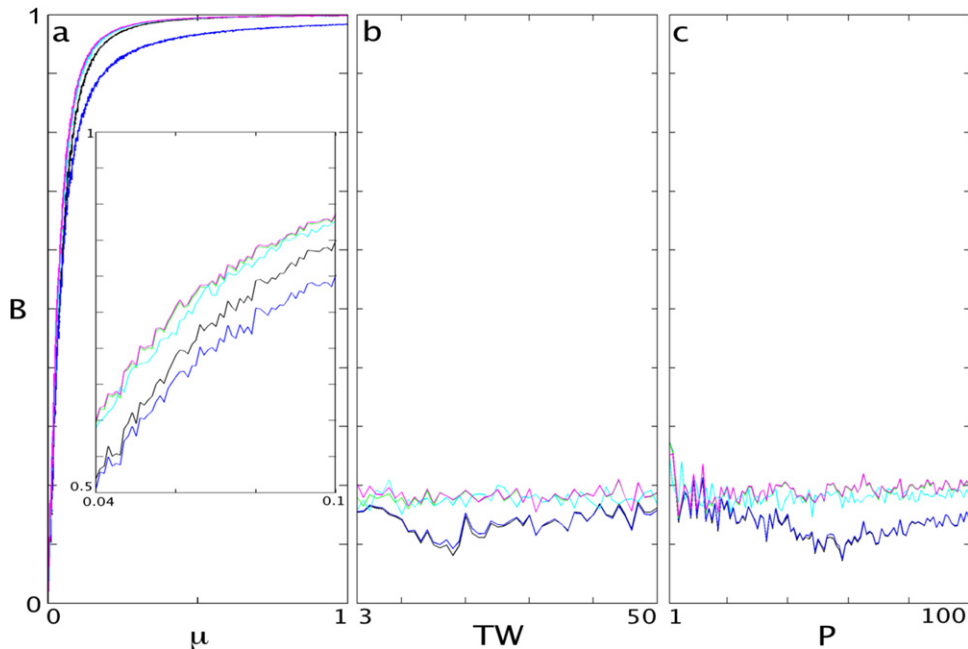
Fig. 2. Performance of GIFA ($\hat{\mu}_{GIFA}$, blue line), CVA ($\hat{\mu}_{CVA}$, black line), Raw ($\hat{\mu}_{raw}$, green line), Regularized ($\hat{\mu}_{reg}$, cyan line) and DFT ($\hat{\mu}_{DFT}$, magenta line) estimates for spatially uncorrelated ($q = 0$) noise: (a) Fidelities (correlations between estimated and true signal, measured from 0 to 1, see text for definition), B_{GIFA}, B_{CVA}, etc., as a function of signal amplitude, μ, at wavenumber $f = 75.5$ with fixed noise amplitude in the frequency band W with $TW = 25$ and $P = 20$. (b) Fidelities as a function of TW with $\mu = 0.01$, $P = 20$ and fixed noise amplitude. (c) Fidelities as a function of P with $\mu = 0.01$, $TW = 25$ and fixed noise amplitude. These harmonic estimates are made where the noise in the AR(2) process is maximum. Note that the raw estimate (magenta line) overlaps the DFT estimate (green line) for most values of the ordinate. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)
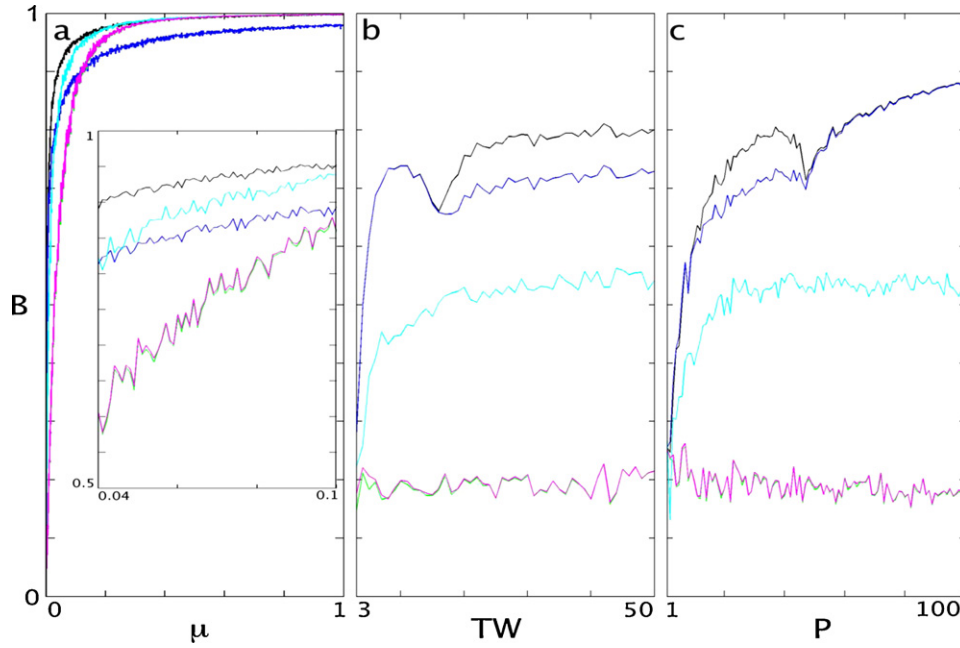
Fig. 3. Performance of GIFA, CVA, Raw, Regularized and DFT estimates for spatially correlated ($q = 3$) noise. Except for the spatial correlation in the noise, all parameters and conditions are as in Fig. 1. Note the huge improvements in the multivariate estimate fidelities in spatially correlated noise.

total of 1800). At an index of 200, the spectrum of timelike eigenvectors became approximately white and we used this feature as the statistical criterion for truncation. Below, we will refer to the fidelities for the different methods as B_{GIFA}, B_{CVA}, B_{reg}, etc.

In spatially uncorrelated noise, estimate fidelities improved as the amplitude of the sinusoid increased (Fig. 2a). Eventually, as μ increased, all estimates reached a fidelity of unity, however, it took considerably higher sinusoid amplitudes for B_{CVA} to reach unity relative to the other methods (data for $\mu > 1$ not shown). The univariate raw multitaper and DFT fidelities, B_{DFT} and B_{raw}, showed no dependence on TW (Fig. 2b) and P (Fig. 2c), as expected, since these methods do not depend on bandwidth or pixel number. B_{reg} also showed little dependence on TW and P. B_{GIFA} and B_{CVA} depended more strongly on both TW and P. In particular, there is a dip in the fidelity where the number of tapers is equal to the number of pixels. This is expected since the multitaper estimates become degenerate when the number of measurements (i.e. the number of tapers) equals the number of variables (i.e. the number of pixels) to be estimated and therefore the estimates are unstable.

The results in Fig. 2 demonstrate that the DFT gave the best estimates in spatially uncorrelated noise simulations. Slightly less accurate estimates are obtained with B_{raw} and B_{reg}. The multivariate estimates, B_{GIFA} and B_{CVA}, are less accurate than any of the univariate estimates. In general, for spatially uncorrelated noise, all estimates had fidelities that were within about 25% of each other. Note that all fidelities in Fig. 2b and c are quite poor. We used a low value of $\mu = 0.01$ for the simulations in these two panels in order to contrast these results with those obtained when the noise is spatially correlated.

3.2. Spatially correlated noise

In order to study a harmonic signal in spatially correlated noise, we generated noise with the AR(2) process, $\eta(t)$ (29), given above. To introduce spatial correlation into the dataset, we first calculated its singular value decomposition

$$\eta(t) = \sum_n \mathbf{p}_n \Delta_n \zeta(t)_n \qquad (31)$$

where $\mathbf{p}_n$ are left eigenvectors, Δ_n are singular values and $\zeta(t)_n$ are right eigenvectors. We then rescaled the singular values $\Delta_n' = \Delta_n/n^q$, where q is an exponent that we will vary below to change the amount of spatial correlation in the simulated data. This changed the statistically white (i.e. flat) eigenvalue spectrum of $\eta(t)$ to a colored spectrum (i.e. non-flat spectrum with some spatial modes having more variance than others), which simulated spatially correlated noise with randomly assigned principal axes,

$$\eta'(t) = \sum_n \mathbf{p}_n \Delta_n' \zeta(t)_n. \qquad (32)$$

We then normalized $\eta'(t)$ such that the sum of the variances in the spatially correlated noise was equal to that in the original spatially uncorrelated AR(2) process, i.e. the trace of the covariance (the total power in the noise) is unchanged. For the results below, the scaling exponent $q = 3$. Note that for $q = 0$, we revert to the spatially uncorrelated noise case.

We combined the signal and noise as in the previous section to form the simulated dataset, $\mathbf{x}(t) = \mathbf{s}(t) + \eta'(t)$. For the data containing spatially correlated noise, $\eta'(t)$, the average SNR per pixel was the same as the average SNR per pixel for the spatially uncorrelated noise studied in the previous section. The results in Figs. 2 and 3 may therefore be compared. As we will see below, the fidelities of the estimates from the various methods become relatively different from the uncorrelated noise case.

In spatially correlated noise, particularly at low amplitudes (see inset Fig. 3a and compare to Fig. 2a), GIFA and CVA gave higher fidelity estimates than the raw multitaper and DFT methods. A similar dip in fidelity was seen in $\hat{\mu}_{GIFA}$ and $\hat{\mu}_{CVA}$ as a function of TW and P at the points of degeneracy. Nevertheless, these estimates still gave better fidelities overall than the other estimates. The fidelities of $\hat{\mu}_{raw}$ and $\hat{\mu}_{DFT}$ were constant and almost identical as a function of TW and P, but less than the multivariate estimates for the spatially correlated noise case. The fidelity of $\hat{\mu}_{reg}$ was better than the univariate estimates, but not as good as the GIFA and CVA estimates.

Fig. 3 shows the advantage of using multivariate, multitaper estimates for low amplitude sinusoids in spatially correlated noise. The fidelity of the GIFA estimate is higher than that of the DFT estimate and improves as TW and P become greater.
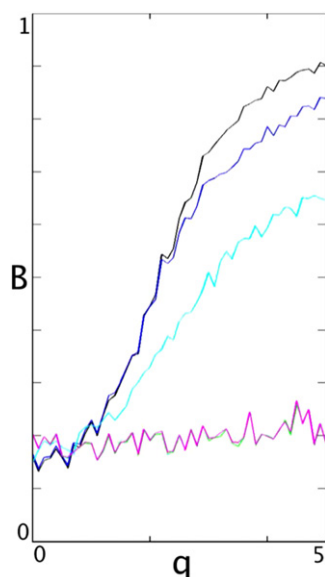
Fig. 4. Dependence of GIFA, CVA, Raw, Regularized and DFT harmonic estimates on the spatial correlation in the noise. Fidelities as a function of the amount of spatial correlation, q, at wavenumber $f = 75.5$ with $\mu = 0.01$, $TW = 25$ and $P = 20$. Note the fourfold improvement in signal estimation for the GIFA method by $q = 3$.

An interesting aspect of the GIFA and CVA estimates is that when there are more tapers than pixels (Fig. 3b) or (from the converse viewpoint) fewer pixels than tapers (Fig. 3c), the GIFA estimate gives greater fidelity, however, at the other limit of more pixels than tapers, the estimate fidelities are indistinguishable (see Section 4).

The improvements in fidelity between Figs. 2 and 3 make it clear that, at some amount of spatial correlation, $\hat{\mu}_{GIFA}$, $\hat{\mu}_{CVA}$ and $\hat{\mu}_{reg}$ must begin to give higher fidelities versus $\hat{\mu}_{raw}$ and $\hat{\mu}_{DFT}$. In Fig. 4, we plot fidelities as a function of spatial correlation in the noise. Fig. 4 demonstrates that the transition in fidelity between the univariate and multivariate methods occurs at $q \approx 1$. At $q = 3$ the GIFA method gives a fourfold improvement in signal estimation relative to the univariate methods.

3.3. Application to periodic stimulation of cat cortex

To demonstrate the effectiveness of our methods, we show results from all five methods as applied to optical imaging data of the intrinsic signal in cat primary visual cortex. All surgical and experimental procedures were performed under institutional animal care guidelines. Experimental details are presented in Everson et al. (1998) and Sirovich and Uglesich (2004). In brief, the cat was anaesthetized with a mixture of propofol (4 mg/kg per h; Diprivan, AstraZeneca) and sufentanil (0.05 μg/kg per h). The pupils were dilated and the eyes were refracted and fit with corrective lenses to improve visual acuity. Visual cortex (Areas 17 and 18) was exposed and illuminated with 605 nm light. At this wavelength, absorption is due primarily to deoxy-hemoglobin and the differential response is well-known to correlate with neuronal activity (Bonhoeffer and Grinvald, 1991; Blasdel, 1992; Sirovich et al., 1996).

Images of light scattered from the cortex were acquired with a CCD camera (PXL, Photometrics, Tucson, AZ, 536 × 389 pixels) at a frame rate of 15 Hz. The camera was fit with a back-to-back macro lens system with 1× magnification (Ratzlaff and Grinvald, 1991) and focused 500 μm beneath the surface of the cortex at the approximate depth of cortical layer IV. Due to the curvature of the cortex, only a portion of our images was truly focused at this depth. Visual stimuli in the form of high-contrast, black and white square-wave gratings drifting in one direction were presented to the animal on a 12 bit video monitor. The spatial frequency of the bars was

0.3 cycles per degree drifting at 2 cycles per second resulting in a temporal frequency of 6.67° per second.

In order to obtain measurements of many orientations, we used a rapid stimulus presentation technique. This method was meant to approximate the periodic stimulation experimental design used by Kalatsky and Stryker (2003) and Kalatsky et al. (2005). A grating of a given orientation was presented for one second (15 frames), then the orientation of the grating was incremented by 5°. Between each stimulus presentation, there was an intervening period of half a second during which a neutral contrast flat field was presented. Upon reaching 180°, the orientation was reset to zero. This gave a period of 54 s for the original orientation to reappear. The orientation was incremented in this way periodically for 50 rotations through 180°. Due to the intervening neutral contrast field and the fact that we incremented the orientation in finite steps, our stimulus was not truly continuous; however, this rapid stimulus presentation technique allowed us to measure the response to 36 different orientations with good signal-to-noise ratio. Images were subsequently averaged in groups of 15 frames, resulting in one frame per orientation which combined to give a set of 1800 images. Therefore, in our analyses, we were looking for a statistically significant harmonic at a wavenumber, $k = 1800/36 = 50$. Note that because it takes 2–3 s for the mapping signal to achieve full amplitude, the phase of the response lags the phase of the stimulus by approximately 6°.

For a multitaper harmonic analysis to be valid, the model that there is a single harmonic within the bandwidth W must apply. As one approaches zero frequency, the bandwidth for the positive frequency harmonic and the negative frequency harmonic that together make up the real-valued sinusoid come within each other's bandwidth and the model is no longer valid. Thus, at minimum, the harmonic of interest must be a full bandwidth W from the origin. In order to determine the number of data points one needs, we note that if the wavenumber (i.e. number of stimulus periods within the dataset) of the stimulus is V, then the bandwidth of the estimate is $[V + TW, V - TW]$. The minimum number of data points will depend on the required statistical power of the estimate (remember that the statistical significance is calculated from the F-statistic with $2(M - 1)$, 2 degrees of freedom). The number of tapers that we use is $M = 2TW - 3$. Take the time-bandwidth (i.e. wavenumber) of the estimate to be $TW = 5$. This gives a 7 taper set of eigenestimates and reasonable statistical power. If we require about 10 time points per wavelength, then we would need approximately 50 datapoints as a minimum requirement for our harmonic estimate. In our results, we had the luxury of being able to use $TW = 20$ resulting in 37 eigenestimates and very good statistical power. Our dataset contained 1800 time points and took an hour to acquire. With just a one hour experiment we were able to estimate the response to 36 different orientations.

We see in Fig. 5 that Hotelling's T^2-statistic exceeded the 99% significance level at a number of peaks. The series of peaks below wavenumber 20 was due to the breakdown of the signal model as described in the previous paragraph. Since the estimates for this analysis used $TW = 20$, the positive frequency sinusoid began to overlap with the bandwidth of the negative frequency peak below this wavenumber (see Eq. 1). This caused spurious peaks. Above wavenumber 20, a few statistically significant peaks were found. In particular, a peak was found at wavenumber 50. We interpreted this to be the response to the stimulus and investigated the real and imaginary parts of the images (Fig. 5b). The signal estimates at wavenumber 50 showed the characteristic spatial pattern of the orientation columns in primary visual cortex. Each phase (color) in the signal estimate represents the response to a particular oriented drifting grating that was presented as a visual stimulus. Loci around which all orientations may be found are called "pinwheels". Inspection of the phase and amplitude plots shows a yellow vascular
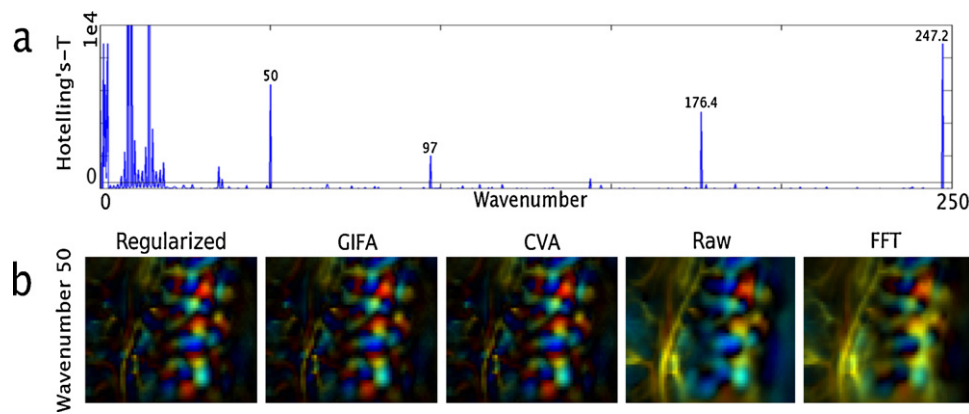
Fig. 5. Orientation response in cat visual cortex. Multitaper harmonic estimates were calculated with $TW = 20$. (a) Hotelling's T^2 statistic as a function of wavenumber (blue) and 99% confidence level (black). 1250 harmonic estimates were performed at 0.2 wavenumber increments from wavenumber 0 to wavenumber 250. (b) Amplitude (intensity) and phase (color, colormap is periodic) images of $\hat{\mu}_{reg}$, $\hat{\mu}_{GIFA}$, $\hat{\mu}_{CVA}$, $\hat{\mu}_{raw}$ and $\hat{\mu}_{DFT}$, the statistically significant mean estimates at wavenumber $k = 50$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

structure in the univariate estimates that is absent in the multivariate estimates. Although we do not know the ground truth for this data and cannot know for sure, we interpret this to mean that the yellow structure represented variance (not the mean) at wavenumber 50. Therefore, it was represented in the noise matrix, $\hat{K}$. The multivariate methods, thus, projected it out of the estimate while optimizing their respective criterion functions. Also, the pattern of pinwheels associated with orientation columns is not the same for the univariate methods, nor is it as distinct, as in the multitaper estimates for the same reason (i.e. the orientation columns contained all modes in the frequency subspace in the raw estimate, but modes associated with spatially correlated variance were rejected due to the optimization procedures in the multivariate methods).

4. Discussion

In our exploration of the accuracy of the DFT, raw, regularized, CVA and GIFA signal estimates as a function of the parameters P, TW and μ for a known signal, it became clear that in datasets with spatially uncorrelated noise the DFT and raw multitaper estimates always had a higher fidelity than the regularized, CVA or GIFA harmonic estimates for all of the parameters that we investigated. This makes sense since one can expect no improvement from multivariate analysis methods when there is no spatial structure to the noise. As spatial correlation in the noise increased, GIFA outperformed both CVA and the other estimates overall.

One apparent problem with the GIFA method is the necessity of selecting α a priori. However, the fidelity of the GIFA estimates was relatively insensitive to the value of α (data not shown). As a rule of thumb for setting the threshold for detection, we use the same criterion used for univariate, multitaper harmonic analysis, i.e. $\alpha = 1/T$ (Mitra and Pesaran, 1999). For this value, there will be, on average, one false harmonic detection across all frequencies in a dataset of length T.

The multitaper methods presented here can be somewhat expensive as regards computation time, but not as expensive as one might think. A DFT analysis requires approximately $2PT$ multiplications and the same number of additions (Note that for an arbitrary frequency that is not on a grid of frequencies that are multiples of the Rayleigh wavelength, an FFT must be padded. This can lead to longer computation times and greater memory requirements). Our multitaper methods require M DFTs. Solution of the multivariate optimization problem is then performed on a roughly $M \times M$ matrix. Typically, M is from 5 to 20 or 25, so this diagonalization takes a negligible amount of time. Therefore, the analysis will take about 5–25 times the time of the DFT analysis. The fine grid of

harmonic analyses that we performed for Fig. 5 took approximately 1/2 an hour, whereas the DFT analysis took a minute or two. Investigating a single frequency takes a negligible amount of time for all analyses. We note that an implementation on a modern GPU should be able to improve on our results due to the GPU's parallel capabilities.

An additional issue to consider is memory constraints. For large datasets, storage of $m = 25$ times the memory size of the original dataset could be a problem. This is the reason that we gave details of our dimension reduction methods in Appendices. For instance, for the data analyzed in Section 3, instead of calculating a DFT on the timecourses of over 10,000 pixels, one can losslessly calculate the DFT on the 1800 timelike eigenvectors from the SVD.

A great advantage of our multitaper methods is that they allow statistical detection of a harmonic signal and, as we have seen, they can give greatly improved (up to fourfold, see Fig. 4) estimates in the spatially correlated noise that is common in brain imaging data. In the cortical imaging data that we analyzed, the estimated response was not a great surprise, since orientation columns have been imaged many times and their structure is well-known. However, when searching for an unknown response, for instance previously unimaged areas in the visual cortex hierarchy, with unknown structure, it can be very useful to be able to appeal to a T^2-statistic for confidence that the structure that one is seeing is truly due to structured response, not structured noise. If one decides to use the multitaper methods, we note that it is not overly expensive to calculate all three multitaper estimates at the same time, since the raw estimate has fixed cost for a given dataset and the GIFA and CVA estimates cost only the extra time required to diagonalize two small ($M \times M$) matrices.

The numerical results that we presented in Fig. 4 indicate that there is a relatively large difference between DFT and raw versus multitaper estimates (regularized, CVA, GIFA) for high spatial correlations, but less difference for low spatial correlations. This points to a useful means of identifying spatial correlation in experimental data[2]: If the estimates are similar and correlate well, it is likely that there is minimal spatial correlation in the data. However, as in the cortical imaging data that we presented, if the DFT and multitaper estimates are very different, our results indicate that this is likely to be due to spatial correlation in the data.

In the parameter range $P > TW$, both the simulation results (Fig. 3) and the cortical imaging results show that there is

---

[2] We thank Lynne Seymour of the Statistics Department at the University of Georgia for pointing this out.

virtually no difference between CVA and GIFA methods. The advantage from the GIFA method comes primarily when $TW < P$. In this regime, GIFA gave a 10% to 15% improvement in fidelity relative to CVA. Thus, MEG/EEG and multielectrode array data would benefit most from our GIFA method.

The results plotted in Figs. 2 and 3 imply that both CVA and GIFA methods perform less well when the pixel number P and the number of degrees of freedom $M − 1$ of the eigenestimate matrix $\mathbf{j}_m$ become commensurate. This implies that one should be careful with the analysis of data with moderate numbers of pixels when using the multivariate optimization methods (GIFA and CVA). If possible, by increasing or decreasing the bandwidth used in these methods, this region of parameters should be avoided for the most accurate results.

5. Conclusions

We have presented three multitaper approaches to the detection and estimation of multivariate harmonic content in noisy data: one, the raw estimate, gave comparable fidelities to the DFT for a harmonic signal in spatially uncorrelated noise for all pixel numbers, P, as well as being accompanied by a statistic measuring the significance of a harmonic signal; while two auxiliary, multivariate methods, the generalized indicator function analysis (GIFA) and canonical variate analysis (CVA) methods, always gave better fidelities in spatially correlated noise.

Our methods generalize the classical multitaper univariate harmonic detection and estimation method to the multivariate context. Of the two multivariate, multitaper analysis methods that we describe, one, CVA, is a well-known classical method, the other, GIFA, has been introduced recently and is often more robust than CVA. The robustness of GIFA estimates in spatially correlated noise has been reconfirmed in the results that we presented here.

Our multivariate methods may be used to detect and more significantly more accurately reconstruct periodic responses in optical and neural imaging data with its characteristic, spatially correlated noise relative to the standard DFT estimates used to date.

Acknowledgements

Appendix A.

A.1. Dimensional reduction for the case $P \geq M − 1$

When $P \geq M − 1$ it becomes necessary to reduce the dimensionality of $\hat{\mathbf{K}}_s$ and $\hat{\mathbf{K}}$ to avoid inverting a singular matrix (Friston et al., 1995). Even when inversion is not required, dimensional reduction can speed up the computations that go into the calculation of the multivariate harmonic estimates. We use a singular value decomposition of the $P \times M$ dimensional matrix $[\mathbf{J}_m]$ to give

$$[\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_M] = \mathbf{\Psi D A}, \tag{33}$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_M]$ is an orthogonal matrix. Note carefully that this formula deviates slightly from the standard definition of the SVD where the columns $\{\mathbf{a}_i\}$ are the right singular eigenvectors.

The above definition, in order to make the following expressions as simple as possible, is such that the *rows* of $\mathbf{A}$ are the right singular eigenvectors. Therefore, the vectors $\{\mathbf{a}_i\}$, also orthonormal, are indexed by singular value and take the place of pixels in the dimensionally reduced formulation. In these coordinates, the mean estimate becomes

$$\hat{\boldsymbol{\mu}} = \mathbf{\Psi D} \hat{\boldsymbol{\mu}}', \tag{34}$$

where

$$\hat{\boldsymbol{\mu}}' \equiv \sum_{m=1}^{M} \frac{\mathbf{a}_m H_m}{\mathcal{H}}. \tag{35}$$

With these expressions, we may now write the signal matrix as,

$$\hat{\mathbf{K}}_s = \mathbf{\Psi D} \mathcal{H} \hat{\boldsymbol{\mu}}' \hat{\boldsymbol{\mu}}'^{\dagger} \mathbf{D} \mathbf{\Psi}^{\dagger} \tag{36}$$

and the noise matrix

$$\hat{\mathbf{K}} = \mathbf{\Psi D} \hat{\mathbf{Z}} \mathbf{D} \mathbf{\Psi}^{\dagger} \tag{37}$$

with $\hat{\mathbf{Z}}$ defined by

$$\hat{\mathbf{Z}} \equiv \sum_{m=1}^{M} (\mathbf{a}_m - \hat{\boldsymbol{\mu}}' H_m)(\mathbf{a}_m - \hat{\boldsymbol{\mu}}' H_m)^{\dagger}. \tag{38}$$

Note that Rank($\mathbf{Z}$) = $M − 1$, whereas the $\{\mathbf{a}_i\}$ in the above expression are of length M. Therefore, if this matrix needs to be inverted, the procedure for finding a pseudo-inverse given in the text with Eqs. (12) and (13) should be used.

Using the above expressions, the CVA eigenvalue problem becomes

$$\mathcal{H} \hat{\mathbf{Z}}^{-1} \hat{\boldsymbol{\mu}}' \hat{\boldsymbol{\mu}}'^{\dagger} \boldsymbol{\zeta} = \rho \boldsymbol{\zeta} \tag{39}$$

in the singular value basis (33), and the relation $\boldsymbol{\phi} = \mathbf{\Psi D}^{-1} \boldsymbol{\zeta}$ gives the estimated multivariate signal components in the original P-dimensional coordinates.

The GIFA eigenvalue problem becomes

$$\mathbf{D} \left[\mathcal{H} \hat{\boldsymbol{\mu}}' \hat{\boldsymbol{\mu}}'^{\dagger} - \tau_\alpha^2 \hat{\mathbf{Z}} \right] \mathbf{D} \boldsymbol{\xi} = \gamma \boldsymbol{\xi} \tag{40}$$

where $\boldsymbol{\phi} = \mathbf{\Psi} \boldsymbol{\xi}$ is the transformation back into the original P-dimensional coordinates.

A.2. Dimensional reduction when $P \gg T$

In a typical application to a large multivariate dataset, such as experimental data from an imaging experiment, if P, the number of pixels, is greater than T the number of time samples, the raw data may be reduced in dimension via a singular value decomposition before the application of one of our multivariate harmonic analysis methods. In this case, the raw data, a $P \times T$ matrix, $\mathbf{x}(t)$, may be expressed as

$$\mathbf{x}(t) = \mathbf{\Theta Q R} \equiv \mathbf{\Theta B} \tag{41}$$

with $\mathbf{\Theta} = [\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_T]$ a matrix of left singular eigenvectors, $\mathbf{Q}$ a diagonal matrix of singular values, $\mathbf{R} = \mathbf{r}(t) = [\mathbf{r}_1, \ldots, \mathbf{r}_T]$ a matrix with *rows* the right singular eigenvectors (as above), giving $\mathbf{B} = \mathbf{b}(t) = [\mathbf{b}_1, \ldots, \mathbf{b}_T] = [Q_{11} \mathbf{r}_1, \ldots, Q_{TT} \mathbf{r}_T]$.

In this basis,

$$\mathbf{J}_m = \mathbf{\Theta L}_m \tag{42}$$

where

$$\mathbf{L}_m = \sum_t \mathbf{b}(t) h_m(t) e^{-i\omega t}. \tag{43}$$

At this point, we apply the multitaper methods presented above in Section 2 on the transformed M-dimensional data $\{\mathbf{L}_m\}$. After

we obtain the raw, CVA and GIFA estimates in the M-dimensional space, we transform back to the original P-dimensional space using the expression

$$\hat{\boldsymbol{\mu}} = \Theta \hat{\boldsymbol{\mu}}'. \qquad (44)$$

Note that the lossless dimensional reduction described here may be made lossy by truncating the first singular value decomposition (41) and retaining a number, $Q < T$, of singular eigenvectors for the analysis. For very large datasets, this may become necessary to reduce computation time. It is important to note that once one truncates the first singular value decomposition, a certain number of eigenvectors have been discarded. Because SVD eigenvectors are typically ordered in terms of the amount of covariance they represent in the data, the discarded eigenvectors represent the least spatially correlated dimensions in the dataset. Therefore, an effective regularization has been performed and spatially correlated dimensions have been retained. This means that via the truncation, the "raw" harmonic estimate can be improved versus the untruncated estimate when the noise is spatially correlated. We show this phenomenon in Section 3. Note that care should be taken if the data is truncated, since the signal could possibly be thrown away with the truncated modes.

References

Anderson T. An introduction to multivariate statistical analysis. New York: John Wiley and Sons; 1984.

Blasdel G. Differential imaging of ocular dominance and orientation selectivity in monkey striate cortex. J Neurosci 1992;12:3115–38.

Bonhoeffer T, Grinvald A. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. Nature 1991;353:267–94.

Box G, Jenkins G. Time series analysis: forecasting and control. San Francisco: Holden-Day; 1976.

de Munck J, Huizenga H, Waldorp L, Heethaar R. Estimating stationary dipoles from meg/eeg data contaminated with spatially and temporally correlated background noise. IEEE Trans Signal Process 2002;50:1565–72.

Ennis D, Johnson N. Noncentral and central chi-square, f, and beta distributions as special cases of the distribution function of an indefinite quadratic form. Commun Statist Theory Methods 1993;22:897–905.

Everson R, Prashanth A, Gabbay M, Knight B, Sirovich L, Kaplan E. Representation of spatial frequency and orientation in the visual cortex. PNAS 1998;95:8334–8.

Friston K, Frith C, Frackowiak R, Turner R. Characterizing dynamic brain responses with fMRI: a multivariate approach. NeuroImage 1995;2:166–72.

Giri N. On the complex analogues of t^2 and r^2 tests. Ann Math Statist 1965;36:664–70.

Kalatsky V, Polley D, Merzenich M, Schreiner C, Stryker M. Fine functional organization of auditory cortex revealed by Fourier optical imaging. Proc Natl Acad Sci USA 2005;102:13325–30.

Kalatsky V, Stryker M. New paradigm for optical imaging: temporally encoded maps of intrinsic signal. Neuron 2003;38:529–45.

Mitra P, Pesaran B. Analysis of dynamic brain imaging data. Biophys J 1999;76:691–708.

Mitra P, Bokil. Observed brain dynamics. Oxford University Press; 2008.

Percival D, Walden A. Spectral analysis for physical applications. Cambridge University Press; 1993.

Pesaran B, Sornborger A, Nishimura N, Kleinfeld D, Mitra P. Spectral analysis for dynamical imaging data. NY: Cold Spring Harbor Laboratory Press; 2004.

Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA. Temporal structure in neuronal activity during working memory in macaque parietal cortex. Nat Neurosci 2002;5:805–11.

Ratzlaff E, Grinvald A. A tandem-lend epifluorescence macroscope: hundred-fold brightness advantage for wide-field imaging. J Neurosci Methods 1991;36:127–37.

Sirovich L, Everson R, Kaplan E, Knight B, O'Brien E, Orbach D. Modeling the functional organization of the visual cortex. Physica D 1996;96:355–66.

Sirovich L, Uglesich R. The organization of orientation and spatial frequency in primary visual cortex. Proc Natl Acad Sci USA 2004;101:16941–6.

Sornborger A, Sailstad C, Kaplan E, Knight B, Sirovich L. Spatio-temporal analysis of optical imaging data. NeuroImage 2003a;18:610–21.

Sornborger A, Sirovich L, Morley G. Extraction of periodic multivariate signals: Mapping of voltage dependent dye fluorescence in mouse heart. IEEE Trans Med Imaging 2003b:1537–49.

Sornborger A, Yokoo T, Delorme A, Sailstad C, Sirovich L. Extraction of average and differential dynamical responses in stimulus-locked experimental data. J Neurosci Methods 2005;141:223–9.

Tauxe L. Sedimentary records of relative paleointensity of the geomagnetic field: theory and practice. Rev Geophys 1993;31:319–54.

Thomson D. Spectrum estimation and harmonic analysis. In: Proceedings of the IEEE 70; 1982. p. 1055–96.

Thomson D. Multitaper analysis of nonstationary and nonlinear time series data. In: Fitzgerald WJ, Smith RL, Walden AT, Young PC, editors. Nonlinear and nonstationary signal processing. Cambridge University Press; 2000.

Thomson D, Schild R. Time delay estimates for Q0957+561 A, B. London: Chapman and Hall; 1997.

Ulrych T, Bishop T. Maximum entropy spectral analysis and autoregressive decomposition. Rev Geophys Space Phys 1975;13:183–200.

Walden A. A unified view of multitaper multivariate spectral estimation. Biometrika 2000;87:767–87.

Yokoo T, Knight B, Sirovich L. An optimization approach to signal extraction from noisy multivariate data. NeuroImage 2001;14:1309–26.