

## COMPLEX-VALUED TIME SERIES MODELING FOR IMPROVED ACTIVATION DETECTION IN FMRI STUDIES

BY DANIEL W. ADRIAN\*, RANJAN MAITRA<sup>†,1</sup> AND DANIEL B. ROWE<sup>‡,2</sup>

*Grand Valley State University\**, *Iowa State University<sup>†</sup>* and *Marquette University<sup>‡</sup>*

A complex-valued data-based model with  $p$ th order autoregressive errors and general real/imaginary error covariance structure is proposed as an alternative to the commonly used magnitude-only data-based autoregressive model for fMRI time series. Likelihood-ratio-test-based activation statistics are derived for both models and compared for experimental and simulated data. For a dataset from a right-hand finger-tapping experiment, the activation map obtained using complex-valued modeling more clearly identifies the primary activation region (left functional central sulcus) than the magnitude-only model. Such improved accuracy in mapping the left functional central sulcus has important implications in neurosurgical planning for tumor and epilepsy patients. Additionally, we develop magnitude and phase detrending procedures for complex-valued time series and examine the effect of spatial smoothing. These methods improve the power of complex-valued data-based activation statistics. Our results advocate for the use of the complex-valued data and the modeling of its dependence structures as a more efficient and reliable tool in fMRI experiments over the current practice of using only magnitude-valued datasets.

**1. Introduction.** Functional magnetic resonance imaging (fMRI) is a popular noninvasive method for studying brain function. Since its inception 25 years ago, fMRI has led to important developments in neurosurgical planning, cognitive neuroscience, and the perception of mental illness, among other fields [Rosen and Savoy (2012)]. The imaging modality is based on the fact that when neurons fire in response to a stimulus or a task, the blood oxygen level in neighboring vessels changes, affecting the magnetic resonance (MR) signal on the order of 0.5–3% due to the differing magnetic susceptibilities of oxygenated and deoxygenated hemoglobin [Lazar (2008)]. This difference is behind the so-called Blood Oxygen Level Dependent (BOLD) contrast [Belliveau et al. (1991), Kwong et al. (1992), Ogawa et al. (1990), Bandettini et al. (1993)] which is utilized as a surrogate for neural activity and is used to acquire time-course sequences of images.

---

Received December 2016; revised September 2017.

<sup>1</sup>Supported in part by the National Science Foundation CAREER Grant # DMS-0437555 and by the National Institutes of Health (NIH) award #R21EB016212.

<sup>2</sup>Supported in part by the National Institutes of Health (NIH) award #R21NS087450.

*Key words and phrases.* Area under the ROC curve, contrast-to-noise ratio, finger-tapping motor experiment, hemodynamic response function, Kronecker product, neurosurgical planning guide, phase information, signal-to-noise ratio, structured covariance matrix.

The stimulus time course commonly consists of a block design, in which stimulus and resting periods are alternated regularly, or is event related, in which case the stimulus is applied instantaneously at several time points with the subject at rest at other times.

Each MR image is obtained in a series of steps from the  $k$ -space data [Brown, Kincaid B. M and Ugurbil (1982), Ljunggren (1983), Tveit (1983)] which encodes different frequency contributions to each voxel. The varied frequencies result from different magnetic field gradients [Jezzard and Clare (2001)] and need to be inverted to localize measurements at each voxel. This step is implemented by applying the inverse Fourier transform [Jain (1989)] on the  $k$ -space data and results in a complex-valued observation at each voxel and each time-point. Thus, acquired fMRI data at each voxel and time point can, in reality, be written in terms of its real and imaginary (or equivalently, magnitude and phase) components.

Acquired MR datasets have typically used only the magnitude measurements at each voxel for display and analysis. This practice of using only the magnitude data while discarding the phase at each voxel has carried over to fMRI practice so much that most statistical analyses routinely ignore the phase data and base their inferences on only the magnitude time series at each voxel [Rowe and Logan (2004)]. Thus, in a literal sense, it could be said that fMRI analyses have almost exclusively thrown away half of its collected data, and though the magnitude values unarguably contain more information than do the phase measurements about the BOLD response of interest, we argue and demonstrate in this paper that the phase contains information also important to the analysis.

The general strategy in so-called “magnitude-only” statistical analysis of fMRI data is to fit, at each voxel, a model—commonly a general linear model [Friston et al. (1995)]—to the time series observations against a transformation of the input stimulus. This transformation is the expected BOLD response and is effectively modeled in terms of a convolution of the stimulus time course with the hemodynamic response function (HRF), which measures the delay and dispersion of the BOLD response to an instantaneous neuronal activation [Friston, Jezzard and Turner (1994), Glover (1999)]. This provides the setting for the application of the Statistical Parametric Mapping (SPM) technique of Friston et al. (1990), in which the time series at each voxel is reduced to a test statistic that summarizes the association between each voxel time course and the expected BOLD response [Bandettini et al. (1993)]. The resulting map is then thresholded to identify voxels that are significantly activated [Genovese, Lazar and Nichols (2002), Worsley et al. (1996), Logan and Rowe (2004)]. Commonly, the images at each time point are spatially smoothed before each voxel time series is modeled and the test statistics are computed. Smoothing can improve the contrast-to-noise ratio (CNR) by averaging away the noise in neighboring voxels, as long as the activation is not smoothed away as well [Smith (2001), Lazar (2008)].

A different approach using both magnitude and phase data (i.e., the complete complex-valued data) in the analysis has several advantages. For instance, the

phase time series can be useful in the so-called “brain or vein” problem, that is, in determining whether a voxel showing a task-related magnitude change represents (desired) grey matter or (an undesired) draining vein. This is because it has been demonstrated that voxels containing large, coherently oriented vessels (i.e., draining veins) can exhibit a task-related phase change [Hoogenrad et al. (1998), Menon (2002)]; in contrast, according to theory, voxels containing small, randomly oriented vessels exhibit little phase change. Recent findings also indicate that phase changes can occur outside of voxels containing draining veins as well, and thus (while complicating the “brain or vein” problem) phase data can provide additional physiological information [Feng et al. (2009), Zhao et al. (2007)]. In addition, preliminary evidence indicates that phase information may be important in neuronal current MRI (ncMRI), which aims to directly detect neuron firings (instead of using the BOLD surrogate) but its development is still in its early stages [Bandettini, Petridou and Bodurka (2005)]. Lastly, simulations [Rowe and Logan (2004)] show that magnitude-only data-based activation tests have a substantial drop of power at low signal-to-noise ratios (SNRs) while corresponding complex-valued data-based tests have constant (and higher) power over all SNRs. This is a potentially important result for future fMRI studies incorporating greater spatial resolution, as voxel volume is inversely proportional to SNR.

1.1. *Statistical models for complex-valued voxel time series.* We now introduce notation for complex-valued fMRI time series. Focusing on a single voxel (and suppressing voxel-related subscripts), the complex-valued observation at time  $t$  can be denoted by  $y_{Rt} + iy_{It}$  in terms of real/imaginary components or by  $r_t e^{i\phi_t} = r_t(\cos \phi_t + i \sin \phi_t)$  in terms of magnitude/phase components. Thus, it holds that  $y_{Rt} = r_t \cos \phi_t$ ,  $y_{It} = r_t \sin \phi_t$ ,  $r_t = (y_{Rt}^2 + y_{It}^2)^{\frac{1}{2}}$ , and  $\phi_t = \arctan_4(y_{It}, y_{Rt})$ , the 4-quadrant arctangent [see Glisson (2011), page 348] corresponding to  $\arctan(y_{It}/y_{Rt})$ . For statistical analysis, we consider the complex number  $y_{Rt} + iy_{It}$  to be the bivariate random vector  $(y_{Rt}, y_{It})'$  and also denote the (real-valued) real, imaginary, magnitude, and phase time series vectors, respectively, by  $\mathbf{y}_R = (y_{R1}, \dots, y_{Rn})'$ ,  $\mathbf{y}_I = (y_{I1}, \dots, y_{In})'$ ,  $\mathbf{r} = (r_1, \dots, r_n)'$ , and  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)'$ , with  $n$  denoting the number of scans. Most statistical models for complex-valued voxel time series are based on the fact that  $y_{Rt}$  and  $y_{It}$  are independent normal random variables with the same variance [Wang and Wei (1994)]. These models may be categorized as whether the magnitude and phase are uncoupled [Lai and Glover (1997), Lee et al. (2007)] or (correctly) coupled, as in Nan and Nowak (1999), Rowe and Logan (2004), and Rowe (2005). Here we specifically present the Rowe and Logan (2004) model

$$(1.1) \quad \begin{pmatrix} \mathbf{y}_R \\ \mathbf{y}_I \end{pmatrix} = \begin{pmatrix} \mathbf{X}\boldsymbol{\beta} \cos \theta \\ \mathbf{X}\boldsymbol{\beta} \sin \theta \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_R \\ \boldsymbol{\eta}_I \end{pmatrix}$$

because of its simplicity and because our investigation (see Section 4.2) indicates its adequacy for our dataset. The model couples the expected response of the magnitude time series  $\mathbf{X}\boldsymbol{\beta}$  with the phase location parameter  $\theta$  that is assumed constant

throughout the time series. Here the columns of the design matrix  $\mathbf{X}$  represent features of the fMRI signal including the baseline signal and the expected BOLD response. The errors  $(\eta'_R, \eta'_I)' \sim N(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{\Phi})$ , where  $\mathbf{\Sigma}$  and  $\mathbf{\Phi}$  are matrices of order 2 and  $n$  representing the real/imaginary and temporal covariance, respectively, and  $\otimes$  represents the direct (Kronecker) product. This model generalizes the design matrix and covariance structure of Nan and Nowak (1999), and was originally proposed by Rowe and Logan (2004) who, however, proceeded by assuming independence (i.e., by assuming that  $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_2$  and  $\mathbf{\Phi} = \mathbf{I}_n$ , where  $\mathbf{I}_k$  is the identity matrix of order  $k$ ) when deriving maximum likelihood estimates (MLEs) and a likelihood ratio test (LRT) for activation. Rowe (2005) further generalized model (1.1) by introducing time-dependent phase location parameters according to a linear model. Under this updated model, Rowe (2005) derived LRTs for task-related phase changes such as those caused by draining veins, as discussed previously.

1.2. *Overview of contributions of this paper.* This work offers three major contributions to complex-valued voxel-wise modeling of fMRI time series. First, we develop methodology to correct for instability in the magnitude and phase signals via a “complex-valued running line,” which fits a running-line smoother to the magnitude and phase time series simultaneously via a complex-valued model. Correcting for nuisance signals in magnitude time series which, whether caused by scanner drift or aliased physiological effects [Lazar (2008)], are a well-known processing step in the analysis pipeline. Less known is that instability of phase time series can cause complex-valued data-based detection methods to become less powerful than their magnitude-only counterparts. Consequently, methods have been developed to correct the phase for inhomogeneities in the static magnetic field [Hahn, Nencka and Rowe (2009)], physiological noise [Petridou et al. (2009)], and head motion [Hahn, Nencka and Rowe (2012), Hahn and Rowe (2012)]. In our case, preprocessing via the complex-valued running line controls for nonlinear nuisance magnitude signals, allows for the assumption of constant phase location parameter in model (1.1), and provides a better model fit.

Second, we apply a more general covariance structure to model (1.1) than previously considered in Rowe and Logan (2004), which assumed that the real and imaginary errors were independent with the same variance ( $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_2$ ) and also temporally independent ( $\mathbf{\Phi} = \mathbf{I}_n$ ). Based on work in magnitude fMRI time series, it is widely realized that the temporal independence assumption is not supported in reality for several reasons. For one, the hemodynamic response to a single neural activation takes six seconds to peak and 15–20 seconds to return to baseline and so the measured response occurs over several scans [Friston, Jezzard and Turner (1994)]. The subject’s cardiac and respiratory cycles provide another source of autocorrelation [Friston et al. (2000)]. Following Bullmore et al. (1996), who extended the linear model for magnitude-only time series by adding autoregressive (AR) errors, we apply AR( $p$ ) temporal dependence in the real and imaginary errors

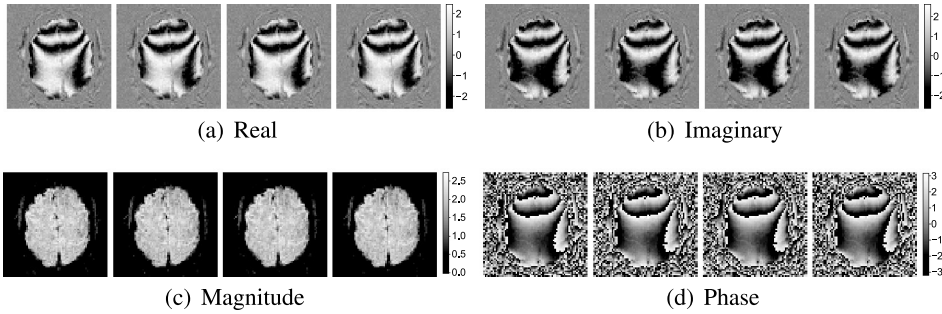


FIG. 1. Images of the (a) real, (b) imaginary, (c) magnitude, and (d) phase components of the middle slice for time points 1, 8, 15, and 22 (moving left to right).

of (1.1). We also generalize the real/imaginary covariance  $\Sigma$ , allowing separate variances for the real and imaginary errors as well as real/imaginary correlation.

Third, we study the effects of spatial smoothing on complex-valued fMRI data, including the correlations induced and their effect on activation detection. Nencka, Hahn and Rowe (2009) developed a theoretical framework for understanding the statistical effects of spatial smoothing (and other preprocessing operations), which was expanded upon in Bruce, Karaman and Rowe (2011) to examine the correlations induced by acquisition methods utilizing multiple receiver coils. However, to our knowledge, the effects of spatial smoothing upon complex-valued data-based activation detection have not previously been investigated.

The remainder of this paper is organized as follows. Section 2 introduces our showcase application, a complex-valued fMRI dataset acquired from a finger-tapping experiment, the type of which is utilized as a noninvasive neurosurgical planning tool. In this context, we introduce the complex-valued running line methodology and derive an LRT statistic for activation based on the extension of model (1.1) including temporal and real/imaginary correlation (Section 3). We compare this test statistic to a magnitude-only  $AR(p)$  model-based LRT statistic, analyzing the finger-tapping experiment dataset under different amounts of smoothing in Section 4 and running simulation studies in Section 5 to explore the nature and validity of the results. We conclude with some discussion in Section 6. This paper also has an online supplement [Adrian, Maitra and Rowe (2018)] providing further details on methodology, performance evaluations, data analysis, and simulation studies. Sections, figures, and tables in the supplement referred to in this paper are labeled with the prefix “S-.”

**2. Motivating example: Aiding neurosurgical planning via accurate activation detection in a finger-tapping experiment.** Our showcase application for this paper comes from a commonly performed unilateral sequential finger-tapping experiment. In this case, the MR images were acquired with  $TR = 1$  s while the (normal healthy male) volunteer was instructed through visual cues to either lie at

rest or to rapidly tap fingers of his right hand (hence unilateral) sequentially in the order of index, middle, ring, and little fingers. The experiment followed a block design with 16 epochs of 15s of tapping alternating with 15s of rest concluding with an additional 10s of rest at the end of the experiment. Hence, the dataset contains a temporal sequence of  $n = 490$  images for use in our analysis. Each image was composed of seven 2.5 mm thick  $96 \times 96$  axial slices with a 24.0 cm FOV. [The data are the last 490 time points analyzed in Karaman, Bruce and Rowe (2015), where it was inadvertently referred to as a bilateral finger-tapping experiment.] For this dataset, the phase components of the time series images were not discarded but stored along with the magnitude images used in traditional fMRI analysis. The data processing flow included Nyquist ghost removal and correction for global zero-order off-resonance using three navigator echos [Jesmanowicz, Wong and Hyde (1993), Nencka, Hahn and Rowe (2008)], image reconstruction from  $k$ -space by inverse Fourier transform [Kumar, Welte and Ernst (1975), Rowe (2016)], and estimation and correction of the dynamic field using temporal off-resonance alignment of single-echo timeseries (TOAST) [Hahn, Nencka and Rowe (2009, 2012)]. A binary mask of voxels above 15% of the maximum voxel signal magnitude was generated from the first image of the Karaman, Bruce and Rowe (2015) dataset, representing voxels within the brain.

Figure 1 shows images of the real, imaginary, magnitude, and phase components of the middle slice at time points one, eight, 15, and 22. (See Figure S-1 for displays of the other slices.) These images appear largely unchanged over time because the BOLD stimulus response is very small compared to the overall MR signal. Figure 2 displays the real, imaginary, magnitude, and phase time series for one of the voxels showing the most activation. It is evident that the real, imaginary, and magnitude time series at this voxel oscillate with a period similar to the stimulus time course; in contrast, the phase time series has relatively constant location (though the fit of the constant phase assumption has some room for improvement, as we will see).

Figure 2 also provides some pointers as to why the fMRI community has found it difficult to recognize that discarding half of the data (the phase) is not necessarily a prudent practice. This is because the magnitude data have the best definition, while the phase data have high variability and do not seemingly provide much visual content. However, the phase data have additional information on variability in the MR signal, and we contend and demonstrate that harnessing this additional information can improve specificity in fMRI analysis.

Paradigms involving hand stimulation, such as in the finger-tapping experiment introduced here, have deeper clinical significance, especially in the context of utilizing fMRI as a noninvasive neurosurgical planning tool for tumor and epilepsy patients. Such experiments have shown that the central sulcus is the location of hand function in the sensorimotor cortex for normal healthy adults [Rumeau et al. (1994)]. At risk of confusion, it is important to distinguish between the *functional* central sulcus and the *anatomical* central sulcus, as the two do not in general

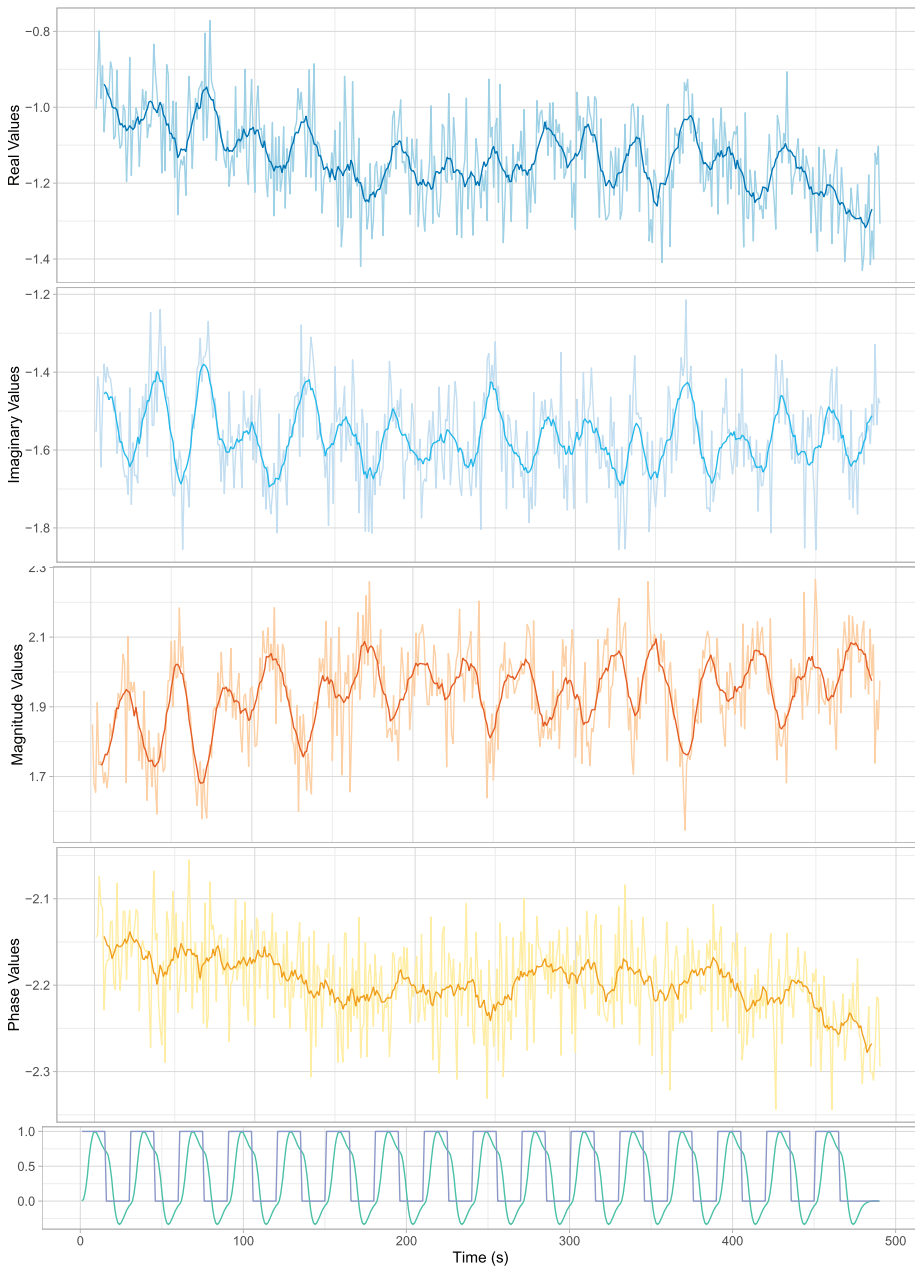


FIG. 2. Time series of (from top) the real, imaginary, magnitude, and phase observations for one of the voxels showing the most activation. Lighter lines in each display represents the raw time series, while the darker lines show the result after applying a simple, central moving average filter with 5 nearest neighbors. The bottom display is of the block design of the stimulus and its convolution with the HRF.

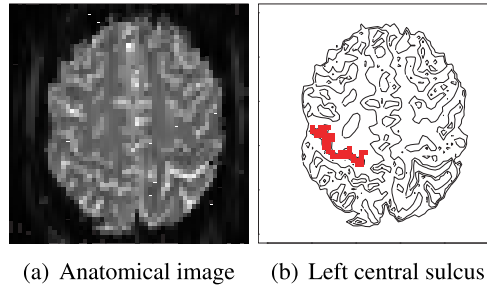


FIG. 3. (a) *Anatomical image of the fifth slice and (b) corresponding contour plot identifying the left central sulcus.*

agree. The functional central sulcus is the brain activation region of hand function determined by the fMRI experiment [Lee et al. (1999)]. On the other hand, the anatomical central sulcus refers to brain anatomy instead of function, in this case the sulcus, or fold, in the cerebral cortex which separates the frontal lobe from the parietal lobe. As a neurosurgical planning tool, such fMRI experiments seek to identify the functional central sulcus in order to better assess the risk/benefit of surgery through knowledge of the relative locations of the proposed surgical target and of specific functional areas on the cortical surface [Lee et al. (1999)]. Traditionally, determining such knowledge required an invasive procedure, but the usage of fMRI allows it to be determined noninvasively and preoperatively.

For a normal healthy subject, the functional central sulcus should align with its anatomical twin, but for many studies, “functional activation may extend over two or three sulci,” which results in “ambiguous identification of the (functional) central sulcus” [see Lee, Jack and Riederer (1998), who also assert that “any level of ambiguity is unacceptable for surgical mapping”]. Thus, methodology that reduces this ambiguity as far as possible is a desirable clinical goal. Figure 3 shows the location of the left anatomical central sulcus [Harnsberger et al. (2007)] on the  $T_2^*$ -weighted anatomical image of our subject’s brain used to create the mask. (We focus on the left central sulcus because its functional version will be activated by right-hand finger tapping due to the decussation of nerve fibers in the upper slices of the brain.) Therefore, the experiment here provides a test case for evaluating not only our methodology but also for the potential applicability of fMRI as a noninvasive clinical tool for the planning of neurosurgeries.

### 3. Methodological development.

3.1. *Trend removal via a complex-valued running line.* Our methodology for the complex-valued running line was motivated by approaches to remove trend in magnitude-data fMRI analysis, of which there are several. In our experience, global polynomial fitting methods are not sufficiently flexible to model the gamut of nonlinear trends, and frequency filtering methods can infer spurious trends at



the ends of the time series. A good alternative is a running line as introduced by Marchini and Ripley (2000), which is similar to a moving average except that a least squares linear fit is applied to the  $k$  nearest neighbors of every time point. For periodic designs, Marchini and Ripley (2000) suggest letting  $k$  equal twice the period, and we adopt this practice here. Using this value of  $k$  removes nuisance trends while preserving task-related magnitude and phase changes.

To extend this methodology to complex-valued time series, we fit the Rowe (2005) model in each neighborhood. Specifically we find the value of the complex-valued running line at  $t^*$  by fitting the model

$$(3.1) \quad \begin{pmatrix} y_{Rt} \\ y_{It} \end{pmatrix} = \begin{pmatrix} \rho_t \cos \theta_t \\ \rho_t \sin \theta_t \end{pmatrix} + \begin{pmatrix} \eta_{Rt} \\ \eta_{It} \end{pmatrix}$$

for  $t$  in the neighborhood  $N(t^*) = \{t : \max(t^* - k, 1) \leq t \leq \min(t^* + k, n)\}$ . In (3.1), the magnitude  $\rho_t$  and the phase  $\theta_t$  are linear functions of time: specifically,  $\rho_t = \beta_0 + s_t \beta_1$ , where  $s_t$  is the  $t$ th entry of an arithmetic sequence, and  $\theta_t = \gamma_0 + g(s_t \gamma_1)$ , where  $g(\cdot) = 2 \arctan(\cdot)$ . The adaptation of the link function  $g(\cdot)$  follows from Fisher and Lee (1992)'s methodology for regression with angular responses that was applied to phase-only fMRI time series in Rowe, Meller and Hoffman (2007) and allows unique ML estimation of  $\gamma_1$ . (It is recommended that the arithmetic sequence be centered around zero for faster convergence of the MLE of  $\gamma_1$ .) Details on the ML estimation of the parameters are given in Section S-2.1. Using these MLEs, the magnitude and phase of the complex-valued running line at time  $t^*$  are given by  $\hat{\rho}_{t^*} = \hat{\beta}_0 + s_{t^*} \hat{\beta}_1$  and  $\hat{\theta}_{t^*} = \hat{\gamma}_0 + g(s_{t^*} \hat{\gamma}_1)$ . (To save computer time, we can perform ML estimation for every  $j$  time points and use linear interpolation to obtain the values in between these time points.) The detrended complex-valued time series is then given by  $\tilde{r}_t e^{i \tilde{\phi}_t}$ ,  $t = 1, \dots, n$ , where  $\tilde{r}_t = r_t - \hat{\rho}_t + \bar{\rho}$ , with  $\bar{\rho} = (1/n) \sum_{t=1}^n \hat{\rho}_t$ , and  $\tilde{\phi}_t = \phi_t - \hat{\theta}_t + \theta_0$ , where  $\theta_0$  is an analyst-specified central phase that is the same at each voxel. Centering each phase time series around  $\theta_0$  establishes phase coherence which is needed because the phase does not generally match between slices—see Figure S-1(d) and Figure 6 of Jesmanowicz, Nencka and Hyde (2014). Phase coherence of the voxel time series is desirable for the application of spatial smoothing, which we discuss next.

*3.2. Spatial smoothing of complex-valued fMRI data.* Following detrending with the complex-valued running line, we smooth the real and imaginary images (separately) at each time point with a three-dimensional discrete isotropic Gaussian filter [Bruce, Karaman and Rowe (2011), Nencka, Hahn and Rowe (2009)]. We vary the amounts of smoothing by applying filters with different bandwidths expressed in terms of the full-width-at-half-maximum (FWHM). As a result, we produce a total of four datasets: three that are smoothed with FWHMs = 2, 4, and 6 voxels and one in which no smoothing is performed. We place special focus on the results of the smoothed dataset with FWHM = 4 voxels because this

represents a moderate amount of smoothing and is the default value used by the Analysis of Functional Neuroimaging (AFNI) software [Cox (1996), Cox (2012), Cox and Hyde (1997)].

3.3. *Overview of activation detection.* We compare the activation detected in these datasets by LRT statistics under complex-valued and magnitude-only data-based models, both of which assume  $AR(p)$  errors; the statistics are denoted by  $\Lambda_{C,p}$  and  $\Lambda_{M,p}$ , respectively. For now, the AR order  $p$  is assumed to be known; in practice,  $p$  will be detected via the methodology to be described in Section 3.5. As mentioned in Section 1, the complex-valued model is an extension of the Rowe and Logan (2004) model,

$$(3.2) \quad \begin{pmatrix} y_R \\ y_I \end{pmatrix} = \begin{pmatrix} X\boldsymbol{\beta} \cos \theta \\ X\boldsymbol{\beta} \sin \theta \end{pmatrix} + \begin{pmatrix} \eta_R \\ \eta_I \end{pmatrix},$$

where  $X$  contains columns for the baseline and expected BOLD response and we assume a constant phase location parameter  $\theta$ . (Due to the complex-valued running line, it is not necessary for  $X$  to contain additional columns, such as polynomials, to model the magnitude trend; further, the constant phase assumption is strengthened.) The errors  $(\eta'_R, \eta'_I)' \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi})$ , where

$$(3.3) \quad \boldsymbol{\Sigma} = \text{Cov}(y_{Rt}, y_{It}) = \begin{pmatrix} \sigma_R^2 & \rho\sigma_R\sigma_I \\ \rho\sigma_R\sigma_I & \sigma_I^2 \end{pmatrix}$$

and  $\boldsymbol{\Phi} = \mathbf{R}_n$  is such that  $\text{Cov}(\eta_R) = \sigma_R^2 \mathbf{R}_n$  and  $\text{Cov}(\eta_I) = \sigma_I^2 \mathbf{R}_n$ ; thus the model assumes that the autocorrelation of the real and imaginary errors are determined by the same AR coefficients  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)'$ . The magnitude-only model is given by  $\mathbf{r} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ , where  $X$  is the same as in (3.2) and  $\boldsymbol{\epsilon}$  follows a  $p$ th order autoregressive structure. (Note that the corresponding parameters in the magnitude-only and complex-valued models are not necessarily equivalent, but we do not distinguish them notationally for the sake of simplicity.)

Under either model, the activation test may be posed generally as  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$  vs.  $H_a : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0}$ , and the LRT statistics have  $\chi_m^2$  asymptotic null distributions, where  $m = \text{rank}(\mathbf{C})$ . We derive the complex-valued model-based LRT statistic (LRTS) in Section 3.4; we relegate magnitude-only LRTS derivations to the supplement (Section S-2.2). At each voxel, a  $p$ -value is computed from the obtained LRTS. We use cluster thresholding [Forman et al. (1995), Logan, Geliazkova and Rowe (2008)] on these  $p$ -values to determine activated voxels by tagging clusters of a minimum size where voxels have  $p$ -values below a specified significance level.

3.4. *Modeling for  $AR(p)$  complex-valued fMRI time series.* Denoting the model parameters by  $\boldsymbol{\tau} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \theta, \sigma_R^2, \sigma_I^2, \rho)$  and utilizing the properties of Kronecker products, model (3.2) has log-likelihood function given by

$$\log L(\boldsymbol{\tau} | \mathbf{y}_R, \mathbf{y}_I) = -n/2 \log[\sigma_R^2 \sigma_I^2 (1 - \rho^2)] - \log |\mathbf{R}_n| - h/[2(1 - \rho^2)],$$

where

$$(3.4) \quad h = \begin{pmatrix} \eta_R \\ \eta_I \end{pmatrix}' \begin{pmatrix} \frac{1}{\sigma_R^2} \mathbf{R}_n^{-1} & -\frac{\rho}{\sigma_R \sigma_I} \mathbf{R}_n^{-1} \\ -\frac{\rho}{\sigma_R \sigma_I} \mathbf{R}_n^{-1} & \frac{1}{\sigma_I^2} \mathbf{R}_n^{-1} \end{pmatrix} \begin{pmatrix} \eta_R \\ \eta_I \end{pmatrix}.$$

The maximum likelihood estimate (MLE) of  $\beta$  is

$$(3.5) \quad \hat{\beta} = \frac{\hat{\beta}_R [\frac{\cos \hat{\theta}}{\hat{\sigma}_R^2} - \frac{\hat{\rho} \sin \hat{\theta}}{\hat{\sigma}_R \hat{\sigma}_I}] + \hat{\beta}_I [\frac{\sin \hat{\theta}}{\hat{\sigma}_I^2} - \frac{\hat{\rho} \cos \hat{\theta}}{\hat{\sigma}_R \hat{\sigma}_I}]}{\frac{\cos^2 \hat{\theta}}{\hat{\sigma}_R^2} + \frac{\sin^2 \hat{\theta}}{\hat{\sigma}_I^2} - 2 \frac{\hat{\rho} \sin \hat{\theta} \cos \hat{\theta}}{\hat{\sigma}_R \hat{\sigma}_I}},$$

where  $\hat{\beta}_R = (\mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{X})^{-1} \mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{y}_R$ ,  $\hat{\beta}_I = (\mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{X})^{-1} \mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{y}_I$ , and  $\hat{\mathbf{R}}_n^{-1}$  is a function of  $\hat{\alpha}$ , the MLE of  $\alpha$ , according to the  $(2p + 1)$ -diagonal matrix given in Pourahmadi (2001), Section 4.4. [The sparseness of the matrix  $\hat{\mathbf{R}}_n^{-1}$  is a computational advantage of the AR( $p$ ) model.] Further, the MLE of  $\theta$  is

$$(3.6) \quad \hat{\theta} = [\arcsin(c/\sqrt{a^2 + b^2}) - \arctan_4(b, a)]/2,$$

where  $a$ ,  $b$ , and  $c$  (defined in terms of  $d$ ,  $e$ , and  $f$  to be defined a little later) are given by  $a = d/\hat{\sigma}_I^2 - e/\hat{\sigma}_R^2$ ,  $b = -f(1/\hat{\sigma}_R^2 + 1/\hat{\sigma}_I^2) - \hat{\rho}(d + e)/(\hat{\sigma}_R \hat{\sigma}_I)$ , and  $c = \hat{\rho}(d - e)/(\hat{\sigma}_R \hat{\sigma}_I) + f(1/\hat{\sigma}_R^2 - 1/\hat{\sigma}_I^2)$ . The terms  $d$ ,  $e$ , and  $f$  are given by  $d = B_{RR}/\hat{\sigma}_R^4 + \hat{\rho}^2 B_{II}/(\hat{\sigma}_R^2 \hat{\sigma}_I^2) - 2\hat{\rho} B_{RI}/(\hat{\sigma}_R^3 \hat{\sigma}_I)$ ,  $e = B_{II}/\hat{\sigma}_I^4 + \hat{\rho}^2 B_{RR}/(\hat{\sigma}_R^2 \hat{\sigma}_I^2) - 2B_{RI}\hat{\rho}/(\hat{\sigma}_R \hat{\sigma}_I^3)$ , and  $f = (1 + \hat{\rho}^2)B_{RI}/(\hat{\sigma}_R^2 \hat{\sigma}_I^2) - \hat{\rho}/(\hat{\sigma}_R \hat{\sigma}_I)(B_{RR}/\hat{\sigma}_R^2 + B_{II}/\hat{\sigma}_I^2)$ , where we have  $B_{RR} = \hat{\beta}'_R (\mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{X}) \hat{\beta}_R$ ,  $B_{II} = \hat{\beta}'_I (\mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{X}) \hat{\beta}_I$ , and  $B_{RI} = \hat{\beta}'_R (\mathbf{X}' \hat{\mathbf{R}}_n^{-1} \mathbf{X}) \hat{\beta}_I$ . The MLEs for  $\sigma_R^2$ ,  $\sigma_I^2$ , and  $\rho$  are given by

$$(3.7) \quad \hat{\sigma}_R^2 = (\mathbf{y}_R - \mathbf{X} \hat{\beta} \cos \hat{\theta})' \hat{\mathbf{R}}_n^{-1} (\mathbf{y}_R - \mathbf{X} \hat{\beta} \cos \hat{\theta})/n;$$

$$(3.8) \quad \hat{\sigma}_I^2 = (\mathbf{y}_I - \mathbf{X} \hat{\beta} \sin \hat{\theta})' \hat{\mathbf{R}}_n^{-1} (\mathbf{y}_I - \mathbf{X} \hat{\beta} \sin \hat{\theta})/n;$$

$$(3.9) \quad \hat{\rho} = (\mathbf{y}_R - \mathbf{X} \hat{\beta} \cos \hat{\theta})' \hat{\mathbf{R}}_n^{-1} (\mathbf{y}_I - \mathbf{X} \hat{\beta} \sin \hat{\theta})/(n \hat{\sigma}_R \hat{\sigma}_I).$$

Lastly, we obtain  $\hat{\alpha}$  by solving the system of equations [Miller (1995)]

$$(3.10) \quad \hat{d}_{0k} = \sum_{j=1}^p [\hat{d}_{ij} + (j/n) \hat{d}_{0,|j-k|}] \hat{\alpha}_j,$$

for  $k = 1, \dots, p$ , with  $\hat{d}_{ij} = \hat{d}_{ij}^{(RR)}/\hat{\sigma}_R^2 + \hat{d}_{ij}^{(II)}/\hat{\sigma}_I^2 - \hat{\rho}/(\hat{\sigma}_R \hat{\sigma}_I) [\hat{d}_{ij}^{(RI)} + \hat{d}_{ij}^{(IR)}]$ ,  $0 \leq i, j \leq p$ . In the preceding development, we have  $\hat{d}_{ij}^{(\iota\zeta)} = \sum_{t=1}^{n-i-j} \hat{\eta}_{\iota, t+i} \hat{\eta}_{\zeta, t+j}$  with  $\iota, \zeta \in R, I$ , wherein  $\hat{\eta}_{Rt} = \mathbf{y}_{Rt} - \mathbf{x}'_t \hat{\beta} \cos \hat{\theta}$ ,  $\hat{\eta}_{It} = \mathbf{y}_{It} - \mathbf{x}'_t \hat{\beta} \sin \hat{\theta}$ , and  $\mathbf{x}'_t$  is the  $t$ th row of  $\mathbf{X}$ ,  $t = 1, \dots, n$ . In practice, starting values for  $\hat{\beta}$  and  $\hat{\theta}$  can be computed under the assumption of independence [see equation (2.10) in Rowe and

Logan (2004)]. ML estimation then consists of iteratively updating (3.5)–(3.10) until convergence. The LRT statistic is

$$(3.11) \quad \Lambda_{C,p} = n \log \left( \frac{\tilde{\sigma}_R^2 \tilde{\sigma}_I^2 (1 - \tilde{\rho}^2)}{\hat{\sigma}_R^2 \hat{\sigma}_I^2 (1 - \hat{\rho}^2)} \right) - 2 \log(|\tilde{\mathbf{R}}_p^{-1}| / |\hat{\mathbf{R}}_p^{-1}|),$$

where  $\tilde{\sigma}_R^2$ ,  $\tilde{\sigma}_I^2$ , and  $\tilde{\rho}$  are restricted MLEs under  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$  (see Section S-2.3),  $\mathbf{R}_p$  is the  $p \times p$  correlation matrix such that  $\text{Cov}(\eta_{R1}, \dots, \eta_{Rp}) = \sigma_R^2 \mathbf{R}_p$  and  $\text{Cov}(\eta_{I1}, \dots, \eta_{Ip}) = \sigma_I^2 \mathbf{R}_p$  and  $\hat{\mathbf{R}}_p^{-1}$  and  $\tilde{\mathbf{R}}_p^{-1}$  are functions of  $\hat{\boldsymbol{\alpha}}$  and  $\tilde{\boldsymbol{\alpha}}$ , respectively, as in Pourahmadi (2001), Section 4.4. [Equation (3.11) relies on the equality  $|\mathbf{R}_n| = |\mathbf{R}_p|$ , which is another computational advantage of the AR model.]

**3.5. Choosing the order of the AR model.** We suggest a sequential testing approach to decide on the AR order  $p$ . Starting with  $k = 1$ , and for increasing  $k$ , we posit  $H_0 : p = k - 1$  vs.  $H_a : p \geq k$  (or, in terms of the AR coefficients,  $H_0 : \forall j \geq k, \alpha_j = 0$  vs.  $H_a : \exists j \geq k : \alpha_j \neq 0$ ). The estimated AR order is then  $\hat{p} = k' - 1$ , where  $k'$  is the first  $k$  in the sequence of tests for which  $H_0$  can not be rejected. Note that, in the procedure, the significance level controls the probability of over detection of the order (i.e., the order detected is greater than the true order) in the sense described in Section S-2.4. For both the magnitude-only and the complex-valued models, we employ a LRT-based test statistic given by  $2(\hat{\ell}_k - \hat{\ell}_{k-1})$ , where  $\hat{\ell}_k$  is the optimized log-likelihood for the AR( $k$ ) model. From standard results, this test statistic is asymptotically  $\chi_1^2$ -distributed under  $H_0 : p = k - 1$ . We denote the orders detected by the magnitude-only and complex-valued model test statistics by  $\hat{p}_M$  and  $\hat{p}_C$ , respectively.

**4. Application to fMRI study.** We discuss the application of the methodology to the finger-tapping dataset described in Section 2.

**4.1. Analysis pipeline and specifications.** We analyzed the complex-valued (CV) and magnitude-only (MO) datasets separately, yet in tandem. First, the CV running line was applied to each CV voxel time series, and a regular running line was applied to each magnitude voxel time series. The central phase of each detrended CV time series was set to be  $\theta_0 = \pi/4$  (for reasons discussed in Section S-3.3). Then, three datasets were created from both the CV and MO datasets by smoothing the real, imaginary, and magnitude images at each time point with 3-dimensional Gaussian filters with FWHMs = 2, 4, and 6 voxels; thus, including the spatially unsmoothed datasets, there were a total of four CV and MO datasets. The CV and MO models were then applied to each voxel time series. For both models, the design matrix  $\mathbf{X}$  had  $n = 490$  rows and  $q = 2$  columns: one column was for an intercept modeling the baseline MR signal and the other was the waveform modeling the expected BOLD response [Glover (1999)]. Consequently, with  $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ , the activation tests posited  $H_0 : \beta_1 = 0$  vs.  $H_a : \beta_1 \neq 0$  with the

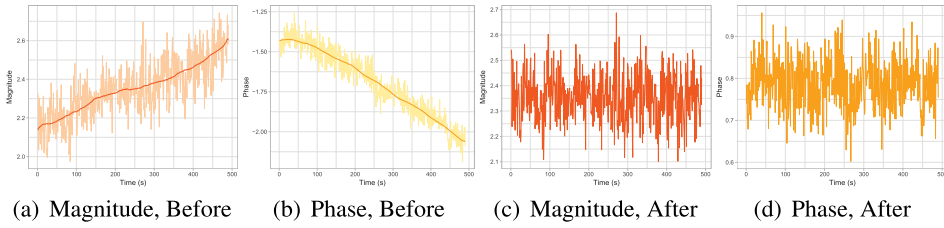


FIG. 4. A voxel's (a), (c) magnitude and (b), (d) phase time series (a), (b) before and (c), (d) after correction by the complex-valued running line. Figures (a), (b) also show the CV running-line fit.

LRT statistics being  $\chi_1^2$ -distributed under  $H_0$ . For both the CV and MO models, we computed two test statistics, one using the detected AR order at each voxel and, for comparison, also with  $p \equiv 0$ . Thus, we compared the performance of four LRT statistics, denoted by  $\Lambda_{C,\hat{p}_C}$ ,  $\Lambda_{M,\hat{p}_M}$ ,  $\Lambda_{C,0}$ , and  $\Lambda_{M,0}$ . For each LRT statistic (LRTS) and smoothing level, we computed activation maps using cluster thresholding with third-order neighborhood structure, a size threshold of 11 voxels [Forman et al. (1995)], and a significance level threshold of 0.001 [as recommended by Woo, Krishnan and Wager (2014)].

**4.2. Addressing model fit and model assumptions.** We first examined the performance of the CV running line. Figure 4 shows a voxel's magnitude and phase time series before and after correction by the CV running line and suggests that nuisance trends were well removed in both time series. For a global assessment, we fit model (3.1) to the entirety of each voxel time series (not just the nearest neighbors) and computed an LRTS for  $H_0 : \gamma_1 = 0$  vs.  $H_a : \gamma_1 \neq 0$ . Using a Bonferroni-corrected significance level of 0.05, 63.3% of the voxels rejected  $H_0$ . This suggests that even when with the application of advances such as TOAST [Hahn, Nencka and Rowe (2009, 2012)] (done here, see Section 2) to stabilize the phase data, a phase detrending preprocessing step such as the CV running line may be needed to allow for the constant phase location assumption of model (3.2). This preprocessing step is important because phase instability can degrade the power of complex-valued data-based activation tests [Hahn, Nencka and Rowe (2009)], making them less powerful than their magnitude-only data-based counterparts.

We refer to Section S-3.1 for additional assessments of complex-valued model fit and assumptions. These include assessments of the assumptions of phase coupling and the equality of the AR coefficients for the real and imaginary errors, an assessment of the autoregressive fit to the autocorrelation, and a comparison of the computational times for the CV and MO data-based analyses.

**4.3. Activation maps.** Figure 5 displays activation maps under the four LRTSs for the fifth slice of the dataset smoothed three-dimensionally with  $\text{FWHM} = 4$  voxels. We display the fifth slice in order to compare the regions detected to be

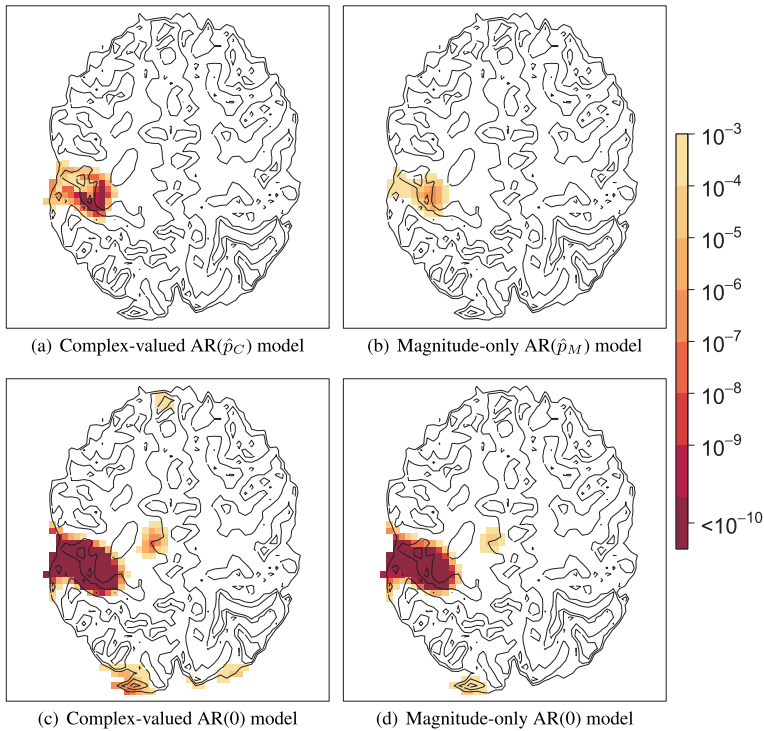


FIG. 5. Activation maps for the finger-tapping experiment after cluster thresholding. The  $p$ -value of the observed test statistic for each voxel identified as activated under each model is overlaid on the anatomic slice under investigation.

functionally active with the anatomical left central sulcus in Figure 3(b). Although all four statistics show functional activation in the anatomical left central sulcus (as it should for our normal subject), the activation map using  $\Lambda_{C, \hat{p}_C}$  identifies it with the least ambiguity. For instance, the  $\Lambda_{M, \hat{p}_M}$  map shows less significant  $p$ -values in the left central sulcus, suggesting that it is less powerful than  $\Lambda_{C, \hat{p}_C}$ . Further, the independent error model-based maps detect brain regions outside the left central sulcus, and detection inside the region is more diffuse than the  $\Lambda_{C, \hat{p}_C}$  map. Thus, the superior ability of the complex-valued  $AR(p)$  model to more accurately identifying the correct activation region points to its applicability in fMRI finger stimulation experiments for neurosurgical planning in tumor and epilepsy patients, where the goal is to locate the functional central sulcus accurately (vide our discussion in Section 2).

For a broader view, Section S-3.2 presents activation images for three slices and using smoothing levels of  $FWHMs = 2, 4,$  and  $6$  voxels and no smoothing. The better performance of the complex-valued  $AR(p)$  model-based statistic held over all slices and smoothing kernels. However, with spatially unsmoothed data, there was much less difference between the maps of the four activation statistics,

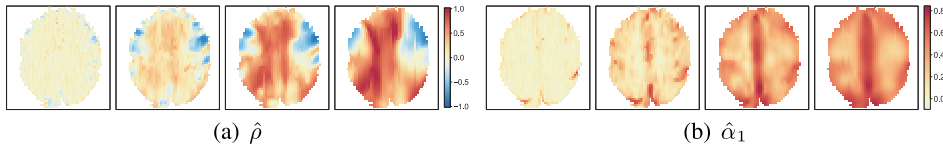


FIG. 6. Voxel-wise parameter estimates (a)  $\hat{\rho}$  and (b)  $\hat{\alpha}_1$  for (moving left to right within each subfigure) spatial smoothing with FWHM = 0, 2, 4, and 6 voxels. (FWHM = 0 means no spatial smoothing.)

and it was difficult to determine a clear winner. We explain this result in terms of the effect of smoothing (or none) on model parameter estimates in the next section. Section S-3.2 also displays maps obtained by false discovery rate (FDR) thresholding and assuming  $p = 1$ . Results (see Figure S-7) were similar to those in Figure 5.

4.4. *Effect of spatial smoothing on parameter estimates.* We examined the effect of smoothing on the complex-valued AR( $p$ ) model parameter estimates and the activation maps. (Recall from the previous section that the  $\Lambda_{C, \hat{\rho}_C}$  maps were favored when spatial smoothing was performed at any level, but there was not much difference between the maps for the unsmoothed data.) For easier interpretation of the extent of temporal correlation, we assumed  $p = 1$  for each voxel. (Spatial and frequency distributions of the detected AR orders are given in Figures S-9, S-10 and and S-11.) Figure 6 contains images of the real/imaginary correlation  $\hat{\rho}$  and the temporal AR parameter  $\hat{\alpha}_1$  for the fifth slice (corresponding images for the other parameters are given in Figure S-12). The unsmoothed data,  $\hat{\rho}$  and  $\hat{\alpha}_1$  are close to zero, indicating that the independence assumptions applied in Rowe and Logan (2004) may be adequate. This explains why the independent model-based statistics had similar activation maps to their counterparts that assumed an AR model for the unsmoothed data. Further, simulations in Rowe and Logan (2004) indicated that at the SNRs of the finger-tapping dataset, there were very small differences between the activation detection power between the complex-valued and magnitude-only model-based statistics under the assumptions of temporal and real/imaginary independence. However, we found that increased smoothing was associated with greater values of  $|\hat{\rho}|$  and  $\hat{\alpha}_1$ . This result may be unexpected: while it would be expected that the spatial smoothing would induce spatial correlation [Nencka, Hahn and Rowe (2009)], it may be puzzling that it would also induce temporal correlation—in this case, both within and between the real and imaginary time series. Figure 7 displays the real and imaginary time series for a representative voxel [see Figure 8(a)] for the unsmoothed and smoothed (FWHM = 4 voxels) data and provides another view of our remarks above. For the unsmoothed data, the real and imaginary time series resemble white noise and the real/imaginary components are not correlated; however, with spatial smoothing, the path of the time series resembles an AR process with correlated real and

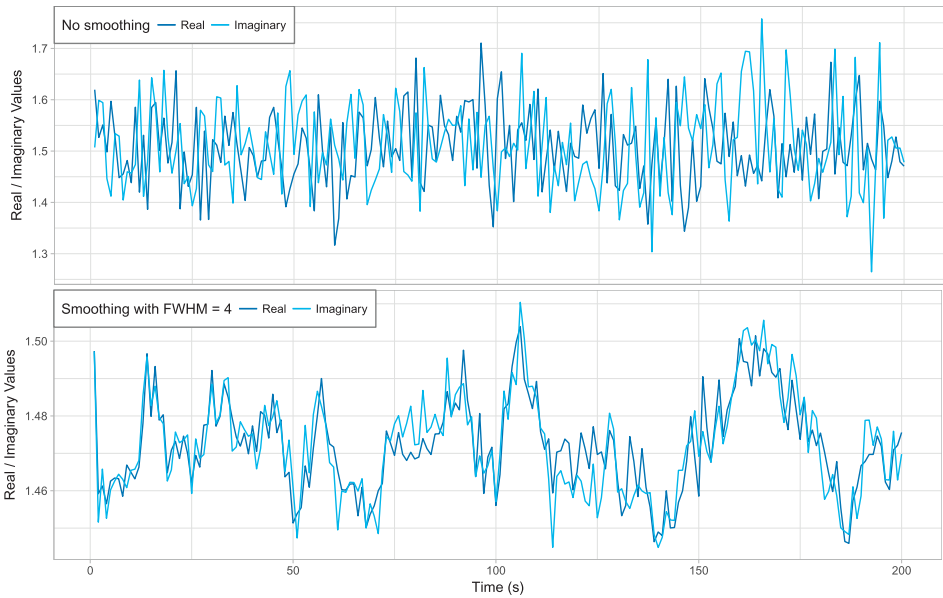


FIG. 7. Time series plots of the real and imaginary values—showing only the first 200 s for ease of viewing—for the representative voxel in Figure 8(a) for the (top) unsmoothed and (bottom) smoothed data with  $FWHM = 4$  voxels.

imaginary components. This may be explained by the application of smoothing on spatio-temporal data [Karaman et al. (2014)]. In summary, the assumption of temporal and real/imaginary independence is increasingly untenable with increasing smoothing (also see Figure S-13). This may be a reason behind the more spurious and diffuse activation displayed in the independent model-based activation maps. It may also explain the apparent reduced power of the magnitude-only  $AR(p)$  model-based map, as the real/imaginary correlation cannot be estimated using magnitude-only data. We report our tests of these assertions with simulation studies in Section 5.

By way of explaining the effect of spatial smoothing on the parameter estimates, we offer the following spatiotemporal model for complex-valued fMRI data following detrending via the CV running line. Letting  $y_{Rst}$  and  $y_{Ist}$  be the real and imaginary measurements at voxel  $s$  and time  $t$ , respectively, the model states that

$$(4.1) \quad \begin{pmatrix} y_{Rst} \\ y_{Ist} \end{pmatrix} = (\mathbf{x}'_t \boldsymbol{\beta}_s + v_{st}) \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} + \begin{pmatrix} w_{Rst} \\ w_{Ist} \end{pmatrix},$$

where  $\mathbf{x}'_t \boldsymbol{\beta}_s$  is the mean response and  $\theta$  is the central phase as before (and is equal for all  $s$  following the application of the CV running line), and  $v_{st}$  and  $(w_{Rst}, w_{Ist})'$  represent two sources of variation. Thus, (4.1) suggests that the real and imaginary errors in model (3.2) can be further decomposed into



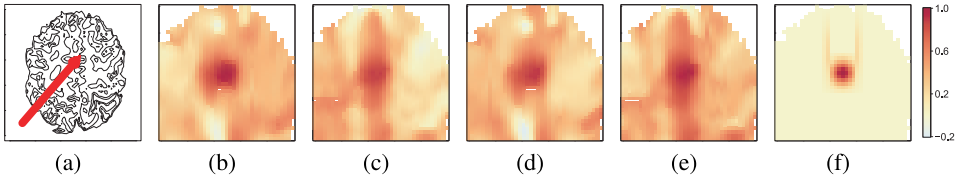


FIG. 8. (a) Location of voxel in the fifth slice from which our representative time series was chosen. Correlation between the real component of the time series at the voxel and each (b) real or (c) imaginary time series in a  $37 \times 37$  spatial neighborhood centered at the voxel in (d). Corresponding correlations between the imaginary time series at the voxel in (a) and the (d) real and (e) imaginary time series. (f) Correlation expected in (b) and (e) due to spatial smoothing assuming independence. [Such correlations in (c) and (d) are expected to be zero because the real and imaginary component images are smoothed separately.]

$\eta_{Rst} = v_{st} \cos \theta + w_{Rst}$  and  $\eta_{Ist} = v_{st} \sin \theta + w_{Ist}$ . This decomposition is supported by the fMRI literature, where it is well established that the (magnitude) noise contains multiple components, including “both white [e.g., thermal (Johnston) noise] and colored components [e.g., pulsatile motion of the brain caused by cardiac cycles or local modulation of the static magnetic field ( $B_0$ ) by respiratory movement]” [Friston et al. (2000)]. [See also Purdon and Weisskoff (1998), who model fMRI noise time series as an AR(1) time series plus added white noise.] In our model,  $v_{st}$  represents the colored (i.e., autocorrelated) and  $(w_{Rst}, w_{Ist})$  represent the white components. In addition to being autocorrelated, the model assumes that  $v_{st}$  are spatially correlated [Zarahn, Aguirre and D’Esposito (1997)]; for example, their covariance could be modeled as a Kronecker product of spatial and temporal AR(1) covariance matrices, as in Kruggel and von Cramon (1999). This fits the framework of Karaman et al. (2014) for describing the effects of processing such as spatial smoothing on spatiotemporally uncorrelated, temporally correlated, spatially correlated, and spatiotemporally correlated data. Figure 8 demonstrates that the spatial correlations in the finger-tapping dataset are much more extensive than would be expected by spatial smoothing alone. In contrast,  $w_{Rst}, w_{Ist}$  are assumed to be mutually, temporally, and spatially independent; the independence of the real and imaginary thermal noise components is well established [Gudbjartsson and Patz (1995)].

We now discuss how model (4.1) supports the observed dependence of the parameter estimates on spatial smoothing. First, the fact that the temporal and real/imaginary cross-correlations are near-zero for the unsmoothed data suggests that the variances of  $w_{Rst}$  and  $w_{Ist}$  are larger than  $v_{st}$ . However, under spatial smoothing, the variance due to  $(w_{Rst}, w_{Ist})$  decreases—more rapidly with increasing FWHM—due to the averaging of the independent random variables. On the other hand, because the  $v_{st}$ s are spatially correlated, their fluctuations are less affected by such averaging. The net effect is that the temporal correlation increases with spatial smoothing. Also, our choice of  $\theta_0 = \pi/4$  for all voxels in applying the CV

running line means that the real and imaginary time series share a common  $v_{st}$  in model (4.1). Thus, the fact that smoothing reduces the variation of  $(w_{Rst}, w_{Ist})$  more than  $v_{st}$  means that it increases the real/imaginary cross-correlation as well. To clarify the above discussion, we present some derivations under simplifying assumptions in Section S-3.3.1.

To conclude, it is interesting to note that the parameter estimates also depend on the choice of the central phase  $\theta_0$ . We explore this fact in Section S-3.3.2, including how it is expected under model (4.1) and that the choice of  $\theta_0 = \pi/4$  supports the assumption that the real and imaginary errors share the same AR coefficients.

**5. Simulations.** Analysis of the finger-tapping dataset in Section 4 yielded the observation that if the dataset was smoothed, the complex-valued AR( $p$ )-model-based activation map identified the appropriate region with the least ambiguity. With unsmoothed data however, the maps provided by the four competing statistics  $\Lambda_C, \hat{\rho}_C, \Lambda_M, \hat{\rho}_M, \Lambda_{C,0}$ , and  $\Lambda_{M,0}$  were very similar. We attributed this to the fact that the estimated temporal and real/imaginary correlations were negligible without smoothing but sizable with smoothing. To understand and clarify these results in settings of known ground truth, we performed simulation studies to examine (1) the effect of the real/imaginary covariance, including the real/imaginary correlation and the relative sizes of the real and imaginary variances (Section 5.2), (2) the effect of temporal correlation (Section 5.3), and (3) the effect of spatial smoothing through simulating a region of interest (ROI) (Section 5.4).

5.1. *Simulation setup.* We simulated complex-valued voxel time series according to (3.2) with  $\mathbf{X}$  specified as in Section 4.1. From properties of the conditional multivariate normal distribution, these simulated real and imaginary time series, denoted by  $\mathbf{y}_R^*$  and  $\mathbf{y}_I^*$ , respectively, can be generated as

$$(5.1) \quad \begin{aligned} \mathbf{y}_R^* &\sim N(\mathbf{X}\boldsymbol{\beta} \cos \theta, \sigma_R^2 \mathbf{R}_n), \\ \mathbf{y}_I^* | \mathbf{y}_R^* &\sim N(\mathbf{X}\boldsymbol{\beta} \sin \theta + \rho \sigma_I / \sigma_R (\mathbf{y}_R^* - \mathbf{X}\boldsymbol{\beta} \cos \theta), \sigma_R^2 (1 - \rho^2) \mathbf{R}_n). \end{aligned}$$

Because we generated magnitude and phase data without nuisance trends, we did not apply the CV running line. Further, we did not apply spatial smoothing to the simulated data, but instead quantified its effect through the simulation parameters. Otherwise, our simulation studies mimicked the finger-tapping data analysis pipeline in using the same  $\mathbf{X}$ , order detection procedure,  $H_0$  and  $H_a$ , and specifications for cluster thresholding. The corresponding magnitude time series were computed from these simulated complex-valued time series, resembling fMRI data collection in practice. In practice fMRI data are unitless, so we parameterized the simulations through the SNR  $\equiv \beta_0/\sigma$  and the CNR  $\equiv \beta_1/\sigma$ , where  $\sigma^2 = (\sigma_R^2 + \sigma_I^2)/2$ , making the results applicable to any fMRI dataset. Unless otherwise specified, the simulation parameters were SNR = 190, CNR = 0.05,  $\theta = \pi/4$ , and  $\sigma = 0.011$ , which are typical estimates from the dataset smoothed

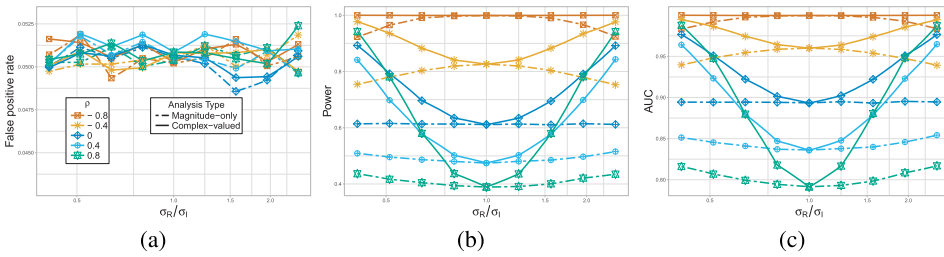


FIG. 9. Plots showing the (a) false positive rate, (b) power and (c) AUC for the complex-valued and magnitude-only activation statistics in terms of  $\sigma_R/\sigma_I$  and  $\rho$ . Color and line types indicate the same settings across all displays.

with  $\text{FWHM} = 4$  voxels. Other parameters varied according to the study performed. At each set of parameters, we replicated 100,000 time series (Sections 5.2 and 5.3) and 10,000 ROIs (Section 5.4).

We evaluated the activation statistics using measures that differed according to the simulation study. In Sections 5.2 and 5.3, we generated voxel time series without a spatial context and used standard hypothesis testing methodology. In these studies, we calculated the true and false positive rates as the proportions of truly activated and nonactivated time series detected as activated, using the  $\chi_1^2$  null distribution. For a test statistic to be of any practical use, its false positive rate must (at least approximately) be equal to the significance level. In addition, we calculated the area under the receiver operating characteristic (ROC) curve (AUC) as per calculations detailed in Section S-4.1. For evaluating performance in obtaining activation maps in Section 5.4, so we used the Jaccard (1901) Index (JI) to measure the percent overlap between the estimated and true activation maps for different statistics. The JI, or *modified overlap*, introduced by Maitra (2010) in fMRI as a preferable alternative to the more-commonly used *overlap* or Dice coefficient [Dice (1945), Sørensen (1948)], is defined between the  $i$ th and  $j$ th activation maps as  $m\omega_{ij} = V_{ij}/(V_i + V_j - V_{ij})$ , where  $V_i$ ,  $V_j$ , and  $V_{ij}$  are the number of voxels declared activated in map  $i$ , map  $j$ , and both maps  $i$  and  $j$ , respectively. Thus  $m\omega_{ij}$  can be interpreted as the proportion of voxels that are activated in both maps  $i$  and  $j$  from among those activated in either of them.

5.2. *Effect of covariance of real and imaginary errors.* To examine the effect of the covariance  $\Sigma$ , we varied the ratio of standard deviations  $\sigma_R/\sigma_I$  and the correlation coefficient  $\rho$ , with the former varied such that  $\sigma$  is constant so the SNR and CNR remain unaffected. To isolate this effect, we assumed temporal independence when simulating time series and calculating activation statistics. Figure 9(a) shows that the false positive rates of both the CV- and MO-model-based statistics are approximately equal to the significance of 0.05 at all covariance parameters, so both statistics are potentially useful. However, the powers and AUCs of  $\Lambda_{C,0}$  is greater than  $\Lambda_{M,0}$  whenever the real and imaginary variances are unequal, and

increasingly so as  $\sigma_R/\sigma_I$  gets farther from one. In view of this result, we note that the distributions of  $\hat{\sigma}_R/\hat{\sigma}_I$  for the finger-tapping dataset at various smoothing levels are given in Figure S-12(f). We see that  $\hat{\sigma}_R/\hat{\sigma}_I$  tends to be close to one for the unsmoothed data but is farther from one for the smoothed data. This is a potential reason that the  $\Lambda_{C, \hat{p}_C}$  map showed more power than the  $\Lambda_{M, \hat{p}_M}$  map for the smoothed data, but the two maps were similar for the unsmoothed data.

Interestingly also, the power and AUC functions in Figure 9 are greater for negative  $\rho$  than for positive  $\rho$ . In fact, this is part of a larger issue where the specific form of the power/AUC functions in terms of  $\sigma_R/\sigma_I$  and  $\rho$  depends on the value of  $\theta$ . This is discussed in detail in Section S-4.2. No matter the values of  $\theta$  and  $\Sigma$ , however, the power and AUC of the CV model-based activation statistic are always greater than or equal to its MO model-based counterpart.

**5.3. Varying the AR(1) coefficient.** To examine the effect of the temporal dependence upon the activation statistics, we simulated time series with AR(1) errors with coefficients varying from 0 to 0.8. To isolate this effect, we let  $\sigma_R = \sigma_I$  and  $\rho = 0$ . (A separate simulation with  $\rho = 0.5$ , not shown here, gave nearly identical results.) We first studied the ability of each model to detect the correct AR order. Figure 10(a) demonstrates that the complex-valued model detects the true order (1) for a greater proportion of simulated time series than the magnitude-only model. This greater power to reject  $H_0 : \alpha_1 = 0$  may be attributed to the fact that the complex-valued dataset contains twice as many autocorrelated quantities as the magnitude-only data. Note also that most of the errors are of under-detection of the order since the significance level of 0.05 controls over-detection errors.

Zarahn, Aguirre and D’Esposito (1997) found that the false positive rates of (magnitude-only data-based) activation statistics under the assumption of independence are inflated. To see whether this result carries over to the context of the complex-valued data, we computed the activation statistics  $\Lambda_{C, \hat{p}_C}$  and  $\Lambda_{M, \hat{p}_M}$

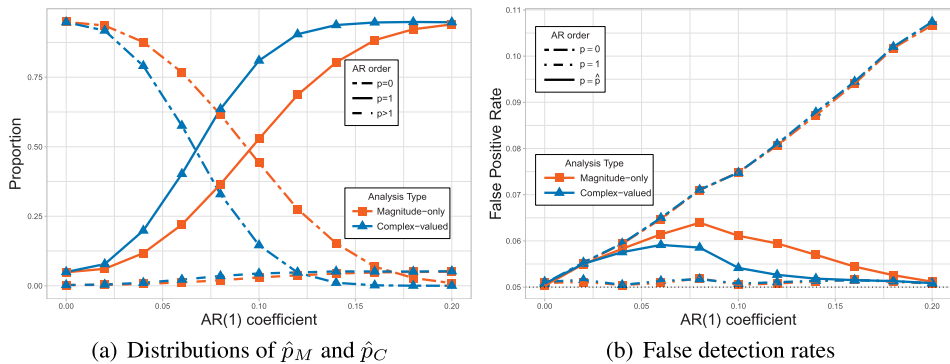


FIG. 10. Effect, for the complex-valued and magnitude-only fits, of increasing the AR(1) coefficient on (a) order detection and (b) the false positive rate.

(based on the detected orders) and also  $\Lambda_{C,p}$  and  $\Lambda_{M,p}$  for  $p = 0, 1, 2$ . As shown in Figure 10(b), compared to the significance level of 0.05, the false positive rates of the LRTSs under independence are indeed inflated, and increasingly so as the AR(1) coefficient increases. (Remarkably, the false positive rate balloons to 0.38 for  $\alpha_1 = 0.8$ .) Thus, it is evident that independent-error-based activation statistics are unreliable even with small amounts of autocorrelation. With the correct order, however, the LRTSs properly control the false positive rates. This control also holds when the order is overspecified; note that the lines for  $p = 2$  are not shown because they coincide with those for  $p = 1$ . Our results here are thus similar to those in Section 4 where the spurious activation regions detected under the independence-assumed maps for the smoothed data could be considered false positives.

The false positive rates of  $\Lambda_{M,\hat{p}_M}$  and  $\Lambda_{C,\hat{p}_C}$  are in between those of the correct order-based statistics and the independence-based statistics because of order detection errors. However, the false positive rate of  $\Lambda_{C,\hat{p}_C}$  is closer to 0.05 than  $\Lambda_{M,\hat{p}_M}$ , which points to another advantage of the complex-valued model. To explain why, let  $P_f(\cdot)$  be the false positive rate, which can be expressed as the weighted average

$$(5.2) \quad P_f(\Lambda_{D,\hat{p}_D}) = \sum_k P(\hat{p}_D = k) P_f(\Lambda_{D,k}),$$

where  $D = C$  or  $M$  for the complex-valued or magnitude-only data models, respectively. In this case, we can approximate the average over all orders  $k$  to include only  $k = 0$  and 1 for two reasons: first, the proportion of voxels with  $\hat{p} > 1$  is small, and second,  $P_f(\Lambda_{D,k}) \approx P_f(\Lambda_{D,1})$  for  $k > 1$ . Thus, we have the approximation

$$(5.3) \quad P_f(\Lambda_{D,\hat{p}_D}) \approx P(\hat{p}_D = 0) P_f(\Lambda_{D,0}) + P(\hat{p}_D \geq 1) P_f(\Lambda_{D,1}).$$

The fact that  $P_f(\Lambda_{M,\hat{p}_M}) > P_f(\Lambda_{C,\hat{p}_C})$  then follows from that  $P_f(\Lambda_{D,0})$  is greater than  $P_f(\Lambda_{D,1})$ , for  $D = C, M$ , and that  $P(\hat{p}_M = 0) > P(\hat{p}_C = 0)$ .

Section S-4.3 reports results of simulations to study the effect of autoregressive dependence for lags more than one. To conclude, we have shown that order under-detection inflates the false activation detection rate but over-detection does not, so we recommend using a larger significance level for sequentially detecting the order.

5.4. *Examining a region of interest (ROI).* To perform a simulation study incorporating the spatial aspect of fMRI analysis, we simulated from the small region (ROI) of the finger-tapping dataset shown in Figure 11(a). The voxel time series within the ROI were simulated using model (3.2) with  $p = 1$  and the four sets of parameter estimates in Table 1, each of which is representative of these estimates under one of the four levels of spatial smoothing.

For each parameter set, we used the same parameters for each voxel in the replicate ROIs except for  $\beta_1$ , which was zero or positive depending on whether the voxel was part of the anatomical left central sulcus in Figure 3(b). We computed

TABLE 1  
Parameter values used in simulating the ROIs in Section 5.4

FWHM	$\beta_0$	$\beta_1$	$\theta$	$\sigma_R$	$\sigma_I$	$\rho$	$\alpha_1$
0	1.89	0.0090	$\pi/4$	0.0887	0.0882	0.011	0.036
2	1.89	0.0027	$\pi/4$	0.0225	0.0225	0.133	0.177
4	1.89	0.0019	$\pi/4$	0.0113	0.0118	0.343	0.373
6	1.89	0.0018	$\pi/4$	0.0086	0.0094	0.394	0.473

four activation maps for each replicate ROI based on the four activation statistics, using cluster thresholding on each map. Figures 11(b), (c), (e), (f) show the true and false activation rates over the replicated ROIs for the parameter set corresponding to the smoothed data with FWHM = 4 voxels. (See Figure S-18 for similar plots for the other parameter sets.) The false positive rates in Figure 11(e, f) for the independent error-based activation statistics are greater than their counterparts which have autocorrelated errors. As in Section 5.3, autocorrelation in the simulated time series makes the independent-model-based activation statistics fail to follow the  $\chi_1^2$  null distribution under  $H_0$ , as illustrated in Figure S-19. In fact, the false positive rates increase with amount of smoothing. However, they are similar

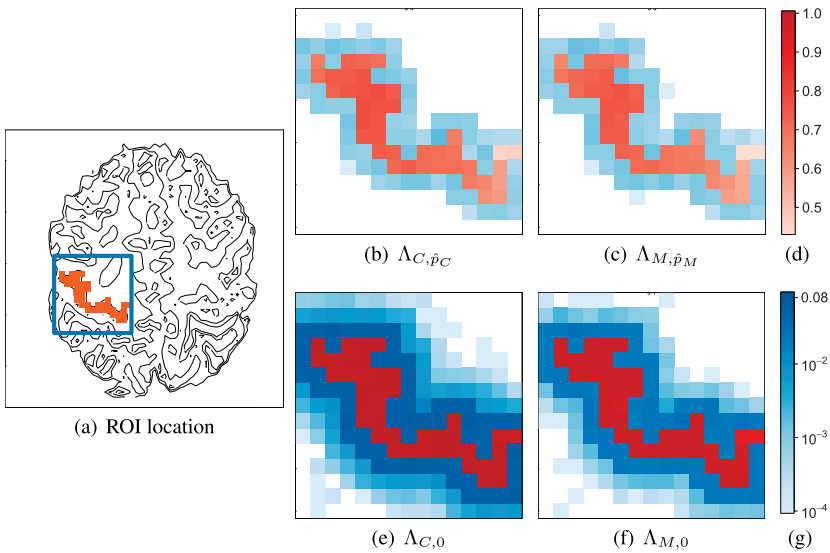


FIG. 11. (a) Location of simulated ROI (blue box) with voxels (dark orange) chosen to be activated according to the anatomical left central sulcus in Figure 3. (b), (c), (e), (f) Images of true activation detection rates (legend in d) and false positive rates (legend in g) for the simulated ROIs for the parameter values corresponding to spatial smoothing with FWHM = 4 voxels.

TABLE 2

*Mean Jaccard similarity coefficients over replications of simulated ROIs for each LRT statistic, measuring the proportion of overlap between the statistic-based activation maps and the truth, when different amounts of spatial smoothing is applied (determined by FWHM, with 0 denoting no spatial smoothing)*

	FWHM = 0	FWHM = 2	FWHM = 4	FWHM = 6
Magnitude-only model	0.844	0.730	0.678	0.717
Complex-valued model	0.843	0.741	0.700	0.744

for all four activation statistics with the unsmoothed data, which can be attributed to its near absence of autocorrelation.

The above findings once again suggest that independence model-based activation statistics are inapplicable with spatial smoothing because of their inflated false positive rates. Therefore, we remove them from consideration when computing Jaccard indices to measure the agreement between the simulation-based activation maps and the known truth. Table 2 gives the mean Jaccard indices over ROI replications; see Figure S-20 for kernel density estimates of the distributions. On the whole, the Jaccard indices are similar under the parameter set representing the unsmoothed data, but with smoothing, the complex-valued model-based indices are greater than their the magnitude-only model-based counterparts, indicating that activation maps obtained via complex-valued  $AR(p)$  modeling have higher agreement with the truly activated voxels. This conforms with our observations for the finger-tapping dataset. (It should be noted that the Jaccard indices in Table 2 are not meant to be compared between parameter sets, but rather only between complex-valued and magnitude-only models within the same parameter set.)

**6. Discussion.** In this paper, we have expanded and further developed complex-valued time series analysis of fMRI data. As explained in the Introduction, such datasets are truly complex-valued when collected, but most analysis methods routinely discard the phase information, utilizing only the magnitude images in the data analysis. In this paper therefore, we have proposed a general model for complex-valued time series, extending the independence assumptions of Rowe and Logan (2004), and have derived a LRTS for detecting activated brain voxels under this model framework. We compared its performance to a statistic similarly derived under a Gaussian-assumed magnitude-only linear model. We applied this methodology to a unilateral finger-tapping dataset of the type used for neurosurgical planning, in which the purpose is clear identification of the left functional central sulcus. Our analysis showed that the complex-valued  $AR(p)$  model-based activation map more clearly and accurately identifies this brain region than the magnitude-only model and the independence-based complex-valued

model in cases where the images were spatially smoothed, as commonly done in practice. For the unsmoothed data, the activation maps performed similarly. To explain this result, we found that temporal and real/imaginary correlation is induced by smoothing and proposed a model to interpret these findings. For further clarification, we performed several simulation studies. In one, we found that the complex-valued model has greater detection power than the magnitude-only model because it estimates the variances and correlation between of the real and imaginary errors, which are not estimable under the magnitude-only model. In another, we found that independence-based activation statistics can have greatly inflated false positive rates when even small amounts of autocorrelation are present in the time series. Further, we found that the complex-valued model is better able to discriminate between activated and nonactivated voxels, as measured by both the AUC and the Jaccard index. We also emphasize that these results occurred at SNR values typical of fMRI data (on the order of 50 for the unsmoothed data and higher for the smoothed data), which represents an improvement over findings in previous complex-valued fMRI studies [Nan and Nowak (1999), Rowe and Logan (2004)] that only showed substantial improvements in complex-valued over magnitude-only model-based activation detection at SNRs below five.

There are many areas that could benefit from further attention. In this paper, we have shown improved activation detection with complex-valued AR modeling of fMRI time series and shown its potential in the context of neurosurgical planning by accurately mapping the (left) functional central sulcus. Note, however, that the objective of neurosurgical planning is preoperative brain mapping that allows for the preservation of cerebral function during brain tumor resection. Such planning also may need to map tactile, visual, language or other functional areas which will involve other experimental paradigms, for which complex-valued fMRI time series analysis of the kind we have recommended in this paper can potentially provide greater accuracy. Further, while we hope that the development of this paper will encourage statisticians to recommend that this data not be routinely discarded but analyzed more thoroughly for more accurate activation results, we note that vast numbers of legacy magnitude-valued fMRI datasets exist (and continue to be collected). Especially at lower SNRs, the observations in these datasets are Rice-distributed, but they have all been analyzed using Gaussian AR models. Therefore, it would be of interest to develop methodology for Rice-distributed time series.

**Acknowledgements.** The authors are very grateful to the Editor, an Associate Editor and three reviewers whose very detailed and insightful comments on earlier versions of this manuscript greatly improved its content and presentation.

#### SUPPLEMENTARY MATERIAL

**Supplement to “Complex-valued time series modeling for improved activation detection in fMRI studies”** (DOI: 10.1214/17-AOAS1117SUPP; .pdf).



Section S-1 displays the real, imaginary, magnitude, and phase components of the data for all slices of the finger-tapping dataset used in this paper. Section S-2 provides additional details and derivations on our methodology. Further details on the analysis of the finger-tapping dataset are provided in Section S-3 while more simulation-based analyses are in Section S-4.

## REFERENCES

- ADRIAN, D. W, MAITRA, R. and ROWE, D. B. (2017). Supplement to “complex-valued time series modeling for improved activation detection in fMRI studies.” DOI:10.1214/17-AOAS1117SUPP.
- BANDETTINI, P. A., PETRIDOU, N. and BODURKA, J. (2005). Direct detection of neuronal activity with MRI: Fantasy, possibility, or reality? *Appl. Magn. Reson.* **29** 65–88.
- BANDETTINI, P. A., JESMANOWICZ, A., WONG, E. C. and HYDE, J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med.* **30** 161–173.
- BELLIVEAU, J. W., KENNEDY, D. N., MCKINSTRY, R. C., BUCHBINDER, B. R., WEISSKOFF, R. M., COHEN, M. S., VEVEA, J. M., BRADY, T. J. and ROSEN, B. R. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* **254** 716–719.
- BROWN, KINCAID B. M, T. R. and UGURBIL, K. (1982). NMR chemical shift imaging in three dimensions. *Proc. Natl. Acad. Sci. USA* **79** 3523–3526.
- BRUCE, I. P., KARAMAN, M. M. and ROWE, D. B. (2011). A statistical examination of SENSE image reconstruction via an isomorphism representation. *Magn. Reson. Imag.* **29** 1267–1287.
- BULLMORE, E., BRAMMER, M., WILLIAMS, S. C. R., RABE-HESKETH, S., JANOT, N., DAVID, A., MELLERS, J., HOWARD, R. and SHAM, P. (1996). Statistical methods of estimation and inference for function MR image analysis. *Magn. Reson. Med.* **35** 261–277.
- COX, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res. Internat. J.* **29** 162–173.
- COX, R. W. (2012). AFNI: What a long strange trip it has been. *NeuroImage* **62** 743–747.
- COX, R. W. and HYDE, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR Biomed.* **10** 171–178.
- DICE, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology* **26** 297–302.
- FENG, Z., CAPRIHAN, A., BLAGOEV, K. B. and CALHOUN, V. D. (2009). Biophysical modeling of phase changes in BOLD fMRI. *NeuroImage* **47** 540–548.
- FISHER, N. I. and LEE, A. J. (1992). Regression models for an angular response. *Biometrics* **48** 665–677. MR1187598
- FORMAN, S. D., COHEN, J. D., FITZGERALD, M., EDDY, W. F., MINTUN, M. A. and NOLL, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magn. Reson. Med.* **33** 636–647.
- FRISTON, K. J., JEZZARD, P. and TURNER, R. (1994). Analysis of functional MRI time-series. *Hum. Brain Mapp.* **1** 153–171.
- FRISTON, K. J., FRITH, C. D., LIDDLE, P. F., DOLAN, R. J., LAMMERTSMA, A. A. and FRACKOWIAK, R. S. J. (1990). The relationship between global and local changes in PET scans. *J. Cereb. Blood Flow Metab.* **10** 458–466.
- FRISTON, K. J., HOLMES, A. P., WORSLEY, K. J., POLINE, J.-B., FRITH, C. D. and FRACKOWIAK, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2** 189–210.
- FRISTON, K. J., JOSEPHS, O., ZARAHN, E., HOLMES, A. P., ROUQUETTE, S. and POLINE, J.-B. (2000). To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. *NeuroImage* **12** 196–208.

- GENOVESE, C. R., LAZAR, N. A. and NICHOLS, T. E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* **15** 870–878.
- GLISSON, T. H. (2011). *Introduction to Circuit Analysis and Design*. Springer, The Netherlands.
- GLOVER, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* **9** 416–429.
- GUDBJARTSSON, H. and PATZ, S. (1995). The Rician distribution of noisy data. *Magn. Reson. Med.* **34** 910–914.
- HAHN, A. D., NENCKA, A. S. and ROWE, D. B. (2009). Improving robustness and reliability of phase-sensitive fMRI analysis using temporal off-resonance alignment of single-echo timeseries (TOAST). *NeuroImage* **44** 742–752.
- HAHN, A. D., NENCKA, A. S. and ROWE, D. B. (2012). Enhancing the utility of complex-valued functional magnetic resonance imaging detection of neurobiological processes through postacquisition estimation and correction of dynamic B(0) errors and motion. *Hum. Brain Mapp.* **33** 288–306.
- HAHN, A. D. and ROWE, D. B. (2012). Physiologic noise regression, motion regression, and TOAST dynamic field correction in complex-valued fMRI time series. *NeuroImage* **59** 2231–2240.
- HARNSBERGER, H. R., OSBORN, A. G., ROSS, J. S., MOORE, K. R., SALZMAN, K. L., CARASCO, C. R., HALMITON, B. E., DAVIDSON, H. C. and WIGGINS, R. H. (2007). *Diagnostic and Surgical Imaging Anatomy: Brain, Head and Neck, Spine*, 3rd ed. Amirsys, Salt Lake City, Utah.
- HOOGENRAD, F. G., REICHENBACH, J. R., HAACKE, E. M., LAI, S., KUPPUSAMY, K. and SPRENGER, M. (1998). In vivo measurement of changes in venous blood-oxygenation with high resolution functional MRI at 95 tesla by measuring changes in susceptibility and velocity. *Magn. Res. Med.* **39** 97–107.
- JACCARD, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull. Soc. Vaudoise Sci. Nat.* **37** 547–579.
- JAIN, A. K. (1989). *Fundamentals of Digital Image Processing*. Prentice Hall, New York.
- JESMANOWICZ, A., NENCKA, A. and HYDE, J. S. (2014). Direct radiofrequency phase control in MRI by digital waveform playback at the larmor frequency. *Magn. Reson. Imag.* **71** 846–852.
- JESMANOWICZ, A., WONG, E. C. and HYDE, J. S. (1993). Phase correction for EPI using internal reference lines. In *Proceedings from the International Society of Magnetic Resonance in Medicine* **12** 1239.
- JEZZARD, P. and CLARE, S. (2001). Principles of nuclear magnetic resonance and MRI. In *Functional MRI: An Introduction to Methods* (P. Jezzard, P. M. Matthews and S. M. Smith, eds.) **3** 67–92. Oxford Univ. Press, New York.
- KARAMAN, M., BRUCE, I. P. and ROWE, D. B. (2015). Incorporating relaxivities to more accurately reconstruct MR images. *Magn. Reson. Imag.* **33** 374–384.
- KARAMAN, M., NENCKA, A. S., BRUCE, I. P. and ROWE, D. B. (2014). Quantification of the statistical effect of spatiotemporal processing of nontask fMRI data. *Brain Connectivity* **4** 649–661.
- KRUGGEL, F. and VON CRAMON, D. Y. (1999). Modeling the hemodynamic response in single-trial functional MRI experiments. *Magn. Res. Med.* **42** 787–797.
- KUMAR, A., WELTI, D. and ERNST, R. R. (1975). NMR Fourier zeugmatography. *J. Magn. Res.* **18** 69–83.
- KWONG, K. K., BELLIVEAU, J. W., CHESLER, D. A., GOLDBERG, I. E., WEISSKOFF, R. M., PONCELET, B. P., KENNEDY, D. N., HOPPEL, B. E., COHEN, M. S., TURNER, R., CHENG, H.-M., BRADY, T. J. and ROSEN, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. In *Proceedings from the National Academy of Sciences, USA* **89** 5675–5679.

- LAI, S. and GLOVER, G. (1997). Detection of BOLD fMRI signals using complex data. In *Proceedings from the International Society of Magnetic Resonance in Medicine* **5** 1671.
- LAZAR, N. A. (2008). *The Statistical Analysis of Functional MRI Data*. Springer, Berlin.
- LEE, C. C., JACK, C. R. and RIEDERER, S. J. (1998). Mapping of the central sulcus with functional MR: Active versus passive activation tasks. *Amer. J. Neurorad.* **19** 847–852.
- LEE, C. C., WARD, H. A., SHARBROUGH, F. W., MEYER, F. B., MARSH, W. R., RAFFEL, C., SO, E. L., CASCINO, G. D., SHIN, C., XU, Y., RIEDERER, S. J. and JACK, C. R. (1999). Assessment of functional MR imaging in neurosurgical planning. *Amer. J. Neurorad.* **20** 1511–1519.
- LEE, J., SHAHRAM, M., SCHWARTZMAN, A. and PAULY, J. M. (2007). Complex data analysis in high-resolution SSFP fMRI. *Magn. Reson. Med.* **57** 905–917.
- LJUNGGREN, S. (1983). A simple graphical representation of Fourier-based imaging methods. *J. Magn. Res.* **54** 338–343.
- LOGAN, B. R., GELIAZKOVA, M. P. and ROWE, D. B. (2008). An evaluation of spatial thresholding techniques in fMRI analysis. *Hum. Brain Mapp.* **29** 1379–1389.
- LOGAN, B. R. and ROWE, D. B. (2004). An evaluation of thresholding techniques in fMRI analysis. *NeuroImage* **22** 95–108.
- MAITRA, R. (2010). A re-defined and generalized percent-overlap-of-activation measure for studies of fMRI reproducibility and its use in identifying outlier activation maps. *NeuroImage* **50** 124–135.
- MARCHINI, J. L. and RIPLEY, B. D. (2000). A new statistical approach to detecting significant activation in functional MRI. *NeuroImage* **12** 366–380.
- MENON, R. S. (2002). Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI. *Magn. Res. Med.* **47** 1–9.
- MILLER, J. W. (1995). Exact maximum likelihood estimation in autoregressive processes. *J. Time Series Anal.* **16** 607–615.
- NAN, F. Y. and NOWAK, R. D. (1999). Generalized likelihood ratio detection for fMRI using complex data. *IEEE Trans. Med. Imag.* **18** 320–329.
- NENCKA, A. S., HAHN, A. D. and ROWE, D. B. (2008). The use of three navigator echoes in Cartesian EPI reconstruction reduces Nyquist ghosting. In *Proceedings from the International Society of Magnetic Resonance in Medicine* **16** 3032.
- NENCKA, A. S., HAHN, A. D. and ROWE, D. B. (2009). A mathematical model for understanding the STatistical effects of k-space (AMMUST-k) preprocessing on observed voxel measurements in fcMRI and fMRI. *J. Neurosci. Met.* **181** 268–282. DOI:10.1016/j.jneumeth.2009.05.007.
- OGAWA, S., LEE, T. M., NAYAK, A. S. and GLYNN, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Reson. Med.* **14** 68–78.
- PETRIDOU, N., SCHAFFER, A., GOWLAND, P. and BOWTELL, R. (2009). Phase vs. magnitude information in functional magnetic resonance imaging time series: Toward understanding the noise. *Magn. Reson. Imag.* **27** 1046–1057.
- POURAHMADI, M. (2001). *Foundations of Time Series Analysis and Prediction Theory*. Wiley, New York. MR1849562
- PURDON, P. L. and WEISSKOFF, R. M. (1998). Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. *Hum. Brain Mapp.* **6** 239–249.
- ROSEN, B. R. and SAVOY, R. L. (2012). fMRI at 20: Has it changed the world? *NeuroImage* **62** 1316–1324.
- ROWE, D. B. (2005). Modeling both the magnitude and phase of complex-valued fMRI data. *NeuroImage* **25** 1310–1324.
- ROWE, D. B. (2016). Image Reconstruction in Functional MRI. In *Handbook of Neuroimaging Data Analysis* 205–232 Chapman & Hall/CRC, London.

- ROWE, D. B. and LOGAN, B. R. (2004). A complex way to compute fMRI activation. *NeuroImage* **23** 1078–1092.
- ROWE, D. B., MELLER, C. P. and HOFFMAN, R. G. (2007). Characterizing phase-only fMRI data with an angular regression model. *J. Neurosci. Met.* **161** 331–341.
- RUMEAU, C., TZOURIO, N., MURAYAMA, N., PERETTI-VITON, P., LEVRIER, O., JOLIOT, M., MAZOYER, B. and SALAMON, G. (1994). Location of hand function in the sensorimotor cortex: MR and functional correlation. *Amer. J. Neurorad.* **15** 567–572.
- SMITH, S. M. (2001). Preparing fMRI data for statistical analysis. In *Functional MRI: An Introduction to Methods* Chapter 12. Oxford Univ. Press, Oxford.
- SØRENSEN, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab* **5** 1–34.
- TWEIG, D. B. (1983). The  $k$ -trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. *Med. Phys.* **10** 610–21.
- WANG, T. and WEI, T. (1994). Statistical analysis of MR imaging and its applications in image modeling. In *Proceedings of the IEEE International Conference on Image Processing and Neural Networks* **1** 866–870.
- WOO, C.-W., KRISHNAN, A. and WAGER, T. D. (2014). Cluster-extent based thresholding in fMRI analyses. *NeuroImage* **91** 412–419.
- WORSLEY, K. J., MARRETT, S., NEELIN, P., VANDAL, A. C., FRISTON, K. J. and EVANS, A. C. (1996). A unified statistical approach for determining significant voxels in images of cerebral activation. *Hum. Brain Mapp.* **4** 58–73.
- ZARAHN, E., AGUIRRE, G. K. and D’ESPOSITO, M. (1997). Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage* **5** 179–197.
- ZHAO, F., JIN, T., WANG, P., HU, X. and KIM, S.-G. (2007). Sources of phase changes in BOLD and CBV-weighted fMRI. *Magn. Reson. Med.* **57** 520–527.

D. W. ADRIAN  
DEPARTMENT OF STATISTICS  
GRAND VALLEY STATE UNIVERSITY  
ALLENDALE, MI 49401-9403  
USA  
E-MAIL: adriand1@gvsu.edu

R. MAITRA  
DEPARTMENT OF STATISTICS  
IOWA STATE UNIVERSITY  
AMES, IOWA 50011-1090  
USA  
E-MAIL: maitra@iastate.edu

D. B. ROWE  
DEPARTMENT OF MATHEMATICS, STATISTICS  
AND COMPUTER SCIENCE  
MARQUETTE UNIVERSITY  
MILWAUKEE, WI 53233  
USA  
E-MAIL: daniel.rowe@marquette.edu