ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 23 (2004) 1078–1092

A complex way to compute fMRI activation

Daniel B. Rowe[a,*] and Brent R. Logan[b]

[a]Department of Biophysics, Medical College of Wisconsin, Milwaukee, WI 53226, USA
[b]Division of Biostatistics, Medical College of Wisconsin, Milwaukee, WI, USA

Received 23 February 2004; revised 24 May 2004; accepted 5 June 2004
Available online 29 September 2004

In functional magnetic resonance imaging, voxel time courses after Fourier or non-Fourier "image reconstruction" are complex valued as a result of phase imperfections due to magnetic field inhomogeneities. Nearly all fMRI studies derive functional "activation" based on magnitude voxel time courses [Bandettini, P., Jesmanowicz, A., Wong, E., Hyde, J.S., 1993. Processing strategies for time-course data sets in functional MRI of the human brain. Magn. Reson. Med. 30 (2): 161-173 and Cox, R.W., Jesmanowicz, A., Hyde, J.S., 1995. Real-time functional magnetic resonance imaging. Magn. Reson. Med. 33 (2): 230-236]. Here, we propose to directly model the entire complex or bivariate data rather than just the magnitude-only data. A nonlinear multiple regression model is used to model activation of the complex signal, and a likelihood ratio test is derived to determine activation in each voxel. We investigate the performance of the model on a real dataset, then compare the magnitude-only and complex models under varying signal-to-noise ratios in a simulation study with varying activation contrast effects.

Keywords: Functional magnetic resonance imaging; Magnetic field inhomogeneities; Voxel time courses

Introduction

In magnetic resonance imaging, we aim to image the effective density of "spinning" protons in a real valued physical object. The equations of physics work out that the Fourier transform (FT) of the effective proton spin density (PSD) is a spatial frequency spectrum. We will obtain the spatial frequency spectrum and perform an inverse Fourier transform (IFT) to obtain the effective proton spin density. This is done by taking successive measurements in time of a real valued signal, a voltage in a wire. The time axis is transformed to the spatial frequency or *k*-space axis. This physical signal or voltage is real valued, but it is "complex demodulated." In measuring the signal, there can be either one or two analog to digital (A to D) converters. If there is a single A to D converter, successive signal measurements are alternately multiplied by either a cosine or a sine to obtain real (inphase) and imaginary (quadrature) parts. These two measurements are then shifted either half a step forward or backward to temporally align them. If there are two A to D converters, two measurements are then taken at the same time with one multiplied by a cosine and the other by a sine. This discretely measured complex valued signal is the discrete FT of the PSD. A discrete IFT is applied to the discretely measured signal. The original object or PSD is real valued, but due to phase imperfections, a complex image of PSDs is produced (Haacke et al., 1999).

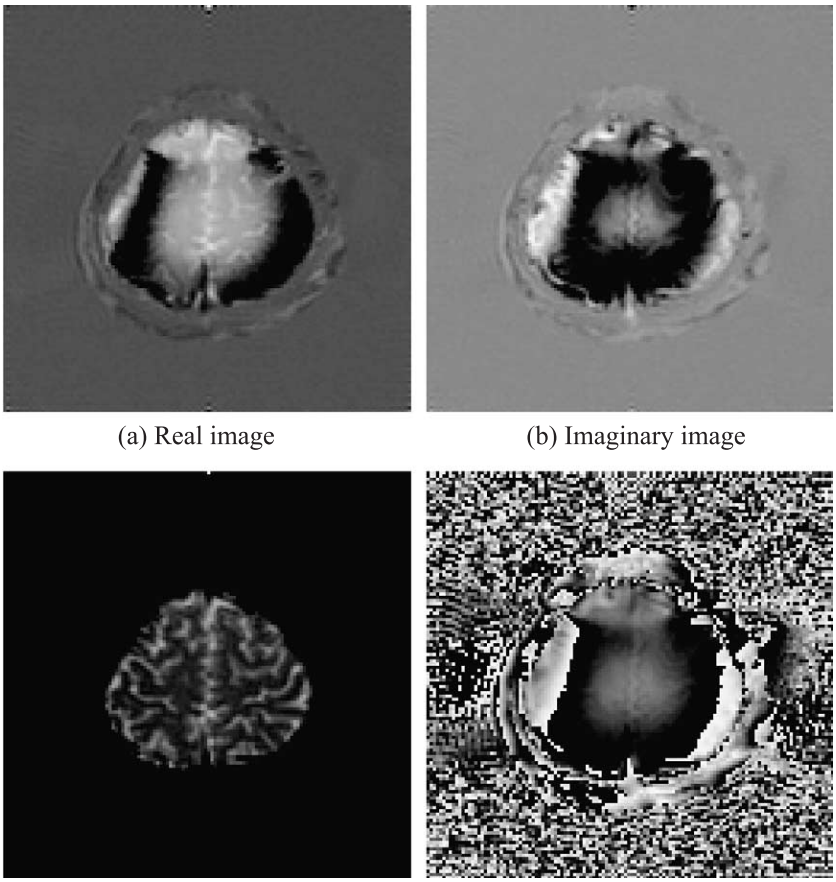
After Fourier or non-Fourier image reconstruction, each voxel contains a time course of real and imaginary components of the measured PSD. Magnitude images are produced by taking the square root of the sum of squares of the real and imaginary parts of the measured PSD in each voxel at each time point. Nearly all fMRI studies obtain a statistical measure of functional activation based on magnitude-only image time courses. When this is done, phase information in the data is discarded. This is illustrated in Fig. 1, where the real, imaginary, magnitude, and phase images are shown at a single point in time, for the example dataset discussed later.

Magnitude models typically assume normally distributed errors; alternatively, one can assume that the original real and imaginary components of the PSD have normally distributed errors. Independent normally distributed errors on the measured complex signal or equivalently complex PSD translate to a Ricean distributed magnitude-only image that is approximately normal for large signal-to-noise ratios (SNR).

When computing magnitude-only image time courses and activations, the signal-to-noise ratio (SNR) may not be large enough for this approximate normality to hold. This is increasingly true with higher voxel resolutions and in voxels with a large degree of signal dropout. In addition, phase information or half of the numbers are discarded. A more accurate model should properly model the noise and use all the information contained in the real and imaginary components of the data. However, SNR exhibits an approximate linear increase and contrast-to-noise ratio (CNR) an approximate super linear increase with main magnetic field strength, but these gains may be mediated by such things as increased physiologic noise,

* Corresponding author. Department of Biophysics, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226. Fax: +1 414 456 6512.
E-mail address: dbrowe@mcw.edu (D.B. Rowe).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.06.042



(a) Real image (b) Imaginary image

(c) Magnitude image (d) Phase image

Fig. 1. Real, imaginary, magnitude, and phase images at a fixed point in time.

more complicated signal features in the proximity of strong susceptibility gradients, and changes in intrinsic relaxation times (Kruger et al., 2001).

Previous analyses of complex fMRI data have been proposed, including nonmodel-based exploratory Independent Component Analysis (ICA) (Calhoun et al., 2002) as well as directly modeling the complex activation data (Lai and Glover, 1997; Scharf and Friedlander, 1994; Nan and Nowak, 1999). Previous simple linear regression models by Scharf and Friedlander (1994) and Lai and Glover (1997) did not accurately model the phase, while we correctly account for it through a nonlinear multiple regression model. A subsequent model by Nan and Nowak (1999) correctly assumed that the phase imperfections for the baseline and signal were the same but were limited to a single baseline and signal model because of their model parameterization. In addition, their model did not directly estimate the regression coefficients or phase angle. We reparameterize and extend the model proposed by Nan and Nowak (1999) to a multiparameter baseline and signal model. Additionally, we formulate the hypothesis test in terms of contrasts, which allows for more elaborate hypothesis testing such as deconvolution and comparisons between multiple task conditions. Finally, our parameterization allows us to estimate the phase angle directly instead of the sine and cosine of the phase angle. We compare the results of the proposed model to a strict magnitude-only model in terms of thresholded activation maps on a real dataset. Finally, simulations are performed comparing our model to a magnitude-only model for various signal-to-noise ratios and task-related contrast effects.

Model

In MRI/fMRI, we aim to image a real valued physical object $\rho(x,y)$ and obtain a measured object $\rho_m(x,y)$ by measuring a 2D complex valued signal $s_m(k_x,k_y)$ at spatial frequencies (k_x,k_y). This signal consists of a true complex valued signal $s(k_x,k_y)$ plus a random complex noise term $\delta(k_x,k_y)$ with real and imaginary components that are assumed to be independent and identically normally distributed. Even if there were no phase imperfections, it is necessary to observe the imaginary parts of this signal because we phase-encode for proper image formation. After image reconstruction, we obtain a complex valued measured object plus complex valued noise.

Neglecting the voxel location and focusing on a particular voxel, the complex valued image measured over time in a given voxel is

$$\rho_{mt} = [\rho_{Rt} + \eta_{Rt}] + i[\rho_{It} + \eta_{It}]$$

where $(\eta_{Rt},\eta_{It})' \sim \mathcal{N}(0,\Sigma)$ and $\Sigma = \sigma^2 I_2$. The distributional specification is on the real and imaginary parts of the image and not on the magnitude.

A nonlinear multiple regression model is introduced individually for each voxel that includes a phase imperfection θ in which, at time t, the measured effective proton spin density is given by

$$\rho_{mt} = [x_t'\beta\cos\theta + \eta_{Rt}] + i[x_t'\beta\sin\theta + \eta_{It}] \tag{2.1}$$

where $\rho_t = x_t'\beta = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_q x_{qt}$. The phase imperfection in Eq. (2.1) is a fixed and unknown quantity, which may be estimated voxel by voxel. Just as in Nan and Nowak (1999), we have also found this phase specification to be reasonable.

In fMRI, we take repeated measurements over time while a subject is performing a task. In each voxel, we compute a measure of association between the observed time course and a preassigned reference function that characterizes the experimental paradigm.

Magnitude activation

The typical method to compute activations (Bandettini et al., 1993; Cox et al., 1995) is to use only the nonunique magnitude $|\rho_{mt}|$ which is denoted by y_t and written as

$$y_t = \left[(x_t'\beta\cos\theta + \eta_{Rt})^2 + (x_t'\beta\sin\theta + \eta_{It})^2\right]^{\frac{1}{2}}. \tag{2.2}$$

The magnitude-only model in Eq. (2.2) discards any information contained in the phase, given by

$$\phi_t = \tan^{-1}\left[\frac{\rho_{It} + \eta_{It}}{\rho_{Rt} + \eta_{Rt}}\right].$$

The magnitude is not normally distributed but is Ricean-distributed. Both the magnitude and the phase are approximately normal for large SNRs (Gudbjartsson and Patz, 1995; Rice, 1944) as outlined in Appendix A. The special case of the Ricean distribution where there is no signal is known as the Rayleigh distribution. It is known (Haacke et al., 1999) that a histogram of noise outside the brain without any signal is Rayleigh-distributed.

The Ricean distribution is approximately normal for large signal-to-noise ratios (small relative error variance). This can be shown by completing the square in Eq. (2.2) and proceeding as follows

$$\begin{aligned} y_t &= \left\{[x_t'\beta]^2 + \left[\eta_{Rt}^2 + \eta_{It}^2\right] + 2[x_t'\beta]\ [\eta_{Rt}\cos\theta + \eta_{It}\sin\theta]\right\}^{\frac{1}{2}} \\ &= [x_t'\beta]\left\{1 + \frac{2[\eta_{Rt}\cos\theta + \eta_{It}\sin\theta]}{[x_t'\beta]} + \frac{\left[\eta_{Rt}^2 + \eta_{It}^2\right]}{[x_t'\beta]^2}\right\}^{\frac{1}{2}} \\ &\approx x_t'\beta + \epsilon_t \end{aligned} \tag{2.3}$$

where $\epsilon_t = \eta_{Rt}\cos\theta + \eta_{It}\sin\theta \sim N(0,\sigma^2)$. Again, the $\cos\theta$ and $\sin\theta$ arise from phase imperfections. If there were no phase imperfections, then $\theta = 0$. In this derivation, the approximation $\sqrt{1+u} \approx 1 + u/2$ was used for $|u| \ll 1$. This model can also be written as

$$\underset{n\times 1}{y} = \underset{n\times(q+1)}{X}\ \underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\epsilon} \tag{2.4}$$

where $\epsilon \sim \mathcal{N}(0,\sigma^2\Phi)$ and Φ is the temporal correlation matrix, often taken to be $\Phi = I_n$ after suitable preprocessing of the data.

The unconstrained maximum likelihood estimates of the vector of regression coefficients $\hat{\beta}$ and the error variance $\hat{\sigma}^2$ are given by

$$\begin{aligned} \hat{\beta} &= (X'X)^{-1}X'y, \\ \hat{\sigma}^2 &= \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right)/n. \end{aligned} \tag{2.5}$$

To construct a generalized likelihood ratio test of the hypothesis H_0:$C\beta = 0$ vs. H_1:$C\beta \neq 0$, we maximize the likelihood under the constrained null hypothesis. This leads to constrained MLEs

$$\begin{aligned} \tilde{\beta} &= \Psi\hat{\beta}, \\ \tilde{\sigma}^2 &= \left(y - X\tilde{\beta}\right)'\left(y - X\tilde{\beta}\right)/n, \end{aligned} \tag{2.6}$$

where

$$\Psi = I_{q+1} - (X'X)^{-1}C'\left[C(X'X)^{-1}C'\right]^{-1}C. \tag{2.7}$$

Then the likelihood ratio statistic magnitude model is given by

$$-2\log\ \lambda_M = n\ \log\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right). \tag{2.8}$$

This has an asymptotic χ_r^2 distribution, where r is the fill row rank of C, and is asymptotically equivalent to the usual t or F tests associated with statistical parametric maps. For example, with a model with β_0 representing an intercept, β_1 representing a linear drift over time, and β_2 representing a contrast effect of a stimulus. Then to test whether the coefficient for the reference function or stimulus is 0, set $C = (0,0,1)$, so that the hypothesis is H_0:$\beta_2 = 0$. The LR test has an asymptotic χ_1^2 distribution and is asymptotically equivalent to the usual t tests for activation given by

$$t_2 = \frac{\hat{\beta}_2}{SE\left(\hat{\beta}_2\right)}.$$

We use the χ^2 representation for ease of comparability with the complex activation model. However, note that the χ^2 distribution is valid only asymptotically and as a result may be inaccurate in the extreme tails such as might be needed for a Bonferroni adjustment. Alternatively, permutation resampling techniques may be used, which the authors found to give similar results in the example used later.

Complex activation

Alternatively, we can represent the observed data at time point t as a 2×1 vector instead of as a complex number

$$\begin{pmatrix} y_{Rt} \\ y_{It} \end{pmatrix} = \begin{pmatrix} x_t'\beta\cos\theta \\ x_t'\beta\sin\theta \end{pmatrix} + \begin{pmatrix} \eta_{Rt} \\ \eta_{It} \end{pmatrix}, \quad t = 1,\ldots,n.$$

This model can also be written as

$$\underset{2n\times 1}{y} = \underset{2n\times 2(q+1)}{\begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}}\ \underset{2(q+1)\times 1}{\begin{pmatrix} \beta\cos\theta \\ \beta\sin\theta \end{pmatrix}} + \underset{2n\times 1}{\eta} \tag{2.9}$$

where it is specified that the observed vector of data $y = (y_R',y_I')'$ is the vector of observed real values stacked on the vector of observed complex values and the vector of errors $\eta = (\eta_R',\eta_I')' \sim \mathcal{N}(0,\Sigma \otimes \Phi)$ is similarly defined. Here, we assume that $\Sigma = \sigma^2 I_2$ and $\Phi = I_n$.

Due to the multiparameter baseline and signal model in Eq. (2.9), this is a generalization of the simple linear regression model by Nan and Nowak (1999) where there is only a mean and signal reference function. Previous simple linear regression models by Scharf and Friedlander (1994) and Lai and Glover (1997) did not accurately model the phase, while we correctly account for it through a nonlinear multiple regression model. Our generalization allows for more elaborate hypothesis testing frameworks, such as deconvolution and comparisons between task conditions.

As with the magnitude-only model, we can obtain unrestricted maximum likelihood estimates of the parameters as derived in Appendix B to be

$$\hat{\theta} = \frac{1}{2}\tan^{-1}\left[\frac{2\hat{\beta}_R'(X'X)\hat{\beta}_I}{\hat{\beta}_R'(X'X)\hat{\beta}_R - \hat{\beta}_I'(X'X)\hat{\beta}_I}\right]$$

$$\hat{\beta} = \hat{\beta}_R\cos\hat{\theta} + \hat{\beta}_I\sin\hat{\theta},$$

$$\hat{\sigma}^2 = \frac{1}{2n}\left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \hat{\beta}\cos\hat{\theta} \\ \hat{\beta}\sin\hat{\theta} \end{pmatrix}\right]' \times \left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \hat{\beta}\cos\hat{\theta} \\ \hat{\beta}\sin\hat{\theta} \end{pmatrix}\right], \tag{2.10}$$

where

$$\hat{\beta}_R = (X'X)^{-1}X'y_R,$$

$$\hat{\beta}_I = (X'X)^{-1}X'y_I.$$

Note that the estimate of the regression coefficients is a linear combination or "weighted" average of estimates from the real and imaginary parts. The regression coefficients of this model may also be estimated using principal components. Also, note that, although the ML estimate of σ^2 is biased, the degree of bias is generally small ($E(\hat{\sigma}^2) = (2n - q - 2)/(2n) \times \sigma^2$) because n is large relative to q.

The maximum likelihood estimates under the constrained null hypothesis H_0:$C\beta = 0$ are derived in Appendix B and given by

$$\tilde{\theta} = \frac{1}{2}\tan^{-1}\left[\frac{2\hat{\beta}_R'\boldsymbol{\Psi}(X'X)\hat{\beta}_I}{\hat{\beta}_R'\boldsymbol{\Psi}(X'X)\hat{\beta}_R - \hat{\beta}_I'\boldsymbol{\Psi}(X'X)\hat{\beta}_I}\right]$$

$$\tilde{\beta} = \boldsymbol{\Psi}\left[\hat{\beta}_R\cos\tilde{\theta} + \hat{\beta}_I\sin\tilde{\theta}\right],$$

$$\tilde{\sigma}^2 = \frac{1}{2n}\left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \tilde{\beta}\cos\tilde{\theta} \\ \tilde{\beta}\sin\tilde{\theta} \end{pmatrix}\right]' \times \left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \tilde{\beta}\cos\tilde{\theta} \\ \tilde{\beta}\sin\tilde{\theta} \end{pmatrix}\right], \tag{2.11}$$

where $\boldsymbol{\Psi}$ is as defined in Eq. (2.7) for the magnitude model.

This formulation of the model requires us to correctly deal with the phase angle. An alternative formulation is to let $\alpha_1 = \cos\theta$ and $\alpha_2 = \sin\theta$. Then the model is

$$y = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \alpha_1\beta \\ \alpha_2\beta \end{pmatrix} + \eta, \quad \alpha_1^2 + \alpha_2^2 = 1. \tag{2.12}$$

With the model formulation in Eq. (2.12), we can identify it as a reduced rank regression model (Reinsel and Velu, 1998) with a sum of squares equal to 1 constraint on the α coefficients. In the same way as before, the parameters can be estimated under the unconstrained model as derived in Appendix B to be

$$\hat{\alpha}_1 = \hat{\beta}'(X'X)\hat{\beta}_R \Big/ \left[\left(\hat{\beta}'(X'X)\hat{\beta}_R\right)^2 + \left(\hat{\beta}'(X'X)\hat{\beta}_I\right)^2\right]^{1/2}$$

$$\hat{\alpha}_2 = \hat{\beta}'(X'X)\hat{\beta}_I \Big/ \left[\left(\hat{\beta}'(X'X)\hat{\beta}_R\right)^2 + \left(\hat{\beta}'(X'X)\hat{\beta}_I\right)^2\right]^{1/2}$$

$$\hat{\beta} = \hat{\alpha}_1\hat{\beta}_R + \hat{\alpha}_2\hat{\beta}_I,$$

$$\hat{\sigma}^2 = \frac{1}{2n}\left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \hat{\alpha}_1\hat{\beta} \\ \hat{\alpha}_2\hat{\beta} \end{pmatrix}\right]' \left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \hat{\alpha}_1\hat{\beta} \\ \hat{\alpha}_2\hat{\beta} \end{pmatrix}\right]. \tag{2.13}$$

Again, note that the estimate of the regression coefficients is a linear combination or "weighted" average of estimates from the real and imaginary parts.

Similarly, the maximum likelihood estimates under the constrained null hypothesis H_0:$C\beta = 0$ are derived in Appendix B and given by

$$\tilde{\alpha}_1 = \tilde{\beta}'(X'X)\hat{\beta}_R \Big/ \left[\left(\tilde{\beta}'(X'X)\hat{\beta}_R\right)^2 + \left(\tilde{\beta}'(X'X)\hat{\beta}_I\right)^2\right]^{1/2}$$

$$\tilde{\alpha}_2 = \tilde{\beta}'(X'X)\hat{\beta}_I \Big/ \left[\left(\tilde{\beta}'(X'X)\hat{\beta}_R\right)^2 + \left(\tilde{\beta}'(X'X)\hat{\beta}_I\right)^2\right]^{1/2}$$

$$\tilde{\beta} = \boldsymbol{\Psi}\left(\tilde{\alpha}_1\hat{\beta}_R + \tilde{\alpha}_2\hat{\beta}_I\right),$$

$$\tilde{\sigma} = \frac{1}{2n}\left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \tilde{\alpha}_1\tilde{\beta} \\ \tilde{\alpha}_2\tilde{\beta} \end{pmatrix}\right]' \left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \tilde{\alpha}_1\tilde{\beta} \\ \tilde{\alpha}_2\tilde{\beta} \end{pmatrix}\right] \tag{2.14}$$

In computing maximum likelihood estimates, an iterative maximization known as the Iterative Conditional Modes (ICM) algorithm (Lindley and Smith, 1972; Rowe, 2001, 2003) is used.

Then, for either formulation [Eq. (2.9) or (2.12)], the generalized likelihood ratio statistic for the complex fMRI activation model is

$$-2\log\lambda_C = 2n\log\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right). \tag{2.15}$$

This statistic has an asymptotic χ_r^2 distribution similar to the magnitude-only model statistic in Eq. (2.8) with the same caveats as mentioned previously for the magnitude-only model. Note that, when $r = 1$, one-sided testing can be done using the signed likelihood ratio test (Severini, 2001) given by

$$Z_C = \mathrm{Sign}\left(C\hat{\beta}\right)\sqrt{-2\log\lambda_C},$$

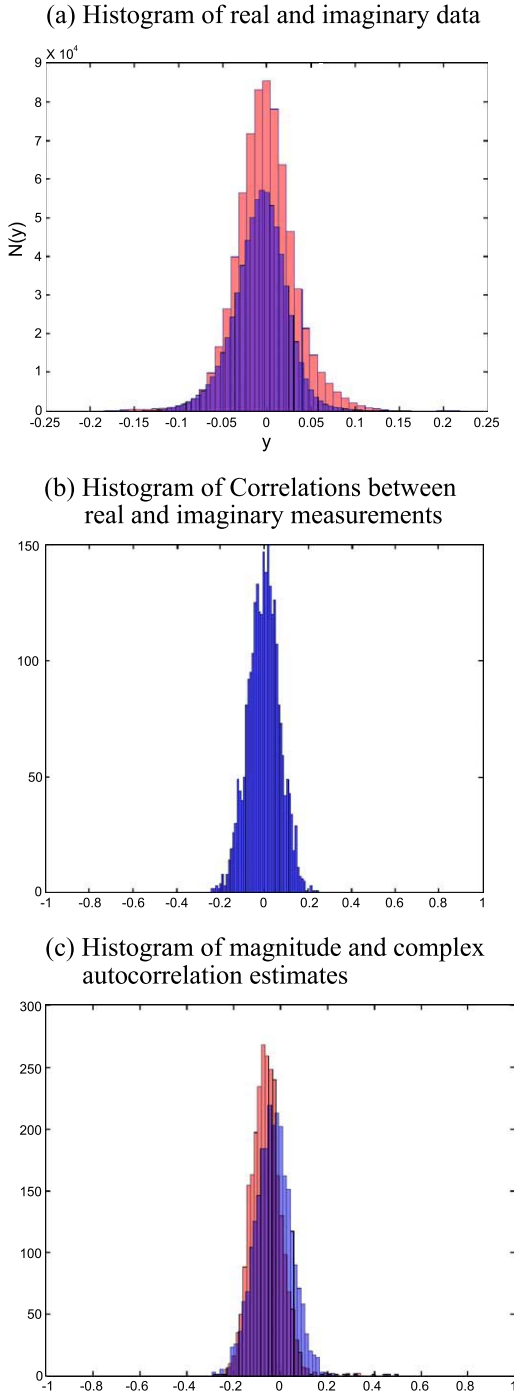


Fig. 2. No signal assessment of complex model.

which has an approximate standard normal distribution under the null hypothesis.

Application to fMRI dataset

A bilateral sequential finger-tapping experiment was performed in a block design with 16-s off followed by eight epochs of 16-s on and 16-s off. Scanning was performed using a 1.5 T GE Signa in which five axial slices of size 96 × 96 were acquired. In image reconstruction, the acquired data was zero filled to 128 × 128. After Fourier reconstruction, each voxel has dimensions in mm of 1.5625 × 1.5625 × 5, with TE = 47 ms. Observations were taken every TR = 1000 ms, so that there are 272 in each voxel. Data from a single axial slice through the motor cortex were selected for analysis. Preprocessing using an ideal 0/1 frequency filter (Gonzales and Woods, 1992; Press et al., 1992) was performed to remove respiration and low frequency physiological noise in addition to the removal of the first three points to omit machine warm-up effects.

First, we checked the validity of the complex model assumptions by examining observations outside the brain where there is no task-related activation. Proper modeling of the noise is essential before modeling the signal. The phase angle was plotted against time to investigate stability of the phase over time; this was relatively constant over time, confirming the observation of Nan and Nowak (1999), and is omitted for brevity. Next histograms of the real (blue) and imaginary (red) components are constructed separately and superimposed on one another in Fig. 2a. These appear to both be approximately normally distributed with similar variances. Fig. 2b contains histograms of the correlations between the real and imaginary components, again computed from the time courses outside the brain. These are distributed closely around 0, indicating that the assumption of independence of the real and imaginary components is reasonable. Finally, an autoregressive order 1 [AR(1)] model was fit to

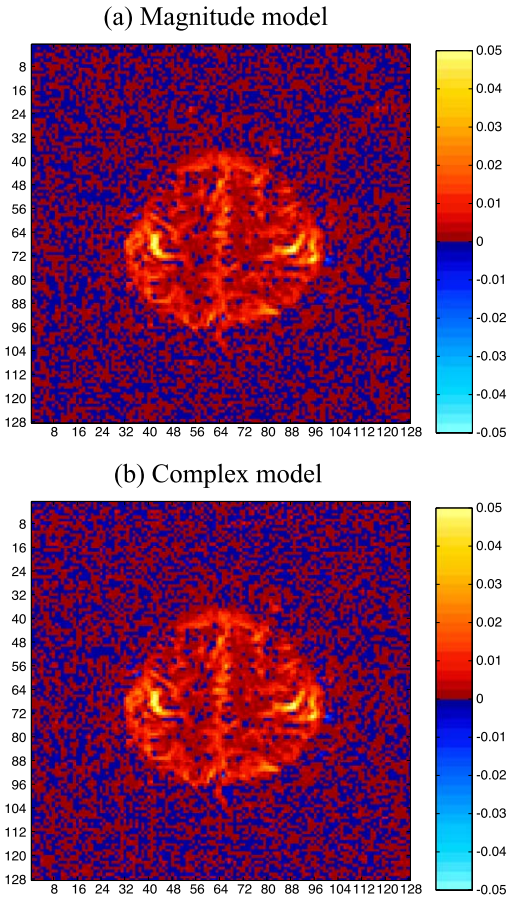


Fig. 3. Estimates of the reference function coefficients, β_2's.

the outside the brain time series, and the autocorrelation parameter ρ for the complex (red) and magnitude (blue) data was estimated from the data and histograms superimposed on one another. These are presented in Fig. 2c where a 5% Bonferroni-adjusted threshold is $\rho_{DW} = 0.26$ for the magnitude-only model and $\rho_{DW} = 0.19$ for the complex model. These thresholds were determined via Monte Carlo simulation by generating datasets of the same length and model. Most of the correlations are between -0.2 and 0.2, indicating little temporal autocorrelation of the data without signal. The autocorrelations may need to be accounted for in practice using the techniques described in Appendix C; however, the procedures appear to be robust to mild departures from the assumptions as seen here.

After verifying the model assumptions, we next model the signal and compare the results of fitting the complex and magnitude-only models. The linear magnitude-only and nonlinear complex multiple regression models were fit to the data with an intercept, a zero mean time trend and a ± 1 square wave reference function. Parameter estimates of the task-related activation β_2 are given in Fig. 3 for (a) the magnitude-only model and (b) the complex model. These coefficient estimates are visually similar between the two models but are numerically different.

As previously noted, the estimated β_2 coefficients for the complex model in Fig. 3b under the alternative hypothesis are a linear combination or "weighted" average between the estimated value from the real and imaginary parts. This "weighting" is displayed in Fig. 4 where the α_1 weights are in Fig. 4a, the estimated coefficient values from the real part β_{R2} are in Fig. 4b, the α_2 weights are in Fig. 4c, and the estimated coefficient values from the real part β_{I2} are in Fig. 4d.

After fitting the model, residuals were used to reevaluate the assumptions for the complex model. Histograms of the real (blue) and imaginary (red) components are constructed from the residual time courses for all voxels separately and superimposed on one another in Fig. 5a. These appear to both be approximately normally distributed with similar variances. Fig. 5b contains histograms of the correlations between the real and imaginary components, computed from the residual time courses for all voxels. These are distributed closely around 0, indicating that the assumption of independence of the real and imaginary components is reasonable. An autoregressive order 1 [AR(1)] model was fit to the residual time courses in every voxel, and the autocorrelation parameter ρ for the complex (red) and magnitude-only (blue) data was estimated from the data and histograms superimposed on one another. These are presented in Fig. 5c where again 5% Bonferroni-adjusted magnitude and complex thresholds can be applied as previously described. The temporal autocorrelation present in the residuals is similar for the magnitude-only and

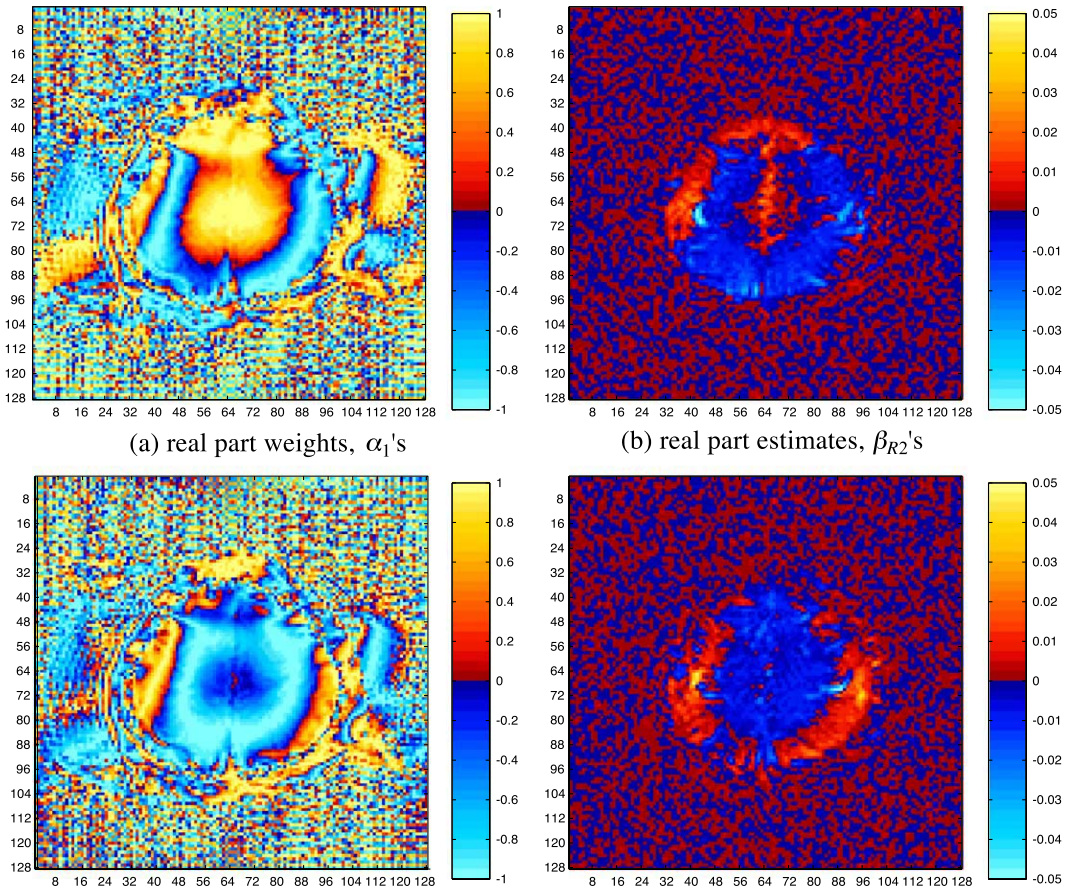


Fig. 4. Complex real and imaginary estimated β'_2s and weights α's.

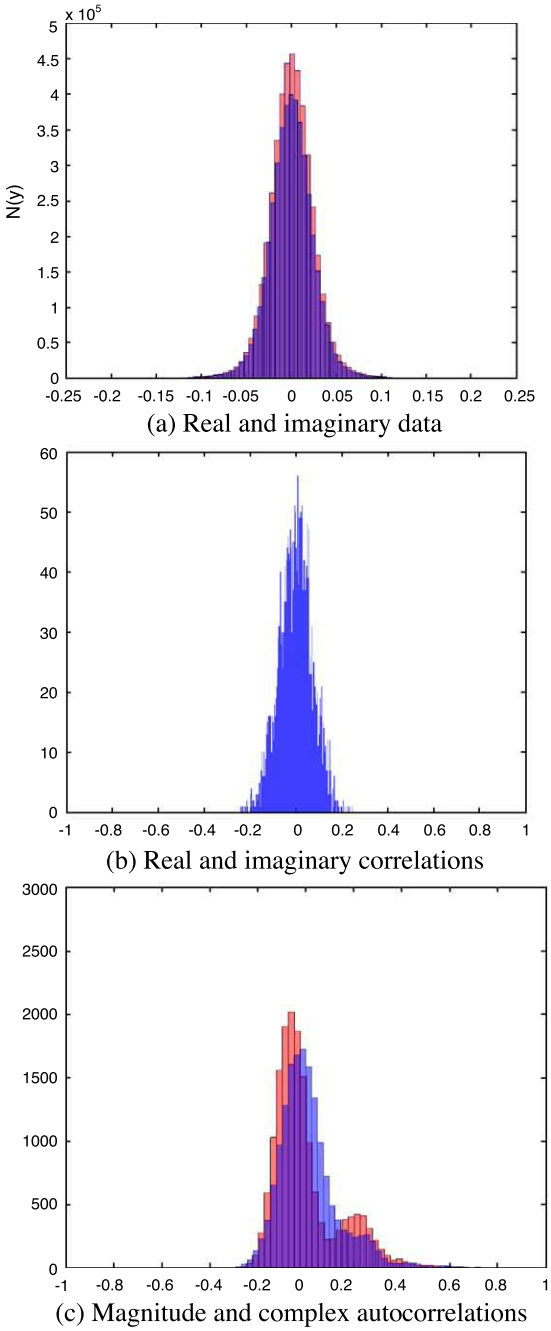


Fig. 5. Residual histogram assessment of complex model.

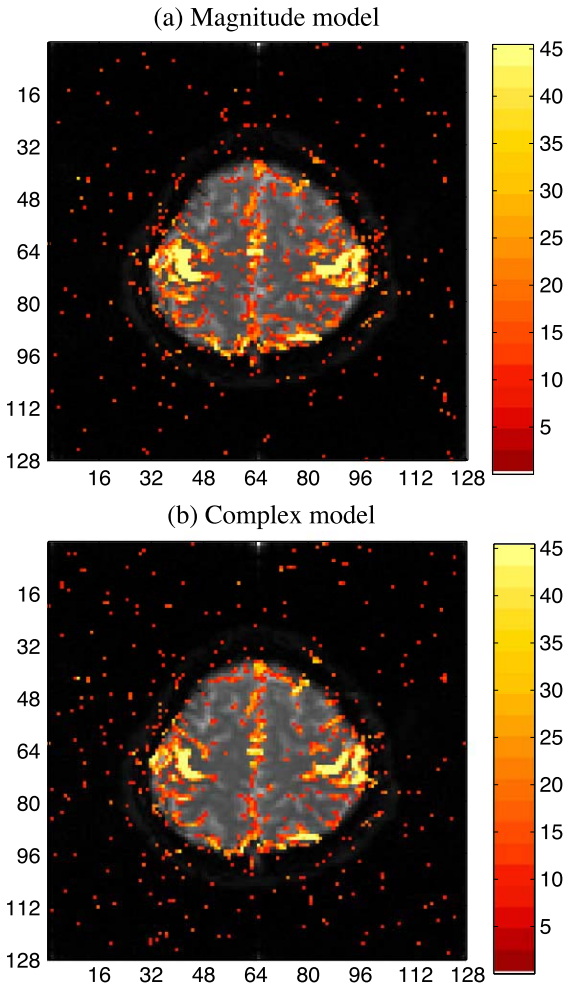


Fig. 6. Activation images using the LR test, thresholded at a 5% false discovery rate.

complex models. Most of the correlations are between −0.20 and 0.20, indicating that most of the voxels have little temporal autocorrelation.

Next, we looked for statistically significant task-related activation using a 5% false discovery rate (FDR) threshold. This was done by applying the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995; Genovese et al., 2002; Logan and Rowe, 2004) to the voxel P values obtained from the χ^2_1 approximation from the likelihood ratio statistic. Images of statistically significant activation are given in Fig. 6 for the magnitude-only and complex models. As previously mentioned, resampling techniques that permuted the complex-valued residuals to determine false discovery rate and Bonferroni-thresholded statistical parametric maps were applied and found to be virtually identical to assuming the approximate χ^2 distribution.

While the activation images are similar, note that the complex model appears to have sharper or more well-defined activation

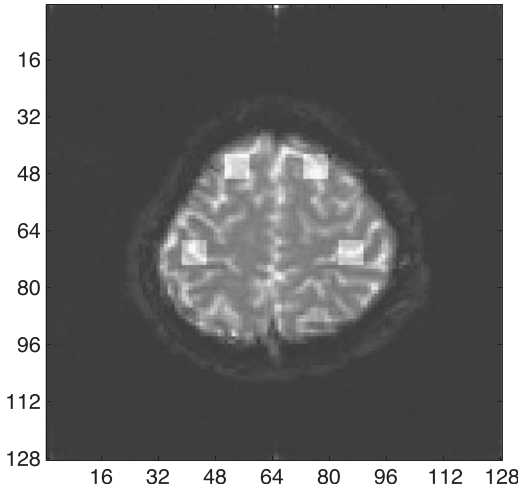


Fig. 7. Anatomical with ROIs.

regions which align better with the gray matter at which the activation is supposed to occur.

fMRI simulation

Data are generated to simulate the same bilateral finger-tapping fMRI block design experiment with $n = 269$ points where the true activation structure is known so that the two activation methods can be evaluated. A 128×128 slice is selected for analysis within which four 7×7 ROIs as lightened in Fig. 7 are designated to have activation.

For this slice, simulated fMRI data are constructed according to a multiple regression model which consists of an intercept, a time trend for all voxels and also a reference function x_{2t} for voxels in each ROI which is related to a block experimental design. This model dictates that for voxel i at time t,

$$y_{it} = [(\beta_0 + \beta_1 t + \beta_{2i} x_{2t})\alpha_{1i} + \eta_{Rit}] + i[(\beta_0 + \beta_1 t + \beta_{2i} x_{2t})\alpha_{2i} + \eta_{Iit}],$$

where η_{Rit}, η_{Iit} are i.i.d. $N(0,\sigma^2)$. In this simulation study, α_1 and α_2 are voxel-dependent and taken from the estimated values for the example dataset shown in Figs. 4a and c, while $\beta_1 = 0.00001$ and $\sigma = 0.04909$ are assumed constant across voxels with values taken from a "highly active" voxel in the activation region of the sample

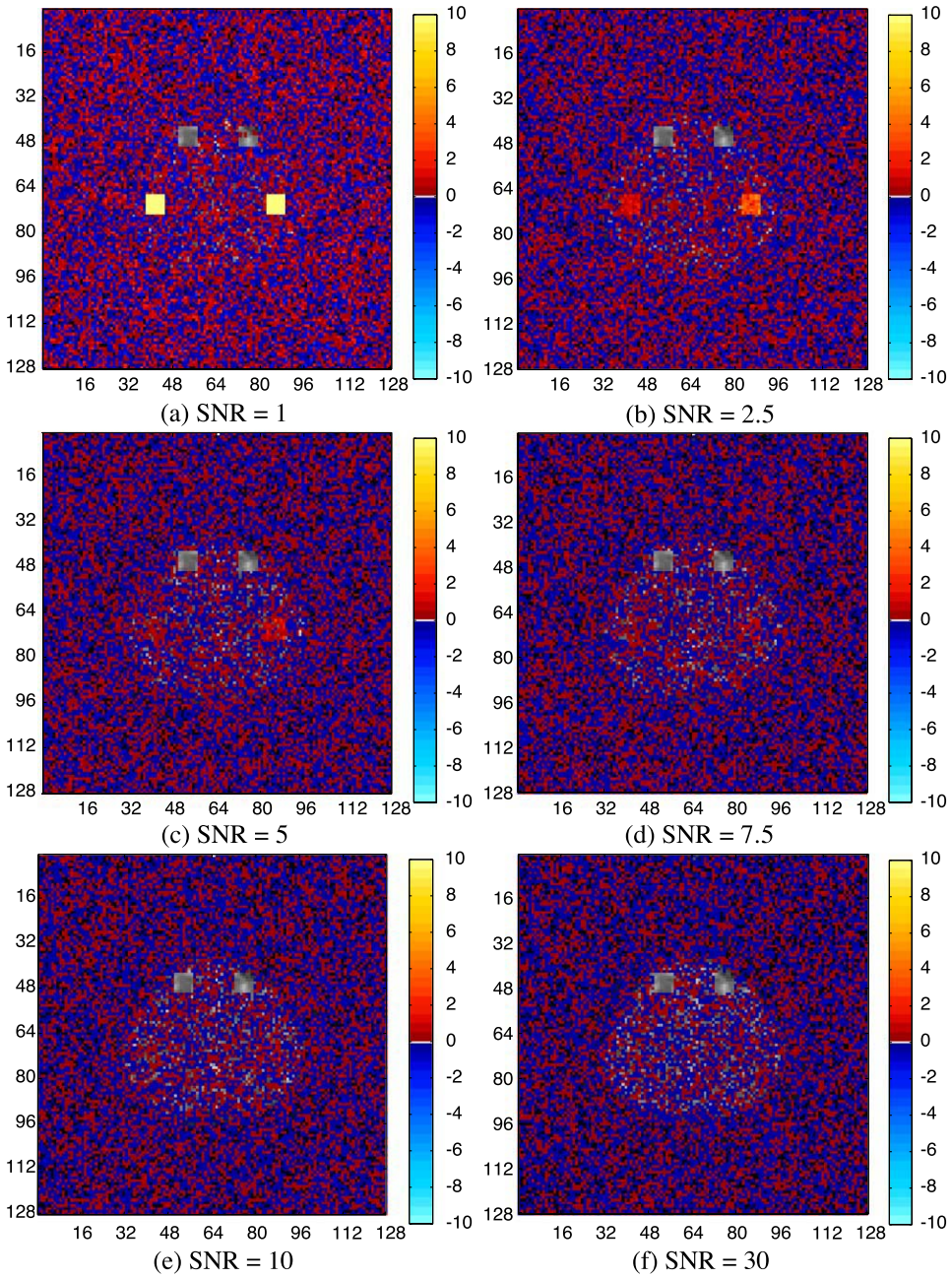


Fig. 8. Differences in power between the models varying SNR, 5% PCE threshold.

dataset. The coefficient for the reference function β_2 is zero outside the ROI. Inside each ROI, β_2 has constant value determined by a contrast-to-noise ratio (CNR = β_2/σ) of 1, 0.5, 0.25, 0.125, going from left to right and top to bottom. To investigate the contrast effect of the signal-to-noise ratio (SNR) typically defined to be the mean divided by the standard deviation of a voxel time course. Note that the magnitude of β_0 observed in the real dataset is generally much larger than β_1 or β_2, indicating that it is the dominant feature in the SNR in addition to being the time course mean. Therefore, since the variance is held fixed, we parameterize the SNR by varying β_0, so that the ratio SNR = β_0/σ takes on values between 1 and 30, where 30 is approximately the value of SNR found in "highly active" voxels for the example dataset, and other values represent decreasing SNR.

In each voxel for a given model and SNR, 1000 simulated images were generated and thresholded using an unadjusted threshold with a 5% type I or per comparison error (PCE) rate, the Benjamini-Hochberg procedure with a 5% false discovery rate (FDR), and the Bonferroni procedure with a 5% familywise error (FWE) rate. For each thresholding method, the power or relative frequency over the 1000 simulated images with which each voxel was detected as active was recorded. Differences in power between the complex and magnitude-only models were calculated for each voxel, mapped to a color scale, and shown in Figs. 8–10 for the three thresholding procedures. Voxels with zero difference in power were assigned the voxel anatomical gray scale.

Note that there are little differences between the complex and magnitude-only models for the CNR = 0.5 to 1 range;

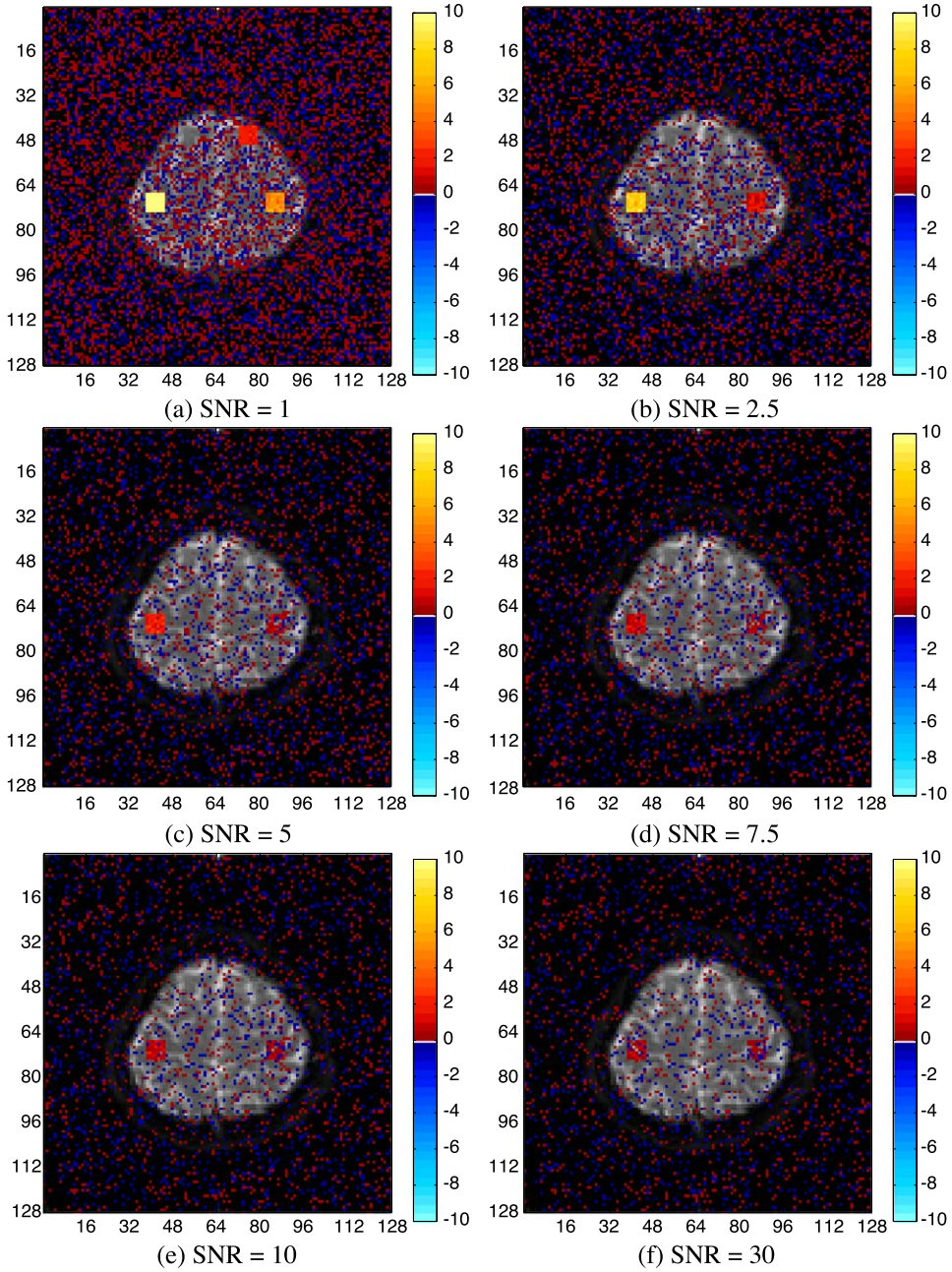


Fig. 9. Differences in power between the models varying SNR, 5% FDR threshold.

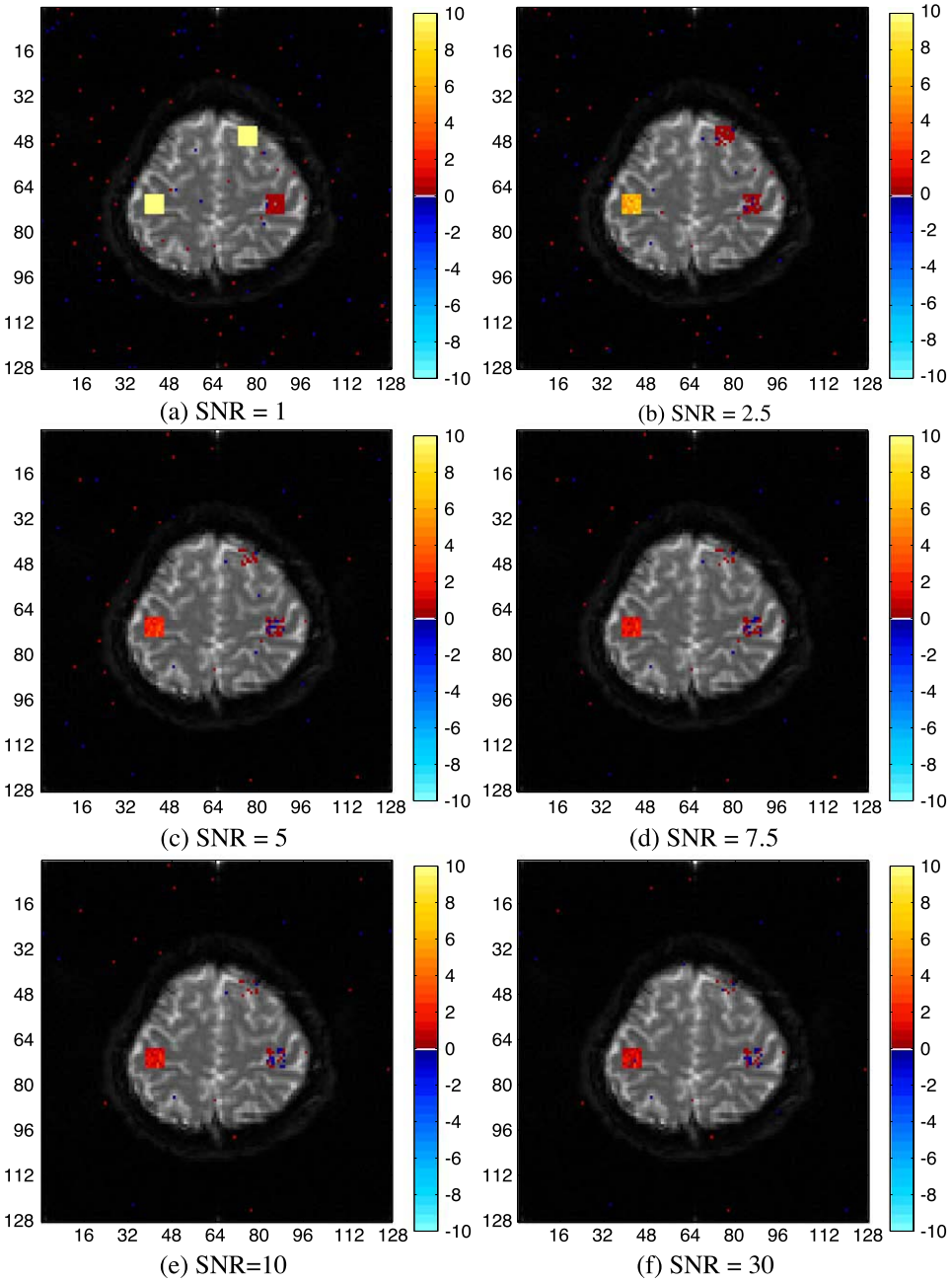
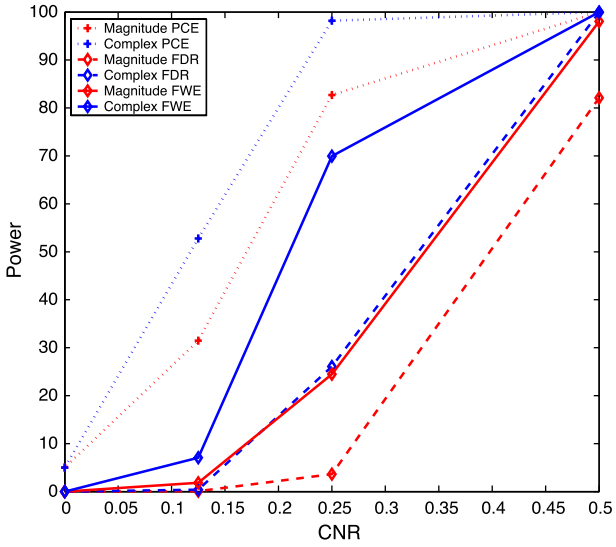


Fig. 10. Differences in power between the models varying SNR, 5% FWE threshold.

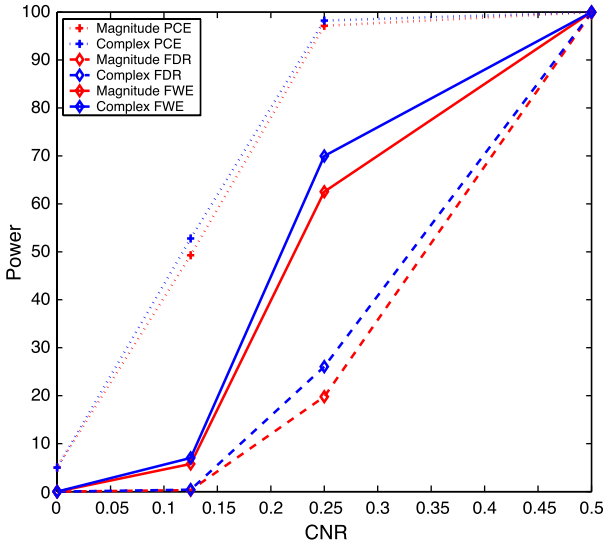
however, this is because the power is approximately 1. For less strong related contrast effects, the differences are sensitive to the SNR, and the complex model is generally useful for low SNR.

To further illustrate the power improvement of the complex model over the magnitude-only model for low SNRs, we plotted the power curves (that average within each ROI) as a function of the contrast-to-noise ratios. This was done for the three thresholding procedures (5% PCE, 5% FDR, 5% FWE) and the complex (blue) and magnitude-only (red) models. These power curves are given in Fig. 11 where for all CNRs, the curves are from top to bottom for the 5% PCE (dotted), 5% FDR (solid), and 5% FWE (dashed) thresholds. These power curves illustrate similar results as before, that the complex model power curve is higher than the magnitude-only model power curve for low SNR, but the lines are quite close for higher SNR.

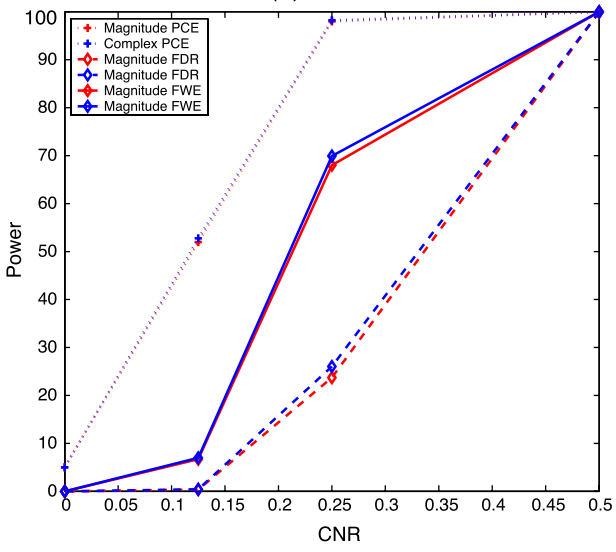
To reiterate the advantage of the complex model over the magnitude-only model for low SNRs, in Fig. 12, we plotted power versus SNR (0.5, 1, 2.5, 5, 7.5, 10) for the four CNRs. The dotted curves represent the 5% PCE threshold, the solid curves the 5% FDR threshold, and the dashed curves the 5% FWE threshold. The curve marker + denotes a CNR of 1, the ○ marker a CNR of 0.5, the * a CNR of 0.25, and the × marker a CNR of 0.125. For example, in Fig. 12c, the solid blue line with an asterisk [-*-] is the complex model FDR power curve for a CNR of 0.25, while the solid red line with an asterisk [-*-] is the corresponding magnitude-only model FDR power curve for a CNR of 0.25. It is evident for a given CNR that the complex model power curve is constant



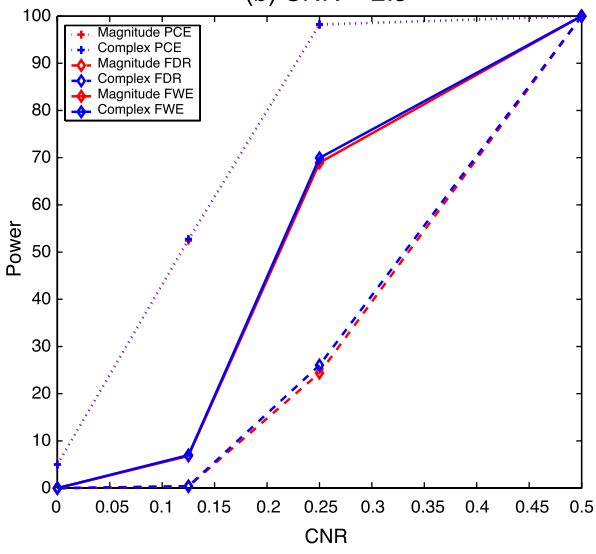
(a) SNR = 1



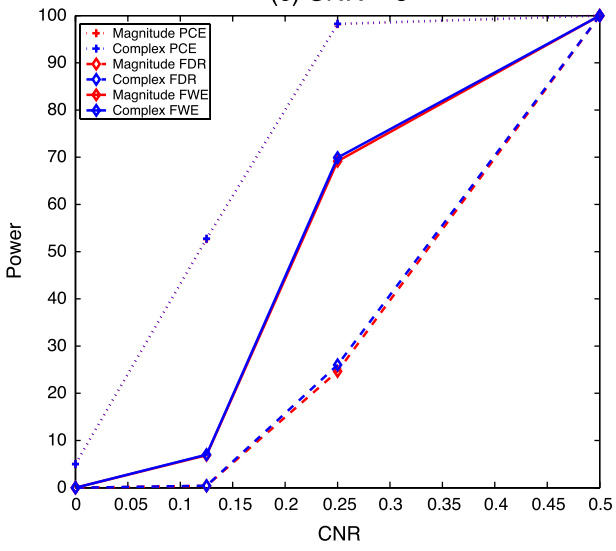
(b) SNR = 2.5



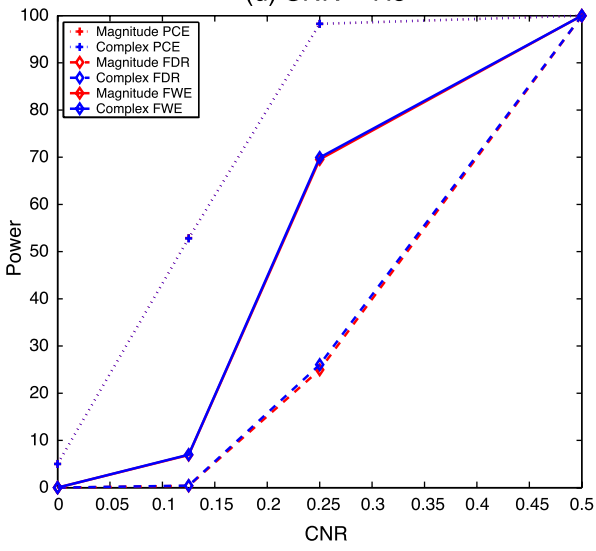
(c) SNR = 5



(d) SNR = 7.5



(e) SNR = 10



(f) SNR = 30

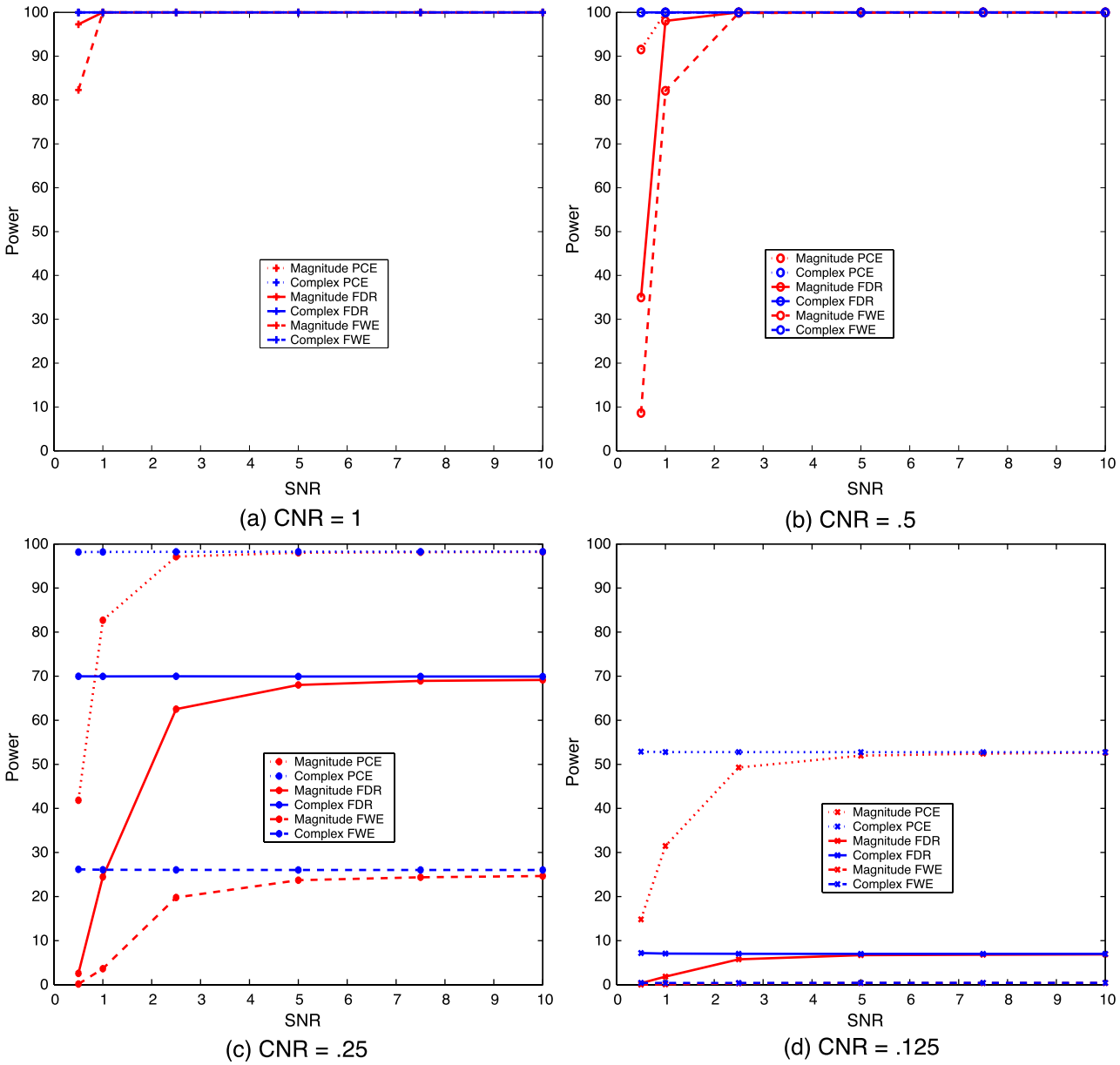


Fig. 12. Power versus SNR for complex (blue) and magnitude-only (red) models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

irrespective of SNR, while the magnitude-only model power curve falls rapidly as the SNR decreases.

Conclusions

A complex data fMRI activation model was presented as an alternative to the typical magnitude-only data model. Activation statistics were derived from generalized likelihood ratio tests for both models. Activation from both models were presented for real fMRI data, then simulations were performed to compare the power to detect activation regions between the two models for several signal-to-noise ratios with varying task-related contrast effects.

It was found that, for large signal-to-noise ratios, both models were comparable. However, for smaller signal-to-noise ratios, the complex activation model demonstrated superior power of detection over the magnitude-only activation model. This strongly indicates that modeling the complex data may become more useful as voxel sizes get smaller, since this decreases the SNR.

Appendix A. Magnitude and phase distributions

The distribution of the magnitude and phase can be derived as follows. Let $y_R = \rho\cos\theta + \eta_R$ and $y_I = \rho\sin\theta + \eta_I$ where η_R and

Fig. 11. Power versus CNR for complex (blue) and magnitude-only (red) models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

η_I are normally distributed with mean zero and variance σ^2. Then, make a change of variable from (y_R,y_I) to polar coordinates $r^2 = y_R^2 + y_I^2$ and $\phi = \tan^{-1}(y_I/y_R)$ or $y_R = r\cos\phi$ and $y_I = r\sin\phi$. The Jacobian of this transformation is $J(y_R,y_I \rightarrow r,\phi) = r$. The joint distribution of r and ϕ using trigonometric identities becomes

$$p(r,\phi|\rho,\theta,\sigma^2) = \frac{r}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left[r^2 + \rho^2 - 2\rho r\cos(\phi - \theta)\right]}.$$

Magnitude distribution

The marginal distribution of the magnitude r is found by integrating out the phase ϕ

$$p(r|\rho,\theta,\sigma^2) = \frac{r}{\sigma^2} e^{-\frac{1}{2\sigma^2}\left[r^2+\rho^2\right]} \int_{\phi=-\pi}^{\pi} \frac{1}{2\pi} e^{\frac{1}{\sigma^2}\rho r\cos(\phi - \theta)} d\phi$$

where the integral factor often denoted $I_o(r\rho/\sigma^2)$ is the zeroth order modified Bessel function of the first kind. The normal limiting distribution for large SNR or $\rho \rightarrow \infty$, is found by using the asymptotic form $I_o(r\rho/\sigma^2) \approx \exp(r\rho/\sigma^2)/\sqrt{(2\pi r\rho/\sigma^2)}$ of the Bessel function (Anderson and Kirsch, 1996). Additionally, in this limit, it is assumed that the exponential form of the normal distribution drops off more rapidly compared to the variation in the ratio $\sqrt{r/\rho}$ left as a factor. The distribution of the magnitude becomes the normal distribution with mean ρ and variance σ^2. The Rayleigh limiting distribution for zero SNR or $\rho = 0$ is found by noting that $I_o(0) = 1$. The distribution of the magnitude becomes

$$p(r|\rho,\sigma^2) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}.$$

Phase distribution

The marginal distribution of the phase ϕ is found by integrating out the magnitude r

$$p(\phi|\rho,\theta,\sigma^2) = \frac{e^{-\frac{\rho^2}{2\sigma^2}}}{2\pi}\left[1 + \frac{\rho}{\sigma}\sqrt{2\pi}\cos(\phi-\theta)e^{\frac{\rho^2\cos^2(\phi-\theta)}{2\sigma^2}} \times \int_{r=-\infty}^{\frac{\rho\cos(\phi-\theta)}{\sigma}} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz.\right]$$

The normal limiting distribution for large SNR or $\rho \rightarrow \infty$ is found by multiplying through, noting that the first term is approximately zero, that the difference between ϕ and θ is small so that the cosine of their difference is approximately one, and the sine of their difference is approximately their difference. The distribution of the phase becomes the normal distribution with mean θ and variance $(\sigma/\rho)^2$.

The uniform limiting distribution for zero SNR or $\rho = 0$ is found by noting that the integral factor goes to unity. The distribution of the phase becomes uniform on $[-\pi,\pi]$. The distribution of the magnitude and phase for intermediate values of SNR can be found by numerical integration or Monte Carlo simulation. The complex model presented in this paper does not make these large SNR approximations.

Appendix B. Generalized likelihood ratio test

Complex model with θ

In applications using multiple regression including fMRI, we often wish to test linear contrast hypothesis (for each voxel) such as

$$\begin{array}{lll} H_0 : C\beta = \gamma & \text{vs} & H_1 : C\beta \neq \gamma \\ \sigma^2 > 0 & & \sigma^2 > 0, \end{array}$$

where C is an $r \times (q+1)$ matrix of full row rank and γ is an $r \times 1$ vector.

The likelihood ratio statistic is computed by maximizing the likelihood $p(y|\beta,\sigma^2,X)$, with respect to β, and σ^2 under the null and alternative hypotheses. Denote the maximized values under the null hypothesis by $(\tilde{\beta},\ \tilde{\sigma}^2)$ and those under the alternative hypothesis as $(\hat{\beta},\ \hat{\sigma}^2)$. These maximized values are then substituted into the likelihoods and the ratio taken. With the aforementioned distributional specifications, the likelihood of the model is

$$p(y|X,\beta,\theta,\sigma^2) = \left(2\pi\sigma^2\right)^{-\frac{2n}{2}} e^{-\frac{h}{2\sigma^2}} \tag{B.1}$$

where

$$\begin{aligned} h &= \frac{1}{2n}\left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \beta\cos\theta \\ \beta\sin\theta \end{pmatrix}\right]'\left[y - \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}\begin{pmatrix} \beta\cos\theta \\ \beta\sin\theta \end{pmatrix}\right] \\ &= \beta'(X'X)\beta - 2\beta'(X'X)\left[\hat{\beta}_R\cos\theta + \hat{\beta}_I\sin\theta\right] + \hat{\beta}_R'(X'X)\hat{\beta}_R \\ &\quad + \hat{\beta}_I'(X'X)\hat{\beta}_I \\ &\quad + y_R'\left[I_n - X(X'X)^{-1}X'\right]y_R + y_I'\left[I_n - X(X'X)^{-1}X'\right]y_I \end{aligned}$$

Unrestricted MLEs

Maximizing this likelihood with respect to the parameters is the same as maximizing the logarithm of the likelihood with respect to the parameters. In the case of β and θ, it is the same as minimizing the h term in the exponent.

$$\frac{\partial}{\partial\beta}h\Big|_{\beta=\hat{\beta},\,\theta=\hat{\theta},\,\sigma^2=\hat{\sigma}^2} = 2(X'X)\hat{\beta} - 2(X'X)\left[\hat{\beta}_R\cos\hat{\theta} + \hat{\beta}_I\sin\hat{\theta}\right]$$

$$\frac{\partial}{\partial\theta}h\Big|_{\beta=\hat{\beta},\,\theta=\hat{\theta},\,\sigma^2=\hat{\sigma}^2} = -2\hat{\beta}(X'X)\left[\left(-\sin\hat{\theta}\right)\hat{\beta}_R + \left(\cos\hat{\theta}\right)\hat{\beta}_I\right]$$

$$\frac{\partial}{\partial\sigma^2}\log\left[p(y|X,\beta,\theta,\sigma^2)\right]\Big|_{\beta=\hat{\beta},\,\theta=\hat{\theta},\,\sigma^2=\hat{\sigma}^2} - \frac{2n}{2}\frac{1}{\hat{\sigma}^2} + \frac{\hat{h}}{2}\frac{1}{(\hat{\sigma}^2)^2}$$

where is $\hat{h}$ with h with MLEs substituted in. By setting these derivatives equal to zero and solving, we get the MLEs under the unrestricted model given in Eq. (2.10).

Restricted MLEs

Maximizing this likelihood with respect to the parameters is the same as maximizing the logarithm of the likelihood with respect to the parameters. In the case of β and θ, it is the same as minimizing the h term in the exponent with the restriction in the form of a Lagrange multiplier as

$$h = \beta'(X'X)\beta - 2\beta'(X'X)\left[\hat{\beta}_R\cos\theta + \hat{\beta}_I\sin\theta\right] + \hat{\beta}_R'(X'X)\hat{\beta}_R + \hat{\beta}_I'(X'X)\hat{\beta}_I$$

$$+ y_R'\left[I_n - X(X'X)^{-1}X'\right]y_R + y_I'\left[I_n - X(X'X)^{-1}X'\right]y_I$$

$$+ 2\psi'(C\beta - \gamma).$$

Note that the maximization is performed by Lagrange multipliers and the appropriate term has been added to h

$$\frac{\partial}{\partial\beta}h\Big|_{\beta=\tilde{\beta},\,\theta=\tilde{\theta},\,\psi=\tilde{\psi},\,\sigma^2=\tilde{\sigma}^2} = 2(X'X)\tilde{\beta} - 2(X'X)\left[\tilde{\beta}_R\cos\tilde{\theta} + \tilde{\beta}_I\sin\tilde{\theta}\right] + 2C\tilde{\psi}$$

$$\frac{\partial}{\partial\theta}h\Big|_{\beta=\tilde{\beta},\,\theta=\tilde{\theta},\,\psi=\tilde{\psi},\,\sigma^2=\tilde{\sigma}^2} = -2\tilde{\beta}'(X'X)\left[(\sin\tilde{\theta})\tilde{\beta}_R + (\cos\tilde{\theta})\tilde{\beta}_I\right]$$

$$\frac{\partial}{\partial\psi}h\Big|_{\beta=\tilde{\beta},\,\psi=\tilde{\psi},\,\psi=\tilde{\psi},\,\sigma^2=\tilde{\sigma}^2} = 2(C\tilde{\beta} - \gamma)$$

$$\frac{\partial}{\partial\sigma^2}\log\left[p\left(y|X,\beta,\theta,\sigma^2\right)\right]\Big|_{\beta=\tilde{\beta},\,\theta=\tilde{\theta},\,\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} - \frac{2n}{2}\frac{1}{\tilde{\sigma}^2} + \frac{\tilde{h}}{2}\frac{1}{(\tilde{\sigma}^2)^2}$$

where $\tilde{h}$ is h with MLEs substituted in. By setting these derivatives equal to zero and solving, we get the MLEs under the restricted model given in Eq. (2.11).

Note that $\hat{\sigma}^2 = \hat{h}/(2n)$ and $\tilde{\sigma}^2 = \tilde{h}/(2n)$. Then the generalized likelihood ratio is

$$\lambda = \frac{p\left(y|\tilde{\beta},\tilde{\sigma}^2,\tilde{\theta},X\right)}{p\left(y|\hat{\beta},\hat{\sigma}^2,\hat{\theta},X\right)} = \frac{(\tilde{\sigma}^2)^{-2n/2}e^{-2\tilde{h}n/(2\tilde{h})}}{(\hat{\sigma}^2)^{-2n/2}e^{-2\hat{h}n/(2\hat{h})}}, \tag{B.2}$$

and Eq. (2.15) for the GLRT follows.

Complex Model with α_1 and α_2

Alternatively, the model can be written with $\alpha_1 = \cos\theta$ and $\alpha_2 = \sin\theta$.

Unrestricted MLEs

The term in the exponent is

$$h = \beta'(X'X)\beta - 2\beta'(X'X)\left[\hat{\beta}_R\alpha_1 + \hat{\beta}_I\alpha_2\right] + \hat{\beta}_R'(X'X)\hat{\beta}_R + \hat{\beta}_I'(X'X)\hat{\beta}_I$$

$$+ y_R'\left[I_n - X(X'X)^{-1}X'\right]y_R + y_I'\left[I_n - X(X'X)^{-1}X'\right]y_I - 2\delta\left(\alpha_1^2 + \alpha_2^2 - 1\right).$$

Note the Lagrange multiplier constraint that $\alpha_1^2 + \alpha_2^2 = 1$.

Maximizing this likelihood with respect to the parameters is the same as maximizing the logarithm of the likelihood with respect to the parameters. In the case of β, α_1, and α_2, it is the same as minimizing the h term in the exponent.

$$\frac{\partial}{\partial\beta}h\Big|_{\beta=\hat{\beta},\,\alpha_1=\hat{\alpha}_1,\,\alpha_2=\hat{\alpha}_2,\,\delta=\hat{\delta},\,\sigma^2=\hat{\sigma}^2} = 2(X'X)\hat{\beta} - 2(X'X)\left[\hat{\beta}_R\hat{\alpha}_1 + \hat{\beta}_I\hat{\alpha}_2\right]$$

$$\frac{\partial}{\partial\alpha_1}h\Big|_{\beta=\hat{\beta},\,\alpha_1=\hat{\alpha}_1,\,\alpha_2=\hat{\alpha}_2,\,\delta=\hat{\delta},\,\sigma^2=\hat{\sigma}^2} = -2\hat{\beta}'(X'X)\hat{\beta}_R - 2\hat{\delta}(2\hat{\alpha}_1)$$

$$\frac{\partial}{\partial\alpha_2}h\Big|_{\beta=\hat{\beta},\,\alpha_1=\hat{\alpha}_1,\,\alpha_2=\hat{\alpha}_2,\,\delta=\hat{\delta},\,\sigma^2=\hat{\sigma}^2} = -2\hat{\beta}'(X'X)\hat{\beta}_I - 2\hat{\delta}(2\hat{\alpha}_2)$$

$$\frac{\partial}{\partial\delta}h\Big|_{\beta=\hat{\beta},\,\alpha_1=\hat{\alpha}_1,\,\alpha_2=\hat{\alpha}_2,\delta=\hat{\delta},\,\sigma^2=\hat{\sigma}^2} = -2\left(\hat{\alpha}_1^2 + \hat{\alpha}_2^2 - 1\right)$$

$$\frac{\partial}{\partial\sigma^2}\log\left[p\left(y|X,\beta,\alpha_1,\alpha_2,\sigma^2\right)\right]\Big|_{\beta=\hat{\beta},\,\alpha_1=\hat{\alpha}_1,\,\alpha_2=\hat{\alpha}_2,\,\delta=\hat{\delta},\,\sigma^2=\hat{\sigma}^2} - \frac{2n}{2}\frac{1}{\hat{\sigma}^2} + \frac{\hat{h}}{2}\frac{1}{(\hat{\sigma}^2)^2}$$

where is $\hat{h}$ is h with MLEs substituted in. By setting these derivatives equal to zero and solving, we get the MLEs under the unrestricted model given in Eq. (2.13).

Restricted MLEs

The term in the exponent is

$$h = \beta'(X'X)\beta - 2\beta'(X'X)\left[\hat{\beta}_R\alpha_1 + \hat{\beta}_I\alpha_2\right] + \hat{\beta}_R'(X'X)\hat{\beta}_R + \hat{\beta}_I'(X'X)\hat{\beta}_I$$

$$+ y'_R\left[I_n - X(X'X)^{-1}X'\right]y_R + y_I'\left[I_n - X(X'X)^{-1}X'\right]y_I$$

$$- 2\delta\left(\alpha_1^2 + \alpha_2^2 - 1\right) + 2\psi'(C\beta - \gamma).$$

Maximizing this likelihood with respect to the parameters is the same as maximizing the logarithm of the likelihood with respect to the parameters. In the case of β, α_1, and α_2, it is the same as

minimizing the h term in the exponent. Note that the maximization is performed by Lagrange multipliers and the appropriate term has been added to h

$$\frac{\partial}{\partial \beta} h\Big|_{\beta=\tilde{\beta},\,\alpha_1=\tilde{\alpha}_1,\,\alpha_2=\tilde{\alpha}_2,\,\delta=\tilde{\delta},\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} = 2(X'X)\tilde{\beta} - 2(X'X)\left[\tilde{\beta}_R\tilde{\alpha}_1 + \tilde{\beta}_I\tilde{\alpha}_2\right] + 2C\tilde{\psi}$$

$$\frac{\partial}{\partial \psi} h\Big|_{\beta=\tilde{\beta},\alpha_1=\tilde{\alpha}_1,\alpha_2=\tilde{\alpha}_2,\delta=\tilde{\delta},\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} = 2\left(C\tilde{\beta} - \gamma\right)$$

$$\frac{\partial}{\partial \alpha_1} h\Big|_{\beta=\tilde{\beta},\alpha_1=\tilde{\alpha}_1,\alpha_2=\tilde{\alpha}_2,\delta=\tilde{\delta},\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} = -2\tilde{\beta}'(X'X)\tilde{\beta}_R - 2\tilde{\delta}(2\tilde{\alpha}_1)$$

$$\frac{\partial}{\partial \alpha_2} h\Big|_{\beta=\tilde{\beta},\alpha_1=\tilde{\alpha}_1,\alpha_2=\tilde{\alpha}_2,\delta=\tilde{\delta},\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} = -2\tilde{\beta}(X'X)\tilde{\beta}_I - 2\tilde{\delta}(2\tilde{\alpha}_2)$$

$$\frac{\partial}{\partial \delta} h\Big|_{\beta=\tilde{\beta},\alpha_1=\tilde{\alpha}_1,\alpha_2=\tilde{\alpha}_2,\delta=\tilde{\delta},\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} = -2\left(\tilde{\alpha}_1^2 + \tilde{\alpha}_2^2 - 1\right)$$

$$\frac{\partial}{\partial \sigma^2} \log\left[p\left(y|X,\beta,\alpha_1,\alpha_2,\sigma^2\right)\right]\Big|_{\beta=\tilde{\beta},\alpha_1=\tilde{\alpha}_1,\alpha_2=\tilde{\alpha}_2,\delta=\tilde{\delta},\psi=\tilde{\psi},\sigma^2=\tilde{\sigma}^2} = -\frac{2n}{2}\frac{1}{\tilde{\sigma}^2} + \frac{\tilde{h}}{2}\frac{1}{(\tilde{\sigma}^2)^2}$$

where $\hat{h}$ is h with MLEs substituted in. By setting these derivatives equal to zero and solving, we get the MLEs under the restricted model given in Eq. (2.14).

Appendix C. Prewhitening

In many applications of regression, the errors may be temporally autocorrelated resulting in correct estimation of the regression coefficients but inflated estimation of the residual error variance.

In the multiple regression complex model, the observation error covariance matrix may not be the identity matrix. A common practice is to estimate Φ with $\hat{\Phi}$, prewhiten, then repeat the analysis. For example, an AR(1) temporal autocorrelation (Markov) matrix with autocorrelation parameter ρ_R for the real part and ρ_I for the imaginary part is estimated by $\hat{\rho}_R$ and $\hat{\rho}_R$, their average taken to obtain $\hat{\rho}$ and $\hat{\Phi}$ formed. Estimation of temporal autocorrelation parameters may be done using pseudogeneralized least squares (Bullmore et al., 1996). Then, by obtaining the factorization $\hat{\Phi} = PP'$ and premultiplying

$$\begin{pmatrix} Py_R \\ Py_I \end{pmatrix} = \begin{pmatrix} PX & 0 \\ 0 & PX \end{pmatrix}\begin{pmatrix} \beta\cos\theta \\ \beta\sin\theta \end{pmatrix} + \begin{pmatrix} P\eta_R \\ P\eta_I \end{pmatrix}$$

$$\begin{pmatrix} y_{R*} \\ y_{I*} \end{pmatrix} = \begin{pmatrix} X_* & 0 \\ 0 & X_* \end{pmatrix}\begin{pmatrix} \beta\cos\theta \\ \beta\sin\theta \end{pmatrix} + \begin{pmatrix} \eta_{R*} \\ \eta_{I*} \end{pmatrix} \quad \text{(C.1)}$$

Now, $\eta_* = (\eta_{R*},\eta_{I*})' \sim \mathcal{N}(0, \Sigma \otimes I_n)$, and the data are analyzed according to the complex multiple regression model.

References

Anderson, A.H., Kirsch, J.E., 1996. Analysis of noise in phase contrast MR imaging. Med. Phys. 23 (6), 857–869.

Bandettini, P., Jesmanowicz, A., Wong, E., Hyde, J.S., 1993. Processing strategies for time-course data sets in functional MRI of the human brain. Magn. Reson. Med. 30 (2), 161–173.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc., B 57, 289–300.

Bullmore, E., Brammer, M., Williams, S., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. Magn. Reson. Med. 35, 261–277.

Calhoun, V.D., Adali, T., Pearlson, G.D., van Zijl, P.C.M., Pekar, J.J., 2002. Independent component analysis of fMRI data in the complex domain. Magn. Reson. Med. 48 (1), 80–192.

Cox, R.W., Jesmanowicz, A., Hyde, J.S., 1995. Real-time functional magnetic resonance imaging. Magn. Reson. Med. 33 (2), 230–236.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15, 772–786.

Gonzales, R.C., Woods, R.E., 1992. Digital Image Processing. Addison-Wesley Publishing Company, Reading, MA.

Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy data. Magn. Reson. Med. 34 (6), 910–914.

Haacke, E.M., Brown, R., Thompson, M., Venkatesan, R., 1999. Magnetic Resonance Imaging: Principles and Sequence Design. John Wiley and Sons, New York.

Kruger, G., Kastrup, A., Glover, G.H., 2001. Neuroimaging at 1.5T and 3.0T: comparison of oxygenation-sensitive magnetic resonance imaging. Magn. Reson. Med. 45 (4), 595–604.

Lai, S., Glover, G.H., 1997. Detection of BOLD fMRI signals using complex data. Proc. ISMRM, 1671.

Lindley, D.V., Smith, A.F.M., 1972. Bayes estimates for the linear model. J. R. Stat. Soc., B 34 (1).

Logan, B.R., Rowe, D.B., 2004. An evaluation of thresholding techniques in fMRI analysis. NeuroImage 22 (1), 95–108.

Nan, F.Y., Nowak, R.D., 1999. Generalized likelihood ratio detection for fMRI using complex data. IEEE Trans. Med. Imag. 18 (4), 320–329.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in C, (second ed.). Cambridge University Press, Cambridge, UK.

Reinsel, G.C., Velu, R.P., 1998. Multivariate Reduced-Rank Regression: Theory and Applications, Lect. Notes Stat. vol. 136. Springer Verlag, New York, pp. 136.

Rice, S.O., 1944. Mathematical analysis of random noise. Bell Syst. Tech. J. 23, 282. (Reprinted by N. Wax, Selected papers on Noise and Stochastic Process, Dover Publication, 1954. QA273W3).

Rowe, D.B., 2001. Bayesian source separation for reference function determination in fMRI. Magn. Reson. Med. 45 (5), 374–378.

Rowe, D.B., 2003. Multivariate Bayesian Statistics. CRC Press, Boca Raton, FL, USA.

Scharf, L.L., Friedlander, B., 1994. Matched subspace detectors. IEEE Transactions on Signal Processing 42 (8), 2146–2157.

Severini, T.A., 2001. Likelihood Methods in Statistics. Oxford University Press, Oxford, UK.