Magnitude and Phase Modeling for fMRI Brain Activation

Daniel B. Rowe and Andrew S. Nencka
Department of Biophysics
Medical College of Wisconsin
8701 Watertown Plank Road
Milwaukee, WI 53226

KEY WORDS: fMRI, complex data, Rowe-Logan

Abstract:

In fMRI, the objective is to image the dynamically changing effective proton spin density of a real-valued object. This is performed by selecting a slice of tissue then applying gradients to encode then measure the complex-valued Fourier transform of the effective proton spin density. Due to the magnetic field "irregularities," the inverse Fourier transform reconstructed object is complex-valued as are the voxel time courses. Nearly all fMRI studies derive functional "activation" based on magnitude-only voxel time courses. Here the entire complex or bivariate data are modeled rather than just the magnitude-only data. A general nonlinear multiple regression model is used to describe task related magnitude and/or phase changes within the complex-valued signal, and likelihood ratio tests are derived to determine statistically significant activation in each voxel.

1. Introduction

With the assumption that the noise in fMRI is primarily due to thermal fluctuations in the scanner hardware, the real and imaginary channels in spatial frequency k-space are independently corrupted by normally distributed noise. The inverse Fourier reansform reconstruction procedure is linear and thus the complex voxel time series are also normally distributed. If the k-space measurements are statistically independent then the voxels are also independent in image space. If the k-space measurements are correlated then so are the voxels in image space. Thus, it is often assumed that the noise in each pixel can be considered to be derived from a bivariate normal distribution with mean zero and standard deviation σ (Henkelman, 1985,1986). Although the object being imaged is real, phase irregularities cause flaws in the encoding of proton spins by magnetic field gradients and the resulting reconstructed images are complex (Bernstein et al., 1989; Macovski, 1996; and Haacke et al., 1999). The complex observation in every voxel at every time point is generally transformed from rectangular coordinates of real-imaginary to equivalent information polar coordinates of magnitude-phase. The unique nonlinear transformation of the real-imaginary complex data with bivariate normal distribution to equivalent magnitude-phase data now has joint distribution in Eqn. 1.1

$$p(r,\phi) \quad = \quad \frac{r}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}\left[r^2+\rho^2-2\rho r\cos(\phi-\theta)\right]} \quad (1.1)$$

as indicicated by Rowe and Logan (2004). This joint distribution is marginalized to the Ricean distribution for the magnitude-only data given in Eqn. 1.2 and complicated distribution for the phase-only data given in Eqn. 1.3 (Rice, 1944; Gudbjartsson and Patz, 1995; Rowe and Logan, 2004). In most analyses the phase portion of the data is discarded and the magnitude-only data with probability distribution function in Eqn. 1.2 is used.

$$p(r) \quad = \frac{r}{\sigma^2}e^{-\frac{(r^2+\rho^2)}{2\sigma^2}}\int_{\phi=-\pi}^{\pi}\frac{1}{2\pi}e^{\frac{\rho r\cos(\phi-\theta_0)}{\sigma^2}}d\phi \quad (1.2)$$

$$p(\phi) \quad = \frac{e^{-\frac{\rho^2}{2\sigma^2}}}{2\pi}\left[1+\frac{\rho}{\sigma}\sqrt{2\pi}\cos(\phi-\theta_0)e^{\frac{\rho^2\cos^2(\phi-\theta_0)}{2\sigma^2}} \times\int_{z=-\infty}^{\frac{\rho\cos(\phi-\theta_0)}{\sigma}}\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz\right] \quad (1.3)$$

If the signal-to-noise ratio (SNR) in a voxel is large, then the Ricean distribution in Eqn. 1.2 is well approximated by the normal distribution with mean ρ and variance σ^2. However, if the SNR is *a priori* not known to be sufficiently large or is questionable, a more accurate approximation to the Ricean distribution can be utilized (Rowe, 2005a). Most functional brain activation techniques utilize only the magnitude data and the magnitude marginal normal distribution thus discarding all temporal phase information. Initial magnitude-only activation methods which simplified the complex image problem included appropriate assumptions for the given image resolutions (Bandettini et al., 1993; Cox et al., 1995; Friston et al., 1995). The high SNR from the large voxels allowed the magnitude's Ricean noise to be accurately approximated by assuming the noise was normally distributed. Voxels have become smaller and their SNR decreased approximately linearly with their volume (Kruger et al., 2001). Tasks have become more specific and intrinsic signal response decreased thus SNR is decreased. Furthermore, the previously discarded phase information generally contained little pertinent information as it was significantly corrupted by relatively large magnetic field gradients compared to voxel size, caused by poor magnetic field shimming and air-tissue interfaces. Additionally, subject motion and physiologic noise further compounded phase measurements due to the large magnetic field gradients across voxels. At this time, the phase is much better and useful information can potentially be discerned from it.

2. Complex Models

Neglecting the voxel location and treating them individually, the complex-valued image measured over time in a given voxel is

$$y_t = [\rho_t \cos\theta_t + \eta_{Rt}] + i[\rho_t \sin\theta_t + \eta_{It}] \ , \quad t = 1, \ldots, n \tag{2.1}$$

where $(\eta_{Rt}, \eta_{It})' \sim \mathcal{N}(0, \sigma^2 I_2)$. The true population magnitude is ρ_t and phase is θ_t. The distributional specifications in Eqn. 2.1 are on the real and imaginary parts of the measured image and not on the magnitude or phase. The temporally varying magnitude is generally described in terms of a linear model $\rho_t = x_t'\beta = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_q x_{qt}$ where x_t is the t^{th} row of an $n \times (q+1)$ design matrix X and β is a $(q+1) \times 1$ vector of magnitude regression coefficients. Alternatively, we can represent the observed data at time point t as a 2×1 vector instead of as a complex number

$$\begin{pmatrix} y_{Rt} \\ y_{It} \end{pmatrix} = \begin{pmatrix} x_t'\beta \cos\theta_t \\ x_t'\beta \sin\theta_t \end{pmatrix} + \begin{pmatrix} \eta_{Rt} \\ \eta_{It} \end{pmatrix} , \quad t = 1, \ldots, n \ . \tag{2.2}$$

The model described in Eqn. 2.2 is a very general nonlinear multiple regression model. Rowe and Logan (2004) introduced a precursor to this model in which the phase was temporally constant $\theta_t = \theta$ in each voxel but varied in space from voxel to voxel. The phase is a fixed and unknown quantity, which may be estimated voxel by voxel. Using this magnitude activation in complex data with a constant phase model, Rowe and Logan were able to obtain closed form analytic maximum likelihood parameter estimators under the unrestricted alternative hypothesis $H_1 : C\beta \neq 0$ to be

$$\begin{aligned} \hat{\theta} &= \tfrac{1}{2}\tan^{-1}\left[\frac{2\hat{\beta}_R'(X'X)\hat{\beta}_I}{\hat{\beta}_R'(X'X)\hat{\beta}_R - \hat{\beta}_I'(X'X)\hat{\beta}_I}\right] \\ \hat{\beta} &= \hat{\beta}_R \cos\hat{\theta} + \hat{\beta}_I \sin\hat{\theta}, \\ \hat{\sigma}^2 &= \tfrac{1}{2n}\left[y - \begin{pmatrix} X\hat{\beta}\cos\hat{\theta} \\ X\hat{\beta}\sin\hat{\theta} \end{pmatrix}\right]'\left[y - \begin{pmatrix} X\hat{\beta}\cos\hat{\theta} \\ X\hat{\beta}\sin\hat{\theta} \end{pmatrix}\right] \end{aligned} \tag{2.3}$$

where $\hat{\beta}_R = (X'X)^{-1}X'y_R$, $\hat{\beta}_I = (X'X)^{-1}X'y_I$, and C is a $d \times (q+1)$ contrast matrix. Note that the estimate of the regression coefficients is a linear combination or "weighted" average of estimates from the real and imaginary parts.

Rowe and Logan were also able to obtain closed form analytic maximum likelihood parameter estimators under the restricted null hypothesis $H_0 : C\beta = 0$ to be

$$\begin{aligned} \tilde{\theta} &= \tfrac{1}{2}\tan^{-1}\left[\frac{2\hat{\beta}_R'\Psi(X'X)\hat{\beta}_I}{\hat{\beta}_R'\Psi(X'X)\hat{\beta}_R - \hat{\beta}_I'\Psi(X'X)\hat{\beta}_I}\right] \\ \tilde{\beta} &= \Psi[\hat{\beta}_R \cos\tilde{\theta} + \hat{\beta}_I \sin\tilde{\theta}], \\ \tilde{\sigma}^2 &= \tfrac{1}{2n}\left[y - \begin{pmatrix} X\tilde{\beta}\cos\tilde{\theta} \\ X\tilde{\beta}\sin\tilde{\theta} \end{pmatrix}\right]'\left[y - \begin{pmatrix} X\tilde{\beta}\cos\tilde{\theta} \\ X\tilde{\beta}\sin\tilde{\theta} \end{pmatrix}\right] \end{aligned} \tag{2.4}$$

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

These two variance estimators were combined to form a generalized likelihood ratio statistic with large sample χ^2 distribution under the null hypothesis. This was verified through permutation resampling of complex-valued residuals. It was also noted that this generalized likelihood ratio statistic could be turned into a z statistic via a signed likelihood ratio statistic provided $d = 1$ (Rowe and Logan, 2004).

A subsequent complex-valued model was described by Rowe and Logan (2005) in which the phase was assumed to be unique at each time point $\theta_t \neq \theta_{t'}$. In this magnitude activation in complex data with an unrestricted phase model, Rowe and Logan were able to obtain closed form analytic maximum likelihood parameter estimators under the unrestricted alternative hypothesis $H_1 : C\beta \neq 0$ to be

$$\begin{aligned} \hat{\theta}_t &= \tan^{-1}\left(\frac{y_{It}}{y_{Rt}}\right), \quad t = 1, \ldots, n \\ \hat{\beta} &= (X'X)^{-1}X'r, \\ \hat{\sigma}^2 &= \frac{1}{2n}\left[y - \begin{pmatrix} \hat{A}_1 X\hat{\beta} \\ \hat{A}_2 X\hat{\beta} \end{pmatrix}\right]'\left[y - \begin{pmatrix} \hat{A}_1 X\hat{\beta} \\ \hat{A}_2 X\hat{\beta} \end{pmatrix}\right] , \end{aligned} \tag{2.5}$$

where $\hat{A}_1$ and $\hat{A}_2$ are diagonal matrices with $\cos\hat{\theta}_t$ and $\sin\hat{\theta}_t$ as the t^{th} diagonal element. The vector r

is the observed magnitude with t^{th} element r_t given by $\sqrt{y_{Rt}^2 + y_{It}^2}$.

Rowe and Logan were also able to obtain closed form analytic maximum likelihood parameter estimators under the restricted null hypothesis $H_0 : C\beta = 0$ to be

$$\begin{aligned}
\tilde{\theta}_t &= \tan^{-1}\left(\frac{y_{It}}{y_{Rt}}\right), \quad t = 1, \ldots, n \\
\tilde{\beta} &= \Psi\hat{\beta}, \\
\tilde{\sigma}^2 &= \frac{1}{2n}\left[y - \begin{pmatrix} \tilde{A}_1 X\tilde{\beta} \\ \tilde{A}_2 X\tilde{\beta} \end{pmatrix}\right]'\left[y - \begin{pmatrix} \tilde{A}_1 X\tilde{\beta} \\ \tilde{A}_2 X\tilde{\beta} \end{pmatrix}\right],
\end{aligned} \tag{2.6}$$

where Ψ is as previously defined, $\tilde{A}_1$ and $\tilde{A}_2$ are diagonal matrices with $\cos\tilde{\theta}_t$ and $\sin\tilde{\theta}_t$ as the t^{th} diagonal element. The restricted regression coefficients can also be shown to be equivalent to the magnitude-only model because the multiplicative factor Ψ is identical in both cases.

The likelihood ratio statistic with some algebra was shown to be written as

$$F = \frac{(n-q-1)}{d}\frac{\hat{\beta}'C'[C(X'X)^{-1}C']^{-1}C\hat{\beta}}{2n\hat{\sigma}^2} . \tag{2.7}$$

The F statistic and equivalent likelihood ratio statistic is identical to the one from the magnitude-only model. In either case the F statistic follows the same distribution. If the signal-to-noise ratio is large so that r_t is approximately normal, then F follows an $F_{d,n-q-1}$ distribution under the null hypothesis, where d is the full row rank of C. Otherwise, one might use the Ricean distribution (Rice, 1944; Gudbjartsson and Patz, 1995; Rowe, 2005a) to derive the proper distribution of the F statistic. In either case, the estimates of β and the likelihood ratio test depend only on the magnitude data.

This magnitude activation in complex data with an unrestricted phase model was shown to be equivalent to the magnitude-only model in terms of regression coefficients and activation statistics. In essence, deriving the magnitude-only model from complex data.

The previously described two magnitude activation models represent the extremes in terms of temporal phase change parametrization assumptions. These being constant and completely unspecified over time. There are many possible parameterizations between these two including describing the temporal phase changes in terms of a linear model as in Eqn. 2.7 so that both magnitude and phase change linearly over time

$$\begin{aligned}
y_t &= [\rho_t\cos\theta_t + \eta_{Rt}] + i[\rho_t\sin\theta_t + \eta_{It}] \\
\rho_t &= x_t'\beta = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_{q_1}x_{q_1 t} \\
\theta_t &= u_t'\gamma = \gamma_0 + \gamma_1 u_{1t} + \cdots + \gamma_{q_2}u_{q_2 t}, \\
& \qquad t = 1, \ldots, n
\end{aligned} \tag{2.8}$$

where x_t' is the t^{th} row of a design matrix X for the magnitude and u_t' is the t^{th} row of a design matrix U for the phase. The last columns of X and U are (different) task related reference functions. In this model, we can determine task related magnitude and/or phase changes within the observed complex-valued time course that has a component related to the reference function. This can be accomplished with magnitude and phase contrast matrices $C = (0, ..., 0, 1)$ and $D = (0, ..., 0, 1)$.

There are four readily visible hypotheses for testing
$H_a : C\beta \neq 0, \quad D\gamma \neq 0$
$H_b : C\beta = 0, \quad D\gamma \neq 0$
$H_c : C\beta \neq 0, \quad D\gamma = 0$
$H_d : C\beta = 0, \quad D\gamma = 0$
that can combined in different ways to form specific meaningful hypothesis pairs. The constant and unrestricted phase models and hypotheses are supported within this framework. Other hypotheses are possible including one sided or interval hypotheses.

The log likelihood using the joint magnitude-phase distribution was written as

$$\begin{aligned}
LL &= -n\log(2\pi) - \textstyle\sum_{t=1}^n \log r_t - n\log\sigma^2 \\
&\quad - \tfrac{1}{2\sigma^2}\textstyle\sum_{t=1}^n \left[r_t^2 + (x_t'\beta)^2 - 2(x_t'\beta)\underbrace{r_t\cos(\phi_t - u_t'\gamma)}_{r_{*t}}\right] \\
&= -n\log(2\pi) - \textstyle\sum_{t=1}^n \log r_t - n\log\sigma^2 \\
&\quad -\tfrac{1}{2\sigma^2}\left[(r - X\beta)'(r - X\beta) + 2(r - r_*)'X\beta\right] \\
&\quad +\psi'(C\beta - 0) + \delta'(D\gamma - 0)
\end{aligned}$$

and iteratively maximized under the unconstrained null and constrained alternative hypotheses to yield estimates $H_0 : (\tilde{\beta}, \tilde{\gamma}, \tilde{\sigma}^2)$ and $H_1 : (\hat{\beta}, \hat{\gamma}, \hat{\sigma}^2)$. The appropriate Lagrange constraints are retained or omitted appropriately. These maximum likelihood estimates are utilized in a generalized likelihood ratio test statistic.

3. Application to fMRI dataset

A bilateral finger tapping experiment was performed in a block design with 16s off followed by eight epochs of 16s on and 16s off. Scanning was performed using a 1.5T GE Signa in which 5 axial slices of size 96×96 were acquired. In image reconstruction, the acquired data was zero filled to 128×128. After Fourier image reconstruction, each voxel has dimensions in mm of $1.5625 \times 1.5625 \times 5$, with TE= 47ms. Observations were taken every TR= 1000ms so that there are 272 in each voxel. Data from a single axial slice through the motor cortex was selected for analysis. Pre-processing using an ideal 0/1 frequency filter was performed to remove respiration and low frequency physiological noise in addition to the removal of the first three points to omit signal equilibration effects. Where necessary, the phase time courses were unwrapped for jumps greater than π between successive observations.

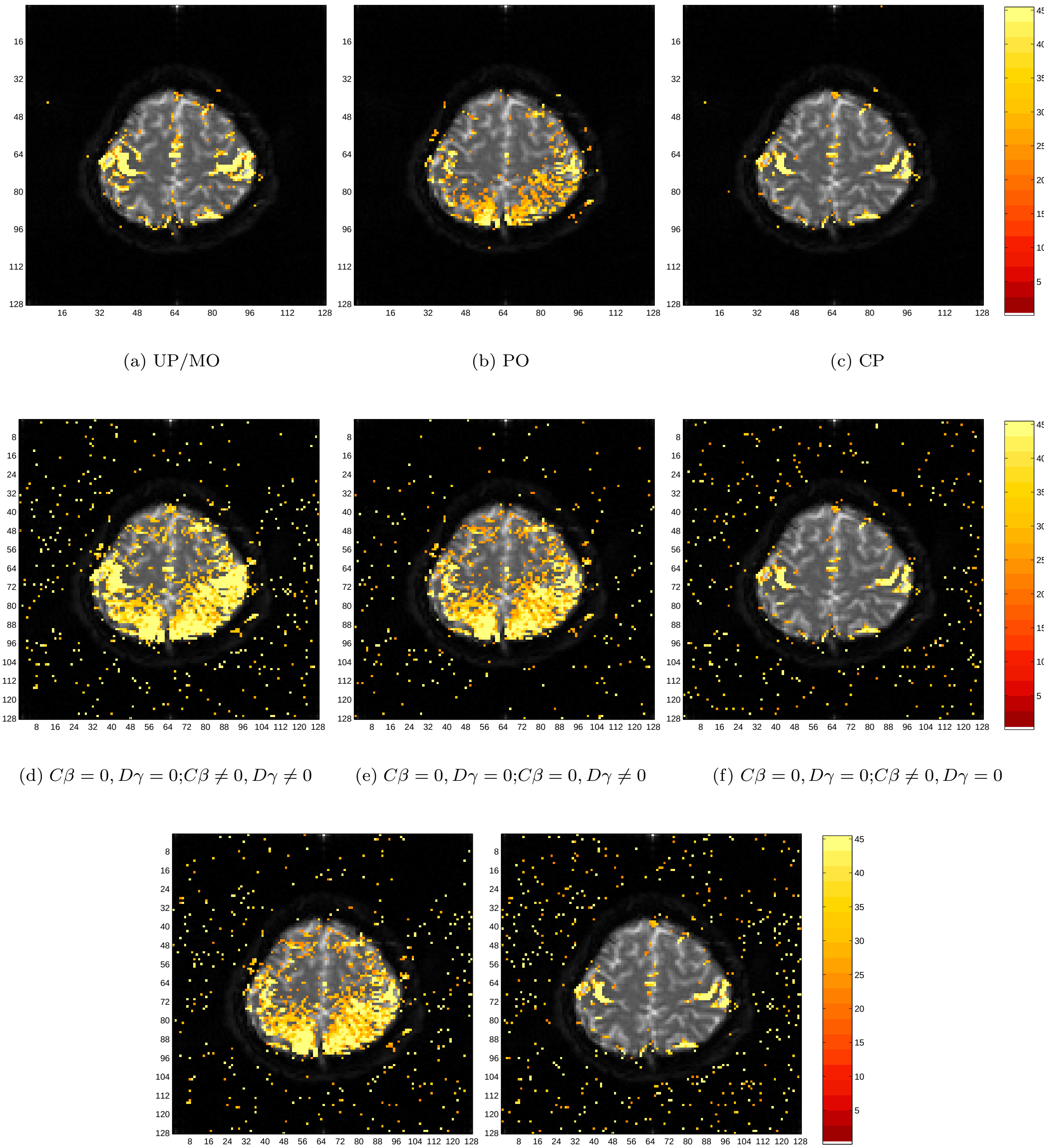
The abovementioned activation models were fit to the data with an intercept, a zero mean time trend, and a ± 1 square wave reference function. In Fig. 1(a)-(c) are 5% Bonferroni familywise error (FWE) rate thresholded χ^2-statistic activation maps with real fMRI data for (a) the complex unrestricted phase (UP) or usual magnitude-only data model; (b) and a phase-only (PO) data model (activation from phase-only data assuming normality); (c) the Rowe-Logan complex constant phase (CP) activation model. In Fig. 1(d)-(h) are 5% Bonferroni familywise error (FWE) rate thresholded χ^2-statistic activation maps from the five hypothesis pairs from the complex linear magnitude and/or phase model for (d) H_d vs H_a, (d) H_d vs H_b, (f) H_d vs H_c, (g)H_c vs H_a, and (h) H_b vs H_a.

The activation maps for voxels within the brain in Figs.1 (c), (f), and (h) are similar, as are those in Figs. 1(b), (e), and (g). The above threshold activated voxels in Fig. 1(c) appear to be a subset of those in Fig. 1(a). The above threshold voxel activation map in Fig. 1(h) appears to be a combination of those in Figs. 1(b) and (c).

4. Conclusions

A very general 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 and applied on real data. It was found that the magnitude-only data model declares voxels as active regardless of any phase changes, phase-only data model declares voxels as active regardless of any magnitude changes, and the five complex linear phase models were sensitive to different (CNR,TRPC) combinations. The complex linear phase model is very general and includes all previously introduced activation models as special cases. Perhaps this model will reach its full potential with other experimental data acquisition methods such as flow tagging or steady state free precession. There are indications that this model might be useful in eliminating unwanted voxels due to venous contributions to the BOLD signal (Rowe, 2005; Nencka and Rowe, 2005).

Figure 1: Bonferroni 5% thresholded activation statistics.



(a) UP/MO

(b) PO

(c) CP

(d) $C\beta = 0, D\gamma = 0; C\beta \neq 0, D\gamma \neq 0$

(e) $C\beta = 0, D\gamma = 0; C\beta = 0, D\gamma \neq 0$

(f) $C\beta = 0, D\gamma = 0; C\beta \neq 0, D\gamma = 0$

(g) $C\beta \neq 0, D\gamma = 0; C\beta \neq 0, D\gamma \neq 0$

(h) $C\beta = 0, D\gamma \neq 0; C\beta \neq 0, D\gamma \neq 0$

References

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

[2] M.A. Bernstein, D.M. Thomasson, and W.H. Perman. Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction. *Med. Phys.*, 16(5):813–817, 1989.

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

[4] H. Gudbjartsson and S. Patz. The Rician distribution of noisy data. *Magn. Reson. Med.*, 34(6):910–914, 1995.

[5] E.M. Haacke, R. Brown, M. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Principles and Sequence Design*. John Wiley and Sons, New York, NY, USA, 1999.

[6] R.M. Henkelman. Erratum: Measurement of signal intensities in the presence of noise in MR images [med. phys. 12,232(1985)]. *Med. Phys.*, 12(2):230–233, 1985.

[7] R.M. Henkelman. Measurement of signal intensities in the presence of noise in MR images. *Med. Phys.*, 12(2):230–233, 1985.

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

[9] A. Macovski. Noise in MRI. *Magn. Reson. Med.*, 36(3):494–497, 1996.

[10] R.S. Menon. Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI. *Magn. Reson. Med.*, 47(1):1–9, 2002.

[11] A. S. Nencka and D. B. Rowe. Complex constant phase method removes venous BOLD component in fMRI. *Proc. Soc. of Magn. Reson. Med.*, 13:495, 2005.

[12] S.O. Rice. Mathematical analysis of random noise. *Bell system Tech. J.*, 23:282, 1944. Reprinted by N. Wax, Selected papers on Noise and Stochastic Process, Dover Publication, 1954. QA273W3.

[13] D.B. Rowe. Complex activation is more focal and concentrated to parenchymal tissue. *Proc. Soc. of Magn. Reson. Med.*, 13:1575, 2005.

[14] D.B. Rowe. Modeling both the magnitude and phase of complex-valued fMRI data. *NeuroImage*, 25(4):1310–1324, 2005.

[15] D.B. Rowe. Parameter estimation in the magnitude-only and complex-valued fMRI data models. *NeuroImage*, 25(4):1124–1132, 2005.

[16] D.B. Rowe and B.R. Logan. A complex way to compute fMRI activation. *NeuroImage*, 23(3):1078–1092, 2004.

[17] D.B. Rowe and B.R. Logan. Complex fMRI analysis with unrestricted phase is equivalent to a magnitude-only model. *NeuroImage*, 24(2):603–606, 2005.