Generalized Likelihood Ratio Detection for Functional MRI Signal Using Complex Data

F.Y.Nan, R.D.Nowak

Dept. of Electrical Engineering,

Michigan State University, E. Lansing, MI, 48824

nanfangy@egr.msu.edu,   nowak@egr.msu.edu

Phone:(517)432-2158,   Fax:(517)353-1980

In preparation for submission to *IEEE Transactions on Medical Imaging*

March 9, 1999

Abstract

The majority of fMRI studies obtain functional information using statistical tests based on the magnitude image reconstructions. Recently, a complex correlation (CC) test was proposed based on the complex image data in order to take advantage of phase information in the signal. However, the CC test ignores additional phase information in the baseline component of the data. In this paper, a new detector for fMRI based on a Generalized Likelihood Ratio Test (GLRT) is proposed. The GLRT exploits the fact that the fMRI response signal as well as the baseline component of the data share a common phase. Theoretical analysis and Monte Carlo simulation are used to explore the performance of the new detector. At relatively low signal intensities, the GLRT outperforms both the standard magnitude data test and the CC test. At high signal intensities, the GLRT performs as well as the standard magnitude data test and significantly better than the CC test.

Index Terms: functional magnetic resonance imaging (fMRI), signal detection, statistics, hypothesis testing

1 Introduction

In functional magnetic resonance imaging (fMRI) a series of MR images of the brain are acquired over time to detect neural activity. Neural activity has been linked to blood oxygenation levels in blood vessels nearby active neurons. This relationship is called the Blood Oxygenation Level Dependent (BOLD) effect [28]. Subtle variations in the magnetic properties of oxygenated and de-oxygenated blood induce changes in the MR signal intensity which can be used to detect neural activity. The BOLD effect can be used to obtain maps of active and non-active regions of the brain. In order to achieve high SNR, the spatial *and* temporal imaging resolution must be limited [15]. Unfortunately, low resolution imaging may lead to a loss in signal information originating in microvasculature [26]. Hence, there is a fundamental trade-off between resolution and SNR in fMRI. It is therefore of great interest to develop reliable detection methods for fMRI in the presence of noise.

Most fMRI imaging methods are based on detecting intensity changes in a sequence of two or more MR images of a certain volume of the brain. When comparing just two images — a "rest state" and "active state" image — a two-sample t-test is routinely used. In repetitive experiments involving a dynamic time sequence of images, a correlation method is common in which the correlation between each voxel time series and a reference signal is used to decide whether or not activity is present [2]. We will focus on the repetitive experiments in this paper, however similar results and conclusions can be drawn in other contexts such as event-related experiments. Many generalizations and extensions of this simple idea have been proposed [12, 11, 17, 22, 25, 26, 27], and under various assumptions and experimental setups the fMRI detection problem is equivalent to well-known statistical tests including t-tests, F-tests, and χ^2-tests. Bayesian strategies have also been recently proposed for fMRI [10], [16].

Almost all fMRI tests are based on the magnitude image data. In standard practice, the raw MRI data is reconstructed and the magnitude is taken to eliminate the (unknown) phase. A common approach to fMRI detection is based on a test statistic obtained by correlating the magnitude voxel time-series data with a known reference signal which is assumed to be representative of the BOLD response. We call this detector the magnitude correlation (MC) test. Recently, however, Lai and Glover proposed a complex correlation (CC) test based on the complex image data (*i.e.,* image data before taking the magnitude of each voxel), in order to take advantage of phase information in the data and improve the detectability of fMRI responses [17].

Here, we show that the CC test statistic is F-distributed and has a constant false-alarm rate (CFAR) property. This means that a specified false-alarm rate, *i.e.,* the probability of deciding a voxel is active when in fact it is not, can be achieved irrespective of the signal-to-noise ratio, which is generally unknown *a priori.* Throughout the paper we denote the false-alarm rate by P_f. Despite the CFAR property, the CC test focuses only on the response component of the data and ignores the constant baseline component of the data. The constant component does not contain information relevant to the response itself, but it does contain important information about the phase. Although the phase is a *nuisance* parameter in the testing problem, more accurate knowledge of the phase can improve the detectability of the fMRI response. In this paper, we propose a new test based on the Generalized Likelihood Ratio Test (GLRT) principle that allows us to incorporate the phase information contained in the constant data component.

Theoretical and Monte Carlo studies are used to show that the GLRT outperforms the CC test. Specifically, we show that for a fixed false-alarm rate P_f, the GLRT's detection rate is higher than that of the CC test. Furthermore, we show that the GLRT also performs significantly better than the MC test at low SNR. The performances of the GLRT and MC test are roughly the same at high SNR, and in such situations both perform better than the CC test.

In this paper we stress very simple voxel-wise testing based on a Gaussian white noise observation model. Voxel-wise testing ignores spatial relationships in fMRI data. Moreover, the white noise model does not capture more complicated disturbances present in fMRI data such as time-correlated noises due to physiologic motions. However, since the focus of this work is to assess the potential benefits of fMRI detection using complex data, we employ a simple data model and testing procedure to explore this basic issue, realizing our assumptions are perhaps too simplistic in many practical cases. However, it is possible to extend our results and conclusions to more elaborate approaches based on more realistic data and/or correlated noise models, potentially accounting for uncertainties in the BOLD response and/or spatial relationships among neighboring voxels. Such extensions are briefly discussed in the conclusions.

The paper is organized as follows. In Section 2, we review a basic model for fMRI data and establish some vector notation that will be used throughout the paper. In Section 3, we examine the standard MC and CC tests and study the statistical properties of each. We also derive the new GLRT for fMRI. The properties of the GLRT are discussed and mathematical analysis is relegated to the Appendices. In Section 4, we compare the performance of the MC test, CC test, and GLRT in various SNR regimes. Although the GLRT statistic does not have a standard distribution, we use exhaustive Monte Carlo simulation to assess the performance of the detector. Our results show that *the GLRT does have a CFAR property and we give a simple rule for choosing the threshold level to achieve a desired* P_f. We also demonstrate the performance of all methods in a simulated fMRI experiment. We close in Section 5 with discussion and conclusions.

2 fMRI Signal Model

The most common reconstruction technique in MRI is to compute the inverse discrete Fourier transform (DFT) of the raw data. Due to phase errors which are difficult to control, the signal component of the measurements appears in both real and imaginary channels [19], [5]. This suggests the following simple model for an fMRI voxel time-series. Let $\mathbf{x}$ denote an $N \times 1$ vector containing the voxel time-series data:

$$\mathbf{x} = (a\mathbf{1} + b\mathbf{r})(\cos\vartheta + i\sin\vartheta) + \sigma\mathbf{n}_c \tag{1}$$

The data vector $\mathbf{x}$ consists of two complex signal components. The first component $a\mathbf{1}$ is a constant baseline component, where $\mathbf{1}$ denotes an $N \times 1$ vector of ones and $a > 0$ is the amplitude of the constant component. This vector represents the average value of the time-series. The second component $b\mathbf{r}$ is the oscillatory[1] response signal. The vector $\mathbf{r}$ is a reference function that models the expected response characteristic, and it is assumed to be known throughout the paper. The amplitude b characterizes the strength of the response. In the absence of activity $b = 0$. The two components share a common phase ϑ. Hence, we model this *phase-coupling* by multiplying both components by the complex number $\cos\vartheta + i\sin\vartheta$, where $i = \sqrt{-1}$. In addition to the two signal components, an additive complex Gaussian white noise component $\sigma\mathbf{n}_c$ models random fluctuations in the fMRI time-series. The term $\mathbf{n}_c$ denotes a standard (zero mean, unit variance) complex Gaussian vector. The factor σ scales the noise resulting in a variance of σ^2. In general, the parameters of this model a, b, ϑ, and σ^2 are unknown and are different for each voxel time-series. We have compared this model to actual fMRI time-series and found that our assumptions are fairly reasonable. In particular, the phase-coupling between the constant and response components has been verified by our experiments. Figure 1 shows the real part, imaginary part and phase of one time series from real fMRI experiment and it is illustrative of the constant phase idea. One limitation of our model, as mentioned earlier, is the white noise assumption. The white noise model is generally an oversimplification of the noises inherent in fMRI (*e.g.,* physiologic motions), but it is a tractable model for exploring the potential of complex data detection schemes. Extensions of this work to more realistic noise models are discussed in Section 5.

[1] We focus on repetitive experiments with (known) oscillatory response signals to illustrate our ideas. However, it may be possible to extend the analysis and results to other situations such as event-related experiments (see

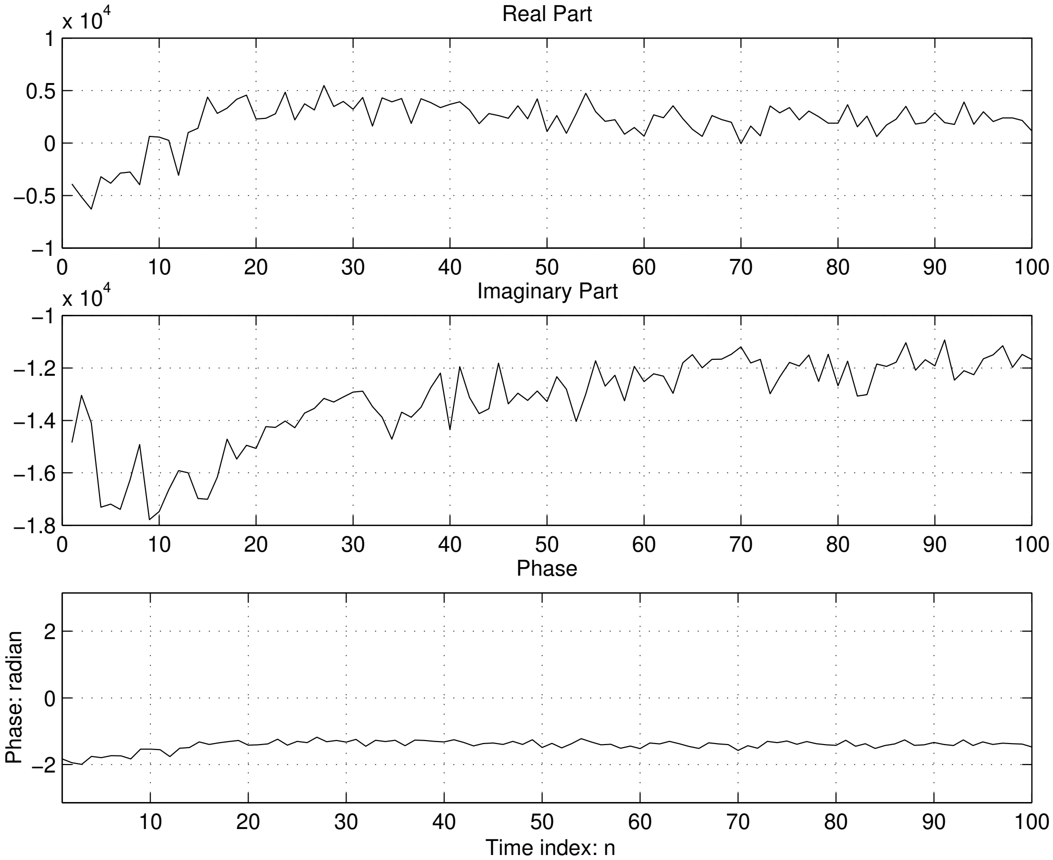


Figure 1. The real part, imaginary part and phase of one time series from a real fMRI experiment, illustrating the the constant phase idea in our model.

A GLRT can be derived directly from this complex model (1), but this form is not well suited for mathematical analysis. Because complex numbers can be interpreted as pairs of real numbers, we re-express the complex model (1) as a $2N \times 1$ dimensional real-valued model:

$$\mathbf{y} = \mathbf{S}\phi + \mu\mathbf{H}\phi + \sigma\mathbf{n} \tag{2}$$

where $\mu = b/a$ and

$$\mathbf{y} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{r} & \mathbf{0} \\ \mathbf{0} & \mathbf{r} \end{bmatrix}, \quad \phi = \begin{bmatrix} a\cos\vartheta \\ a\sin\vartheta \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} \mathbf{n}_{cR} \\ \mathbf{n}_{cI} \end{bmatrix}.$$

The subscripts R and I denote real and imaginary parts, respectively. The phase-coupling in the complex model is manifest in the real model as a nonlinear coupling between the parameter μ and the parameter ϕ.

We point out here that this (nonlinear) model stands in marked contrast to the classical linear regression model

$$\mathbf{y} = \mathbf{S}\phi_1 + \mu\mathbf{H}\phi_2 + \sigma\mathbf{n}. \tag{3}$$

where ϕ_1 and ϕ_2 are independent. In the next section we show that the CC test proposed by Lai and Glover [17] can be derived from the linear model in (3) above. It is our contention that the nonlinear model is a more accurate representation of the physical fMRI problem, and, indeed, the new GLRT based on the nonlinear model outperforms the CC test.

Before moving on, let us establish some basic mathematical conventions and notation that are used throughout the paper. All vector norms are the standard (Euclidean) 2-norm. Let $\mathbf{M}$ denote an $p \times q$ matrix, (*e.g.,* $\mathbf{S}$ or $\mathbf{H}$). Let $\mathbf{P_M}$ denote the matrix that projects a vector onto the subspace spanned by the columns of $\mathbf{M}$, *i.e.,* $\mathbf{P_M} = \mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T$, where the superscript T denotes matrix transposition. Let $\mathbf{P}^{\perp}_{\mathbf{M}}$ denote the matrix projecting a vector onto the complementary subspace that is orthogonal to the subspace spanned by the columns of $\mathbf{M}$, *i.e.,* $\mathbf{P}^{\perp}_{\mathbf{M}} = \mathbf{I} - \mathbf{P_M}$, where $\mathbf{I}$ denotes the $p \times p$ identity matrix.

3 Methods for Functional MRI Detection

In this section we review the basic MC and CC tests and introduce the new GLRT test for fMRI detection. In fact, all three tests may be interpreted as GLRTs based on different data models as we will show. These interpretations illuminate the underlying (although sometimes overlooked) modeling assumptions associated with each method. Before we look at each method, let us briefly review the GLRT principle.

3.1 Generalized Likelihood Ratio Tests

The likelihood ratio test (LRT) [14] is an optimal method for deciding which of two hypotheses (competing data models) best described a set of observed data. The data model corresponding to each hypothesis is a probability density function (pdf). Unfortunately, however, to implement the LRT, the pdf's under each hypothesis must be completely specified. This is not the case in fMRI. In fMRI, we have two hypotheses; H_0, BOLD response absent ($\mu = 0$), and H_1, BOLD response present ($\mu \neq 0$). Under hypothesis H_0, the vector

ϕ and the noise power σ^2 are unknown. Under hypothesis H_1, ϕ, σ^2 and μ are unknown. Due to the unknowns, in fMRI we have what is called a *composite* hypothesis test.

In special cases, it is possible to derive *universally most powerful* (UMP) tests for a composite hypothesis problems. However in the complicated problems with multiple nuisance parameters, UMP tests are not easily derived or more often unavailable. In such complicated cases, such as the fMRI problem at hand, there are two standard approaches to composite hypothesis testing. The *Bayesian* approach prescribes prior pdf's for the unknown parameters themselves and the likelihoods are integrated against these pdf's to eliminate the dependence of the LRT on the unknown parameters. The *generalized likelihood ratio test* (GLRT) is another approach to composite hypothesis testing. The GLRT is often preferable to Bayesian approaches due to its ease of implementation and less restrictive assumptions (specifically, the GLRT does not require the specification of prior probability distributions for the unknown parameters) [14], [23]. For example, the prior pdf's for the unknown parameters are generally unknown to us. For these reasons, we focus on the GLRT in this paper.

The idea of GLRT is motivated by the classical likelihood ratio test. In a simple hypothesis testing problem, the pdf for each hypothesis is completely known. Let $p_{H_i}(\mathbf{x}; \Theta_i)$, $i = 0, 1$, denote the pdf's corresponding to the two hypotheses. Recall that $\mathbf{x}$ denotes the data. The argument Θ_i denotes a *known* parameter that specifies the precise form of the pdf. For example, Θ_i may represent the mean and covariance of a multivariate Gaussian density. The LRT decides H_1 if

$$L(\mathbf{x}) = \frac{p_{H_1}(\mathbf{x}; \Theta_1)}{p_{H_0}(\mathbf{x}; \Theta_0)} > \eta,$$

where η is a user specified threshold, which can be chosen to achieve a desired P_f. The likelihood ratio (LR), $L(\mathbf{x})$, is a function of the data $\mathbf{x}$, and it is called the *test statistic.*

The GLRT is also based on the LR, but in the composite case the parameters are unknown. The key idea in the GLRT is to replace the unknown parameters by their maximum likelihood estimates (MLE's). In general, the GLRT decides H_1 if

$$L(\mathbf{x}) = \frac{p_{H_1}(\mathbf{x}; \widehat{\Theta}_1)}{p_{H_0}(\mathbf{x}; \widehat{\Theta}_0)} > \eta,$$

where $\widehat{\Theta}_1$ is the MLE of Θ_1 assuming H_1 is true, and $\widehat{\Theta}_0$ is the MLE of Θ_0 assuming H_0 is true. The MLE of a parameter is simply the value that makes the observed data most likely (*i.e.,* the value of the parameter that maximizes the corresponding pdf evaluated at $\mathbf{x}$). The GLRT has no optimality property in general, but it asymptotically approaches the UMP test among all parameter-invariant tests [6]. For more details on maximum likelihood estimation and the GLRT, see [14].

3.2 Method 1: Magnitude Correlation Test

Under the assumption of a Gaussian white noise model for the complex data, the magnitude of fMRI data is Rician distributed [19]. However, for large values of the ratio a/σ (ratio of baseline signal intensity to noise standard deviation) the Rician density can be well approximated as Gaussian distribution. To see this, note that x_j, the jth observation in the time series can be written as:

$$x_j = (a + br_j)\cos\theta + \sigma n_{Rj} + i[(a + br_j)\sin\theta + \sigma n_{Ij}]$$

So

$$\begin{aligned} y_j \equiv |x_j| &= \sqrt{[(a + br_j)\cos\theta + \sigma n_{Rj}]^2 + [(a + br_j)\sin\theta + \sigma n_{Ij}]^2} \\ &= \sqrt{(a + br_j)^2 + \sigma^2(n_{Rj}^2 + n_{Ij}^2) + 2(a + br_j)\sigma(n_{Rj}\cos\theta + n_{Ij}\sin\theta)} \\ &= (a + br_j)\sqrt{1 + \frac{2\sigma(n_{Rj}\cos\theta + n_{Ij}\sin\theta)}{a + br_j} + \frac{\sigma^2}{(a + br_j)^2}(n_{Rj}^2 + n_{Ij}^2)} \end{aligned}$$

Note that $n_{Rj}\cos\theta + n_{Ij}\sin\theta$ is nothing but another Gaussian random variable. We denote it as n_j. Also note that n_j and n_k are independent for $j \neq k$. $n_{Rj}^2 + n_{Ij}^2$ is a χ_2^2 random variable. Under the assumptions that $a >> \sigma$ and $\mu = b/a$ is very small, the third term under the square root sign is much smaller than the second one and therefore can be neglected. Then using the binomial expansion

$$\sqrt{1 + x} \approx 1 + \frac{1}{2}x \qquad |x| << 1$$

we arrive at

$$y_j \approx a + br_j + \sigma n_j.$$

Hence, a very common approach to fMRI detection is to use the following Gaussian approximation:

$$\mathbf{y} \approx a\mathbf{1} + b\mathbf{r} + \sigma\mathbf{n} \tag{4}$$

where here $\mathbf{y} = |\mathbf{x}|$, $\mathbf{n} = [n_1 \ \ n_2 \cdots n_N]^T$ is (real) Gaussian distributed as $N(\mathbf{0}, \mathbf{I})$, and with $b = 0$ under H_0 and $b \neq 0$ under H_1. Hence, in this case $\Theta_0 = [a \ \ \sigma^2]$ and $\Theta_1 = [a \ \ b \ \ \sigma^2]$. Bear in mind that this approximation does not accurately model the data in cases in which a/σ is relatively small as we shall see later in some examples.

The GLRT for this problem is based on the following test statistic [23], [24]:

$$t_1(\mathbf{y}) \ = \ (N-1)\frac{\| P_{\mathbf{r}} P_{\mathbf{1}}^{\perp}\mathbf{y} \|^2}{\| P_{\mathbf{r}}^{\perp} P_{\mathbf{1}}^{\perp}\mathbf{y} \|^2} = (N-1)[L_1(\mathbf{y}) - 1] \tag{5}$$

where

$$L_1(\mathbf{y}) = \frac{\| P_{\mathbf{1}}^{\perp}\mathbf{y} \|^2}{\| P_{\mathbf{1r}}^{\perp}\mathbf{y} \|^2}.$$

If $t_1(\mathbf{y}) > \eta_1$, then we decide H_1, otherwise we choose H_0. We call this test the magnitude correlation (MC) test, because the test statistic $t_1(\mathbf{y})$ is proportional to the correlation between the magnitude data $\mathbf{y}$ and the reference signal $\mathbf{r}$.

The test statistic $t_1(\mathbf{y})$, under the assumption that $\mathbf{y}$ is truly Gaussian, is distributed as $F_{1,(N-1)}(\mathrm{SNR})$ [24], where $\mathrm{SNR} = \mu^2 a^2/\sigma^2$. That is SNR is the non-centrality parameter of F-distribution. Unfortunately, the Gaussian approximation (4) is inaccurate when $a/\sigma \leq 3$. In fact, when a/σ is small, we don't know the distribution of MC detector nor whether or not it has CFAR property. So determination of a proper threshold (to obtain a desired

P_f) is theoretically very difficult. Moreover, since the Gaussian approximation is no longer reasonable in this case, one expects the performance of the MC test to suffer. This is indeed the case as we shall see in the next section. How to solve this problem will be explained in next section together with our numerical results.

3.3 Method 2: Complex Correlation Test

Recall the linear model

$$\mathbf{y} = \mathbf{S}\phi_1 + \mu\mathbf{H}\phi_2 + \sigma\mathbf{n}. \tag{6}$$

The GLRT based on this model is well known in the signal processing literature as a matched subspace detector [24], and in fact coincides with the CC test proposed by Lai and Glover [17]. The unknown parameters in this case are $\Theta_0 = [\phi_1^T \;\; \phi_2^T \;\; \sigma^2]$ and $\Theta_1 = [\phi_1^T \;\; \phi_2^T \;\; \mu \;\; \sigma^2]$. The GLRT is given by

$$t_2(\mathbf{y}) = (N-1)[L_2(\mathbf{y}) - 1] = (N-1)\frac{\| P_{\mathbf{H}} P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2}{\| P_{\mathbf{SH}}^{\perp} P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2} = (N-1)\frac{\| P_{\mathbf{H}} P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2}{\| P_{\mathbf{H}}^{\perp} P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2} \tag{7}$$

where

$$L_2(\mathbf{y}) = \frac{\| P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2}{\| \mathbf{y} \|^2 - \| P_{\mathbf{S}}\mathbf{y} \|^2 - \| P_{\mathbf{H}}\mathbf{y} \|^2} = \frac{\mathbf{y}^T P_{\mathbf{S}}^{\perp}\mathbf{y}}{\mathbf{y}^T P_{\mathbf{SH}}^{\perp}\mathbf{y}}. \tag{8}$$

If $t_2(\mathbf{y}) > \eta_2$, then we decide H_1, otherwise choose H_0. This test is called the *complex correlation* (CC) test, because it is equivalent to the test proposed in [17], which is based on the correlation between the reference signal and real and imaginary components of the complex data. The pdf of $t_2(\mathbf{y})$ is non-central $F_{2,2(N-1)}(\text{SNR})$, where again $\text{SNR} = \mu^2 a^2/\sigma^2$, and thus it has the CFAR property (note that the CC test statistic has a standard central $F_{2,2(N-1)}$ distribution under H_0 ($\mu = 0$)). However, the CC test has one drawback. As we pointed out in the previous section, the phase-coupling between the constant and BOLD response components of the data dictates the nonlinear model (2), as opposed to the linear model used to derive the CC test. Therefore, we next propose a novel GLRT based on this more accurate model.

3.4 Method 3: A New GLRT for fMRI

The unknown parameters in model (2) are $\Theta_0 = [\phi^T \quad \sigma^2]$ and $\Theta_1 = [\phi^T \quad \mu \quad \sigma^2]$, under H_0 and H_1, respectively. Recall that the phase coupling introduces a nonlinear coupling between the parameter μ and ϕ under H_1. This nonlinearity makes the MLE's more difficult to compute, but, remarkably, a closed-form expression for the GLRT statistic does exist. Using the GLRT principle and the nonlinear model (2), in Appendix 1 we derive the following test statistic:

$$t_3(\mathbf{y}) = [L_3(\mathbf{y}) - 1](N - 1) \tag{9}$$

where

$$L_3(\mathbf{y}) = \frac{2\| P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2}{\| P_{\mathbf{H}}^{\perp}\mathbf{y} \|^2 + \| P_{\mathbf{S}}^{\perp}\mathbf{y} \|^2 - \sqrt{\| P_{\mathbf{H}}\mathbf{y} \|^4 + \| P_{\mathbf{S}}\mathbf{y} \|^4 + 2\| P_{\mathbf{H}}\mathbf{y} \|^2 \| P_{\mathbf{S}}\mathbf{y} \|^2 \cos 2\varphi}} \tag{10}$$

with

$$\cos\varphi = \frac{(\theta_1, \theta_2)}{\| \theta_1 \| \| \theta_2 \|} = \frac{\mathbf{y}^T \frac{HS^T}{N} \mathbf{y}}{\| P_{\mathbf{H}}\mathbf{y} \| \| P_{\mathbf{S}}\mathbf{y} \|} \tag{11}$$

and θ_1 and θ_2 are two sufficient statistics

$$\theta_1 = \frac{1}{N}\mathbf{H}^T\mathbf{y}, \qquad \theta_2 = \frac{1}{N}\mathbf{S}^T\mathbf{y}. \tag{12}$$

As usual, given a specified threshold η_3, we decide H_1 if $t_3(\mathbf{y}) > \eta_3$, and H_0 otherwise.

An interesting relationship exists between this test and the CC test. Compare (8) with (10). Note that if $\varphi = 0$ in (10) then L_3 and L_2 coincide. It is precisely through the term $\cos 2\varphi$ that the effect of phase coupling comes into play. Furthermore, if the true parameter $\mu \ll 1$ under H_1, which is the case for most fMRI detection problems[2], then we show in Appendix 3 that $L_3(\mathbf{y})$ asymptotically (as N, the length of the time-series, increases) has the same distribution as $L_2(\mathbf{y})$. For example, this means that for long repetitions of an experimental task (large N) the CC test and our new GLRT will have roughly the same performance in terms of detection power at a given false-alarm rate.

Unfortunately, a closed form for the pdf of such a test statistics as $t_3(\mathbf{y})$ is not as easily accessible to us as for the CC test. We can, however, show that the pdf of $t_3(\mathbf{y})$ is a function of μ and a^2/σ^2 alone (see Appendix 2). This is a desirable feature, since in general a test could depend on all the unknown parameters $(a, \mu, \vartheta, \sigma^2)$. Hence, this result shows that the test is only a function of two key variables, instead of four.

[2] The assumption $\mu \equiv \frac{b}{a} \ll 1$ implies that the strength of the response b is much smaller than the baseline intensity a.

Unfortunately though, the dependence of the test on a^2/σ^2 implies that, in general, this test does not have the CFAR property. Hence, selection of a threshold η_3 to achieve a desired P_f is more difficult. However, the relationships between this test and the CC test suggest some possibilities for threshold selection. From (10), the upper bound of $L_3(\mathbf{y})$ is easily seen to be coincident with $L_2(\mathbf{y})$. So, one method of threshold selection is to choose our threshold slightly smaller than that determined for the $F_{2,2(N-1)}$ distribution.

A more accurate method is to determine the exact thresholds via Monte Carlo simulation. Some of the results of our simulations are given in the next section. Here, we summarize the conclusions. Exhaustive Monte Carlo simulation reveals that the GLRT test is also CFAR when $a/\sigma \geq 1$, which is the case for most if not all fMRI experiments. More importantly and more interestingly, *to achieve the desired P_f, the proper threshold of our GLRT detector is almost exactly one half that of the corresponding threshold required for a $F_{1,(N-1)}$ distributed test statistic.* This is not only partially justified theoretically in Appendix 3, but it is also confirmed by extensive Monte Carlo simulation. In mathematical terms, we have that the density of the test statistic T_3 under H_0 is related to the $F_{1,N-1}$ density by the approximation

$$p_{T_3|H_0}(t_3) \approx 2f(2t_3),$$

where f denotes the density of an $F_{1,(N-1)}$ distributed statistic. This implies that a very accurate threshold can be selected using standard $F_{1,(N-1)}$ distribution tables [1]. The factor of 2 on the right-hand-side above is due to the fact that our analysis revealed that the threshold producing a specified false-alarm rate for our new GLRT was almost exactly twice that required for the same rate with an $F_{1,(N-1)}$ statistic.

4 Monte Carlo Analysis of the Tests

4.1 Comparison of the Three Detectors

In order to compare the performances of the three detectors, we have run exhaustive Monte Carlo experiments. Because the GLRT does not generally have the CFAR property, it is necessary to study the performance for different values of a/σ. However, as mentioned above, for $a/\sigma \geq 1$, the Monte Carlo analysis reveals that the GLRT is essentially CFAR.

In Tables 1-3, we compare the detection rates P_d of the three tests under three different false-alarm rate specifications. The false-alarm rates were selected to be representative of those commonly used in fMRI. The most difficult element of the Monte Carlo analysis, except in the CC test case, is the determination of proper thresholds to achieve a desired false-alarm rate with each detector. The CC test is $F_{2,2(N-1)}$ distributed under H_0, and therefore the proper threshold is very easily determined from standard tables [1].

Because the MC test and GLRT are not known to possess the CFAR property, the proper threshold will, in general, depend on a/σ. For a given value of a/σ, the threshold needed to achieve a desired false-alarm rate can be determined via Monte Carlo analysis and trial-and-error over a range of thresholds. This is precisely how the thresholds were determined for the results given in Tables 1-3. Remarkably, however, the Monte Carlo analysis revealed that both the MC test and the GLRT were essentially CFAR so long as $a/\sigma > 1$, which is almost always true in fMRI. Moreover, the Monte Carlo analysis supports the use of some very simple rules for threshold selection.

First, in the case of the MC test, for very large values of a/σ the magnitude data is very well approximated as Gaussian. Therefore, in such situations, the MC test is (approximately) $F_{1,(N-1)}$ distributed under H_0 and the proper threshold can be determined again from standard tables [1]. Because, the Monte Carlo simulations show that for $a/\sigma \geq 1$ the MC test is essentially CFAR, the proper threshold may be determined from $F_{1,(N-1)}$ distribution for all cases in which $a/\sigma \geq 1$. The derivation of equation (4) also proves this, although not strictly.

Second, the similar performances of the GLRT and MC test for large a/σ suggested the possibility of a relationship between the GLRT statistic and the $F_{1,(N-1)}$ statistic. This intuition led to the discovery that the proper threshold for the GLRT can be determined from the $F_{1,(N-1)}$ as well. Specifically, our analysis shows that the proper threshold for the GLRT can be selected as 1/2 of the threshold required to achieve the desired false-alarm rate with a $F_{1,(N-1)}$ distributed statistic. What this implies mathematically is that the "tail" behavior of the GLRT statistic's distribution, which is not known explicitly, must coincide with the tail behavior of the $F_{1,(N-1)}$ distribution.

The results in the three tables show clearly that our GLRT detector performs best for all three (low, medium, high) a/σ cases. The CC detector does perform better than MC detector at low a/σ case. However, as the constant component becomes more and more dominant over the noise, the GLRT and MC test significantly outperform the CC test.

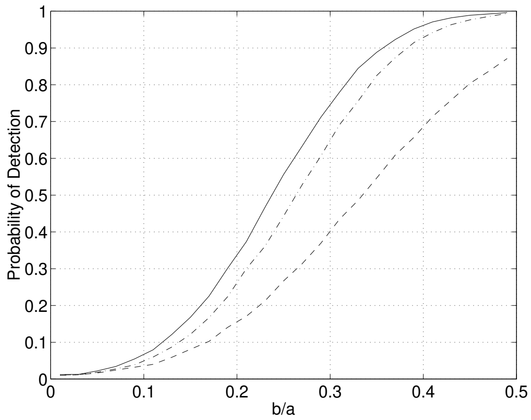
Finally note that the detection rate of the CC test is constant for fixed $\mathrm{SNR} = \mu^2(a/\sigma)^2$.

This is expected because the CC test statistic is non-central $F_{2,2(N-1)}(\text{SNR})$ under H_1. Remarkably, note that the dependence of detection rate of our GLRT detector also depends only on SNR. The same is not true of the MC test, whose performance drops off severely as a/σ decreases.

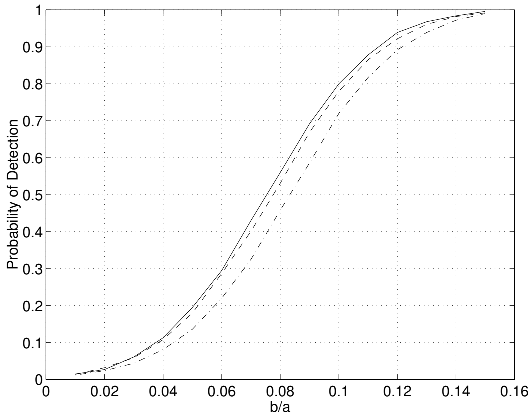
We also illustrate the results graphically in Fig. 2. Here we plot the probability of detection as a function of relative response strength $\mu \equiv \frac{b}{a}$. Fig. 2(a) shows the case with $a/\sigma = 1.0$, which is fairly low, so we expect the performance of the MC detector to be poor, which is indeed the case. Fig. 2(b) shows the case for $a/\sigma = 3.162$, and we see that the performance of the MC detector begins to surpass that of the CC detector but is still inferior to that of our new GLRT. Fig. 2(c) indicates the case for $a/\sigma = 10$, which is quite large, and so the MC detector and GLRT detector have almost the same performance.

4.2 A Simulated fMRI Study

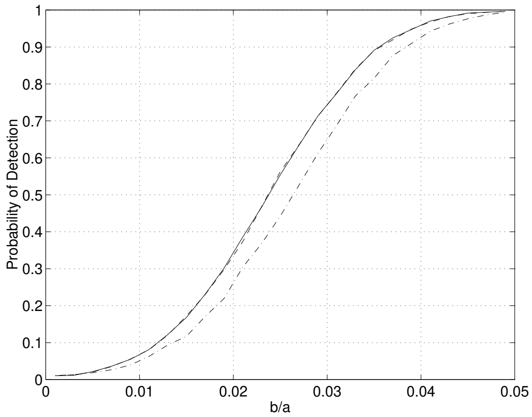
Figure 3 shows one slice image of the brain (64×64 voxels). A 9×9 voxel region in the lower right corner of the brain (indicated in white) was selected to be active in this simulation. For the simulation, a $N = 120$ length time series was simulated for each voxel. The reference signal $\mathbf{r}$ was a square wave with period 10. The fluctuation of reference signal about the constant baseline level was $\pm 10\%$, *i.e.*, $\mu = 0.1$. The noise variance in each time series was set so that $a/\sigma = 3.162$. For each time series, the phase is a constant. But spatially, the phase has a random fluctuation (modeled here as a Gaussian noise with zero mean 0.1 variance) about a constant phase of $\pi/3$. The desired false-alarm rate in this example was chosen to be $P_f = 0.01$, and thus the three thresholds for CC, GLRT and MC



(a)



(b)



(c)

Figure 2. Three performance curves (detection probability P_D v. response strength $\mu = \frac{b}{a}$) with $N = 120, P_f = .01$ while $\mu = b/a$ varies. Therefore the thresholds are chosen as in Table 1. solid $(- - -)$ line for GLRT; dash-dot(-.) line for CC; dashed $(- -)$line for MC. (a) $a/\sigma = 1$. The curve at the top is for GLRT, the middle one for CC, the bottom one for MC;(b) $a/\sigma = 3.162$. The top one is for GLRT, the middle one for MC, the bottom one for CC (c) $a/\sigma = 10$, the GLRT and MC curves coalesce to one in the top while the CC detector remains at the bottom.

detectors are 4.70, 3.43, and 6.85, respectively. The actual false-alarm and detection rates observed in this simulation, given in the caption of Figure 3, are in excellent agreement with the tabulated Monte Carlo results.

Table 1: P_d with $P_f = 0.01, N = 120$

	threshold	4.70	3.43	6.85
a/σ	μ	CC	GLRT	MC
1	.3162	.72	.80	.44
3.162	.1	.72	.80	.78
10	.03162	.72	.80	.80

Table 2: P_d with $P_f = 0.025, N = 120$

	threshold	3.75	2.58	5.15
a/σ	μ	CC	GLRT	MC
1	.3162	.82	.88	.58
3.162	.1	.82	.88	.87
10	.03162	.82	.88	.88

Table 3: P_d with $P_f = 0.05, N = 120$

	threshold	3.03	1.96	3.92
a/σ	μ	CC	GLRT	MC
1	.3162	.88	.93	.69
3.162	.1	.88	.93	.92
10	.03162	.88	.93	.93

5 Discussions and Conclusions

This paper developed a novel GLRT for fMRI in the complex data domain, and compared it to the commonly used MC test and the recently proposed CC test. We derived a very simple, closed-form expression for the GLRT statistic. In fact, all three tests are roughly equal in terms of computational complexity. Although the GLRT statistic does not obey a well known distribution, theoretical analysis established a basic invariance principle for the statistic, and it was shown that the GLRT and the CC test are asymptotically equivalent (as the length fMRI time series increases). Monte Carlo analysis was used to demonstrate that the GLRT performs better than the MC test or CC test overall. Furthermore, the analysis revealed that desired false-alarm rates can be achieved with the GLRT using thresholds determined by well known distribution tables.

There are several avenues for future work within the GLRT framework. First, the noise structure in fMRI is very complicated. For simplicity and the purpose of demonstrating our method and ideas, we assume the noise is white Gaussian. The whiteness assumption doesn't change the problem essentially, since given a *known* time-correlation structure we can always use the Choleksy factorization of the noise covariance to whiten the data [23], producing a model with the same form as that used throughout this paper. Hence, many of our conclusions are easily extended to more realistic noise models that incorporate random fluctuations due to the respiration and cardiac cycle and patient motions [13,18,4], for example. One difficulty that we face, however, is that the noise correlation structure is usually unknown *a priori*, and estimation of the noise covariance is a challenging issue that we are

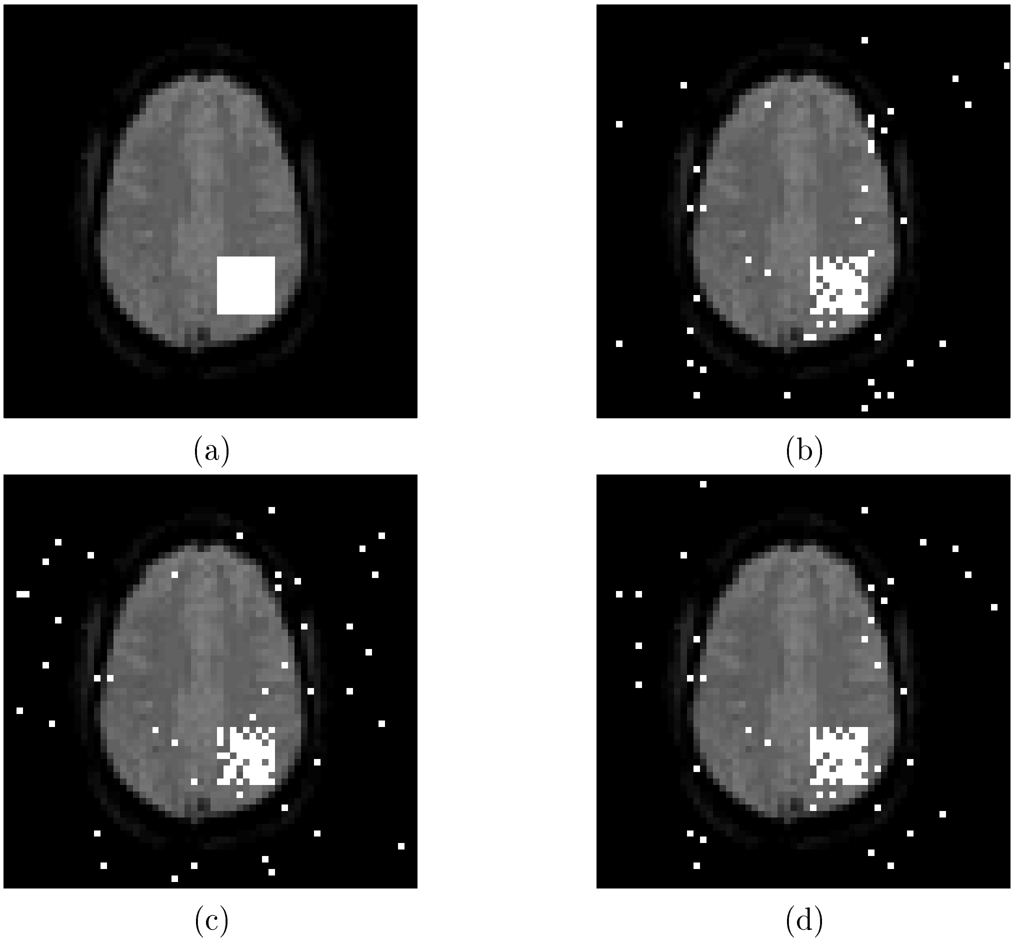


Figure 3. *Simulated fMRI experiment. (a) Brain image with simulated activation region highlighted. The MC test, CC test, and GLRT test are compared in (b)-(d). A threshold was selected for each test to produce a $P_f = 0.01$. (b) MC test results. Detection rate $P_d = 0.77$. (c) CC test results. $P_d = 0.70$. (d) GLRT results. $P_d = 0.79$.*

investigating. Second, more realistic (and necessarily more complicated) signal models can be used in the GLRT framework. For example, multi-parameter models of the reference signal $\mathbf{r}$ could account for uncertainties in the BOLD response. Multi-parameter linear regression models of the response could be used within the GLRT framework to make the test more robust to such uncertainties. However, we caution that more complicated models may or may not improve the performance. It is also possible to deal with unknown delays (different onset latencies in different voxels) by correlating the voxels with shifted versions of the response $\mathbf{r}$ and selecting the shift that produces the maximum correlation. This approach is another instance of a GLRT procedure in which we are effectively computing the MLE of the delay and plugging it into the LRT. We are currently investigating such methods.

Finally, we close with a summary of our conclusions regarding complex data domain fMRI. First, at relatively high baseline signal intensity ($a/\sigma > 3$), the simple MC test, which is very common in practice, performs quite well. Hence, in such regimes there is no compelling reason for testing based on the complex data. This is expected since the magnitude data is approximately Gaussian at high signal intensity, in which case the MC test is nearly optimal. In fact, in most typical fMRI experiments $a/\sigma > 3$ and the MC test is adequate. However, at lower baseline signal intensity the performance of the MC test drops off dramatically, and in such situations complex data tests such as the new GLRT and CC test offer superior performance. Low signal intensity does occur as the spatial and/or temporal resolution of the fMRI study is increased. Most fMRI experiments work with limited resolution in order to avoid the low signal intensity problem. However, high resolution, low signal intensity fMRI may be useful in certain research or clinical paradigms, and in such cases we advocate the GLRT.

Acknowledgment

The authors thank the reviewers for their comments and suggestions.

6 Appendix

6.1 Appendix 1: Derivation of GLRT Statistic

In our derivations, we assume $\mathbf{1}^T\mathbf{r} = 0$ and $\mathbf{r}^T\mathbf{r} = N$, so $\mathbf{S}^T\mathbf{H} = \mathbf{H}^T\mathbf{S} = 0, \mathbf{S}^T\mathbf{S} = \mathbf{H}^T\mathbf{H} = N\mathbf{I}_{2\times 2}$. The second condition can always be satisfied since we can always normalize the reference signal without changing the problem essentially. The first condition is generally not immediately fulfilled, however it is easily achieved by orthogonalization. For example, if $\mathbf{r}$ is not orthogonal to the baseline constant signal $\mathbf{1}$, then we define an orthogonal response signal $\widetilde{\mathbf{r}} \equiv P_{\mathbf{1}}^{\perp}\mathbf{r}$. In fact, this orthogonalization is applied to the data vectors in the test statistics $t_1(\mathbf{y})$, $t_2(\mathbf{x})$, and $t_3(\mathbf{x})$ (Note that in the complex case this corresponds to the projection operator $\mathbf{P}_S^{\perp}\mathbf{r}$). Hence, without loss of generality, we assume $\mathbf{r}$ is orthogonal to $\mathbf{1}$.

Let $\widehat{\sigma}_0$ and $\widehat{\sigma}_1$ denote the MLEs of noise variance under hypotheses H_0 and H_1, respectively. Recall that the GLRT statistic is given by $\frac{p_{H_1}(\mathbf{x};\widehat{\Theta}_1)}{p_{H_0}(\mathbf{x};\widehat{\Theta}_0)}$ It is easy to show that this statistic reduces to

$$L_3(\mathbf{y}) = \min(\widehat{\sigma}_0^2)/\min(\widehat{\sigma}_1^2) \tag{13}$$

Recall that under Gaussian distribution the maximum likelihood estimate is the same as least square estimate. Therefore, calculation of $\min \widehat{\sigma}_0^2$ is straightforward:

$$\min \widehat{\sigma}_0^2 = \| P_{\mathbf{S}}^{\perp} \mathbf{y} \|^2 \tag{14}$$

Determining $\min(\widehat{\sigma}_1^2)$ is much more difficult due to the nonlinear coupling between the two unknowns μ and ϕ. To circumvent this difficulty we first decompose $\mathbf{y}$ and $\mathbf{y} - \mathbf{S}\phi - \mu\mathbf{H}\phi$ into three orthogonal components, *i.e.*,

$$\mathbf{y} = P_{\mathbf{SH}}^{\perp}\mathbf{y} + P_{\mathbf{S}}\mathbf{y} + P_{\mathbf{H}}\mathbf{y}$$

and then

$$\begin{aligned}
\widehat{\sigma}_1^2 &= \| \mathbf{y} - \mathbf{S}\phi - \mu\mathbf{H}\phi \|^2 \\
&= \| P_{\mathbf{SH}}^{\perp}\mathbf{y} + (P_{\mathbf{S}}\mathbf{y} - \mathbf{S}\phi) + (P_{\mathbf{H}}\mathbf{y} - \mu\mathbf{H}\phi) \|^2 \\
&= \| P_{\mathbf{SH}}^{\perp}\mathbf{y} \|^2 + \| P_{\mathbf{S}}\mathbf{y} - \mathbf{S}\phi \|^2 + \| P_{\mathbf{H}}\mathbf{y} - \mu\mathbf{H}\phi \|^2 \\
&= \| P_{\mathbf{SH}}^{\perp}\mathbf{y} \|^2 + \| \mathbf{S}\theta_2 - \mathbf{S}\phi \|^2 + \| \mathbf{H}\theta_1 - \mu\mathbf{H}\phi \|^2 \\
&= \| P_{\mathbf{SH}}^{\perp}\mathbf{y} \|^2 + N[\theta_1^T\theta_1 + \theta_2^T\theta_2 + (1+\mu^2)\phi^T\phi - 2(\theta_2 + \mu\theta_1)^T\phi]
\end{aligned}$$

where θ_1 and θ_2 are given by (12).

Now $\min \widehat{\sigma}_1^2$ is equivalent to $\min J = (1+\mu^2)\phi^T\phi - 2(\theta_2 + \mu\theta_1)^T\phi$.

Setting partial derivatives of J with respect to μ and ϕ to zero:

$$\begin{cases} \frac{\partial J}{\partial \mu} & = & 2\mu\phi^T\phi - 2\theta_1^T\phi & = 0, \\ \frac{\partial J}{\partial \phi} & = & 2(1+\mu^2)\phi - 2(\theta_2 + \mu\theta_1) & = 0. \end{cases}$$

We then get

$$\widehat{\mu} = \frac{\theta_1^T\widehat{\phi}}{\widehat{\phi}^T\widehat{\phi}}, \tag{15}$$

$$\widehat{\phi} = \frac{\theta_2 + \widehat{\mu}\theta_1}{1+\widehat{\mu}^2}. \tag{16}$$

So now

$$\min \widehat{\sigma}_1^2 = \| P_{\mathbf{SH}}^{\perp}\mathbf{y} \|^2 + N(\theta_1^T\theta_1 + \theta_2^T\theta_2) - N(1+\widehat{\mu}^2)\widehat{\phi}^T\widehat{\phi}$$

Furthermore, note that

$$N(\theta_1^T\theta_1 + \theta_2^T\theta_2) = \| P_{\mathbf{H}}\mathbf{y} \|^2 + \| P_{\mathbf{S}}\mathbf{y} \|^2, \quad \| P_{\mathbf{SH}}^{\perp}\mathbf{y} \|^2 + \| P_{\mathbf{S}}\mathbf{y} \|^2 + \| P_{\mathbf{H}}\mathbf{y} \|^2 = \| \mathbf{y} \|^2$$

which gives

$$\min \widehat{\sigma}_1^2 = \| \mathbf{y} \|^2 - N(1+\widehat{\mu}^2)\widehat{\phi}^T\widehat{\phi} = \| \mathbf{y} \|^2 - \frac{N}{1+\widehat{\mu}^2}[\theta_2^T\theta_2 + 2\widehat{\mu}\theta_1^T\theta_2 + \widehat{\mu}^2\theta_1^T\theta_1]$$

Eliminating $\widehat{\phi}$ from (15) and (16) (or setting derivative of the above equation with respect to $\widehat{\mu}$ to zero) shows that $\widehat{\mu}$ must satisfy:

$$\frac{\theta_2^T\theta_2 + 2\widehat{\mu}\theta_1^T\theta_2 + \widehat{\mu}^2\theta_1^T\theta 1}{1+\widehat{\mu}^2} = \frac{\theta_1^T\theta_2 + \widehat{\mu}\theta_1^T\theta_1}{\widehat{\mu}} \tag{17}$$

Using this equation, $\min \widehat{\sigma}_1^2$ can be further simplified:

$$\min \widehat{\sigma}_1^2 = \parallel \mathbf{y} \parallel^2 - N(\theta_1^T\theta_1 + \frac{\theta_1^T\theta_2}{\widehat{\mu}}) \tag{18}$$

This equation is important in our derivation of asymptotic expression of $L_3(\mathbf{y})$ in Appendix 3.

From (17), $\widehat{\mu}$ satisfies quadratic equation $\mu^2 + c\mu - 1 = 0$ with

$$c = \frac{\theta_2^T\theta_2 - \theta_1^T\theta_1}{\theta_1^T\theta_2}$$

Since $\mu_1\mu_2 = -1$, there are two solutions of opposite signs:

$$\widehat{\mu} = -\frac{c}{2} \pm \sqrt{1 + (\frac{c}{2})^2}$$

However, from (18), to make sure $\widehat{\sigma}_1^2$ is minimal, $\widehat{\mu}$ must have the same sign as $\theta_1^T\theta_2$, and so the unique solution for $\widehat{\mu}$ is:

$$\widehat{\mu} = -\frac{c}{2} + \sqrt{1 + (\frac{c}{2})^2}.$$

Substitution of $\widehat{\mu}$ into (18) yields the right solution for min $\widehat{\sigma}_1^2$. Finally from (13), (14) and (18) we get the closed form expression for $L_3(\mathbf{y})$ as given by (10).

Instead of using $L_3(\mathbf{y})$ directly, we use (9) as our test statistic, the main reason is to get a good comparisons between the three different detectors. It will become much clearer when we study the asymptotic property of $t_3(\mathbf{y})$.

6.2 Appendix 2: Invariance of GLRT test statistics

The most difficult part to use invariant theory is to find an appropriate set of transformations which fully exploits the structure of the signal to be detected. Since our problem is nonlinear, finding this set of transformation is not an easy matter. Actually finding and proof come simultaneously.

The following theorem may be regarded as an extension of the results in [24], which is not hard to prove, and therefore we simply state the result.

Theorem: *Consider the model*

$$\mathbf{y} = \mu \mathbf{H}\phi + \mathbf{S}\phi + \sigma \mathbf{n} \tag{19}$$

where $\mathbf{n}$ *is Gaussian distributed as* $N(0, \mathbf{I})$. *The model is invariant to the group of transformations defined by:*

$$G \equiv \{g(\mathbf{y}) : g(\mathbf{y}) = c\mathbf{Q}_{\mathbf{S}}\mathbf{Q}_{\mathbf{H}}\mathbf{y}\} \tag{20}$$

where

$$\begin{cases} \mathbf{Q_S} \equiv \mathbf{U_S} \mathbf{Q} \mathbf{U_S}^T + P_{\mathbf{S}}^{\perp} \equiv \frac{\mathbf{S}}{\sqrt{N}} \mathbf{Q} \frac{\mathbf{S}^T}{\sqrt{N}} + \mathbf{I} - \frac{\mathbf{S}\mathbf{S}^T}{N} \\ \mathbf{Q_H} \equiv \mathbf{U_H} \mathbf{Q} \mathbf{U_H}^T + P_{\mathbf{H}}^{\perp} \equiv \frac{\mathbf{H}}{\sqrt{N}} \mathbf{Q} \frac{\mathbf{H}^T}{\sqrt{N}} + \mathbf{I} - \frac{\mathbf{H}\mathbf{H}^T}{N} \end{cases}$$

c is any constant; $\mathbf{Q}$ is any 2×2 orthogonal matrix; $\mathbf{U_S}$ and $\mathbf{U_H}$ are defined in an obvious way.

The geometrical meaning of this transformation is consecutive rotations of the original signal within the $\mathbf{S}$ plane and then within $\mathbf{H}$ plane followed by scalings that introduces unknown variance.

Under the above transformation, $g(\mathbf{y})$ is explicitly expressed as

$$g(\mathbf{y}) = \mu_1 \mathbf{H} \phi_1 + \mathbf{S} \phi_1 + \sigma_1 \mathbf{n} \tag{21}$$

with the induced transformation $\bar{G}$ given by:

$$\begin{cases} \mu_1 &= \mu \\ \phi_1 &= c\mathbf{Q}\phi \\ \sigma_1 &= c\sigma \end{cases}$$

It follows that $(\mu, \| \phi \|^2 / \sigma^2)$ or more simply $(\mu, a^2/\sigma^2)$ is a set of maximal invariant parameters under $\bar{G}$ (see [6], [21], [23]) for details on invariance principles). And it is easy to verify that $t_3(\mathbf{y})$ is invariant to the transformation group G. Therefore, from standard invariance theory (again see [6], [21], [23]), the pdf of $t_3(\mathbf{y})$ is a function of μ and a/σ alone (instead of all four model parameters a, b, ϑ, and σ).

6.3 Appendix 3: Asymptotics of GLRT test statistics

In detection problem, our concern is low SNR case. In our situation, we assume that the true parameter μ under H_1 is small, i.e., $\mu \to 0$. By the asymptotic property of maximum likelihood estimates, as $N \to \infty$, $\widehat{\mu} \to \mu$. In order to get a more accurate approximation of $\widehat{\mu}$, we use (16) combined with $\widehat{\mu} \to \mu \to 0$ (as $N \to \infty$) and get $\widehat{\phi} \to \theta_2$, so from (15) (as $N \to \infty$),

$$\widehat{\mu} \approx \frac{\theta_1^T \theta_2}{\theta_2^T \theta_2}, \tag{22}$$

which is the maximum likelihood estimate for the corresponding linear model (3).

Substituting the above equation into (18) we have

$$\min \widehat{\sigma}_1^2 = \parallel \mathbf{y} \parallel^2 - N(\theta_1^T \theta_1 + \theta_2^T \theta_2) = \parallel \mathbf{y} \parallel^2 - \parallel P_{\mathbf{S}} \mathbf{y} \parallel^2 - \parallel P_{\mathbf{H}} \mathbf{y} \parallel^2 \tag{23}$$

Noting that

$$\begin{aligned} P_{\mathbf{SH}}^{\perp} &= P_{\mathbf{S}}^{\perp} P_{\mathbf{SH}}^{\perp} P_{\mathbf{S}}^{\perp}, & (24) \\ P_{\mathbf{S}}^{\perp} - P_{\mathbf{SH}}^{\perp} = P_{\mathbf{S}}^{\perp} P_{\mathbf{H}} P_{\mathbf{S}}^{\perp} &= P_{\mathbf{H}} = P_{\mathbf{H}} P_{\mathbf{S}}^{\perp}, & (25) \\ P_{\mathbf{SH}}^{\perp} P_{\mathbf{S}}^{\perp} = P_{\mathbf{H}}^{\perp} P_{\mathbf{S}}^{\perp} &= P_{\mathbf{SH}}^{\perp} = \mathbf{I} - P_{\mathbf{S}} - P_{\mathbf{H}}. & (26) \end{aligned}$$

we get, from equations (13), (14), (23), as $N \to \infty$,

$$L_3(\mathbf{y}) \approx L_2(\mathbf{y})$$

This implies that $t_3(\mathbf{y})$ asymptotically has the same distribution as $t_2(\mathbf{y})$, *i.e.*, non-central $F_{2,2(N-1)}(\text{SNR})$, and thus asymptotically has CFAR property.

References

[1] *Standard Mathematical Tables and Formulae.* 29th Edition, CRC Press, 1991.

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

[3] R.T. Behrens and L. L. Scharf. Signal rocessing applications of obligue projection operators. *IEEE Trans. Signal Processing*, 42(6):1413–1424, 1994.

[4] J.W. Belliveau, D.N. Kennedy, B.R. Buchbinder R.C. McKinstry, R.M. Weisskoff, M. S. Cohen, J. M. Vevea, T.J. Brady, and B. R. Rosen. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, 254(5032):716–719, 11 1991.

[5] M. A. Berstein, 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.

[6] S. Bose and A. O. Steinhardt. A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices. *IEEE Trans. Signal Processing*, 43(9):2164–2175, 1995.

[7] R. T. Constable and R. M. Henkelman. Why MEM does not work in MR image reconstruction. *Magn. Reson. Med.*, 14:12–25, 1990.

[8] W. A. Edelstein, P. A. Bottomley, and L. M. Pfeifer. A signal-to-noise calibration procedure for NMR imaging systems. *Med. Phys.*, 11:180–185, 1984.

[9] W. A. Edelstein, G. Glover, C. Hardy, and R. Redington. The intrinsic signal-to-noise ratio in NMR imaging. *Magn. Reson. Med.*, 3:604–618, 1986.

[10] L. R. Frank, R. B. Buxton, and E. C. Wong. Probabilistic analysis and functional magnetic resonance imaging data. *Magn. Reson. Med.*, 39:132–148, 1998.

[11] C. R. Genovese. A time-course model for fMRI data. In *Proc. Intl. Symp. Magn. Reson. Med. Meeting*, page 1669, 1997.

[12] C. R. Genovese, D. C. Noll, and W. F. Eddy. Estimating test-retest reliability of functional MR imaging i:statistical methodology. *Magn. Reson. Med.*, 38:497–507, 1997.

[13] Xiaoping Hu, Tuong Huu Le, Todd Parrish, and Peter Erhard. Retrospective estimation and correlation of physiological fluctuation in functional MRI. *Magn. Reson. Med.*, 34:201–212, 1995.

[14] S. M. Kay. *Fundamentals of Statistical Signal Processing. Detection Theory.* Prentice-Hall, New Jersey, 1998.

[15] S. Kim, W. Richter, and K. Ugurbil. Limitations of temporal resolution in functional MRI. *Magn. Reson. Med.*, 37:631–636, 1997.

[16] T. Kim, L. Al-Dayeh, P. Patel, and M. Singh. Bayesian processing for fMRI. In *Proc. Intl. Soc. Magn. Reson. Med.*, Sidney, Australia, 1998.

[17] S. Lai and G. H. Glover. Detection of BOLD fMRI signals using complex data. In *Proc. Intl. Soc. Magn. Reson. Med. Meeting*, page 1671, 1997.

[18] Tuong Huu Le and Xiaoping Hu. Retrospective estimation and correlation of physiological artifacts in fMRI by direct extraction of physiological activity from MR data. *Magn. Reson. Med.*, 35:290–298, 1996.

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

[20] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill. Noise and filtration in magnetic resonance imaging. *Med. Phys.*, 12:586–591, 1985.

[21] J. Muirhead. *Aspects of Multivariate Statistical Theory.* Wiley, New York, 1982.

[22] U. E. Ruttimann, N. F. Ramsey, D. W. Hommer, P. Thevanaz, Ch. Lee, and M. Unser. Analysis of functional magnetic resonance images by wavelet decomposition. In *Proc. Intl. Conf. Image Proc.*, volume I, pages 633–636, 1995.

[23] L. L. Scharf. *Statistical Signal Processing. Detection, Estimation, an Time Series Analysis.* Addison-Wesley, Reading, MA, 1991.

[24] L. L. Scharf and B. Friedlander. Matched subspace detectors. *IEEE Trans. Signal Processing*, 42(8):2146–2157, 1994.

[25] K. Sekihara and H. Koizumi. Detecting cortical activities from fMRI time-course data using MUSIC algorithm with forward and backward covariance averaging. *Magn. Reson. Med.*, 35:907–813, 1996.

[26] B. Siewert, B. M. Bly, G. Schlaug, D. G. Darby, V. Thangaraj, S. Warach, and R. Edelman. Comparison of the bold- and epistar-technique for functional brain imaging using signal detection theory. *Magn. Reson. Med.*, 36:249–255, 1996.

[27] S. C. Strother, I. Kanno, and D. A. Rottenberg. Principle component analysis, variance partioning, and functional connectivity. *Journal of Cerebral Blood Flow and Metabolism*, 15:353–360, 1995.

[28] A. Villringer and U. Dirnagl. Coupling of brain activity and cerebral blood flow: Basis of functional neuroimaging. *Cerebrovascular and Brain Metabolism Reviews*, 7:240–276, 1995.