

## On the Probability Density Function of the Test Statistic for One Nonlinear GLR Detector Arising from fMRI

Fangyuan Nan

School of Electronic Information, Wuhan University, Wuhan, China, 430079

Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, BC, Canada

---

**Abstract:** Recently an important and interesting nonlinear generalized likelihood ratio (GLR) detector emerged in functional magnetic resonance imaging (fMRI) data processing. However, the study of that detector is incomplete: the probability density function (pdf) of the test statistic was drawn from numerical simulations without much theoretical support and is therefore, not firmly grounded. This correspondence presents more accurate (asymptotic) closed form of the pdf by resorting to a non-central Wishart matrix and by asymptotic expansion of some integrals. It is then confirmed theoretically that the detector does possess constant false alarm rate (CFAR) property under some practical regimes of signal to noise ratio (SNR) for finite samples and the correct threshold selection method is given, which is very important for real fMRI data processing.

**Key words:** Generalized likelihood ratio test, non-central Wishart matrix, eigenvalue, probability density function, asymptotic analysis, asymptotic expansion, fMRI

---

### INTRODUCTION

In our research on voxelwise detection of fMRI, we encountered the following model and problem<sup>[1]</sup>. Let  $x$  denote an  $N \times 1$  complex random vector:

$$x = (a_1 + br)e^{i\theta} + \sigma n_c = (1 + \mu r)ae^{i\theta} + \sigma n_c, \text{ with } \mu \doteq b/a.$$

The observation data  $x$  represents a complex time series at one voxel and its consists of three complex components. The first component  $a_1$  is a constant (DC) baseline component, where  $1$  denotes an  $N \times 1$  vector of 1's. DC component is always present and typically  $a/\sigma \geq 1$ . The second component  $br$  is a real oscillating signal. The vector  $r$  is a reference function that models the expected response characteristic of activation pattern. We assume  $r$  to be known throughout the study. The third component of additive complex Gaussian white noise  $\sigma n_c$  models errors primarily due to thermal noises in the patient. The term  $n_c$  denotes a standard complex Gaussian vector (with mean zero and covariance matrix  $I^{N \times N}$ ). In general,  $a$ ,  $b$ ,  $\theta$  and  $\sigma^2$  are unknown scalars and are different for each voxel time-series. The detection is with regard to  $b$ : under  $H_0$ ,  $b=0$  indicates there is no brain activity in the corresponding voxel; under  $H_1$ ,  $b \neq 0$  indicates the presence of activity in the voxel. We have compared this model to actual

fMRI time-series and found that our assumptions are in good agreement with actual data<sup>[1]</sup>.

As in<sup>[1]</sup>, we express the complex model as a  $2N \times 1$  dimensional real-valued model:  $y = S\phi + \mu H\phi + \sigma n$ , where

$$y \doteq \begin{bmatrix} x_R \\ x_I \end{bmatrix}, S \doteq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, H \doteq \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}, \phi \doteq \begin{bmatrix} a \cos \theta \\ a \sin \theta \end{bmatrix}, n \doteq \begin{bmatrix} n_{cR} \\ n_{cI} \end{bmatrix} \quad (1)$$

The subscripts R and I denote real and imaginary parts, respectively.

The detector was systematically studied in<sup>[1]</sup>. Since its publication, <sup>[1]</sup> has received many citations, including but not limited to<sup>[2,3]</sup>. In particular, our model and methods were also validated and generalized by some other work<sup>[2]</sup>. Readers are strongly suggested to read<sup>[1]</sup>, because many practical issues (like the orthogonality of  $1$  and  $r$  to be mentioned and numerical results- tables and curves illustrating our detector's performance) are contained in<sup>[1]</sup>. While the methodology and main conclusion of<sup>[1]</sup> are still valid, it is defective in one aspect: the closed form of the test statistic's pdf was achieved by numerical simulations (for only one fixed sample length  $N=120$ ) and thus not firmly grounded. Because of the fundamental importance of threshold selection in detection problems and the above-mentioned importance of<sup>[1]</sup> in fMRI data processing, it is worthwhile to do some rigorous theoretical work to fix the defect.

---

**Corresponding Author:** Fangyuan Nan, Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, BC, Canada, Tel: 604-708-1126

Due to inherent difficulty in dealing with nonlinear problems, we first attack a simpler problem with known noise variance. In the second part, we attack the more complex problem with unknown noise variance. The second problem is more realistic and is our ultimate goal. However, as we will see, the study on the first part provides a lot of insight and serves as a very useful basis for the second part.

Before specializing in individual case, we explicitly write the pdf of observation data  $y$ , which will be used in both cases. It is:

$$\begin{aligned} p_y(y, \Theta) &= (2\pi\sigma^2)^{-\frac{2N}{2}} \exp[-\|y - S\phi - \mu H\phi\|^2 / (2\sigma^2)] \\ &= (2\pi\sigma^2)^{-N} \exp\left[-\frac{y^T y - 2y^T(S\phi + \mu H\phi)}{2\sigma^2}\right] \exp\left[-\frac{(S\phi + \mu H\phi)^T(S\phi + \mu H\phi)}{2\sigma^2}\right] \\ &= \exp\left[\frac{y^T y - 2y^T S\phi - 2y^T \mu H\phi}{2\sigma^2}\right] (2\pi^2)^{-N} \exp\left[-\frac{(S\phi + \mu H\phi)^T(S\phi + \mu H\phi)}{2\sigma^2}\right] \end{aligned} \quad (2)$$

with  $\Theta$  standing for unknown parameters.

We further make some definitions and matrix decompositions, for subsequent use.

$$U_1 \doteq \frac{H}{\sqrt{N}}, U_2 \doteq \frac{S}{\sqrt{N}}, U_{12} \doteq [U_2 \ U_1] P_{[SH]} = U_{12} U_{12}^T, \quad (3)$$

$$P_{[SH]}^\perp \doteq U_3 U_3^T, \ U \doteq [U_{12} \ U_3]$$

$$\theta_1^{2 \times 1} \doteq \frac{U_1^T y}{\sqrt{N}} \sim N(\mu\phi, \frac{\sigma^2}{N}), \ \theta_2^{2 \times 1} \doteq \frac{U_2^T y}{\sqrt{N}} \sim N(\phi, \frac{\sigma^2}{N}) \quad (4)$$

$$\theta_3^{(2N-4) \times 1} \doteq \frac{U_3^T y}{\sqrt{N}} \sim N(0, \frac{\sigma^2}{N})$$

The decomposition of two projection matrices  $P_{[SH]}$  and  $P_{[SH]}^\perp$  follows from this fact: suppose  $P^{n \times n}$  is a projection matrix with rank  $r$ . Then there exists  $U_p \in Q^{n \times r}$ , i.e.,  $U_p^T U_p = I^{r \times r}$ , such that  $P^{n \times n} = U_p U_p^T$ . Matrix analysis theory<sup>[4]</sup> can be used to justify this fact. Note  $U_1^{2N \times 2}, U_2^{2N \times 2}, U_{12}^{2N \times 4}, U_3^{2N \times (2N-4)}$  are all quasi-orthogonal matrices (refer to nomenclature), while  $U^{2N \times 2N}$  is an orthogonal matrix. One should notice that orthogonality between them, that is  $U_i^T U_j = 0, i \neq j$ .

Put  $\theta = [\theta_1^T \ \theta_2^T \ \theta_3^T]^T$ , it follows that  $\theta = \frac{U^T y}{\sqrt{N}}$ .

Therefore, the covariance matrix of  $\theta$  is  $\text{cov}(\theta, \theta) = \frac{U^T}{\sqrt{N}} \sigma^2 I \frac{U}{\sqrt{N}} = \frac{\sigma^2}{N} I$ . In other words,  $\theta_1, \theta_2, \theta_3$  are independent, jointly Gaussian.

Following<sup>[1]</sup>, we assume  $I^T r = 0$  and  $r^T r = N$ . We thus have some preliminary results which will be used very often, like  $S^T H = H^T S = 0, S^T S = H^T H = NI^{2 \times 2}$ ,

$$\|P_{[SH]}^\perp\|^2 + \|P_{[SH]}\|^2 + \|P_{[Hy]}\|^2 = \|y\|^2$$

and  $N(\theta_1^T \theta_1 + \theta_2^T \theta_2) + \|P_{[Hy]}\|^2 + \|P_{[Sy]}\|^2$ , etc. where  $\theta_1, \theta_2, \theta_3$  are defined in equation (4).

**Results on known noise variance**

**GLR test statistic:** Recall GLR test statistic is given by:

$$\begin{aligned} l_1(y) &\doteq \frac{\max_{\Theta_1} p_y(y, \Theta_1)}{\max_{\Theta_0} p_y(y, \Theta_0)} = \frac{\max_{(\mu, \phi)} \exp\left[-\frac{\|y - S\phi - \mu H\phi\|^2}{2\sigma^2}\right]}{\max_{\phi} \exp\left[-\frac{\|y - S\phi\|^2}{2\sigma^2}\right]} \\ &= \frac{\exp\left[-\frac{\min_{\mu, \phi} \|y - S\phi - \mu H\phi\|^2}{2\sigma^2}\right]}{\exp\left[-\frac{\|P_{[S]}^\perp y\|^2}{2\sigma^2}\right]} \end{aligned} \quad (5)$$

Where  $p_y(y, \Theta_1)$  is the pdf of observation  $y$  under  $H_i, i=1,0$  with unknown parameters  $\Theta_i$ . Everything is straightforward except for the minimization in (5). Following the spirit in Appendix I of<sup>[11]</sup> to achieve the optimization, we have the result:

$$l_2 y \doteq 2 \lg l_1(y) = \frac{N}{2\sigma^2} \left[ \frac{\|\theta_1\|^2 - \|\theta_2\|^2 + \sqrt{(\|\theta_1\|^2 - \|\theta_2\|^2)^2 + 4(\theta_1, \theta_2)^2}}{2} \right] \quad (6)$$

**Probability density function of test statistic:** Direct calculation of the pdf of test statistics  $l_2(y)$  from equation (6) is formidable. Instead we simplify it to a form in terms of normalized quantities.

$$l_2(y) = \frac{1}{2} \left[ \frac{\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2}{\sqrt{(\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2)^2 + 4(\theta_{1n}, \theta_{2n})^2}} \right] \quad (7)$$

Where we have introduced the following normalized random vectors to facilitate computation

$$\begin{aligned} \theta_{1n} &= \frac{\sqrt{N}}{\sigma} \theta_1 \sim N\left(\frac{\sqrt{N}}{\sigma} \mu\phi, I\right), \ \theta_{2n} \\ &= \frac{\sqrt{N}}{\sigma} \theta_2 \sim N\left(\frac{\sqrt{N}}{\sigma} \phi, I\right) \\ \theta_{3n} &= \frac{\sqrt{N}}{\sigma} \theta_3 \sim N(0, I^{2N-4}) \end{aligned} \quad (8)$$

Also, although  $\theta_{3n}$  is not immediately used here, for sake of comparison and unity they are prepared for use later together with  $\theta_{1n}$  and  $\theta_{2n}$ . The subscript  $n$  means normalized. Also notice the independence between  $\theta_{1n}, \theta_{2n}$  and  $\theta_{3n}$ .

**Introduction of a non-central Wishart matrix:** To compute the pdf of  $I_2(\mathbf{y})$ , we further define a matrix  $Z \doteq [\theta_{2n} \ \theta_{1n}] \sim N(\mathbf{M}, I^{2 \times 2} \otimes I^{2 \times 2})$  with

$$\mathbf{M} = \begin{bmatrix} \frac{\sqrt{N}}{\sigma} \phi & \frac{\sqrt{N}}{\sigma} \mu \phi \\ \frac{\sqrt{N}}{\sigma} \mu \phi & \frac{\sqrt{N}}{\sigma} \phi \end{bmatrix} = \frac{\sqrt{N}}{\sigma} \phi [1 \ \mu] \text{ and another matrix}$$

$$\mathbf{A} \doteq Z^T Z = \begin{bmatrix} \|\theta_{2n}\|^2 & (\theta_{2n}, \theta_{1n}) \\ (\theta_{1n}, \theta_{2n}) & \|\theta_{1n}\|^2 \end{bmatrix} \doteq \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \sim W_2(2, I^{2 \times 2}, \omega)$$

Where  $W_2(2, I^{2 \times 2}, \omega)$  denotes a non-central Wishart distribution<sup>[5]</sup> with noncentrally matrix  $\omega$ ,

$$\omega = I^{2 \times 2} \left( \frac{\sqrt{N}}{\sigma} \right)^2 \begin{bmatrix} 1 \\ \mu \end{bmatrix} \phi^T \phi [1 \ \mu] = \frac{Na^2}{\sigma^2} \begin{bmatrix} 1 & \mu \\ \mu & \mu^2 \end{bmatrix}.$$

The pdf of matrix  $\mathbf{A}$  is explicitly given by<sup>[5]</sup>

$$f_{a_{11}, a_{12}, a_{22}}(a_{11}, a_{12}, a_{22}) = \frac{\exp\left[-\frac{Na^2}{2\sigma^2}(1 + \mu^2)\right]}{4\sqrt{\det(\mathbf{A})}} \quad (9)$$

$$\exp\left[-\frac{1}{2}\text{tr}(\mathbf{A})\right] {}_0F_1\left(1; \frac{Na^2}{4\sigma^2} \begin{bmatrix} 1 & \mu \\ \mu & \mu^2 \end{bmatrix} \mathbf{A}\right)$$

Partially, this is good news: in theory, equation (9) provides a basis for all other pdf's in question. However, because of the very cumbersome hypergeometric function, equation (9) is still very intractable, particularly under  $H_1$ . The difficulty of our problem lies in the noncentral Wishart distribution, about which little is known. Derivation of the pdf of  $I_2(\mathbf{y})$  from equation (7) is still difficult. To bypass the difficulty, we examine the two eigenvalues  $\lambda_1 > \lambda_2 > 0$  of matrix  $\mathbf{A}$ .

$$\begin{cases} \lambda_1 + \lambda_2 = a_{11} + a_{22} = \|\theta_{1n}\|^2 + \|\theta_{2n}\|^2; \\ \lambda_1 \lambda_2 = \det(\mathbf{A}) = a_{11} a_{22} - a_{12}^2 \\ \lambda_1 - \lambda_2 = \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2} \\ \quad = \sqrt{(\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2)^2 + 4(\theta_{1n}, \theta_{2n})^2}. \end{cases} \quad (10)$$

Using these relationships, the test statistic  $I_2(\mathbf{y})$  is greatly simplified to

$$I_2(\mathbf{y}) = \frac{1}{2}[(\lambda_1 + \lambda_2) + (\lambda_1 - \lambda_2) - 2\|\theta_{2n}\|^2] = \lambda_1 - a_{11}. \quad (11)$$

The joint pdf of  $\lambda_1, \lambda_2$  and  $a_{11}$  of matrix  $\mathbf{A}$  under  $H_0$ , is critically useful both in this section and in next section.

**Joint pdf of two eigenvalues and  $a_{11}$  of matrix  $\mathbf{A}$  under  $H_0$ :** Under  $H_0, \mu=0$ , the pdf of noncentral Wishart matrix  $\mathbf{A}$  in equation (9) is simplified to<sup>[6]</sup>

$$f_{a_{11}, a_{12}, a_{22}}(a_{11}, a_{12}, a_{22}) = \frac{1}{4\pi} \frac{\exp(-\frac{1}{2}\delta) \exp(-\frac{a_{11} + a_{22}}{2})} {\sqrt{a_{11} a_{22} - a_{12}^2}} {}_0F_1\left(1; \frac{1}{4}\delta a_{11}\right) \quad (12)$$

with  $\delta = Na^2 / \sigma^2$ . Make a transformation from  $(a_{11}, a_{12}, a_{22})$  to  $(a_{11}, \lambda_1, \lambda_2)$ , we get that is.

$$\begin{cases} a_{11} = a_{11} \\ \lambda_1 = \frac{1}{2} \left[ a_{11} + a_{22} + \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2} \right] \\ \lambda_2 = \frac{1}{2} \left[ a_{11} + a_{22} - \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2} \right] \end{cases}$$

$$\Rightarrow \begin{cases} a_{11} = a_{11}, \\ a_{12} = \pm \sqrt{a_{11}(-a_{11} + \lambda_1 + \lambda_2) - \lambda_1 \lambda_2} \\ a_{22} = -a_{11} + \lambda_1 + \lambda_2. \end{cases}$$

Evaluating the Jacobin of the inverse transform (from now on, many intermediate results and operations are omitted).

$$P_{a_{11}, \lambda_1, \lambda_2}(a_{11}, \lambda_1, \lambda_2) = \sum f_{a_{11}, a_{12}, a_{22}}(a_{11}, a_{12}, a_{22}) \left| \frac{\partial(a_{11}, a_{12}, a_{22})}{\partial(a_{11}, \lambda_1, \lambda_2)} \right|$$

$$= \frac{1}{4\pi} \frac{\exp(-\frac{1}{2}\delta) \exp(-\frac{\lambda_1 + \lambda_2}{2})} {\sqrt{\lambda_1 \lambda_2}} \quad (13)$$

$${}_0F_1\left(1; \frac{1}{4}\delta a_{11}\right) \frac{\lambda_1 - \lambda_2} {\sqrt{-a_{11}^2 + (\lambda_1 + \lambda_2)a_{11} - \lambda_1 \lambda_2}}$$

There are two terms in the above  $\sum$  operation.

**Closed form of pdf of  $I_2(\mathbf{y})$  at finite samples:** One can perform the standard procedure to calculate the pdf of  $I_2(\mathbf{y})$  under  $H_0$  and the exact closed form turns out to be still quite complicated and is omitted here. It shows that, for finite samples and very small signal to noise ratio, our detector generally doesn't have the property of constant false alarm rate (CFAR). Hence, selection of proper threshold to satisfy a prescribed probability of false alarm  $P_f$  is impossible in general, because it is dependent on unknown parameters.

To proceed, let us take a short-cut from another perspective. It is shown in<sup>[7]</sup> (asymptotic distribution of GLR test statistics) that  $I_2(\mathbf{y}) \sim \chi_1^2$ , for large samples ( $N \rightarrow \infty$ ). Here asymptotic analysis is a statistical term, referring to the behavior of some quantity for large data samples. However, large samples assumption

is not feasible in fMRI problem, due to the temporal resolution limitation in fMRI<sup>[8]</sup>. How can we then design an implementable detector for finite samples? To get around this difficulty, we combine result from previous subsection and asymptotic result. We notice one feature, that is, the pdf of  $l_2(y)$  (under  $H_0$ ) depends on unknown parameters through  $\delta = \frac{Na^2}{\sigma^2}$ . For this

parameter, the contributions from  $N$  and from  $(a/\sigma)^2$  are equivalent. Therefore, with fixed finite samples but at some reasonably large signal to noise ratio  $SNR \doteq \delta$ , it must also have CFAR property.

In practice, we can do numerical simulations to probe some SNR region in which the detector is practically, locally with  $l_2(y)|H_0 \sim \chi_1^2$ . And if that SNR region is commensurate with real application settings, our detector is implement able and is CFAR in that region and we should be content. Invariance property can save us a lot of numerical work in this aspect. Since<sup>[1]</sup> already elaborated the numerical results for the case of unknown noise variance, which is our ultimate goal, we omit corresponding work for this similar part.

**Results on unknown noise variance**

**GLR test statistic:** The GLRT statistic is given by 
$$l_3(y) = \frac{p_y|H_1(y; \hat{\Theta}_1)}{p_y|H_0(y; \hat{\Theta}_0)} = \frac{(2\pi \min(\hat{\sigma}_1^2))^{-N}}{(2\pi \min(\hat{\sigma}_0^2))^{-N}} = \left( \frac{\min \hat{\sigma}_0^2}{\min \hat{\sigma}_1^2} \right)^N$$

Naturally we consider another monotonous function of  $l_3(y), l_4(y) \doteq (l_3)^{\frac{1}{N}} = \frac{\min \hat{\sigma}_0^2}{\min \hat{\sigma}_1^2}$ .

Following<sup>[1]</sup>, instead of using  $l_4(y)$  directly, we use equation (14) as our test statistic.

$$t_3(y) = [l_4(y) - 1](N - 1) = \frac{(N - 1)[\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2 + \sqrt{(\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2)^2 + 4(\theta_{1n}, \theta_{2n})^2}]}{\|\theta_{1n}\|^2 + \|\theta_{2n}\|^2 - \sqrt{(\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2)^2 + 4(\theta_{1n}, \theta_{2n})^2} + 2\|\theta_{3n}\|^2} \quad (14)$$

The detailed derivation was already published in<sup>[1]</sup>.

Similar to previous section, we further use the normalized quantities in lieu of their counterparts, the very complicated form is then drastically simplified into one in terms of eigenvalues (when the noise of variance  $\sigma^2$  is known, no one would doubt feasibility of introduction of the normalized quantities, since they are not only beneficial to mathematical computation, they are also physically implementable. In present case of unknown noise variance, the normalized quantities

are not implementable, introduction of them is purely to benefit mathematical calculation).

$$t_3(y) = \frac{(N - 1)[\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2 + \sqrt{(\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2)^2 + 4(\theta_{1n}, \theta_{2n})^2}]}{\|\theta_{1n}\|^2 + \|\theta_{2n}\|^2 - \sqrt{(\|\theta_{1n}\|^2 - \|\theta_{2n}\|^2)^2 + 4(\theta_{1n}, \theta_{2n})^2} + 2\|\theta_{3n}\|^2} \quad (15) = \frac{(N - 1)[\lambda_1 - \|\theta_{3n}\|^2]}{\lambda_2 + \|\theta_{3n}\|^2} \doteq (N - 1) \frac{l_2}{\lambda_2 + a_{33}}$$

**Probability density function of the test statistic**

**Theorem:** Regarding the statistic defined in equation (15), under  $H_0$  and asymptotically (as  $\delta \rightarrow \infty$ ):  $\lambda_1 \gg \lambda_2, \lambda_1$  and  $a_{11}$  are

mutually independent with  $\lambda_2; l_2 \sim \chi_1^2, \lambda_2 \sim \chi_1^2$ .

Combined with the trivial result that  $a_{33} \doteq \|\theta_{3n}\|^2 \sim \chi_{2N-4}^2$  is independent of the other three,

$$t_3 \Big| H_0 \sim \frac{N - 1}{2N - 3} F_{1, 2N - 3}.$$

**Proof step 1:** The marginal (asymptotic) pdf of  $l_2$  in the numerator is already derived in previous section. We first find the (asymptotic) joint distribution of  $\lambda_1$  and  $\lambda_2$  from equation (13).

$$f_{\lambda_1, \lambda_2}(\lambda_1, \lambda_2) = \frac{1}{4\pi} \exp\left(-\frac{\delta}{2}\right) \exp\left(-\frac{\lambda_1 + \lambda_2}{2}\right) (\lambda_1 - \lambda_2) \sqrt{\lambda_1 \lambda_2} \int_{\lambda_2}^{\lambda_1} \frac{{}_0F_1\left(1; \frac{1}{4} \delta a_{11}\right)}{\sqrt{(\lambda_1 - a_{11})(a_{11} - \lambda_2)}} da_{11} \quad (16)$$

with

$${}_0F_1\left(1; \frac{1}{4} \delta a_{11}\right) = \frac{1}{\pi} \int_0^\pi \exp(\sqrt{\delta a_{11}} \cos \theta) d\theta = I_0(\sqrt{\delta a_{11}}).$$

We first cope with the integral in the above equation. By a series of change of variables,

$$I_1(\lambda_1, \lambda_2) \doteq \int_0^\pi d\theta \int_{\lambda_2}^{\lambda_1} \frac{\exp(\sqrt{\delta a_{11}} \cos \theta)}{\pi \sqrt{(\lambda_1 - a_{11})(a_{11} - \lambda_2)}} da_{11} = \frac{1}{\pi} \int_0^\pi d\theta \int_{\frac{\lambda_1 - \lambda_2}{2}}^{\frac{\lambda_1 + \lambda_2}{2}} \frac{\exp(\sqrt{\delta} \cos \theta \sqrt{x + \frac{\lambda_1 + \lambda_2}{2}})}{\sqrt{(\lambda_1 - \lambda_2 + x)(\lambda_1 - \lambda_2 - x)}} dx \quad x = a_{11} - \frac{\lambda_1 + \lambda_2}{2} = \frac{1}{\pi} \int_0^\pi d\theta \int_{-1}^1 \frac{\exp\left[\sqrt{\delta} \cos \theta \sqrt{\frac{\lambda_1 - \lambda_2}{2} y + \frac{\lambda_1 + \lambda_2}{2}}\right]}{\sqrt{(1 + y)(1 - y)}} dy \quad y = \frac{x}{\frac{\lambda_1 - \lambda_2}{2}} = \frac{1}{\pi} \int_0^\pi d\theta \int_{-\frac{x}{2}}^{\frac{x}{2}} \exp(\sqrt{\delta} \cos \theta \sqrt{\frac{\lambda_1 - \lambda_2}{2}} \sqrt{\sin \phi + \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}}) d\phi \quad y = \sin \phi = \frac{1}{\pi} \int_{-\frac{x}{2}}^{\frac{x}{2}} d\phi \int_0^\pi \exp(\sqrt{\delta} \cos \theta \sqrt{\frac{\lambda_1 - \lambda_2}{2}} \sqrt{\sin \phi + \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}}) d\theta.$$

Now we are going to do some asymptotic expansion to the above integral. Here, asymptotic expansion is completely a mathematical term, referring to the behavior of some integral when some parameter is very large. Specifically, we will use Laplace method for asymptotic expansion several times in this section. The reader is referred to some mathematical book<sup>[9]</sup> for more exposition on Laplace method and other methods for asymptotic expansions.

Repeating Laplace method for asymptotic expansion, we have  $I(\lambda_1, \lambda_2) \approx (\delta)^{\frac{-1}{2}} \frac{\exp(\sqrt{\delta\lambda_1})}{(\lambda_1 - \lambda_2)^{\frac{1}{2}}}$  and

$$f_{\lambda_1\lambda_2}(\lambda_1, \lambda_2) \approx \frac{1}{4\pi} \frac{\exp(-\frac{\delta}{2}) \exp(-\frac{\lambda_1}{2}) \exp(-\frac{\lambda_2}{2}) \sqrt{\lambda_1 - \lambda_2} \exp(\sqrt{\delta\lambda_1})}{\delta^{\frac{1}{2}} \sqrt{\lambda_1} \sqrt{\lambda_2}} \quad (17)$$

**Step 2:** Now we will prove, asymptotically (for large  $\delta$ ),  $\lambda_1 \gg \lambda_2$ . Equivalently, we need to prove  $\frac{\lambda_2}{\lambda_1} \rightarrow 0$ , i.e.,  $P(\frac{\lambda_2}{\lambda_1} > \epsilon) \rightarrow 0$  for any  $\epsilon > 0$ , where P stands for probability. For  $\epsilon \geq 1$ , it is obvious because  $P=0$ . We consider  $0 < \epsilon < 1$ . From equation (17)

$$\begin{aligned} P(\frac{\lambda_2}{\lambda_1} > \epsilon) &= P(\lambda_2 > \epsilon \lambda_1) \\ &= \frac{1}{4\pi} \exp(-\frac{\delta}{2})^{\delta-\frac{1}{2}} \int_0^\infty \frac{\exp(\sqrt{\delta\lambda_1} - \frac{\lambda_1}{2})}{\sqrt{\lambda_1}} d\lambda_1 \int_{\epsilon\lambda_1}^{\lambda_1} \frac{\exp(-\frac{\lambda_2}{2}) \sqrt{\lambda_1 - \lambda_2}}{\sqrt{\lambda_2}} d\lambda_2 \\ &< \frac{1}{4\pi} \exp(-\frac{\delta}{2})^{\delta-\frac{1}{2}} \int_0^\infty \frac{\exp(\sqrt{\delta\lambda_1} - \frac{\lambda_1}{2})}{\sqrt{\lambda_1}} d\lambda_1 \int_0^{\lambda_1} \exp(-\frac{\lambda_2}{2}) \frac{\sqrt{\lambda_1 - \lambda_2}}{\sqrt{\lambda_2}} d\lambda_2 \\ &< \frac{1}{4\pi} \exp(-\frac{\delta}{2})^{\delta-\frac{1}{2}} \int_0^\infty \frac{\exp(\sqrt{\delta\lambda_1} - \frac{\lambda_1}{2})}{\sqrt{\lambda_1}} d\lambda_1 \cdot \pi \end{aligned}$$

due to the fact that the integral

$$\begin{aligned} \int_0^{\lambda_1} \exp(-\frac{\lambda_2}{2}) \frac{\sqrt{\lambda_1 - \lambda_2}}{\sqrt{\lambda_2}} d\lambda_2 &= \int_0^1 \exp(-\frac{r\lambda_1}{2}) \frac{\sqrt{1-r}}{r} dr \\ &= \int_0^{\frac{\pi}{2}} \exp(-\frac{\lambda_1 \sin^2 \theta}{2}) 2 \cos^2 \theta d\theta < e^0 \cdot 2 \cdot 1 \cdot \frac{\pi}{2} = \pi. \end{aligned}$$

Repeatedly using the Laplace method, we get, as  $\delta \rightarrow \infty$ ,

$$\begin{aligned} P(\lambda_2 > \epsilon \lambda_1) &< \frac{\pi}{4\pi} \exp(-\frac{\delta}{2})^{\delta-\frac{1}{2}} \int_0^\infty \frac{\exp(\sqrt{\delta\lambda_1} - \frac{\lambda_1}{2})}{\sqrt{\lambda_1}} d\lambda_1 \\ &< \frac{1}{4} e^{-\frac{\delta}{2}} \delta^{-\frac{1}{2}} \int_0^\infty \exp(\frac{\delta}{2} - \frac{(\lambda_1 - \delta)^2}{8\delta}) \lambda_1^{-\frac{1}{2}} d\lambda_1 \\ &= \frac{1}{4} e^{-\frac{\delta}{2}} \delta^{-\frac{1}{2}} e^{\frac{\delta}{2}} \delta^{-\frac{1}{2}} 2 \cdot \frac{1}{2} \sqrt{\pi 8\delta} = \sqrt{\frac{\pi}{2}} \delta^{-\frac{1}{2}} \rightarrow 0. \end{aligned}$$

Equation (17) can thus be factored as  $f_{\lambda_1\lambda_2}(\lambda_1, \lambda_2) = f_{\lambda_1}(\lambda_1) f_{\lambda_2}(\lambda_2)$  with

$$f_{\lambda_1}(\lambda_1) = \frac{1}{2\sqrt{2\pi}} \exp(-\frac{\delta}{2})^{\delta-\frac{1}{2}} \exp(\sqrt{\delta\lambda_1} - \frac{\lambda_1}{2}),$$

$$f_{\lambda_2}(\lambda_2) = \frac{1}{\sqrt{2\pi}} \frac{\exp(-\frac{\lambda_2}{2})}{\sqrt{\lambda_2}} \sim \chi_1^2, \text{ i.e., } \lambda_1, \lambda_2 \quad \text{are}$$

(asymptotically) independent.

**Step 3:** We are to prove that  $a_{11}, \lambda_2$  are also independent (asymptotically), which is much easier than the above derivation. From equation (13) and since  $\lambda_1 \gg \lambda_2$  which is just established,

$$P_{a_{11}\lambda_2}(a_{11}, \lambda_2) = \frac{1}{4\pi} \exp(-\frac{\delta}{2}) \frac{\exp(-\frac{\lambda_2}{2}) {}_0F_1(1; \frac{1}{4} \delta a_{11})}{\sqrt{\lambda_2} \sqrt{a_{11} - \lambda_2}} \quad (18)$$

$$\int_{a_{11}}^\infty \frac{\exp(-\frac{\lambda_1}{2}) \sqrt{\lambda_1}}{\sqrt{\lambda_1 - a_{11}}} d\lambda_1$$

The integral is only a function of  $a_{11}$ , however, complex it is. And obviously  $a_{11} - \lambda_2 \approx a_{11}$  as  $\delta \rightarrow \infty$  because  $a_{11} \sim \chi_2^2(\delta), \lambda_2 \sim \chi_1^2(0)$ .

Therefore, equation (18) factors and  $a_{11}, \lambda_2$  are independent (asymptotically). And finally our theorem ensues.

**Closed form of GLRT test statistic's pdf at finite samples:** We use the same reasoning as in previous section. We already have the result that

$$t_3(y)|H_0 \sim \frac{N-1}{2N-3} F_{1,2N-3} \text{ for large } e\delta = \frac{Na^2}{\sigma^2}. \text{ For this}$$

parameter, the contributions from N and from  $\left(\frac{a}{\sigma}\right)^2$  are

equivalent. Therefore, for finite sample but at some reasonably high SNR, the detector is practically CFAR

and the correct threshold should be chosen according to  $\frac{N-1}{2N-3} F_{1,2N-3}$ .

In<sup>[1]</sup>, extensive numerical simulations reveals that the GLR test is also CFAR when  $\frac{\sigma}{\sigma} \geq 1$ , 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 GLR detector is almost exactly one half that of the corresponding threshold required for a  $F_{1,(N-1)}$  distributed test statistic.

Equivalently this means that  $t_3 \sim \frac{1}{2} F_{1,N-1}$ . The discrepancy between result in<sup>[1]</sup> and theoretical result here is easily explained. The numerical simulations were only performed for  $N=120$  in<sup>[1]</sup>. In this case  $\frac{N-1}{2N-3} F_{1,2N-3} \approx \frac{1}{2} F_{1,N-1}$ . If  $N$  gets much smaller (in either fMRI or coherent interference detection to be discussed), theoretical instead to numerical result need be used.

### DISCUSSION AND CONCLUSION

This study provides an asymptotic, closed form of the pdf of the test statistic of one non linear GLRT detector with important applications to fMRI, thereby a more accurate threshold selection method is established for finite samples. This is very important for practical fMRI data processing.

As pointed out in introduction, our model is the simplest, but it leaves room for extension to deal with more complex situations such as trends, physiological fluctuations (respiration and cardiac cycle), patient motions etc. Color noise can be whitened provided we know data correlation matrix. Practical issues like these are already mentioned in<sup>[1]</sup>. Solution of the nonlinear problems with general representation of signal subspace and nuisance subspace  $H$  and  $S$  will be presented in the future.

**Nomenclature:** Scalars, vectors and matrices are not distinguished notationally (except in the very beginning part of introduction of the model). In other words, all quantities in the study are regarded as matrices (of commensurate dimensions), following Matlab convention.

$A \otimes B$ :	Kroneker product between matrix $A$ and $B$
$a \doteq b$ :	$a$ is defined through $b$ and vice versa
$A \equiv B$ :	$A$ is identically equal to $B$
$A^T$ :	transport of matrix $A$
$\text{tr}(A)$ :	trace of matrix $A$

$\det(A)$ :	determinant of matrix $A$
$A^{m \times n}$ :	matrix $A$ with dimension $m \times n$
$I^{p \times p}$ :	the $p \times p$ identity matrix
$O^{m \times m}$ :	the set of all orthogonal matrices with dimensions $m \times m$
$Q^{n \times p}$ :	the set of all $n \times p$ ( $p < n$ ) quasi-orthogonal matrices, meaning of $U^T U = I^{p \times p}$ if $U \in Q^{n \times p}$
$P_M$ :	matrix that projects a vector onto the range space of matrix $M$ , $P_M = M(M^T M)^{-1} M^T$
$P_M^\perp$ :	$I - P_M$ , $\chi_n^2$ , $F_{m,n}$ ; $\chi^2$ and $F$ distributions
$(\theta_1, \theta_2) \doteq \theta_1^T \theta_2$ :	the inner product between two vectors $\theta_1, \theta_2$
$\ v\  \doteq \sqrt{v^T v}$ :	2-norm of vector $v$
$\Gamma(n)$ :	Gamma function
$I_n(x)$ :	modified Bessel function of the first kind with order $n$ <sup>[7]</sup>
${}_0F_1(l; a)$ :	hypergeometric function of $a$ , $a$ can be a scalar, or a matrix <sup>[5]</sup>
pdf:	probability density function
GLR:	generalized likelihood ratio
$A \sim p$ :	matrix $A$ has a pdf (the pdf of a matrix is of course defined as the joint pdf of all its elements) of $p$
$N(m, C)$ :	Gaussian pdf with mean $m$ and covariance matrix $C$ .

### REFERENCES

1. Fangyuan, N. and R.D. Nowak, 1999. Generalized likelihood ratio detection for fMRI using complex data. *IEEE Trans. Med. Imag.*, 18: 320-329.
2. Rowe, D.B. and B.R. Logan, 2004. A complex way to compute fMRI activation. *NeuroImage*, 23: 1078-1092.
3. Desai, M., R. Mangoubi, J. Shah, W. Karl, H.P.A. Worth and D. Kennedy, 2002. Functional MRI activity characterization using response time shift estimates from curve evolution. *IEEE Trans. Med. Imag.*, 21: 1402-1412.
4. Scharf, L.L., 1991. *Statistical Signal Processing*, Reading, MA. Addison-Wesley.
5. Muirhead, J., 1982. *Aspects of Multivariate Statistical Theory*. New York, Wiley.
6. Sanchez, L.E. and D.K. Nagar, 2003. Non-central matrix variate Dirichlet distribution. *Taiwanese J. Math.*, 7: 477-491. [www.math.nthu.edu.tw/tjm](http://www.math.nthu.edu.tw/tjm)
7. Kay, S.M., 1998. *Fundamentals of Statistical Signal Processing, Detection Theory*. Englewood Cliffs, NJ, Prentice Hall.
8. Kim, S., W. Richter and K. Ugerbil, 1997. Limitations of temporal resolution in functional MRI. *Magnetic Resonance Medicine*, 37: 631-636.
9. Bender, C.M. and S.A. Orszag, 1999. *Advanced mathematical methods for scientists and engineers: Asymptotic methods and perturbation*. Springer.