ABSTRACT

Since voxel time courses in functional magnetic resonance imaging (fMRI) are mostly produced from complex-valued data by taking the magnitudes, they obey Rician distributions, which can be approximated as Gaussian distributions only when signal-to-noise ratios (SNRs) are high. In this paper, we derive the asymptotic power of our recently developed activation detection statistic for Rician fMRI. The analysis shows that the asymptotic power is dependent only on the ratios of signal parameters to noise parameter of Rician distributed voxel time series, and allows us to better understand the nature of low SNRs in fMRI data analysis. Based on the power analysis, a more general and descriptive definition of SNR is provided than classical one.

Index Terms— Functional MRI, Rician distribution, SNR, activation detection, and asymptotic power analysis.

1. INTRODUCTION

From the inception of neuro-imaging by functional magnetic resonance imaging (fMRI), most attention has focused on the analysis of magnitude voxel time courses with a Gaussian distribution, which in fact obey a Rician distribution. This Gaussian modeling can be justified by the fact that a Rician probability density function (PDF) is well approximated by a Gaussian PDF when signal-to-noise ratio (SNR) is high [1]. Based on Gaussian modeling, general linear model (GLM) has been a dominant framework to model the hemodynamic responses of the human brain and build up detection statistics for activation such as F-statistic.

For more exact modeling of magnitude voxel time courses and more powerful activation detection which hold in the regime of low SNR than classical Gaussian modeling, we developed a Rician detection statistic in the framework of likelihood ratio test (LRT), in which maximum likelihood estimates (MLEs) were obtained by an expectation-maximization (EM) algorithm [2]. To better understand its behavior over various SNRs and obtain insights into the detection power loss by Gaussian modeling in the regime of low SNR, in this article, we derive an asymptotic power function of the developed Rician detection statistic for activation and scrutinize its characteristics. To the best of authors’ knowledge, this paper is the first theoretical power analysis of detection statistics for activation in fMRI data analysis and an exception is our previous paper in a different fMRI context, space-time separability [3]. Some simulation studies for analyzing the detection power of activation statistics in Rician fMRI, e.g. [4], were suggested in the past, typically requiring huge computations.

2. MODEL FORMULATION

Dropping a voxel location index for brevity, we first consider a complex-valued fMRI measurement model and then discuss its magnitude model. For a time \( t = 1, \ldots, n \), we define a complex-valued measurement model as follows [5]:

\[
y_{c,t} = (x_t^T \beta \cos \theta_t + \eta_{r,t}) + j(x_t^T \beta \sin \theta_t + \eta_{i,t}),
\]

where \( \theta_t \) describes phase imperfections, possibly caused by magnetic field inhomogeneity in MR scanner. \( \eta_{r,t} \) and \( \eta_{i,t} \) compose two types of signal components; 1) nuisance signal components, including baseline and temporal drift, and 2) signal component of interest, including the blood oxygenation level dependent (BOLD) response. These signal components are parameterized as

\[
x_t^T \triangleq [1, t, \xi_t^T], \quad \beta^T \triangleq [m, b, f^T],
\]

\[
\xi_t \triangleq [\xi_{1,t}, \ldots, \xi_{L,t}]^T, \quad f \triangleq [f_1, \ldots, f_L]^T,
\]

where \( m \) denotes baseline, \( b t \) means linear temporal drift, and \( \xi_t^T f \) models the BOLD response. \( \xi_t \) contains \( L \) basis functions and \( f \) involves the associated activation amplitude. This representation of the BOLD response covers various modelings, including FIR [6, 7] and Laguerre modeling [8]. The simple modeling of nuisance signal components in (2) can be straightforwardly extended to more complex ones, including higher polynomial terms and/or trigonometric terms.

The magnitude voxel time series can be formed from the complex-valued formation in (1), producing

\[
y_t \triangleq \left( (x_t \cos \theta_t + \eta_{r,t})^2 + (x_t \sin \theta_t + \eta_{i,t})^2 \right)^{\frac{1}{2}},
\]

which obeys a Rician distribution at a time \( t \). Note that fMRI data being analyzed dominantly consist of these real-valued magnitude measurements. At a time \( t \), the PDF of the Rician random variable \( y_t \) has a form of

\[
p(y_t; \beta, \sigma^2) = \frac{y_t}{\sigma^2} \exp \left(-\frac{y_t^2 + (x_t^T \beta)^2}{2\sigma^2}\right) I_0 \left(\frac{y_t x_t^T \beta}{\sigma^2}\right),
\]

where \( I_0(z) \) is the zeroth order modified Bessel function of the first kind. Thanks to temporal independence assumption [2, 4, 5, 9], we have a joint PDF of \( y \triangleq [y_1, \ldots, y_n]^T \) as follows:

\[
p(y; \beta, \sigma^2) = \prod_{t=1}^n p(y_t; \beta, \sigma^2).
\]
Thus a Rician log-likelihood function is provided as
\[
\log L = -n \log \sigma^2 + \sum_{t=1}^{n} \log y_t - \sum_{t=1}^{n} \left\{ \frac{y_t^2 + (x_t^T \beta)^2}{2\sigma^2} - \log I_0 \left( \frac{y_t x_t^T \beta}{\sigma^2} \right) \right\}.
\] (7)

By a Gaussian approximation [1] to a Rician PDF, which holds when SNR is high, one can arrive at the following linear regression formulation:
\[
y = X \beta + \eta,
\] (8)
where \( X \triangleq [x_1, \ldots, x_n]^T \) and \( \eta \sim \mathcal{N}(0, \sigma^2 I_n) \). For voxels whose SNRs are high enough, therefore, a huge amount of simplifications in analysis and model fitting are achievable. This approach based on a Gaussian approximation, widely known as GLM, has been a dominant tool for analyzing fMRI data mainly thanks to its simplicity.

3. DETECTION STATISTIC FOR ACTIVATION

To build up a detection statistic for activation based on the Rician model in (4), we consider the following null and alternative hypotheses:
\[
H_N : C \beta = 0, \quad \text{vs} \quad H_A : C \beta \neq 0,
\] (9)
where the constraint matrix \( C \) is of full rank and has the size of \( r \times (L + 2) \), \( r \) is a design parameter of experimenters. By selecting \( C \), we can specify the hypothesis tested. For example, if we have a signal model with \( L = 3 \), i.e., \( \beta = [m, b, [f_1, f_2, f_3]^T] \), we can choose
\[
C = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix},
\] (10)
where, e.g., the first row is associated with the first activation amplitude, \( f_1 \). The first two columns are for nuisance signal components, \( m \) and \( b \). Thus \( H_N \) states the voxel which we are interested in is not activated by given temporal stimuli and \( H_A \) states that voxel is activated by them.

With parameter estimates from our recently developed Rician-EM algorithm [2], we build up a generalized likelihood ratio test (GLRT) based on Rician distribution, leading to
\[
\begin{align*}
\lambda_R &= 2 \log \Lambda_R = 2n \log \left( \frac{\sigma_n^2}{\sigma_A^2} \right) + \sum_{t=1}^{n} \frac{y_t^2 + (x_t^T \hat{\beta}_N)^2}{\sigma_A^2} - 2 \log \left( I_0(\sigma_n y_t x_t^T \hat{\beta}_A / \sigma_A^2) / I_0(\sigma_n y_t x_t^T \hat{\beta}_N / \sigma_N^2) \right) \\
&= -\sum_{t=1}^{n} \left\{ \frac{y_t^2 + (x_t^T \hat{\beta}_A)^2}{2\sigma_A^2} - \log I_0 \left( \frac{y_t x_t^T \hat{\beta}_A}{\sigma_A^2} \right) \right\} - \log I_0 \left( \frac{y_t x_t^T \hat{\beta}_N}{\sigma_N^2} \right) - 2 \log (\sigma_n / \sigma_A) \quad \text{for } \hat{\beta}_A \text{ and } \hat{\beta}_N \text{ MLEs of } \beta \text{ under } H_A \text{ and } H_N, \text{ respectively}.
\end{align*}
\] (11)

where, e.g., \( \hat{\beta}_A \) and \( \hat{\beta}_N \) are MLEs of \( \beta \) under \( H_A \) and \( H_N \), respectively. \( \Lambda_R \) denotes the ratio of two Rician likelihood functions under \( H_A \) and \( H_N \). For more details of the Rician-EM algorithm and its derivation, the readers are referred to [2, 10].

4. ASYMPTOTIC POWER ANALYSIS

For analyzing an asymptotic power of the Rician activation statistic \( \lambda_R \) in (11), we need an asymptotic distribution of GLRT under \( H_A \). Due to nuisance parameters \((m, b, \sigma^2)\) and non-identically distributed samples along time, we can not use a well-known conventional work on the asymptotic expansion of GLRT statistics for a sequence of local alternatives, requiring independently and identically distributed samples [11]. Instead we apply a more general development for serially correlated and non-identically distributed observations in the presence of nuisance parameters, suggested in [12].

4.1. Asymptotic Expansion of GLRT under a Local \( H_A \)

Let \( \eta \) denote a parameter of interest whose dimension is \( p \) and \( \mu \) denote a nuisance parameter whose dimension is \( q \). We are interested in testing the following composite hypotheses:
\[
H_N : \eta = \eta_0, \quad \text{vs} \quad H_A : \eta \neq \eta_0.
\] (12)

In the example of (10), we have \( \eta = [f_1, f_2, f_3]^T \), \( \eta_0 = [0, 0, 0]^T \), and \( \mu = [m, b, \sigma^2]^T \). We need an asymptotic expansion of the GLRT statistic based on (12) for a sequence of local alternatives, i.e. \( \eta = \eta_0 + \varepsilon / \sqrt{n} \), where \( n' \) is the number of samples and a small \( \varepsilon \) (> 0).

Under some regularity conditions involving the validity of asymptotic expansions of cumulants and the differentiability of log-likelihood function, a GLRT statistic \( T \) has the following asymptotic expansion for a sequence of local alternatives, as shown in [12]:
\[
\Pr (T < t) = \Psi_{p, \Delta}(t) + \frac{1}{\sqrt{n'}} \sum_{j=0}^{3} m_j \Psi_{p+3j, \Delta}(t) + o \left( \frac{1}{\sqrt{n'}} \right),
\] (13)

where, based on the partition of the parameter space, the FIM is partitioned into
\[
\begin{align*}
\mathcal{I} &\triangleq \begin{bmatrix}
\mathcal{I}_{11} & \mathcal{I}_{12} \\
\mathcal{I}_{21} & \mathcal{I}_{22} \end{bmatrix} \\
\mathcal{I}_{11} &\triangleq \begin{bmatrix}
\mathcal{I}_{11(\mu)} & \mathcal{I}_{12(\mu)} \\
\mathcal{I}_{21(\mu)} & \mathcal{I}_{22(\mu)} \end{bmatrix}.
\end{align*}
\] (15)

\( \mathcal{I}_{11(\mu)} \) corresponds to the nuisance parameter and \( \mathcal{I}_{22(\eta)} \) involves the parameter of interest.

4.2. Asymptotic Power Function of \( \lambda_R \)

We apply the general result from the previous section to the activation statistic \( \lambda_R \) in (11). The regularity conditions for the asymptotic expansion in (13) can be straightforwardly checked in our case, but it is tedious to show and thus omitted here. Suppose that we want to test the null hypothesis of \( f = 0 \) against the alternative hypothesis of \( f \neq 0 \). Since we clearly have \( L \) degrees of freedom from the problem construction, we focus on deriving the corresponding non-centrality parameter \( \Delta_R \).

From the Rician log-likelihood function in (7), it can be easily shown that expectations of the second-order derivatives of \( \ell \) (\( \triangleq \log L \)) evaluated at \( f = 0 \) are given by
\[
\begin{align*}
\left[ \frac{\partial^2 \ell}{\partial f_i \partial f_j} \right] &+ \frac{1}{\sigma^2} \sum_{t=1}^{n} (\tilde{A}_t - 1) \xi_{i,t} \xi_{j,t}, \\
\left[ \frac{\partial^2 \ell}{\partial f_i \partial m} \right] &+ \frac{1}{\sigma^2} \sum_{t=1}^{n} (\tilde{A}_t - 1) \xi_{i,t}.
\end{align*}
\] (16)
\[ E \left[ \frac{\partial^2 \ell}{\partial f_i \partial b_j} \right] = \frac{1}{\sigma^2} \sum_{t=1}^{n} (\bar{A}_t - 1) \xi_{i,t}, \]
\[ E \left[ \frac{\partial^2 \ell}{\partial f_i \partial \sigma^2} \right] = \frac{1}{\sigma^3} \sum_{t=1}^{n} (\bar{D}_t(1 - \bar{A}_t) - \bar{B}_t) \xi_{i,t}, \]
where \( f_j \) denotes the \( j \)-th element of activation amplitude \( f \) for \( j = 1, \ldots, L \) and the superscript \( + \) indicates the expectation is evaluated at \( f = 0 \).

For the simple expressions above, we define variables as follows:

\[ \bar{A}_t \triangleq E \left[ \frac{y_t^2}{\sigma^2} \frac{I_1'(a_t)I_0(a_t) - I_2'(a_t)}{I_0^2(a_t)} \right]^+, \]
\[ \bar{B}_t \triangleq E \left[ \frac{y_t}{\sigma} \frac{I_1(a_t)}{I_0(a_t)} \right]^+, \quad \bar{D}_t \triangleq \frac{m + bt}{\sigma}, \]
\[ a_t \triangleq \frac{y_t (m + bt + \xi_t f)}{\sigma^2}, \]
where \( I_1(z) \) denotes the first-order modified Bessel function of the first kind and \( I_1(z) \) is its first-order derivative. It can be easily shown that \( \bar{A}_t \) and \( \bar{B}_t \) depend only on the ratios of nuisance signal parameters to noise parameter, i.e. \( \bar{m} \triangleq m/\sigma \) and \( \bar{b} \triangleq b/\sigma \).

Following the same partition as in (15) with \( \mu = [m, b, \sigma^2]^T \) and \( \eta = f \), the substitution of (16)-(18) to (14) produces the following non-centrality parameter:

\[ \Delta_R = n \bar{f}^T \left[ \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t) \xi_t \xi_t^T - \text{HM}^{-1}\text{H}^T \right] \bar{f}, \]
where we define the activation-to-noise ratio as \( \bar{f} \triangleq f/\sigma \). The symmetric \( 3 \times 3 \) matrix \( \text{M} \) and \( L \times 3 \) matrix \( \text{H} \) are defined by

\[ \text{M}_{11} = \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t), \quad \text{M}_{12} = \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t) t, \]
\[ \text{M}_{13} = \frac{1}{n} \sum_{t=1}^{n} \bar{D}_t(1 - \bar{A}_t) - \bar{B}_t, \quad \text{M}_{22} = \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t)^2 t^2, \]
\[ \text{M}_{23} = \frac{1}{n} \sum_{t=1}^{n} (\bar{D}_t(1 - \bar{A}_t) + \bar{B}_t) t, \]
\[ \text{M}_{33} = \frac{1}{n} \sum_{t=1}^{n} 1 + 2\bar{D}_t - \bar{D}_t^2 \bar{A}_t - 2\bar{D}_t \bar{B}_t, \]
\[ \text{H} = \left[ \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t)^2 \xi_t, \quad \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t) t \xi_t, \quad \frac{1}{n} \sum_{t=1}^{n} (\bar{D}_t(1 - \bar{A}_t) + \bar{B}_t) \xi_t \right] , \]
where \( \text{M}_{i,j}' \) is the entry of \( \text{M} \) at \( (i', j') \).

5. DISCUSSIONS

Since a conventional definition of SNR \( \triangleq \frac{\alpha}{\text{var}(\beta)} \approx \frac{m}{\sigma} \) [5] ignores the variation of voxel time series residing in temporal drift, it might mislead researchers in fMRI to believe that voxels with activation-to-noise ratio close to the high SNR region with high SNR in the non-central chi-square distribution with \( L \) degrees of freedom and \( \alpha \) is a significance level.

From the derivation above, we can draw three conclusions on the asymptotic power of \( \lambda_R \) based on a Rician distribution. First, the asymptotic power depends only on the ratios of signal parameters to noise parameter, \( (\bar{m}, \bar{b}, \bar{f}) \).

For various values of \( \bar{m}, \bar{b} \), we have

\[ \Delta_R \approx n \bar{f}^T \left[ \frac{1}{n} \sum_{t=1}^{n} (1 - \bar{A}_t) \xi_t \xi_t^T - \text{HM}^{-1}\text{H}^T \right] \bar{f}, \]

where \( \Delta_R \) is the non-centrality parameter.

5.1. Behaviors of \( \bar{A}_t \) and \( \bar{B}_t \)

Since \( \bar{A}_t \) and \( \bar{B}_t \) play key roles in the derived non-centrality parameter \( \Delta_R \) in (20), it is useful to investigate the behaviors of these two functions in terms of \( \bar{m} \) and \( \bar{b} \) by identifying the regions with high-low SNRs. From their definitions, we have

\[ \bar{A}_t = \exp \left( -\frac{\bar{D}_t^2}{2} \right) \int_0^\infty \left( I_2(D_tz) + I_0(D_tz) \right) \frac{z^2 \exp \left( -\frac{z^2}{2} \right)}{I_0(D_tz)} \quad dz, \]
\[ \bar{B}_t = \exp \left( -\frac{\bar{D}_t^2}{2} \right) \int_0^\infty I_1(D_tz) z^2 \exp \left( -\frac{z^2}{2} \right) \quad dz. \]
By noting that $I_k(-z) = (-1)^k I_k(z)$ for an integer $k$, $\tilde{A}_t$ is an even function and $\tilde{B}_t$ is an odd function in terms of $\tilde{D}_t$. We can show that $A_t$ is a strictly decreasing function for $D_t > 0$ and strictly increasing function for $D_t < 0$, $A_t = 1$ when $D_t = 0$, and $A_t$ approaches 0 as $|D_t|$ goes to infinity. It can be also shown that $\tilde{B}_t = 0$ when $D_t = 0$ and $\tilde{B}_t \approx D_t$ when $|D_t|$ is large. Example plots of $\tilde{A}_t$ and $\tilde{B}_t$ in the $\bar{m}$-$\bar{b}$ plane are shown in Fig. 1 when $t = 1$.

We conclude, therefore, that the low SNR corresponds to small $|\tilde{D}_t| (\triangleq |\bar{m} + \bar{b}t|)$ for a given time $t$, leading to $\tilde{A}_t \approx 1$ and $\tilde{B}_t \approx 0$, by recognizing when $\tilde{A}_t$ reduces to the non-centrality parameter derived from the Gaussian distributed model in (8). Note that a conventional belief in fMRI data analysis is that high $\bar{m}$ means high SNR, leading to $\text{SNR}_G \triangleq \bar{m}$, whereas our analysis indicates that high $\bar{m}$ and negative $\bar{b}$ may make a Gaussian approximation to a Rician distributed model inaccurate for certain time points.

### 5.2. A New Definition of SNR

Based on the observations in the previous section, we present a more general and descriptive definition of SNR. One possible definition of SNR is

$$\text{SNR}_R = \left| \frac{\text{BNR}}{2} + \frac{n + 1}{2} \right| \text{DNR},$$

(26)

where BNR ($\triangleq \text{SNR}_G$) is baseline-to-noise ratio, DNR ($\triangleq \bar{b}$) denotes drift-to-noise ratio, and $\sum_{i=1}^{n} t/n = (n + 1)/2$. Note that other definitions are possible in the same philosophy. For example, one can use the number of time points $n$ instead of the arithmetic mean of time points. When DNR is relatively small to $n$, we have $\text{SNR}_R \approx |\text{BNR}|$, which is the same as $\text{SNR}_G$.

### 6. CONCLUSIONS

For the first time in fMRI, we provided an asymptotic theoretical analysis of the power of GLRT statistics for activation detection for Rician distributed fMRI. We found that the asymptotic power depends only on the ratios of signal parameters to noise parameter. It was also shown that high SNR is dependent on not only the variation residing in baseline but also the one in temporal drift. Based on the asymptotic power analysis, we extended a conventional definition of SNR to a more general and descriptive one for fMRI voxel time courses.

### 7. REFERENCES


