SPACE-TIME SEPARABILITY IN FUNCTIONAL MRI: ASYMPTOTIC POWER ANALYSIS OF A NEW TEST PROCEDURE

Joonki Noh† *and Victor Solo*‡

†Dept.of Electrical Eng. and Computer Science, University of Michigan, Ann Arbor, USA,
‡School of Electrical Eng., University of New South Wales, Sydney, AUSTRALIA.

ABSTRACT

Space-time separability has been assumed and applied to most of data analysis methods in functional magnetic resonance imaging (FMRI). In a recent work, we developed a procedure for testing space and time separability in the framework of the parametric cepstrum. In this paper, the asymptotic power of the proposed space-time separability test is analyzed. The analysis shows two important properties of the proposed test. The asymptotic power function involves cepstral coefficients only in the non-separable region (parameters of interest). And the non-centrality parameter of the asymptotic power function is a scaled Euclidean metric between the logarithms of a non-separable power spectral density (PSD) and a separable PSD.

Index Terms— FMRI data analysis, space-time separability, the parametric cepstrum, and asymptotic power.

1. INTRODUCTION

In FMRI data analysis, a significant task for a collected dataset is to detect localized activations in the brain induced by given stimuli. Since space-time separability simplifies the problem significantly, in most approaches for the activation study, it has been assumed and accepted without proper justifications. Conceptually, space-time separability implies that spatial and temporal operations can be separately performed for proper activation detection. In SPM, spatial smoothing by Gaussian kernel is first applied to a collected dataset. Then, a general linear model (GLM), leading to temporal filtering, is built up to create an activation map.

For a given dataset, however, the validity of space-time separability assumption is unknown without a proper test. To the best of authors' knowledge, in FMRI, any testing for space and time separability has not been properly treated until now. We recently developed a procedure for testing space and time separability and discussed only its null distribution [1]. In this paper, we analyze the asymptotic power of the proposed test and examine its characteristics to complete the development. This is the first theoretical power analysis in FMRI, especially for detecting space-time separability. Simulation studies for testing space and time separability are computationally very expensive, and are thus intractable.

2. TESTING SPACE-TIME SEPARABILITY

2.1. Measurement Model Formulation

We consider a typical measurement model in FMRI, which has a linear and additive form. For a time point t and a voxel location v, we assume that the measurement has the form of

$$y_{t,v} = X_t^T \beta_v + s_{t,v} + w_{t,v}, \tag{1}$$

where β_v contains nuisance signals, for example, temporally varying drift, and $s_{t,v}$ models the the blood oxygenation level dependent (BOLD) response of the brain. The noise $w_{t,v}$ is a spatiotemporally correlated and stationary Gaussian random field whose mean is zero. For a given stimulus denoted as c_t, the BOLD response can be described linearly as

$$s_{t,v} = \left(\sum_{i=1}^{L} h_{i,t} f_{i,v} \right) * c_t \triangleq \xi_t^T f_v, \tag{2}$$

where $h_{i,t}$ is the i-th basis for modeling temporal responses, $f_{i,v}$ denotes the corresponding activation amplitude to model spatial responses, and L is the number of basis functions. It is shown in [2] that this linear model provides an accurate approximation of the BOLD response to the first order. For integer t and v, we assume $y_{t,v}$ is observed from a rectangle, $\{0, \ldots, T-1\} \times \{0, \ldots, M-1\}$, where T is the number of time points and M is the number of voxels in a ROI.

2.2. Test Statistic for Space-Time Separability

Because the analysis in the spatial and temporal frequency domains allows a substantial amount of simplifications, we perform a test in the frequency domain. Taking the DFT in (1) yields

$$\tilde{y}_{k,l} = \tilde{X}_k^T \tilde{\beta}_l + \tilde{\xi}_k^T \tilde{f}_l + \tilde{w}_{k,l}, \tag{3}$$

where k denotes temporal frequency and l denotes spatial wave-number. Under some regularity conditions involving joint cumulants, for large T and M, central limit theorem (CLT) allows an asymptotic distribution,

$$\frac{1}{\sqrt{TM}} \cdot \tilde{w}_{k,l} \sim \mathcal{N}_c(0, F_{k,l}), \tag{4}$$

where $\mathcal{N}_c$ denotes a complex-valued Gaussian distribution and $F_{k,l}$ is power spectral density (PSD) of the noise [3]. Now, we have the following approximate negative log-likelihood:

$$2\ell(\tilde{\beta}, \tilde{f}, \theta) \triangleq \sum_{k=0}^{T-1} \sum_{l=0}^{M-1} \log F_{k,l} + \frac{\left|\tilde{y}_{k,l} - \tilde{X}_k^T \tilde{\beta}_l - \tilde{\xi}_k^T \tilde{f}_l\right|^2}{TM \cdot F_{k,l}}, \tag{5}$$

where, e.g., $\tilde{f} \triangleq [\tilde{f}_0, \ldots, \tilde{f}_{M-1}]^T$, and θ contains parameters for modeling PSD, that are called the cepstral coefficients.

Space-time separability means the decomposition of the PSD into its pure temporal piece and pure spatial piece, that is, for $\forall(k,l)$, $F_{k,l} = F_k G_l$, where F_k is temporal PSD and G_l is spatial PSD. We therefore have the following non-linear hypothesis testing problem:

$$H_0 : F_{k,l} = F_k G_l \qquad \text{vs} \qquad H_1 : F_{k,l} \neq F_k G_l, \tag{6}$$

where H_0 indicates $w_{t,v}$ is a space-time separable field and H_1 means its alternative. From (6), we construct a likelihood ratio test (LRT), leading to the following test statistic:

$$\mathcal{L}_S \triangleq 2\ell(\widehat{\tilde{\beta}}_1, \widehat{\tilde{f}}_1, \widehat{\theta}_1) - 2\ell(\widehat{\tilde{\beta}}_0, \widehat{\tilde{f}}_0, \widehat{\theta}_0), \tag{7}$$

where all estimates are maximum likelihood estimates (MLEs). For example, $\widehat{F}_{1,k,l}$ is the MLE of the PSD under H_1.

3. REFORMULATION IN THE CEPSTRAL DOMAIN

To make a valid test procedure for space-time separability, we need to control false positive rate (FPR), the probability of type I error. The null distribution of $\mathcal{L}_S$ enables to compute a threshold for α significance label. The cepstrum provides us an useful framework to obtain an asymptotic null distribution of $\mathcal{L}_S$. In fact, in multi-dimensions (3D or 4D in FMRI), noise modeling by the parametric cepstrum has several advantages over AR-based methods [4]. For example, model fitting by the parametric cepstrum is almost linear and can be performed very quickly with FFT [1, 4].

3.1. Parametric Cepstrum

We consider a Fourier series expansion of the logarithm of PSD. By truncating the array of Fourier coefficients, we have a modeling of a PSD by the parametric cepstrum.
For $-\pi \leq \omega, \lambda \leq \pi$,

$$\log F(\omega, \lambda) = \sum_{t=-n}^{n} \sum_{v=-p}^{p} \theta_{t,v} e^{-j(\omega t + \lambda v)}, \tag{8}$$

where n and p denote a temporal order and a spatial order of the model, respectively [4]. Due to a real-valued PSD, note that the cepstrum has a symmetry, $\theta_{t,v} = \theta_{(-t,-v)}$ for $\forall(t,v)$. For non-separable fields, we estimate $R_a(\triangleq 2np+n+p+1)$ cepstral coefficients. One benefit of the cepstral modeling is a linear description of space-time separability. The condition for space-time separability is as follows. For $(t,v) \neq (0,0)$,

$$\theta_{t,v} = \theta_{t,0}\delta_{0,v} + \theta_{0,v}\delta_{t,0}, \tag{9}$$

where $\theta_{t,0}$ is a cepstral coefficient along the temporal axis and $\theta_{0,v}$ is a cepstral coefficient on the spatial plane. For separable fields, we estimate $R_s(\triangleq n+p+1)$ cepstral coefficients.

3.2. Controlling False Positive Rate

Importantly, we have a simple equivalent hypothesis testing problem to (6) in the cepstral domain. For $\forall(t,v) \in \Theta_{ns}$,

$$H_0 : \theta_{t,v} = 0 \qquad \text{vs} \qquad H_1 : \theta_{t,v} \neq 0, \tag{10}$$

where Θ_{ns} is called the non-separable region. We define

$$\begin{aligned} \Theta_a &\triangleq \{|t| \leq n, |v| \leq p\} = \Theta_s \cup \Theta_{ns}, \quad \Theta_{ns} = \Theta_s^c, \\ \Theta_s &\triangleq \{t=0, |v| \leq p\} \cup \{|t| \leq n, v=0\}, \end{aligned} \tag{11}$$

where the whole set of indices Θ_a is partitioned into Θ_{ns} (the non-separable region) and Θ_s (the separable region). From an asymptotic null distribution of LRT [5], the reformulated hypotheses in (10) yields, for α significance level,

$$\gamma(\alpha) = \Phi_{R_{ns}}^{-1}(1-\alpha), \tag{12}$$

where $\Phi_{R_{ns}}(t)$ denotes the cumulative density function (CDF) of a central chi-square distribution with $R_{ns}(\triangleq R_a - R_s = 2np)$ degrees of freedom. Details of the model fitting for the proposed space-time separability test are provided in [1].

4. ASYMPTOTIC POWER ANALYSIS

To analyze the asymptotic power of the proposed procedure for testing space-time separability, an asymptotic distribution of LRT under H_1 is required. Since the proposed test, in the presence of nuisance parameters, e.g., cepstral coefficients in Θ_s and activation amplitudes, involves samples which are non-identically distributed, an asymptotic expansion of the statistic $\mathcal{L}_S$ under H_1 is a non-trivial task. While this kind of problem has long been discussed in the statistics literature, the conventional work, e.g., [6], deals only with independently and identically distributed samples, and is not applicable.

In a recent work, when samples have serial correlation and nuisance parameters exist, an asymptotic expansion of a class of test statistics including LRT is derived for a sequence of local alternatives [5]. Here, the main result in [5] is applied to derive the asymptotic distribution of $\mathcal{L}_S$ for a non-separable field, i.e., under H_1. In fact, some regularity conditions are required for [5], which are briefly described below. These regularity conditions can be straightforwardly checked in our current setups for $\mathcal{L}_S$, but it is tedious to show.

4.1. Asymptotic Expansion of LRT under a Local H_1

Suppose that η denotes a vector containing lexicographically ordered parameters of interest and η_0 is specified by a given null hypothesis. A vector μ is similarly defined for nuisance parameters. We are interested in an asymptotic expansion of a LRT statistic for a sequence of local alternatives, defined as $\eta = \eta_0 + \varepsilon/\sqrt{n'}$, where n' is the number of samples.

According to [5], under regularity conditions involving the validity of asymptotic expansions of cumulants and the differentiability of log-likelihood function, a test statistic T belonging to a class which includes LRT has the following asymptotic expansion for a sequence of local alternatives:

$$P\left(T<t\right) = \Psi_{d,\Delta}(t) + \frac{1}{\sqrt{n'}}\sum_{q=0}^{3} m_q \Psi_{d+2q,\Delta}(t) + o\left(\frac{1}{\sqrt{n'}}\right), \tag{13}$$

where $\Psi_{d,\Delta}(t)$ denotes the CDF of a non-central chi-square distribution with d degrees of freedom and Δ non-centrality parameter. m_qs are computed from asymptotic expansions of the first and second order derivatives of a given log-likelihood function. Remarkably, the degrees of freedom is the same as the length of η and the non-centrality parameter Δ is given in terms of Fisher information matrix (FIM). We have

$$\Delta = \varepsilon^T\left(\mathcal{I}_{11} - \mathcal{I}_{12}\mathcal{I}_{22}^{-1}\mathcal{I}_{21}\right)\varepsilon\ \Bigg|_{\eta=\eta_0}, \tag{14}$$

where, following the partition of the parameter space into two parts, FIM is also partitioned into

$$\mathcal{I}(\eta,\mu) \triangleq \begin{bmatrix} \mathcal{I}_{11}(\eta) & \mathcal{I}_{12}(\eta,\mu) \\ \mathcal{I}_{21}(\mu,\eta) & \mathcal{I}_{22}(\mu) \end{bmatrix}. \tag{15}$$

$\mathcal{I}_{11}(\eta)$ is associated only with parameters of interest and $\mathcal{I}_{22}(\mu)$ involves only nuisance parameters. If matrix $\mathcal{I}_{12}(\eta_0,\mu)$ is a zero, the non-centrality parameter simplifies into

$$\Delta = \varepsilon^T \mathcal{I}_{11}(\eta_0)\varepsilon. \tag{16}$$

The non-centrality parameter is thus independent of nuisance parameter μ, so that the corresponding asymptotic power does not depend on nuisance parameter. In fact, it turns out that this is the case of the proposed test procedure for space-time separability.

4.2. Asymptotic Power of the Test Procedure

We applied the general result in the previous section to the proposed statistic for space-time separability in this section. Modeling by the parametric cepstrum given in (8) allows an expression for the logarithm of a sampled PSD as follows:

$$\log F_{k,l} \triangleq \log F\left(\frac{2\pi k}{T}, \frac{2\pi l}{M}\right) = x_{k,l}^T\theta^{ns} + z_{k,l}^T\theta^{s}, \tag{17}$$

where θ^{ns} denotes a lexicographically ordered column vector containing cepstral coefficients in Θ_{ns} and $x_{k,l}$ is a vector containing associated cosine terms. Note that sine terms in (8) are canceled due to the symmetry of cepstra. θ^s and $z_{k,l}$ are similarly defined for cepstral coefficients in Θ_s. We have $k = 0,\ldots,T-1$ and $l = 0,\ldots,M-1$. Then, by substituting (17) into (5), the following fully parameterized negative likelihood function is obtained:

$$2\ell(\tilde{\beta},\tilde{f},\theta^s,\theta^{ns}) = \sum_{k=0}^{T-1}\sum_{l=0}^{M-1} x_{k,l}^T\theta^{ns} + z_{k,l}^T\theta^s + \frac{1}{TM} \times exp(-x_{k,l}^T\theta^{ns} - z_{k,l}^T\theta^s)\left|\tilde{y}_{k,l} - \tilde{X}_k^T\tilde{\beta}_l - \tilde{\xi}_k^T\tilde{f}_l\right|^2. \tag{18}$$

To simplify the discussion, we first consider the non-centrality parameter Δ for a simplified case without signal components, $\tilde{\beta}$ and $\tilde{f}$. Then, we reconsider those signal components in the general development. In fact, it turns out that results from the simplified case are useful in the general development. After determining Δ, an asymptotic power function can be obtained by combining the central chi-square distribution in (12) and the non-central chi-square distribution in (13) for a given α.

4.2.1. *Case I: without signal components*

To compute the non-centrality parameter Δ, FIM is required in (14). Assuming $\tilde{\beta} = 0$ and $\tilde{f} = 0$, it can be shown that the expected values of the second order derivatives of $2\ell(\theta^s,\theta^{ns})$ are given by, noting that θ_i^{ns} is the i-th element of θ^{ns} and θ_j^s is the j-th element of θ^s,

$$\mathbf{E}\left[\frac{\partial^2 2\ell}{\partial\theta_i^{ns}\partial\theta_m^{ns}}\right] = \sum_{k=0}^{T-1}\sum_{m=0}^{M-1} x_{k,l,i}x_{k,l,m} = 2TM\cdot\delta_{i-m}, \tag{19}$$

$$\mathbf{E}\left[\frac{\partial^2 2\ell}{\partial\theta_j^{s}\partial\theta_n^{s}}\right] = \sum_{k=0}^{T-1}\sum_{m=0}^{M-1} z_{k,l,j}z_{k,l,n} = \begin{cases} TM & \text{if } j=n=1 \\ 2TM\cdot\delta_{j-n} & \text{otherwise} \end{cases}, \tag{20}$$

$$\mathbf{E}\left[\frac{\partial^2 2\ell}{\partial\theta_i^{ns}\partial\theta_n^{s}}\right] = \sum_{k=0}^{T-1}\sum_{m=0}^{M-1} x_{k,l,i}z_{k,l,n} = 0. \tag{21}$$

where $x_{k,l,i}$ denotes the i-th element of $x_{k,l}$ and $z_{k,l,j}$ denotes the j-th element of $z_{k,l}$. Note $\mathbf{E}[|\tilde{y}_{k,l}|^2] = TM\cdot F_{k,l}$ by CLT, that plays a key role to derive (19)-(21). Since $x_{k,l}$ and $z_{k,l}$ involve only harmonic cosine functions, the second equalities in (19)-(21) can be shown for $0 < n \ll T$ and $0 < p \ll M$, which is usually the case in cepstrum modeling. From (19)-(21), we have the following FIM:

$$\mathcal{I}(\theta^{ns},\theta^s) \triangleq \begin{bmatrix} \mathbf{I}_{R_{ns}\times R_{ns}} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \mathbf{I}_{(R_s-1)\times(R_s-1)} \end{bmatrix}, \tag{22}$$

where the identity matrix with a size of $R_{ns}\times R_{ns}$ in the upper block is associated with cepstral coefficients in Θ_{ns}, defining $\mathcal{I}_{11}(\theta^{ns})$. The entry in the center and the identity matrix in the lower block are associated with cepstra in Θ_s, giving $\mathcal{I}_{22}(\theta^s)$. Particularly, the entry in the center corresponds to the cepstral coefficient at the origin, i.e., $(t,v) = (0,0)$. From (14), we have a non-centrality parameter of

$$\Delta_S = TM\cdot(\theta^{ns})^T\mathbf{I}_{R_{ns}\times R_{ns}}(\theta^{ns}). \tag{23}$$

4.2.2. *Case II: general development with signal components*

We now reconsider two signal components. For convenience, $\tilde{\beta}$ and $\tilde{f}$ are decomposed into their real and imaginary parts, $\tilde{\beta} \triangleq \tilde{\beta}^R + j\tilde{\beta}^I$ and $\tilde{f} \triangleq \tilde{f}^R + j\tilde{f}^I$. Re-parameterization of the approximate negative log-likelihood function takes the form of $2\ell(\tilde{\beta}^R, \tilde{\beta}^I, \tilde{f}^R, \tilde{f}^I, \theta^s, \theta^{ns})$. Differentiating this function $2\ell(\cdot)$ with respect to each part of $\tilde{f}_l$ and θ_i^{ns}, and then taking the expectation of it yield the following results:

$$\mathbf{E}\left[\frac{\partial^2 2\ell}{\partial \tilde{f}_l^R \partial \theta_i^{ns}}\right] = 0, \qquad \mathbf{E}\left[\frac{\partial^2 2\ell}{\partial \tilde{f}_l^I \partial \theta_i^{ns}}\right] = 0, \tag{24}$$

where, e.g., $\tilde{f}_l^R$ is the l-th element of $\tilde{f}^R$. Similarly, we obtain the following results for $\tilde{\beta}_l^R$ and $\tilde{\beta}_l^I$:

$$\mathbf{E}\left[\frac{\partial^2 2\ell}{\partial \tilde{\beta}_l^R \partial \theta_i^{ns}}\right] = 0, \qquad \mathbf{E}\left[\frac{\partial^2 2\ell}{\partial \tilde{\beta}_l^I \partial \theta_i^{ns}}\right] = 0. \tag{25}$$

Now, combining (22), (24), and (25) yields the following FIM:

$$\mathcal{I}(\theta^{ns}, \theta^s, \tilde{\beta}, \tilde{f}) \triangleq \begin{bmatrix} \mathbf{I}_{R_{ns}\times R_{ns}} & 0 & 0 \\ 0 & \mathcal{I}(\theta^s) & \mathcal{I}(\theta^s, \tilde{\beta}, \tilde{f}) \\ 0 & \mathcal{I}(\tilde{\beta}, \tilde{f}, \theta^s) & \mathcal{I}(\tilde{\beta}, \tilde{f}) \end{bmatrix}, \tag{26}$$

where $\mathbf{I}_{R_{ns}\times R_{ns}}$ defines $\mathcal{I}_{11}(\theta^{ns})$ and the remaining partition in the lower center block specifies $\mathcal{I}_{22}(\theta^s, \tilde{\beta}, \tilde{f})$. Note that the other partitions are all zero, giving $\mathcal{I}_{12}(\theta^{ns}, \theta^s, \tilde{\beta}, \tilde{f}) = 0$.

From (14), the non-centrality parameter is given by

$$\Delta_S = TM \cdot \sum_{i=1}^{R_{ns}} (\theta_i^{ns})^2, \tag{27}$$

indicating that the non-centrality parameter is independent of signal components ($\tilde{\beta}$ and $\tilde{f}$) and cepstral coefficients in Θ_s for a given R_{ns}. From (8), it can be shown that Δ_S leads to

$$2\Delta_S = \left(\frac{1}{2\pi}\right)^2 \iint (\log F(\omega,\lambda) - \log F_S(\omega,\lambda))^2 \mathrm{d}\omega \mathrm{d}\lambda. \tag{28}$$

Thus, $2\Delta_S$ is an Euclidean metric between the logarithms of a non-separable PSD, $F(\omega,\lambda)$ and a separable PSD, $F_S(\omega,\lambda)$ [7]. Since $\theta_{0,0}$ corresponds to the amplitude of a PSD and other $\theta_{t,v}$s are associated with the shape of that PSD, we can recognize the independence of Δ_S and $\theta_{0,0}$ from (28).

The asymptotic alternative distribution of $\mathcal{L}_S$ is given by

$$\mathcal{L}_S \sim \chi^2_{R_{ns},\Delta_S}, \tag{29}$$

where $\chi^2_{R_{ns},\Delta_S}$ is a non-central chi-square distribution with R_{ns} degrees of freedom and Δ_S non-centrality parameter in (27). We formulate the power, probability that the proposed separability test detects non-separability when H_1 is true. The power function takes the form of

$$P_{Sep}(\theta^{ns}) = 1 - \Psi_{R_{ns},\Delta_S}\left(\Phi^{-1}_{R_{ns}}(1-\alpha)\right). \tag{30}$$

Assuming only one non-zero cepstral coefficient in Θ_{ns}, an example plot of $P_{Sep}(\theta^{ns})$ is placed on Fig.1 for $\alpha = 0.05$, $R_{ns} = 2540$, $T = 99$, and $M = 1435$, the same setup for the human dataset used in [1].

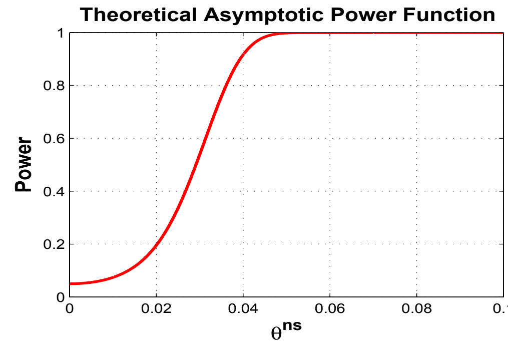

Fig. 1. An example plot of $P_{Sep}(\theta^{ns})$ for $\alpha = 0.5$.

For a given model order (n, p), some important remarks on the asymptotic power can be drawn. Firstly, the asymptotic power is independent of nuisance parameters, e.g., activation amplitudes. Secondly, the asymptotic power is independent of the locations of cepstral coefficients in Θ_{ns} and only affected by the values of cepstral coefficients in Θ_{ns}.

5. CONCLUSIONS

We analyzed the asymptotic power of the recently proposed test for space-time separability. The asymptotic power of the test procedure was not dependent on nuisance parameters, e.g., activation amplitudes. It was only dependent on cepstral coefficients in the non-separable region and was independent of the locations of cepstral coefficients. The non-centrality parameter of the asymptotic power was a scaled Euclidean distance between the logarithms of a non-separable PSD and a separable PSD.

6. REFERENCES

[1] J. Noh and V. Solo, "Testing for space-time separability in functional MRI," in *Proc. IEEE ISBI*, 2007, Washington DC, USA, pp. 412–415.

[2] G.M. Boynton, S.A. Engel, G.G. Glover, and D.J. Heeger, "Linear systems analysis of functional magnetic resonance imaging in Human V1," *Journal of Neuroscience*, vol. 16, No.13, pp. 4207–4221, 1996.

[3] D.R. Brillinger, "Time Series : Data Analysis and Theory," *San Francisco, CA : Holdenday*, 1981.

[4] V. Solo, "Modeling of two-dimensional random fields by parametric cepstrum," *IEEE Trans. on Information Theory*, vol. IT-32, No.6, pp. 743–750, Nov. 1986.

[5] K. Tamaki, "Second order asymptotic properties of a class of test statistics under the existence of nuisance parameters," *Sci. Math. Jpn.*, vol. 61, No.1, pp. 119–143, 2005.

[6] T. Hayakawa, "The likelihood criterion for a composite hypothesis under a local alternative," *Biometrika*, vol. 62, No.2, pp. 451–460, Aug. 1975.

[7] R.J. Martin, "A metric for ARMA processes," *IEEE Trans. on Signal Processing*, vol. 48, No.4, pp. 1164–1170, 2000.