Proceedings of the 4th International Conference on Statistics: Theory and Applications (ICSTA'22)
Prague, Czech Republic – July 28- 30, 2022
Paper No. 115
DOI: 10.11159/icsta22.115

Phase Distributions of Complex Multitaper Transfer Function Estimates

Skyepaphora Griffith[1], Glen Takahara[1], Wesley S. Burr[2]
[1]Queen's University
99 University Ave., Kingston, Canada
skye.griffith@queensu.ca; takahara@queensu.ca
[2]Trent University
1600 W Bank Dr., Peterborough, Canada
wesleyburr@trentu.ca

Abstract – In time series analysis, trends are often studied via sequences of observations taken with respect to, as the name suggests, time. Time series regression models are a classic way of viewing a response time series as a function of one or multiple predictor time series. However, the default assumptions of these models fail to account for the data's temporal trends. By instead building a complex regression model in the frequency domain, these assumptions can be relaxed, providing insight into what coherency is present in the model. The coefficient of the resulting complex regression model is known as a transfer function of frequency. There are several techniques used to estimate transfer functions, but all are subject to the bias-variance trade-off occurring as a byproduct of transformation to the frequency domain. Multitaper spectrum estimation has been shown to minimize spectral leakage (broad-band bias) while providing flexibility in terms of bandwidth selection (variance). Thus, exploration of Multitaper Transfer Function Estimators (MTFEs) is an alluring topic of research. Previous work has explored distribution theory for the modulus of MTFEs, in addition to MTFE variance across simulations, and has revealed effective methods of signal detection more robust to frequency modulation than classic alternatives. However, the distribution of an MTFE's phase has not been explored. This paper demonstrates that for models whose underlying noise is stationary and Gaussian, the MTFE at a given frequency is distributed as a complex Gaussian random variable. From the perspective of the time domain, one can infer parameters of the phase distribution of the MTFE by way of estimating the autocovariance function of the response. This phase distribution provides information which may be useful for signal detection.

***Keywords*:** Multitaper, Spectrum Estimation, Time Series, Complex Distributions, Signal Detection, Regression, Transfer Functions, Stationarity

1. Introduction

The multitaper transfer function model is a modern approach to viewing the world of time series regression – one which examines the frequency structure of its predictors and response without making restrictive assumptions with regards to their stochastic components in the time domain. Examining the series' transfer function model from the frequency domain is an effective way of bypassing the omission of temporal trends in the data, and multitaper transfer function estimators have desirable statistical properties attributed to the multitaper method of spectrum estimation. The resulting complex regression model's coefficient is the estimated multitaper transfer function of frequency, after which the model is named.

The distribution of Multitaper Transfer Function Estimators (MTFEs,) assuming they are obtained from a model whose error terms are Gaussian distributed, has been found to be complex Gaussian. Attributes of MTFEs – such as modulus, phase, and variance across simulations – depend on the periodic signals contained within the time series which make up the model. For instance, MTFE moduli increase on average at frequencies where response eigencoefficients detect a signal, and when the predictor eigencoefficients detect a signal the variance of the MTFE across simulations will contract proportionally. Test statistics have been explored in relation to these behaviours [4], but the phase (the counter-clockwise angle with the real axis) of the complex-valued MTFE has not yet formally been considered.

The behaviour of the MTFE phase distribution shows promise in providing further insight into the utility of transfer function models as a technique for detecting coherency between time series. The extent to which a corresponding test statistic is robust to factors which corrupt more classical methods, and whether said statistic would be viable despite the MTFE's frequency non-stationarity, is the topic of future research.

2. Background

2.1. Time Series Regression, Transfer Functions, and the Multitaper Framework

When working with data taken at equal time increments, it is natural to relate a series of effects to predictive components by performing statistical regression across time ordered observations. The typical time series regression model does exactly this, expressing a response time series $y(t)$ as a function of one or multiple predictor time series $x(t)$, and in doing so assumes error terms are uncorrelated with respect to time. This assumption is inappropriate in the face of temporal trends; noise processes are rarely stationary in the strong sense - certainly not in the presence of any meaningful signal.

For this reason, there are advantages to working with alternatives which can account for such temporal trends in the error structure. One such class is linear filter models, which convolve the predictor series with a sequence of filter coefficients h_τ, as shown in Eq 1. However, the filter coefficients responsible for correcting for the naïve regression model's assumption of temporal uncorrelation are difficult to interpret from the perspective of the time domain. By transforming the linear filter model via the data's spectral representations, a more intuitive model becomes available in the frequency domain: a complex regression model relating transforms of the original data.

$$\boldsymbol{y(t) = \sum_{\tau=-\infty}^{\infty} h_\tau x(t-\tau)} \tag{1}$$

To elaborate, the model of interest relates the Fourier transform of the response time series to that of the predictor time series by way of a complex transfer function [1] of frequency, denoted $H(f)$. The simplest transfer function model applies these Fourier transforms to the original data. However, it is important to note that the frequency domain representations of these data are themselves only estimates and are therefore subject to the bias and variance attributes central to spectrum estimation. Fourier transforms of data are prone to broad-band bias in the form of spectral leakage: large concentrations of energy existing outside a desired frequency bandwidth. This can be accounted for by using a robust approach: the multitaper method [2], which applies strategic tapers (sequences of weights) to time series data before transformation to the frequency domain (Eqs 2-4), has been shown to minimize broad-band spectral leakage whilst maintaining desirable variance properties. This approach uses a family of orthogonal tapers, which in the traditional multitaper algorithm are Slepian sequences [3], denoted as tapers $v_0(t)$ through $v_{K-1}(t)$ for all time points t. This then allows for an estimation of transfer functions through regression of the K complex valued *eigencoefficients* (tapered and transformed spectral objects) of the response against the predictor (Eq 4). In Eqs 2-3, N is the length of the response and predictor time series.

$$\boldsymbol{X_k(f) = \sum_{t=0}^{N-1} v_k(t) x(t) e^{-i2\pi f t}} \tag{2}$$

$$\boldsymbol{Y_k(f) = \sum_{t=0}^{N-1} v_k(t) y(t) e^{-i2\pi f t}} \tag{3}$$

$$Y_k(f) = H(f)X_k(f) \tag{4}$$

The transfer function $H(f)$ itself is a promising avenue of insight into time series regression models whose errors should not be assumed to be uncorrelated. Exploration of these functions with regards to their distributional properties and potential for signal detection has been limited, and largely restricted to cases of sinusoidal signals embedded in Gaussian white noise (see [4]). The remainder of this paper maintains an assumed Gaussian structure of error terms, but where indicated, relaxes the assumption of white noise to include weakly stationary time series.

2.2. Distributional Properties of Multitaper Transfer Function Estimators

Given a time series with underlying noise process distributed as Additive White Gaussian Noise, it can be easily shown that the corresponding set of eigencoefficients are distributed as complex Gaussian (Eq 5). Consider, for instance, the time series $x(t) = \xi_{x,f}(t) + w(t)$, where $\xi_{x,f}$ is a periodic component with frequency f, and w is a white noise process with variance σ_x^2. From Eqs 2-3, the eigencoefficients are complex linear combinations of Gaussian random variables and therefore complex Gaussian themselves. Below, $\gamma_x(h)$ denotes the autocovariance function (ACVF) of $x(t)$ at lag h.

$$X_k(f) \sim N_C\left(\Xi_{x,k}(f), \sigma_{x,k}^2(f), V_{x,k}(f)\right), \tag{5}$$

$$\Xi_{x,k}(f) = \sum_{t=0}^{N-1} v_k(t)\xi_{x,f}(t)e^{-i2\pi ft}, \tag{6}$$

$$\sigma_{x,k}^2(f) = \sum_{\substack{s=0 \\ t=0}}^{N-1} \left(v_k(s)v_k(t)\gamma_x(s-t)\cos\big(2\pi f(s-t)\big)\right), \tag{7}$$

where N_C is the complex Gaussian distribution, and $V_{x,k}(f)$ is the pseudovariance [6] of $X_k(f)$ (this is negligible for $\hat{H}$ and $X_k(f)$, as these objects are roughly circularly symmetric under the assumptions of $x(t)$ [4]). By extension, MTFEs obtained from such series are complex Gaussian as well. Distributional properties of the MTFE such as expectation and variance (Eqs 8-10), have also been derived [4].

$$\hat{H}(f) = \sum_{j=0}^{K-1} X_j^*(f)Y_j(f) \Big/ \sum_{k=0}^{K-1} |X_k(f)|^2 \tag{8}$$

$$E\big[\hat{H}(f)\big] = \sum_{j=0}^{K-1} X_j^*(f)\Xi_{y,j}(f) \Big/ \sum_{k=0}^{K-1} |X_k(f)|^2 \tag{9}$$

$$\mathrm{Var}\left[\hat{\boldsymbol{H}}(f)\right] = \frac{\mathbf{1}}{|X(f)|^2}\left(\sum_{\substack{j=0\\k=0}}^{K-1} X_k^*(f)X_j(f)\sum_{\substack{s=0\\t=0}}^{N-1}\left(v_k(s)v_k(t)\gamma_x(s - t)\cos\big(2\pi f(s-t)\big)\right)\right) \quad (10)$$

where the denominator in Eq 11 represents the dot product of the eigencoefficient vector $X = [X_0, \dots, X_{K-1}]^T$ with its own conjugate transpose.

2.3. Potential Test Statistics for Signal Detection

Some attributes of MTFEs have shown promise for development of test statistics for signal detection, and their distributional properties have therefore been studied as well [4]. In particular, the moduli of MTFEs have been shown to be Rice distributed with large expected values within a chosen bandwidth centered about frequencies at which a signal is present in the response. This phenomenon can be shown empirically by simulating transfer functions and observing their average moduli at Fourier frequencies of interest. In the context of real-world data, the moduli of a single transfer function realization can provide similar insight when averages are instead taken according to a sliding window of frequencies.

A test statistic for response signal presence, denoted T1, was developed as a function of squared MTFE moduli after having transformed the MTFE to a vector of uncorrelated entries [4]. The performance of T1 was then compared to classic alternative tests such as Magnitude Squared Coherence (MSC) and the Harmonic F-test [2]. Notably, it was found to be more robust to the case of frequency structure of the response's stochastic component coinciding with the predictor's purely by chance. While the notion of avoiding "coherent noise" detection may sound paradoxical, the event can be manufactured by producing response eigencoefficients as a convex linear combination of predictor eigencoefficients and complex white noise, weighted according to some chosen factor λ between 0 and 1, inclusive [4]. As λ approaches 1, the noise components of the original series are forced to be increasingly coherent, and this is done without contaminating the series' deterministic components. In this context, detection of the original signal frequency is desired, but the MSC will detect coherency at other frequencies and ultimately struggle to distinguish between coherent signals and coherent noise, especially for large λ. The robustness of T1 to this phenomenon is advantageous relative to MSC in such contexts. Moreover, T1 was shown to be more robust to frequency modulation than the Harmonic F-test: a powerful technique which detects line components via the multitaper method [2, 4].

The variance of MTFEs is also more robust to frequency modulation than the F-test and shows further promise in the development of test statistics for signal detection. It can be shown that the variance of the transfer function is inversely proportional to the spectrum of the predictor time series (Eq 10). This provides a powerful indicator of predictor signals by frequency when sample variance is taken across simulated realizations, as discussed in the context of transfer function moduli. However, unlike in the case of moduli, this approach is problematic to be modified to be accessible from a single realization due to the MTFE's inherent lack of frequency stationarity. In essence, the transfer function estimate at a given frequency is highly dependent on the estimates at surrounding frequencies; while a sliding average of moduli is impervious to this phenomenon, a sliding window's perspective of variance structure is almost inevitably corrupted.

3. Multitaper Transfer Function Phase Estimates

3.1. Distribution of a Complex Gaussian Random Variable's Phase

Consider complex Gaussian random variables (CGRVs). The modulus of a CGRV, in connection with the moduli discussed in section 2, follows a Rice distribution distinguished by the modulus of the CGRV's mean along with the standard deviations of the CGRV's real and imaginary parts. The distribution of a CGRV's phase (also referred to as its argument and denoted by ϕ) is not so straightforward.

To understand the relationship between a CGRV's distribution and that of ϕ, consider the shape and position of the 2-dimensional point-cloud formed by a sample from the CGRV's distribution on the complex plane (as presented in Fig 1). The phases of the points in this cloud are firstly dependent on the location of the cloud's center: the mean of the CGRV. The further the center is from the origin, the further the distribution of ϕ departs from a uniform distribution on -π to π. The distribution of ϕ is secondly influenced by the shape of the cloud, and this influence is particularly prominent when generalizing beyond the spherical cloud associated with CGRVs whose real and imaginary parts are both uncorrelated and of equal variance.

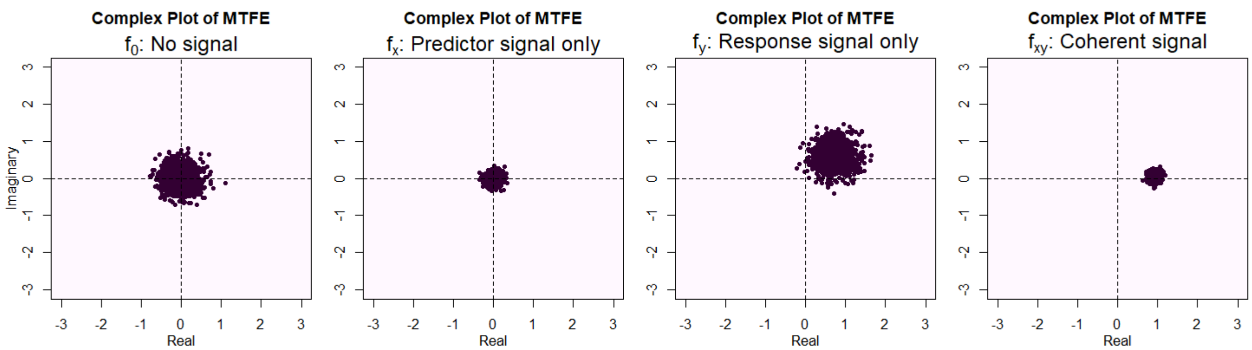

Figure 1: Plotted 2-dimensional point clouds of Complex Gaussian Random Variables. The featured CGRVs are obtained from MTFEs at select frequencies associated with different signal types, including no signal. The MTFE's original predictor and response time series are simulated AR(1) processes with embedded sinusoids at frequencies f_x,and$f_y, respectively$. Both time series additionally contain a signal at f_{xy}.

Let X and Y be the real and imaginary components of a CGRV. When the variances of X and Y are not equal, the cloud becomes elliptical in shape, with major axis running parallel to the real (when Var(X) > Var(Y)) or imaginary (the opposite) axes, respectively. Depending on which quadrant contains the bulk of the cloud, this distortion influences the skewness and variance of ϕ's distribution. Furthermore, the correlation coefficient ρ of X and Y is responsible for a rotation of the cloud about its mean, which is similarly influential on the skewness and variance of ϕ.

Thus, ϕ's distribution can be clearly defined using five parameters: μ_x- the mean of X; μ_y - the mean of Y; σ_x^2- the variance of X; σ_y^2- the variance of Y; and ρ - the correlation between X and Y [5]. For the purposes of this paper, we have assumed ρ to be zero and thus all terms which would include ρ have been omitted. Note that this is not a strong assumption, as approximate independence of the real and imaginary components is the default case, especially for the spectral objects being considered. The pdf $g(\phi)$ corresponding to this distribution is given in Eqs 11-12.

$$g(\boldsymbol{\phi}|\boldsymbol{\mu_x}, \boldsymbol{\mu_y}, \boldsymbol{\sigma_x^2}, \boldsymbol{\sigma_y^2}, \boldsymbol{\rho} = \mathbf{0})$$

$$= \frac{\boldsymbol{\sigma_x \sigma_y} \exp\left\{-\left(\frac{\boldsymbol{\mu_x^2 \sigma_y^2 + \mu_y^2 \sigma_x^2}}{\boldsymbol{2\sigma_x^2 \sigma_y^2}}\right)\right\} \left(\mathbf{1} + \sqrt{\boldsymbol{\pi}}\boldsymbol{A}(\boldsymbol{\phi})\exp\{\boldsymbol{A}(\boldsymbol{\phi})^{\mathbf{2}}\}\text{erfc}[-\boldsymbol{A}(\boldsymbol{\phi})]\right)}{\boldsymbol{2\pi}\left(\boldsymbol{\sigma_y^2 cos^2}(\boldsymbol{\phi}) + \boldsymbol{\sigma_x^2 sin^2}(\boldsymbol{\phi})\right)} \tag{11}$$

$$\boldsymbol{where}\ \ \boldsymbol{A}(\boldsymbol{\phi}) = \frac{\boldsymbol{\mu_x \sigma_y^2 cos}(\boldsymbol{\phi}) + \boldsymbol{\mu_y \sigma_x^2 sin}(\boldsymbol{\phi})}{\boldsymbol{\sigma_x \sigma_y}\sqrt{\mathbf{2}\left(\boldsymbol{\sigma_y^2 cos^2}(\boldsymbol{\phi}) + \boldsymbol{\sigma_x^2 sin^2}(\boldsymbol{\phi})\right)}} \tag{12}$$

Note that erfc is the complementary Gauss error function, hence $\text{erfc}(z)$ is $\text{erf}(z)$ subtracted from 1.

3.2. Distribution of a Multitaper Transfer Function Estimator's Phase

The known distribution of a CGRV's phase can be extended to that of MTFEs. Under the continued assumption that the MTFE is obtained from processes whose stochastic components are Gaussian and weakly stationary, the cloud maintains a spherical shape. The first two parameters of the phase distribution are thus given by the real and imaginary parts of the mean of the MTFE itself, with the third and fourth parameters each half the value of the MTFE's variance (more in Section 3.3). The pdf g of this distribution is notated in terms of these parameters in (Eq 13).

$$\boldsymbol{g}\left(\boldsymbol{\phi}\middle|\boldsymbol{\mu_x} = \text{Re}\{\boldsymbol{E}[\hat{\boldsymbol{H}}(\boldsymbol{f})]\}, \boldsymbol{\mu_y} = \text{Im}\{\boldsymbol{E}[\hat{\boldsymbol{H}}(\boldsymbol{f})]\}, \boldsymbol{\sigma_x^2} = \boldsymbol{\sigma_y^2}\ \frac{\textbf{Var}[\hat{\boldsymbol{H}}(\boldsymbol{f})]}{\mathbf{2}}, \boldsymbol{\rho} = \mathbf{0}\right) \tag{13}$$

A simple simulation study can demonstrate the validity of this derivation in practice. Figure 2 shows M = 10,000 simulations at select frequencies associated with: (1) no signal; (2) a predictor-only signal; (3) a response-only signal; (4) a coherent signal, with the overplotted lines being the theoretical distribution via (Eq 13), and the histograms representing the replicates.

3.3. Insight via Autocovariance in the Time Domain

The parameters defining the ϕ distribution are obtained from the MTFE directly, as this is the CGRV for which the phase distribution is sought. However, in the context of utilizing MTFE attributes to examine periodic trends in time series regression models, a way of determining these parameters exclusively from time-domain objects is desired.

Luckily, as seen in (Eq 9), the mean of an MTFE is proportional to a function of the eigencoefficients of the purely deterministic components of the response time series. Knowledge of what signals exist in the response – estimable via, for instance, the harmonic F-test – is enough to determine the mean of an MTFE, and by extension, the first two parameters of its phase distribution.

The third and fourth parameters are taken from the variance of the MTFE's real and imaginary parts, respectively. Recall that under the assumptions of stationary Gaussian noise, any distortion of the MTFE cloud to an ellipse will be negligible, along with the asymmetry factor ρ. Thus, by the properties of variance for complex random variables, the third and fourth parameters will each be half the value of the MTFE's overall variance. From the time domain, we can determine MTFE variance using the ACVF of the response (Eq 10). The phase distribution obtained from a known (theoretical) response ACVF is compared to histograms of simulated MTFE arguments in (Fig 3, top panels).

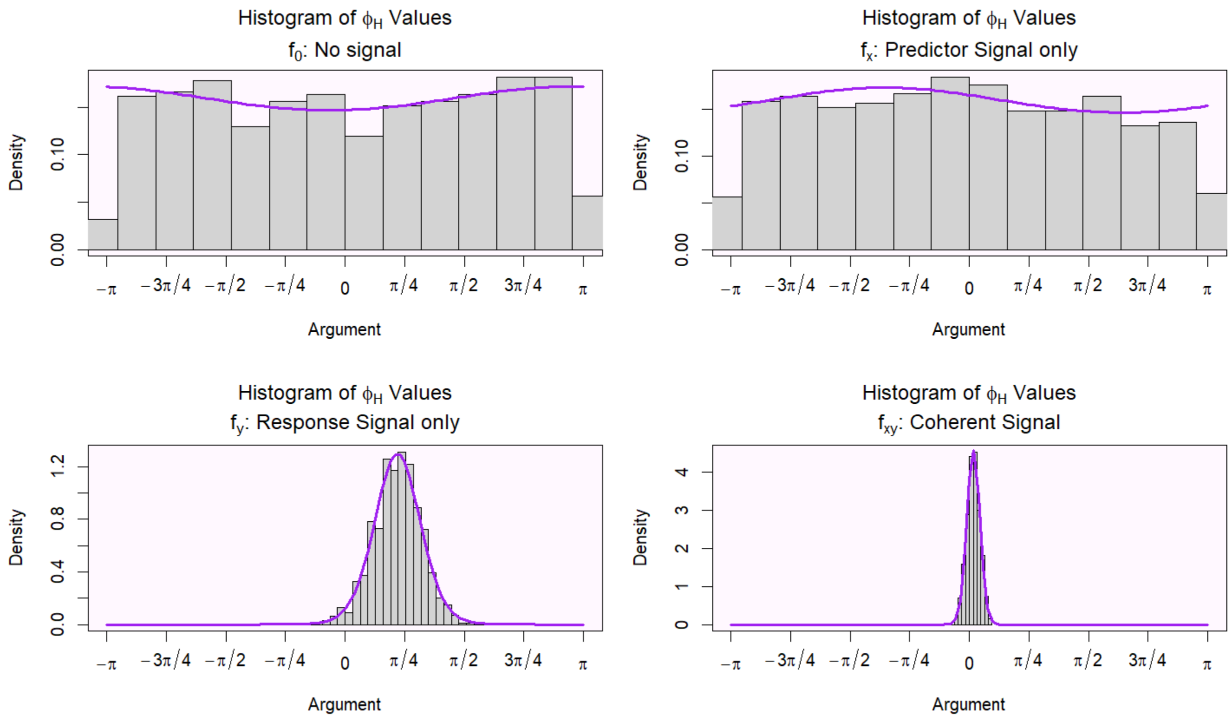

Figure 2 : Probability density function of ϕ, plotted in purple against histograms of phases taken from simulated MTFEs at various frequencies of interest. The parameters for the pdf are taken directly from the MTFEs.

The next concern is whether this technique remains reliable when the deterministic components of the response – and by extension, the response ACVF – are unknown. Figure 3 again compares the same set of histograms to the density function of the phase distribution obtained from the estimated ACVF of the response time series (Fig 3, bottom panels). The agreement is quite good, indicating that the interpretation matches.

3.4. Significance and Potential Utility for Signal Detection

As discussed in Section 2, there has been limited research into the development of MTFE-based test statistics for signal detection. Previous work [4] led to proposal of statistics T1 and T2, functions of the modulus and variance of a MTFE, but neither provided insight into the frequency structure of *predictors*. This work suggests that a new statistic, say, T3, may be able to leverage MTFE phase properties to aid in the detection of coherent signals between the predictor and the response. At response signal frequencies, the MTFE cloud is pulled away from the origin, causing ϕ's distribution to depart from a uniform distribution on $-\pi$ to π, as discussed in Section 3.1. At predictor signal frequencies, the MTFE cloud contracts, thereby reducing the variance of ϕ's distribution. Together, these properties show promise in revealing which frequencies are associated with coherent signals between response and predictor time series.

Unfortunately, as the modulus of H increases, it will likely become more difficult to distinguish between phase distributions of coherent and response-only signals: as the MTFE cloud moves away from the origin, the variance of ϕ's distribution will shrink regardless of the predictor's frequency structure. As it stands, if a coherent signal is detected by the hypothetical T3 statistic, the corresponding frequency of interest could just as easily be attributed to a high amplitude signal in the response and no signal in the predictor. Similarly, while coherency is a sufficient condition for an argument distribution to be centered at zero, it is not *necessary*, and such a distribution can indeed occur at response-only signal frequencies. Thus, unless the frequency structure of either the response or the predictor is already known, there is currently no formal way to distinguish between these phenomena.

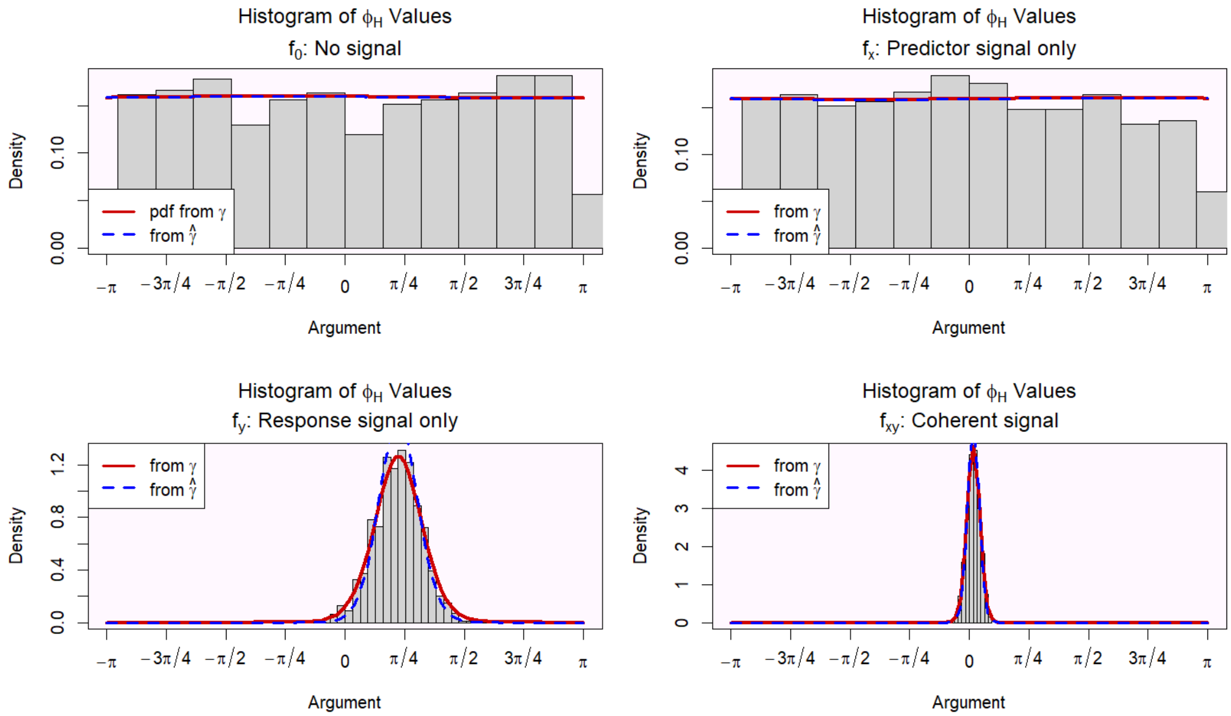

Figure 3: Estimated density functions for ϕ, plotted against the same histograms featured in figure 2. The red curve estimates parameters 3 and 4 using the theoretical ACVF used to produce the MTFE. The blue (dotted) curve estimates parameters 3 and 4 using the estimated ACVF of the response, calculated in the time domain. Both curves use the phase of the MTFE's mean to estimate parameters 1 and 2. ρ is assumed to be zero.

Moreover, it is unclear how robust the hypothetical T3 statistic would be. Further research is required to determine whether T3 would be corrupted by the MTFE's non-stationarity, as was the case in the previously proposed T2. If not, then the effectiveness of using T1's "sliding window" technique to gather expected phase values should be explored. Finally, to justify T3 as a useful alternative, the harmonic F-test (recall that the F-test is capable of detecting coherent line components if applied separately to the predictor and response) performance of T3 must examined in contexts which render the F-test unviable, such as the presence of frequency modulated signals.

4. Conclusion

In general, the residuals of a time series regression model are assumed to be white noise distributed: that is, no correlation is assumed to exist between observations as a function of time. Under these conditions, any periodic components present in the time series are overlooked. The multitaper transfer function model instead takes the form of complex regression in the frequency domain, where the predictors and response are strategically weighted Fourier transforms of the original predictor and response time series, respectively. As per the multitaper method, these transforms are represented by a vector of eigencoefficients, and the regression coefficient is the MTFE itself.

Under the assumptions outlined in this paper, the MTFE is distributed as complex Gaussian. This is on account of the MTFE being a complex linear combination of eigencoefficients, which are themselves complex linear combinations of the Gaussian stochastic components underlying the original time series.

Upon examination, the phase of the MTFE is derived to be exactly that of a complex Gaussian random variable (CGRV) and can be determined by estimating the autocovariance function (ACVF) of the response from the perspective of the time domain. The potential of this behaviour in the context of signal detection, and particularly in relation to alternative, classical techniques, is yet to be explored.

Acknowledgements

Thanks to David J. Thomson for many fruitful discussions of aspects of this work.

References

[1] Brockwell, P.J. and Davis, R.A., 2009. *Time series: theory and methods*. Springer Science & Business Media.

[2] Thomson, D.J. “Spectrum Estimation and Harmonic Analysis.” *Proceedings of the IEEE* 70.9 (1982): 1055-1096.

[3] Slepian, D. "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case." *Bell System Technical Journal* 57.5 (1978): 1371-1430.

[4] Griffith, S. "Multitaper Transfer Function Estimates and the Hunt for Coherent Signals Trespassing on Time Series Regression Models". MSc. thesis, Queen's Univ., Kingston, ON (2020).

[5] Aalo, V.A., Efthymoglou, G.P. and Chayawan, C. "On the envelope and phase distributions for correlated Gaussian quadratures." *IEEE Communications Letters* 11.12 (2007): 985-987.

[6] Gallager, Robert G. "Circularly-symmetric Gaussian random vectors." preprint 1 (2008).