Second Order Statistics of Stationary Complex Valued Time Series and of Real-Complex Valued Mixtures

Alik Mokeichev and Ohad Ben-Shahar

Department of Computer Science and the Zlotowski Center for Neuroscience, Ben-Gurion University, Beer Sheva, Israel

Abstract

Here we study wide sense stationary complex-valued multivariate time series in their frequency domain representation, and provide sufficient descriptors for their second order statistics. For the general case of complex valued multivariate time series, we show that the power spectral densities and cross power spectra of the variates, augmented with a single univariate complementary spectral function, provide a sufficient description of its second order statistics, irrespective of the number of time variates. For the case of mixtures of real valued and complex valued stationary time series, we show that their joint second order statistics are sufficiently described by their by their power spectral density and cross-power spectra, or equivalently, their auto-correlation and cross-correlation functions, similarly to as it is for real valued processes. The proofs are constructive and provide explicit expressions, relating the power spectra and cross-power spectra of such mixtures under their real valued and complex valued representations. Most of our results are readily extended to complex valued infinite time series and continues time signals.

1 Introduction

The analysis of second order statistics of finite univariate or multivariate time series through their correlation functions or spectral functions, have been proven to be fruitful in wide range of disciplines, and few examples include studies in diverse fields such as geophysics [Chang et al. 1997, Shapiro et al. 2005, White and Peterson 1996], neuroscience [Mokeichev et al. 2007, Pillow et al. 2007],and economics [Plerou et al. 1999, Ramchand and Susmel 1998]. In some cases, as for geophysical data analysis [Horel 1984], measurements might be most naturally represented as complex valued time series and examples include the analysis of wind profiles [Mandic et al. 2009] ,ocean surface currents [Kaihatu et al. 1998] or waves [Mooers 1973]. However, in spite of an ongoing progress in probabilistic theory of complex valued random variables and processes, to our best knowledge, a concise treatment of second order statistics of stationary complex valued time series hasnt been provided yet. In the following, we address wide sense stationary complex-valued multivariate time series and provide sufficient descriptors for their second order statistics.

We'll begin with a brief presentation of definitions of second order statistics of real valued random variables and random processes and their recently developed complex valued extensions, to serve as a basis for our subsequent derivations for complex valued time series analysis. For simplicity, throughout the rest of the manuscript we assume that all random vectors and time series have zero mean. For a real valued random vector $x = \{x_1, x_2, ..., x_m\}$, its second order statistics are typically characterized by the covariance matrix $E[xx^T]$, where x^T is the transpose of x [Reinsel 1997]. For a similar treatment of a complex vector-valued random variable, $z = x + iy$, its second order statistics might be captured through a real valued bivariate representation of z where the covariance matrix takes the form [Schreier and Scharf 2003]:

$$E\begin{bmatrix} xx^T & xy^T \\ yx^T & yy^T \end{bmatrix} \in \mathbb{R}^{2m\times 2m}, \tag{1}$$

which express the covariance between the real parts of the variates of z, the imaginary parts of the variates of z, and also between its real and imaginary parts. For some natural phenomena such as waves, a complex valued representation might carry insights which could be obscured by resorting to a real valued representation, and especially when multi-dimensional measurements are analyzed. Therefore, a treatment of second order statistics which is based on complex-valued representation is desirable. However, for a complex valued random vector, $z = \{z_1, z_2, ..., z_m\}$, with asymmetric probability distribution [Neseer and Massey 1993, Ollila 2008, Picinbono and Bondon 1997, Schreier and Scharf 2003](i.e. z and $ze^{i\theta}, \theta \in \{-\pi, .., \pi\}$, have the same second order statistics [Wahlberg and Schreier 2008]) the hermitian covariance matrix $E[zz^*]$, where z^* is the conjugate transpose of z, doesn't carry the complete second order statistic information [Neseer and Massey 1993, Picinbono and Bondon 1997, Schreier and Scharf 2003, Wahlberg and Schreier 2008]. For a complete description of the second order statistics of z, the complementary covariance matrix $E[zz]$ should be considered as well [Schreier and Scharf 2003]. We note about the use of terminology of 'complete' description of second order statistics of complex valued random variables and processes. In a strict sense, like for real valued random vectors, a complete second order statistical description of a complex valued random vector $z = \{z_1, z_2, ..., z_m\}$ is provided by the joint probabilities $p(z_j, z_k)$, $j, k = 1, ..., m$. In this sense, the covariance matrices $E[zz^*]$ and $E[zz]$ provide only partial second order statistical information, as the covariance matrices could be inferred from the joint or conditional probabilities but not vice versa. Still and all, numerous authors [Mandic et al. 2009, Neseer and Massey 1993, Picinbono and Bondon 1997, Schreier and Scharf 2003] adopt the terminology of

’complete’ description of second order statistics, which should be understood as referring to the description provided by the covariance matrix in (Eq. 1). In this respect, the covariance matrix $E[zz^*]$, provides only partial second order information. A more careful choice of terminology would be ’sufficient’ rather than ’complete’ statistical description, implying that unbiased estimates of the covariance matrix $E[zz^*]$ and of the complementary covariance matrix $E[zz]$ are *sufficient statistics* of (1).

For a multivariate continuous stochastic processes, $z(t) = \{z_1(t), z_2(t), ..., z_m(t)\}, t \in \mathbb{R}$, a recent study [Wahlberg and Schreier 2008] has shown that its second order statistics are sufficiently captured by the set of covariance and complimentary covariance matrices:

$$\begin{aligned} &E[z(t_1)z^*(t_2)] \in \mathbb{C}^{m\times m}, \\ &E[z(t_1)z^T(t_2)] \in \mathbb{C}^{m\times m}. \end{aligned} \tag{2}$$

While the above theoretical results about the complete second order statistics of continuous complex valued random processes could be extended for discrete time random processes, their utility in empirical studies might be limited, and that is for several reasons. The applicability of the above results is crucially dependent on the availability of multiple samples of the complex valued process being studied. In practice, only a single finite length realization of a time series might be available [Theiler et al. 1992, Witt et al. 1998], from which the covariance and the complementary covariance matrices might be hard to estimate. Another limitation is that even with the availability of multiple realizations of multivariate discrete time process, $Z(t) = \{z_1(t), z_2(t), ..., z_m(t)\}, t = 1, ..., n$, the second order description provided by the covariance and complementary covariance matrices in (2) is of a very high dimension and requires m^2n^2 scalar values, which might become practically infeasible for very large data sets.

The above limitation are not specific to the analysis of complex-valued time series but are common to real-valued time series as well. For the statistical analysis of real-valued time series, the above limitations could be bypassed if the time series under investigation is assumed to be wide sense stationary. A real valued multivariate time series $X(t) = \{x_1(t), x_2(t), ..., x_m(t)\}, t = 0, ..., n-1$,is said to be wide sense stationary if its variates have constant means,and if its product moments:

$$E[x_j(t), x_l^*(t+d)] = R_{x_j,x_l}(d), \quad j, l = 1, ..., m \tag{3}$$

are functions of d only [Parzen 1961]. In such a case, the functions $R_{x_j,x_l}(s)$, $j, l = 1, ..., m$,termed the cross-correlation (or cross-covariance) functions, provide a second order statistical description of X(t). In the above, $X(t)$ is real valued; however we make use of $x_l^*(t)$ rather than $x_l(t)$ in Eq.(3) as the definition is generally applicable to complex-valued time series as well.

For practical application, a major advantage is that estimates of the cross-correlations function $R_{x_j,x_l}(s)$ might be obtained from a single empirical realization of the time series $X(t)$. Notice also that these function are described with only $\sim m^2n$ scalars, comparing to $\sim m^2n^2$ in (2). Another important advantage, is that by the generalized discrete version of the Wiener-Khintchin theorem [Wiener 1930], the cross-power spectral densities of the variates of $X(t)$ could be expressed as the discrete time Fourier transform of their cross-correlation functions $R_{x_j,x_l}(d)$:

$$P_{x_j,x_l}(k) = \sum_{t=-\infty}^{\infty} R_{x_j,x_l}(t) e^{-2\pi ikt}, \quad j, l = 1, ..., n. \tag{4}$$

The above formulation (4), is very useful as it paves a way for harmonic analysis of stationary time series from its estimated correlation functions [Parzen 1961]. For practical applications, however, where the sampled time series if of finite length, a choice need to be made of how to estimate the correlation functions $R_{x_j,x_l}(d)$. An unbiased estimator is given by [Parzen 1964]:

$$\hat{R}^0_{x_j,x_l}(d) = \frac{1}{n-d} \sum_{t=0}^{n-d-1} x_j(t) x_l^*(t+d) \tag{5}$$

but its variance is greater than that of the biased estimator [Parzen 1964]:

$$\hat{R}^1_{x_j,x_l}(d) = \frac{1}{n} \sum_{t=0}^{n-d-1} x_j(t) x_l^*(t+d). \tag{6}$$

If we further assume that the sampled time series is periodic, than an unbiased estimator is:

$$\hat{R}^2_{x_j,x_l}(d) = \frac{1}{n} \sum_{t=0}^{n-1} x_j(t) x_l^*(t+d) \tag{7}$$

Without loss of generality, to simplify our exposition we'll assume that all time series under consideration are periodic.

As we show in the following, the above auto-correlation and cross-correlation functions of complex-valued wide sense stationary multivariate time series are insufficient, analogously to the case of complex-valued random vectors where their covariance matrix has been shown to provide only partial second order statistical description. Inspired by the previous work on continuous complex valued processes [Wahlberg and Schreier 2008], we define the complementary cross-correlation function, and complementary power-spectrum of finite stationary multivariate time series, and derive the complementary (discrete time) cross-correlation theorem, relating the time domain and frequency domain representations. In the rest, our derivation are done with a frequency domain representation of time series and analogously to previous works on random variables [Neseer and Massey 1993, Picinbono and Bondon 1997, Schreier and Scharf 2003], we show that the power spectrum of a *univariate* complex valued time series,augmented with the complementary power spectrum, provides a sufficient second order description. Interestingly, for a complex valued *multivariate* time series, just single complementary auto-correlation function, augmented to the auto-correlation and cross-correlation functions of its variates, is sufficient to provide the complete information of its second order statistics, irrespective of the number of variates.

Finally, we consider complex-valued multivariate time series and mixtures of complex-valued and real valued time series. In the later case,we show that if the collection of time series contains a real-valued time series, the ensemble second order statistics are entirely captured by the power spectra and cross-power spectra, similarly to the case of real-valued multivariate time series. Explicit expressions, relating the power spectra and cross power spectra of the real-valued multivariate representation and real/complex valued representations are provided.

2 Notations and Definitions

In many studies that employ a frequency domain analysis, the power spectra and cross-power spectra of time series under investigation are being estimated through the fast Fourier transform algorithm (FFT). However, various authors use slightly different definitions and software implementations of the discrete Fourier transform. Here we provide the definitions of discrete Fourier transform to be used in the rest of the manuscript, and corresponding consistent definitions of power spectra and correlation functions.

For a periodic real-valued or complex-valued univariate time series $x(t), t = 0, ..., n-1$, we denote its discrete Fourier transform (DFT) by $F_x(k)$ and its power spectrum by $P_{xx}(k)$ which are defined as:

$$F_x(k) = \sum_{t=0}^{n-1} x(t) e^{\frac{-2\pi i}{n} kt}, \tag{8}$$

And,

$$P_{x,x}(k) = F_x^*(k) F_x(k), \tag{9}$$

respectively. The inverse Fourier transform applied to $F_x(k)$ would be denoted by $\mathcal{F}^{-1}\{F_x(k)\}$ and defined as:

$$\mathcal{F}^{-1}\{F_x(k)\}(t) = \frac{1}{n} \sum_{k=0}^{n-1} F_x(k) e^{\frac{2\pi i}{n} kt} \tag{10}$$

For the auto-correlation function of $x(t)$, we omit the normalization factor $\frac{1}{n}$ from the definition in Eq. 7, and distinguish it with the notation C_{xx}:

$$C_{x,x}(d) = \sum_{t=0}^{n-1} x^*(t) x(t+d) \tag{11}$$

Here the sequence $\tilde{x}(t), t = 0, ..., +\infty$ is understood to be a periodic extension of $x(t)$ such that $\tilde{x}(t) = x(t \mod n)$. In the following we'll omit explicit distinctions between the finite time series $x(t)$ and its extension $\tilde{x}(t)$. With these definitions, we get the discrete time Wiener-Khinchin therem for finite and periodic time series:

$$P_{x,x}(k) = \sum_{t=0}^{n-1} C_{x,x}(t) e^{-2\pi i kt}, \tag{12}$$

Let $y(t)$ be another real-valued or complex-valued univariate time series. Then the corss-power spectrum P_{xy} of $x(t)$ and $y(t)$ is given by:

$$P_{x,y}(k) = F_x^*(k) F_y(k), \tag{13}$$

and the cross-correlations P_{xy} of $x(t)$ and $y(t)$ is given by:

$$C_{x,y}(d) = \sum_{t=0}^{n-1} x^*(t)y(t+d) \qquad (14)$$

3 Second order statistics of real-valued vs. complex valued time series

Here we present the problem of finding a sufficient representation of the second order statistics of wide sense stationary and periodic complex-valued time series. We'll focus our discussion in this section on univariate complex-valued time series and consider their second order description with respect to their equivalent representation as real-valued bivariate time series. In contrast to the real valued representation, as show in section 3.1, if the real valued bivariate time series is being represented as a complex valued (univariate) time series, the power spectrum of the complex valued representation doesn't capture all of the second order statistical information.

A very similar line of reasoning is easily extended to multivariate complex-valued time series, where the power spectra and cross-power spectra of the variates carry only partial second order description when compared to the alternative real-valued multivariate representation. In the subsequent sections 4 and 5 we derive second order statistical descriptions for multivariate complex-valued time series and for complex and real valued mixtures, respectively, which are equivalent to the second order statistics of their real-valued representations.

3.1 Second order statistics of a complex valued time series are not fully captured by their power spectrum

For a real-valued univariate time series, its second order statistics are described by its auto-correlation function [Brillinger 1981], or by the Wiener-Khinchin theorem [Wiener 1930], by its power spectrum. Equivalently, for a real valued (wide sense stationary and periodic) bivariate time series , $X(t) = \{x_1(t), x_2(t)\}$,$t = 0, ..., (n - 1)$,$x_1(t), x_2(t) \in \mathbb{R}$, the second order statistics of $X(t)$ are captured by three correlation functions. C_{x_1,x_1} , C_{x_1,x_2} and C_{x_2,x_2} [Brillinger 1981].

The real-valued bivariate time series $X(t) = \{x_1(t), x_2(t)\}$ could be represented as complex-valued univariate time series $z(t) = x_1(t) + ix_2(t)$, where $i = \sqrt{-1}$. In this case, however, for the complex valued representation, as we show in the following, the auto-correlation function $C_{z,z}(d)$ or equally, the power spectrum $P_{z,z}(k)$, provides only partial second order statistics of $z(t)$:

Proposition 1.

$$P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_1,x_2}(k) \Rightarrow P_{z,z}(k),$$
$$P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_1,x_2}(k) \not\Leftarrow P_{z,z}(k)$$

Proof. The Fourier coefficients $F_z(k)$ of the complex valued time series $z(t) = x_1(t) + ix_2(t)$ could be expressed as a linear combination of the Fourier coefficients $F_{x_1}(k)$ and $F_{x_2}(k)$:

$$\begin{aligned} F_z(k) = \sum_{t=0}^{n-1} \Big[x_1(k) + ix_2(k) \Big] e^{\frac{-2\pi i}{n} kt} = \\ \sum_{t=0}^{n-1} x_1(k) e^{\frac{-2\pi i}{n} kt} + \sum_{t=0}^{n-1} ix_2(k) e^{\frac{-2\pi i}{n} kt} = F_{x_1}(k) + iF_{x_2}(k) \end{aligned} \tag{15}$$

where $F_{x_1}(k)$ and $F_{x_2}(k)$ are the Fourier coefficients of the real valued time series $x_1(t)$ and $x_2(t)$, respectively. Applying (15) to $P_{z,z}(k)$ we get:

$$P_{z,z}(k) = F_z^*(k)F_z(k) = \Big[F_{x_1}(k) + iF_{x_2}(k)\Big]^*\Big[F_{x_1}(k) + iF_{x_2}(k)\Big]$$
$$= \Big[F_{x_1}^*(k) - iF_{x_2}^*(k)\Big]\Big[F_{x_1}(k) + iF_{x_2}(k)\Big] = |F_{x_1}(k)|^2 + |F_{x_2}(k)|^2 + i(F_{x_1}^*(k)F_{x_2}(k) - F_{x_1}(k)F_{x_2}^*(k))$$
$$= P_{x_1,x_1}(k) + P_{x_2,x_2}(k) + i[P_{x_1,x_2}(k) - P_{x_1,x_2}^*(k)] \tag{16}$$

The result in (16) provides explicit expression for the power spectrum $P_{z,z}(k)$ of the complex valued time series $z(t) = x_1(t) + ix_2(t)$ in terms of power and cross power spectra $P_{x_1,x_1}(k)$, $P_{x_2,x_2}(k)$ and $P_{x_1,x_2}(k)$ of the real valued time series $x_1(t)$ and $x_2(t)$, demonstrating that the second order statistical information provided by the power spectrum of a complex valued time series is included also in the power spectra and cross-power spectra of its real-valued bivariate representation. The converse, however, is not true. As it is evident from equation (15), for any $k = 1, ..., n/2$, the Fourier coefficient $F_z(k)$ is determined by two independent variables $F_{x_1}(k)$ and $F_{x_2}(k)$. Therefore,$F_{x_1}(k)$ and $F_{x_2}(k)$ may not be uniquely determined from the value of $F_z(k)$. It follows that also the spectral functions P_{x_1,x_1}, P_{x_1,x_2} and P_{x_2,x_2} may not be uniquely determined from complex valued power spectrum $P_{z,z}$. □

In a similar manner as for the real valued bivariate time series, for a real valued multivariate time series,$X(t) = \{x_1(t), x_2(t), ..., x_{2m}(t)\}$,$t = 1, 2, ..., 2n$ the second order statistical information is fully described by the power spectra $P_{x_i,x_i}(k), i = 1, .., 2m$, and cross power spectra $P_{x_i,x_{i+1}}(k), i = 1, .., 2n - 1$. Note these spectral functions taken together, uniquely determine the cross-power spectrum of any pair of variates $x_r(t)$ and $x_s(t)$, $r, s = 1, ..., 2m$ (this immediately follows from proposition 2 in section 5) ,and therefore provide a sufficient second order description of the multivariate time series. Here again, considering the multivariate complex-valued representation, $Z(t) = \{z_1(t) = x_1(t) + x_2(t), z_2(t) = x_3(t) + x_4(t), ..., z_m(t) = x_{2m-1}(t) + x_{2m}(t)\}$ of the time series, their spectral functions $P_{z_i,z_i}(k), i = 1, .., m$ and cross power spectra $P_{z_i,z_{i+1}}(k), i = 1, .., (m-1)$, do not provide the complete second order statistics, which could be demonstrated with a similar argument as for the univariate case discussed above.

4 Complete description of second order statistics of a complex valued multivariate time series

The partial information that is provided by the power spectrum $P_{z,z}(k)$ of the complex valued time series $z(t)$, and therefore also by its auto-correlation function $C_{z,z}(d)$, is analogous to the case of a complex valued random vector, v, where its covariance matrix $E[vv^*]$ doesn't not provide its complete second order statistics. Several authors have suggested that for a complete of second order statistical description of v, to augment the covariance matrix with a complementary covariance, $E[vv]$. We take a similar approach and define a complementary correlation function, and in the frequency domain, a complementary power spectrum. We show that for a complex valued time series, its power spectrum together with its complementary power spectrum provide a complete description of its second order statistics. Additionally, for a multivariate time series, we show that their power spectra and cross-power spectra when augmented with a single complementary power spectrum irrespective of the number of variates, provide a sufficient second order statistical description.

4.1 complementary correlation and complementary power spectrum of complex valued time series

Let $z_1(t) = x_1(t) + ix_2(t)$ and $z_2(t) = x_3(t) + ix_4(t)$ be complex valued time series. We define complementary correlation function,$\tilde{C}_{z_1,z_2}(d)$ and the complementary power spectrum $\tilde{P}_{z_1,z_2}(k)$ as:

Definition 1.

$$\tilde{C}_{z_1,z_2}(d) = \sum_{t=0}^{n-1} z_1(t) z_2(t+d)$$

Definition 2.

$$\tilde{P}_{z_1,z_2}(k) = F_{z_1}(-k) F_{z_2}(k)$$

In the frequency domain representation, the following relation holds:

Theorem 1. ***(The Complementary Cross-Correlation Theorem):***

$$\tilde{C}_{z_1,z_2}(d) \iff \tilde{P}_{z_1,z_2}(k)$$

Proof.

$$\tilde{C}_{z_1,z_2}(d) = \sum_{t=0}^{n-1} z_1(t) z_2(t+d)$$

With the discrete Fourier transform of $z_1(t)$ and $z_2(t+d)$, we get:

$$\tilde{C}_{z_1,z_2}(d) = \sum_{t=0}^{n-1} \left(\left[\frac{1}{n} \sum_{k_1=0}^{n-1} F_{z_1}(k_1) e^{\frac{2\pi i}{n} k_1 t} \right] \left[\frac{1}{n} \sum_{k_2=0}^{n-1} F_{z_2}(k_2) e^{\frac{2\pi i}{n} k_2 (t+d)} \right] \right) =$$

$$\tfrac{1}{n^2} \sum_{t=0}^{n-1} \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} F_{z_1}(k_1) e^{\frac{2\pi i}{n} k_1 t} F_{z_2}(k_2) e^{\frac{2\pi i}{n} k_2 (t+d)} =$$

$$\tfrac{1}{n^2} \sum_{t=0}^{n-1} \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} F_{z_1}(k_1) F_{z_2}(k_2) e^{\frac{2\pi i}{n} k_1 t} e^{\frac{2\pi i}{n} k_2 (t+d)} =$$

$$\tfrac{1}{n^2} \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} F_{z_1}(k_1) F_{z_2}(k_2) e^{\frac{2\pi i}{n} k_2 d} \left(\sum_{t=0}^{n-1} e^{\frac{2\pi i}{n} (k_1+k_2) t} \right)$$

The sum in parentheses is zero for $k_1 + k_2 \neq 0$, and for $k_1 + k_2 = 0$ it is n. Therefore, the double sum on the indexes k_1 and k_2 may be reduced to a single sum indexed by k

such that $k1 = -k$ and $k2 = k$:

$$\tilde{C}_{z_1,z_2}(d) =$$

$$\tfrac{1}{n^2}\sum_{k=0}^{n-1} F_{z_1}(-k)F_{z_2}(k)e^{\frac{2\pi i}{n}kd}\Big(n\Big) = \frac{1}{n}\sum_{k=0}^{n-1} F_{z_1}(-k)F_{z_2}(k)e^{\frac{2\pi i}{n}kd} =$$

$$\mathcal{F}^{-1}\{F_{z_1}(-k)F_{z_2}(k)\}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform. □

4.2 The complex multivariate correlations theorem

In section (3), we presented the problem of finding a complete description of the second order statistics of a complex-valued multivariate time series.Here we show that the power spectra and cross-power spectra,augmented with a single complementary power function ,provide a complete description of the second order statistics of multivariate complex time series.

We consider a real valued multivariate time series,$X(t) = \{x_1(t), x_2(t), ..., x_{2m}(t)\}$,$t = 0, ..., (n-1)$, or equivalently in a complex valued representation $Z(t) = \{z_1(t) = x_1(t)+x_2(t), z_2(t) = x_3(t)+x_4(t), ..., z_m(t) = x_{2m-1}(t)+x_{2m}(t)\}$. For the real valued multivariate time series, its second order statistics are described by the auto-correlation and cross correlation functions between its variates or equivalently, by the power spectra $P_{x_j,x_j}(k), i = 1, .., 2m$ and cross power spectra $P_{x_j,x_{j+1}}(k), j = 1, .., 2m-1$. In the following we show that an equivalent description of the second order statistics of $Z(t)$, is given by the power spectra $P_{z_i,z_i}(k), i = 1, .., m$ and cross power spectra $P_{z_i,z_{i+1}}(k), i = 1, .., (m-1)$, when augmented with a single complementary power spectrum function, which without loss of generality is chosen to be $\tilde{P}_{z_1,z_1}(k)$.

Theorem 2. ***(The Complex Multivariate Correlations Theorem):***

$$P_{z_j,z_j}(k), j = 1, ..., m,$$
$$P_{z_j,z_{j+1}}(k), j = 1, ..., (m-1)$$
$$\tilde{P}_{z_1,z_1}(k)$$
$$\Longleftrightarrow$$
$$P_{x_j,x_j}(k), j = 1, ..., 2m,$$
$$P_{x_j,x_{j+1}}(k), j = 1, ..., 2m-1$$

Proof. For the first direction of the above equivalence:

$$\begin{gathered} P_{x_j,x_j}(k), j = 1, ..., 2m, \\ P_{x_j,x_{j+1}}(k), j = 1, ..., 2m-1 \Rightarrow \\ P_{z_j,z_j}(k), j = 1, ..., m, \\ P_{z_j,z_{j+1}}(k), j = 1, ..., (m-1), \\ \tilde{P}_{z_1,z_1}(k) \end{gathered} \tag{17}$$

By the result in equation (16), an explicit expression for the power spectrum $P_{z_j,z_j}(k)$ of the complex valued time series $z_j(t) = x_{2j-1}(t) + ix_{2j}(t)$ is given in terms of power and cross power spectra $P_{x_{2j-1},x_{2j-1}}(k)$, $P_{x_{2j},x_{2j}}(k)$ and $P_{x_{2j-1},x_{2j}}(k)$ of the real valued time series $x_{2j-1}(t)$ and $x_{2j}(t)$. Therefore, all of the power spectra of the complex valued variates of Z may be obtained from the power spectra and cross-power spectra of the real valued representation X.

For the cross-power spectra of the complex valued variates, we have:

$$\begin{aligned} P_{z_j,z_{j+1}}(k) &= F^*_{z_j}(k) + iF^*_{z_{j+1}}(k) \\ &= \Big[F_{x_{2j-1}}(k) + iF_{x_{2j}}(k)\Big]^* \Big[F_{x_{2j+1)}}(k) + iF_{x_{2(j+1)}}(k)\Big] \\ &= P_{x_{2j-1},x_{2j+1}} - iP_{x_{2j},x_{2(j+1)}} + iP_{x_{2j-1},x_{2j+1}} + P_{x_{2j},x_{2(j+1)}} \end{aligned} \tag{18}$$

The cross-power spectrum $P_{v,w}$ of two time series $v(t)$ and $w(t)$ may be expressed as:

$$P_{v,w}(k) = |P_{v,w}(k)|e^{i\theta_{v,w}} = \sqrt[+]{P_{v,v}(k)P_{w,w}(k)}e^{i\theta_{v,w}} \tag{19}$$

or by rearrangement we get:

$$e^{i\theta_{v,w}} = \frac{P_{v,w}(k)}{|P_{v,w}(k)|} \tag{20}$$

where $\theta_{v,w}(k) = \theta_w(k) - \theta_v(k)$, is the phase difference between $F_w(k) = |F_w(k)|e^{i\theta_w(k)}$, and $F_v(k) = |F_v(k)|e^{i\theta_v(k)}$.

From equation (20), it follows that by summing phase differences between consecutive variates in X, we may express the cross-power spectrum of any pair of variates in X. For example, for $x_r(t)$ and $x_s(t)$, $r < s \leq 2m$, we get that the cross-power spectrum $P_{r,s}(k)$ is expressed as:

$$P_{r,s}(k) = |P_{r,s}(k)|e^{i\theta_{r,s}(k)} = \sqrt[+]{P_{r,r}P_{s,s}(k)}\prod_{j=r}^{s-1} e^{i\theta_{j,j+1}} \tag{21}$$

Therefore, all of the terms in equation (18) , $P_{x_{2j-1},x_{2(j+1)}}(k)$, $P_{x_{2j-1},x_{2(j+1)}}(k)$ etc., may be explicitly expressed in terms of the given cross-power spectra,$P_{x_j,x_{j+1}}(k), j = 1, ..., 2m-1$, and we get explicit expression also for $P_{z_j,z_{j+1}}(k)$.

To complete the first part of our proof we'll provide an expression for $\tilde{P}_{z_1,z_1}(k)$ in terms of the power spectra and cross-power spectra of X.

$$\begin{gathered}\tilde{P}_{z_1,z_1}(k) = F_{z_1}(-k)F_{z_1}(k) = \Big[F_{x_1}(-k) + iFx_2(-k)\Big]\Big[F_{x_1}(k) + iFx_2(k)\Big] = \\ \Big[F^*_{x_1}(k) + iF^*x_2(k)\Big]\Big[F_{x_1}(k) + iFx_2(k)\Big] = \\ P_{x_1,x_1} + iPx_2, x_1 + iP_{x_1,x_2} - Px_2, x_2,\end{gathered} \tag{22}$$

which completes the first part of our proof.

For the second direction of the equivalence:

$$
\begin{gathered}
P_{z_j,z_j}(k), j = 1, ..., m, \\
P_{z_j,z_{j+1}}(k), j = 1, ..., (m-1), \\
\tilde{P}_{z_1,z_1}(k) \Rightarrow \\
P_{x_j,x_j}(k), j = 1, ..., 2m, \\
P_{x_j,x_{j+1}}(k), j = 1, ..., 2m-1
\end{gathered} \tag{23}
$$

our proof is structured as follows. First, we show that $P_{x1,x1}(k)$, $P_{x2,x2}(k)$ and $P_{x1,x2}(k)$ may be derived from $P_{z_1,z_1}(k)$ and $\tilde{P}_{z_1,z_1}(k)$. This entails that the power-spectrum together with cross-power spectrum of a complex-valued univariate time series are sufficient descriptors of its second order statistics. Second,we show that the cross-power spectrum $P_{x_1,z_j}(k), j = 1, ..., m$ may be derived from the given left hand side of (23), which together with the real-complex correlation theorem (*Theorem 3*) we provide in the subsequent section will complete our proof.

Applying equation (15) to $\tilde{P}_{z_1,z_1}(k)$ together with the symmetry property of the Fourier transform of real valued functions, we get:

$$
\begin{gathered}
\tilde{P}_{z_1,z_1}(k) = F_{z_1}(-k)F_{z_1}(k) = \\
\Big[F_{x_1}(-k) + iF_{x_2}(-k)\Big]\Big[F_{x_1}(k) + iF_{x_2}(k)\Big] = \Big[F^*_{x_1}(k) + iF^*_{x_2}(k)\Big]\Big[F_{x_1}(k) + iF_{x_2}(k)\Big] = \\
P_{x_1,x_1}(k) + iPx_1, x_2(k) + iP^*x_1, x_2(k) - P_{x_2,x_2}(k) = \\
P_{x_1,x_1}(k) - P_{x_2,x_2}(k) + i\Big[Px_1, x_2(k) + P^*x_1, x_2(k)\Big]
\end{gathered} \tag{24}
$$

Noticing that the value in the brackets of the most right hand of (24) is real valued (a sum of a variable and its complex conjugate), and at the same time, the power spectra $P_{x_1,x_1}(k)$ and $P_{x_2,x_2}(k)$ are both real valued we get:

$$
P_{x_1,x_1}(k) - P_{x_2,x_2}(k) = \mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right) \tag{25}
$$

where $\mathfrak{Re}\,(z)$ is the real part of z .

Again, by the symmetry property of the Fourier transform, similar to our calculations in (16), considering the power spectra of the negative frequencies $P_{z,z}(-k)$ we obtain:

$$\begin{aligned}P_{z_1,z_1}(-k) = F^*_{z_1}(-k)F_{z_1}(-k) = \Big[F_{x_1}(-k) + iF_{x_2}(-k)\Big]^*\Big[F_{x_1}(-k) + iF_{x_2}(-k)\Big] \\ = \Big[F^*_{x_1}(-k) - iF^*_{x_2}(-k)\Big]\Big[F_{x_1}(-k) + iF_{x_2}(-k)\Big] \\ = |F_{x_1}(-k)|^2 + |F_{x_2}(-k)|^2 + i(F^*_{x_1}(-k)F_{x_2}(-k) - F_{x_1}(-k)F^*_{x_2}(-k)) \\ = P_{x_1,x_1}(-k) + P_{x_2,x_2}(-k) + i[P_{x_1,x_2}(-k) - P^*_{x_1,x_2}(-k)] \\ = P_{x_1,x_1}(k) + P_{x_2,x_2}(k) + i[P^*_{x_1,x_2}(k) - P_{x_1,x_2}(k)]\end{aligned} \tag{26}$$

Adding (16) to (26) we get:

$$P_{z_1,z_1}(k) + P_{z_1,z_1}(-k) = 2\Big[P_{x_1,x_1}(k) + P_{x_2,x_2}(k)\Big] \tag{27}$$

And by rearrangement:

$$P_{x_1,x_1}(k) + P_{x_2,x_2}(k) = \tfrac{1}{2}\Big[P_{z_1,z_1}(k) + P_{z_1,z_1}(-k)\Big] \tag{28}$$

Summing up equations (28) and (25), we get expression for $P_{x_1,x_1}(k)$:

$$P_{x_1,x_1}(k) = \tfrac{1}{4}\Big[P_{z_1,z_1}(k) + P_{z_1,z_1}(-k)\Big] + \tfrac{1}{2}\,\mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right) \tag{29}$$

And for $P_{x_2,x_2}(k)$:

$$P_{x_2,x_2}(k) = \tfrac{1}{4}\Big[P_{z_1,z_1}(k) + P_{z_1,z_1}(-k)\Big] - \tfrac{1}{2}\,\mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right) \tag{30}$$

Rearranging (24), and substituting back the values of $P_{x_1,x_1}(k)$ and $P_{x_2,x_2}(k)$ from (29) and (30)we get:

$$P_{x_1,x_2}(k) + P^*_{x_1,x_2}(k) = \frac{1}{i}\Big[\tilde{P}_{z_1,z_1}(k) + P_{x_2,x_2}(k) - P_{x_1,x_1}(k)\Big] = \\ \frac{1}{i}\Big[\tilde{P}_{z_1,z_1}(k) - \mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right)\Big] \tag{31}$$

Subtracting (26) from (16) we obtain:

$$P_{z_1,z_1}(k) - P_{z_1,z_1}(-k) = 2i(P_{x_1,x_2}(k) - P^*_{x_1,x_2}(k)) \tag{32}$$

Or equivalently:

$$P_{x_1,x_2}(k) - P^*_{x_1,x_2}(k) = \frac{1}{2i}(P_{z_1,z_1}(k) - P_{z_1,z_1}(-k)) \tag{33}$$

And by summing (33) and (31) we get an explicit expression for $P_{x_1,x_2}(k)$:

$$P_{x_1,x_2}(k) = \\ \frac{1}{4i}\Big[P_{z_1,z_1}(k) - P_{z_1,z_1}(-k)\Big] + \frac{1}{2i}\Big[\tilde{P}_{z_1,z_1}(k) - \mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right)\Big] \tag{34}$$

With (29) and (34) we may get an expression for $P_{x_1,z1}(k)$:

$$P_{x_1,z_1}(k) = F^*_{x_1}(k)\Big[F_{x_1}(k) + iF_{x_2}(k)\Big] = P_{x_1,x_1} + iP_{x1,x2} = \\ \frac{1}{4}\Big[P_{z_1,z_1}(k) + P_{z_1,z_1}(-k)\Big] + \frac{1}{2}\,\mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right) + \\ i\left(\frac{1}{4i}\Big[P_{z_1,z_1}(k) - P_{z_1,z_1}(-k)\Big] + \frac{1}{2i}\Big[\tilde{P}_{z_1,z_1}(k) - \mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right)\Big]\right) = \\ \frac{1}{4}\Big[P_{z_1,z_1}(k) + P_{z_1,z_1}(-k)\Big] + \frac{1}{2}\,\mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right) + \\ \frac{1}{4}\Big[P_{z_1,z_1}(k) - P_{z_1,z_1}(-k)\Big] + \frac{1}{2}\Big[\tilde{P}_{z_1,z_1}(k) - \mathfrak{Re}\left(\tilde{P}_{z_1,z_1}(k)\right)\Big] = \\ \frac{1}{2}\Big[P_{z_1,z_1}(k) + \tilde{P}_{z_1,z_1}(k)\Big] \tag{35}$$

By equations (29), (30) and (34) we have shown:

$$P_{z_1,z_1}(k), \tilde{P}_{z_1,z_1}(k) \Rightarrow P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_1,x_2}(k) \tag{36}$$

With (19) and (35) we may get an expression for P_{x_1,z_2}. Together with (29),(30) and (34) we have:

$$
\begin{gathered}
P_{z_j,z_j}(k), j = 1, ..., m, \\
P_{z_j,z_{j+1}}(k), j = 1, ..., (m-1), \\
\tilde{P}_{z_1,z_1}(k) \Rightarrow \\
P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_1,x_2}(k), \\
P_{x_1,x_1}(k), P_{x_1,z_2}, P_{z_j,z_j}(k), j = 2, ..., m, P_{z_j,z_{j+1}}(k), j = 2, ..., (m-1)
\end{gathered}
\tag{37}
$$

Let $U(t)$ be the multivariate series which is a mixture of the real and complex valued variates:

$$
U(t) = \{x_1(t), z_2(t), ..., z_m(t)\}
$$

Then bottom most line in (37),

$$
P_{x_1,z_2}, P_{z_j,z_j}(k), j = 2, ..., m, P_{z_j,z_{j+1}}(k), j = 2, ..., (m-1)
$$

express the second order statistics of U. By Theorem 3 (the real complex valued correlations theorem) we present in the next section, we have:

$$
\begin{gathered}
P_{x_1,x_1}(k), P_{x_1,z_2}, \\
P_{z_j,z_j}(k), j = 2, ..., m, P_{z_j,z_{j+1}}(k), j = 2, ..., (m-1) \Rightarrow \\
P_{x_1,x_1}(k), P_{x_1,x_3}(k) \\
P_{x_j,x_j}(k), j = 3, ..., 2m, \\
P_{x_j,x_{j+1}}(k), j = 3, ..., 2m-1
\end{gathered}
\tag{38}
$$

Taking together relations (36), (37) and (38) completes our proof.

$\square$

5 The real-complex multivariate correlations theorem

Here we provide a proof that for a multivariate mixture of real and complex valued time series, the joint second order statistics are fully captured by their power and cross-power spectra, much like for real-valued multivariate processes.For a real-valued multivariate time series $X(t)$:

$$X(t) = \{x_1(t), x_2(t), ..., x_m(t)\}, t = 0, ..., n-1$$

its second order statistics are given by the power spectra $P_{x_j,x_j}(k), j = 1, ..., m$, and by the cross-power spectra $P_{x_j,x_{j+1}}(k), j = 1, ..., m-1$. Let $U(t)$ be a multivariate time series, which is a mixture of real-valued and complex-valued univariate time series composed of the variates of $X(t)$. We'll show in the following that we may assume without the loss of generality, that the r lower indexes of the time-variates of $U(t)$ are real valued, and the rest are complex valued:

$$U(t) = \{u_1(t), u_2(t), ..., u_{r+(m-r)/2)}(t)\} =$$
$$\{x_1(t), x_2(t), ..., x_r(t), z_1(t), z_2(t), ..., z_{(m-r)/2}(t)\}$$

where, $x_1(t), ..., x_r(t) \in \mathbb{R}$, and $z_1(t) = x_{r+1}(t) + ix_{r+2}(t), ..., z_{(m-r)/2}(t) = x_{m-1}(t) + ix_m(t) \in \mathbb{C}$. We'll show that the power spectra,$P_{u_j,u_j}(k), j = 1, ..., (r + (m-r)/2)$ of the real/complex valued mixture $U(t)$, and its cross-power spectra $P_{u_j,u_{j+1}}(k), j = 1, ...(r + \frac{m-r}{2} - 1)$ provide a complete description of the second order statistics of $X(t)$.

Theorem 3. ***(The Real/Complex Correlations Theorem):***

$$P_{x_j,x_j}(k), j = 1, ..., m,$$

$$P_{x_j,x_{j+1}}(k), j = 1, ..., m-1$$

$$\Longleftrightarrow$$

$$P_{u_j,u_j}(k), j = 1, ..., (r + \tfrac{m-r}{2}),$$

$$P_{u_j,u_{j+1}}(k), j = 1, ...(r + \tfrac{m-r}{2} - 1)$$

Proof. We'll start by considering a simpler case, of a real-valued trivariate time series $\tilde{X}(t) = \{x_1(t), x_2(t), x_3(t)\}, x_j(t) \in \mathbb{R}, j = 1, 2, 3$, or in the real/complex valued representation $\tilde{U}(t) = \{x_1(t), z(t) = x_2(t) + ix_3(t)\}, z(t) \in \mathbb{C}$, and show that the second order statistics of $\tilde{X}(t)$ are equally described by the power spectra and cross-power spectra of the variates of $\tilde{U}(t)$.

Lemma (1)**.**

$$P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_3,x_3}(k), P_{x_1,x_2}(k), P_{x_2,x_3}(k) \iff P_{x_1,x_1}(k), P_{z,z}(k), P_{x_1,z}(k)$$

Proof. We'll start by showing:

$$P_{x_1,x_1}(k), P_{z,z}(k), P_{x_1,z}(k) \Rightarrow P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_3,x_3}(k), P_{x_1,x_2}(k), P_{x_2,x_3}(k)$$

As $P_{x_1,x_1}(k)$ is already given on the left side, we are left with providing explicit expressions for $P_{x_2,x_2}(k)$, $P_{x_3,x_3}(k)$, $P_{x_1,x_2}(k)$,and $P_{x_2,x_3}(k)$ in terms of $P_{x_1,x_1}(k)$, $P_{z,z}(k)$,and $P_{x_1,z}(k)$.

Applying (15) to $P_{x_1,z}(k)$ we get:

$$P_{x_1,z}(k) = F^*_{x_1}(k)F_z(k) = F^*_{x_1}(k)\Big[F_{x_2}(k) + iF_{x_3}(k)\Big] \tag{39}$$

Applying (15) to $P_{x_1,z}(-k)$, together with the symmetry property of the Fourier transform of real valued functions we get:

$$P_{x_1,z}(-k) = F^*_{x_1}(-k)F_z(-k) =$$
$$F^*_{x_1}(-k)\Big[F_{x_2}(-k) + iF_{x_3}(-k)\Big] = F_{x_1}(k)\Big[F^*_{x_2}(k) + iF^*_{x_3}(k)\Big] \tag{40}$$

Multiplying $P_{x_1,z}(k))$ by $P_{x_1,z}(-k)$, together with (39) and (40), we get:

$$\begin{gathered}
P_{x_1,z}(k)P_{x_1,z}(-k) = \\
\Big(F^*_{x_1}(k)\Big[F_{x_2}(k) + iF_{x_3}(k)\Big]\Big)\Big(F^*_{x_1}(-k)\Big[F_{x_2}(-k) + iF_{x_3}(-k)\Big]\Big) = \\
\Big(F^*_{x_1}(k)\Big[F_{x_2}(k) + iF_{x_3}(k)\Big]\Big)\Big(F_{x_1}(k)\Big[F^*_{x_2}(k) + iF^*_{x_3}(k)\Big]\Big) = \\
\Big(F^*_{x_1}(k)F_{x_2}(k) + iF^*_{x_1}(k)F_{x_3}(k)\Big)\Big(F_{x_1}(k)F^*_{x_2}(k) + iF_{x_1}F^*_{x_3}(k)\Big) = \\
P_{x_1,x_1}(k)P_{x_2,x_2}(k) - P_{x_1,x_1}(k)P_{x_3,x_3}(k) + iP_{x_1,x_1}(k)P_{x_2,x_3}(k) + iP_{x_1,x_1}(k)P^*_{x_2,x_3}(k) = \\
P_{x_1,x_1}\Big(P_{x_2,x_2}(k) - P_{x_3,x_3}(k) + i[P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k)]\Big)
\end{gathered} \tag{41}$$

And by rearranging the above, we get:

$$P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k) = \tfrac{1}{i}\Big(\tfrac{P_{x_1,z}(k)P_{x_1,z}(-k)}{P_{x_1,x_1}} - \Big[P_{x_2,x_2}(k) - P_{x_3,x_3}(k)\Big]\Big) \tag{42}$$

Similarly, considering $P^*_{x_1,z}(k)P^*_{x_1,z}(-k)$:

$$\begin{gathered}
P^*_{x_1,z}(k)P^*_{x_1,z}(-k) = \\
P_{x_1,x_1}(k)\Big(P_{x_2,x_2}(k) - P_{x_3,x_3}(k) - i\Big[P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k)\Big]\Big)
\end{gathered}$$

And by rearrangement:

$$P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k) = \tfrac{1}{i}\Big[P_{x_2,x_2}(k) - P_{x_3,x_3}(k) - \tfrac{P^*_{x_1,z}(k)P^*_{x_1,z}(-k)}{P_{x_1,x_1}(k)}\Big] \tag{43}$$

Adding (42) and (43):

$$P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k) = \frac{1}{2i}\Big[\frac{P_{x_1,z}(k)P_{x_1,z}(-k)}{P_{x_1,x_1}} - \frac{P^*_{x_1,z}(k)P^*_{x_1,z}(-k)}{P_{x_1,x_1}(k)}\Big] = \frac{1}{2iP_{x_1,x_1}}\Big[P_{x_1,z}(k)P_{x_1,z}(-k) - P^*_{x_1,z}(k)P^*_{x_1,z}(-k)\Big] \tag{44}$$

Substituting P_{z_1,z_1}, P_{x_1,x_1} and P_{x_2,x_2} in (33) with $P_{z,z}$, P_{x_2,x_2} and P_{x_3,x_3} respectively, and summing up with (44), we get an expression for $P_{x_2,x_3}(k)$:

$$P_{x_2,x_3}(k) = \frac{1}{4i}\Big[P_{z,z}(k) - P_{z,z}(-k)\Big] + \frac{1}{4iP_{x_1,x_1}}\Big[P_{x_1,z}(k)P_{x_1,z}(-k) - P^*_{x_1,z}(k)P^*_{x_1,z}(-k)\Big] \tag{45}$$

Rearranging (43) we get:

$$P_{x_2,x_2}(k) - P_{x_3,x_3}(k) = \frac{P^*_{x_1,z}(k)P^*_{x_1,z}(-k)}{P_{x_1,x_1}(k)} + i\Big[P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k)\Big] \tag{46}$$

Substituting $P_{x_2,x_3}(k) + P^*_{x_2,x_3}(k)$ in (46) with its value in (44):

$$\begin{aligned} P_{x_2,x_2}(k) - P_{x_3,x_3}(k) = \\ \frac{P^*_{x_1,z}(k)P^*_{x_1,z}(-k)}{P_{x_1,x_1}(k)} + \frac{1}{2P_{x_1,x_1}}\Big[P_{x_1,z}(k)P_{x_1,z}(-k) - P^*_{x_1,z}(k)P^*_{x_1,z}(-k)\Big] = \\ \frac{1}{P_{x_1,x_1}(k)}\Big(P^*_{x_1,z}(k)P^*_{x_1,z}(-k) + \frac{1}{2}\Big[P_{x_1,z}(k)P_{x_1,z}(-k) - P^*_{x_1,z}(k)P^*_{x_1,z}(-k)\Big]\Big) = \\ \frac{1}{2P_{x_1,x_1}(k)}\Big(P_{x_1,z}(k)P_{x_1,z}(-k) + P^*_{x_1,z}(k)P^*_{x_1,z}(-k)\Big) \end{aligned} \tag{47}$$

Substituting P_{z_1,z_1}, P_{x_1,x_1} and P_{x_2,x_2} in (28) with $P_{z,z}$, P_{x_2,x_2} and P_{x_3,x_3} receptively, and adding to (47), we get an expression for $P_{x_2,x_2}(k)$:

$$P_{x_2,x_2}(k) = \frac{1}{4}\Big(P_{z,z}(k) + P_{z,z}(-k) + \frac{P_{x_1,z}(k)P_{x_1,z}(-k) + P^*_{x_1,z}(k)P^*_{x_1,z}(-k)}{P_{x_1,x_1}(k)}\Big) \tag{48}$$

Substituting P_{z_1,z_1}, P_{x_1,x_1} and P_{x_2,x_2} in (28) with $P_{z,z}$, P_{x_2,x_2} and P_{x_3,x_3} receptively, and subtracting (48), we get an expression for $P_{x_3,x_3}(k)$:

$$P_{x_3,x_3}(k) = \frac{1}{4}\Big(P_{z,z}(k) + P_{z,z}(-k) - \frac{P_{x_1,z}(k)P_{x_1,z}(-k)+P^*_{x_1,z}(k)P^*_{x_1,z}(-k)}{P_{x_1,x_1}(k)}\Big) \tag{49}$$

Thus, so far we have provided explicit expressions for $P_{x_1,x_1}(k)$, $P_{x_2,x_2}(k)$,$P_{x_3,x_3}(k)$ and $P_{x_2,x_3}(k)$. To complete the first part of our proof, we are left with providing an expression for $P_{x_1,x_2}(k)$ in terms of $P_{x_1,x_1}(k)$,$P_{z,z}(k)$ and $P_{x_1,z}(k)$. From $P_{x_1,z}(k)$ we may extract the phase difference $\theta_{x_1,z}(k) = \theta_z(k) - \theta_{x_1}(k)$, between the Fourier components $F_{x_1}(k)$ and $F_z(k)$:

$$e^{i\theta_{x_1,z}(k)} = \frac{P_{x_1,z}(k)}{|P_{x_1,z}(k)|} \tag{50}$$

We may express the the phase difference $\theta_{z,x_2}(k)$ between $F_z(k)$ and $F_{x_2}(k)$ as:

$$e^{i\theta_{z,x_2}(k)} = \frac{F^*_z(u,v)F_{x_2}(k)}{|F^*_z(u,v)F_{x_2}(k)|} \tag{51}$$

Where,

$$F^*_z(k)F_{x_2}(k) = [F_{x_2}(k) + iF_{x_3}(k)]^*F_{x_2}(k) = F^*_{x_2}(k)F_{x_2}(k) - iF^*_{x_3}(k)F_{x_2}(k) = P_{x_2,x_2}(k) - iP^*_{x_2,x_3}(k) \tag{52}$$

Substituting (52) into (51) we get:

$$e^{i\theta_{z,x_2}(k)} = \frac{P_{x_2,x_2}(k) - iP^*_{x_2,x_3}(k)}{|P_{x_2,x_2}(k) - iP^*_{x_2,x_3}(k)|} \tag{53}$$

Therefore, the phase difference $i\theta_{x_1,x_2}$ between F_{x_1} and F_{x_2} may be expressed as:

$$e^{i\theta_{x_1,x_2}(k)} = e^{i\theta_{x_1,z}(k)}e^{i\theta_{z,x_2}(k)} = \frac{P_{x_1,z}(k)}{|P_{x_1,z}(k)|}\frac{P_{x_2,x_2}(k) - iP^*_{x_2,x_3}(k)}{|P_{x_2,x_2}(k) - iP^*_{x_2,x_3}(k)|} \tag{54}$$

With (54) we now can express P_{x_1,x_2} as:

$$P_{x_1,x_2}(k) = F^*_{x_1}(k)F_{x_2}(k) = |F^*_{x_1}(k)F_{x_2}(k)|e^{i\theta_{x_1,x_2}(k)} = \sqrt[+]{P_{x_1}(k)P_{x_2}(k)}e^{i\theta_{x_1,x_2}(k)}$$

$$\sqrt[+]{P_{x_1}(k)P_{x_2}(k)}\left[\frac{P_{x_1,z}(k)}{|P_{x_1,z}(k)|}\right]\left[\frac{P_{x_2,x_2}(k)-iP^*_{x_2,x_3}(k)}{|P_{x_2,x_2}(k)-iP^*_{x_2,x_3}(k)|}\right] \tag{55}$$

Which completes the first part of the proof, where we have provided explicit expressions for P_{x_1,x_1}, P_{x_2,x_2},P_{x_3,x_3},P_{x_1,x_2} and P_{x_2,x_3} in terms of P_{x_1,x_1},$P_{z,z}$ and $P_{x_1,z}$.

For the second part of the proof, we'll show the opposite direction:

$$P_{x_1,x_1}(k), P_{x_2,x_2}(k), P_{x_3,x_3}(k), P_{x_1,x_2}(k), P_{x_2,x_3}(k) \Rightarrow P_{x_1,x_1}(k), P_{z,z}(k), P_{x_1,z}(k)$$

By equation 16 in section 3.1, the power spectrum of the complex valued time series $z(t) = x_2(t) + ix_3(t)$, is given by:

$$P_{z,z}(k) = P_{x_2,x_2}(k) + P_{x_3,x_3}(k) + i[P_{x_2,x_3}(k) - P^*_{x_2,x_3}(k)]$$

Therefore, to complete the proof, we are left with providing an expression for $P_{x1,z}$.

The phase difference $\theta_{x_1,x_3}(k)$ between $F_{x_1}(k)$ and $F_{x_3}(k)$ is related to the phase difference $\theta_{x_1,x_2}(k)$ between $F_{x_1}(k)$ and $F_{x_2}(k)$ and $\theta_{x_2,x_3}(k)$ between $F_{x_2}(k)$ and $F_{x_3}(k)$ as: $\theta_{x_1,x_3}(k) = \theta_{x_1,x_2}(k) + \theta_{x_2,x_3}(k)$.

Therefore:

$$e^{i\theta_{x_1,x_3}(k)} = e^{i\theta_{x_1,x_2}(k)}e^{i\theta_{x_2,x_3}(k)} = \left[\frac{P_{x_1,x_2}(k)}{|P_{x_1,x_2}(k)|}\right]\left[\frac{P_{x_2,x_3}(k)}{|P_{x_2,x_3}(k)|}\right] \tag{56}$$

From (56) we may derive now an expression for cross power spectrum $P_{x1,x_3}(k)$:

$$P_{x1,x_3}(k) = F_{x_1}(k)F^*_{x_3}(k) = (|F_L(u,v)||F^*_b(u,v)|e^{i\theta_{L,b}(u,v)} = \\ \sqrt[+]{P_{x1,x_1}(k)P_{x3,x_3}(k)}\left[\frac{P_{x1,x_2}(k)}{|P_{x1,x_2}(k)|}\right]\left[\frac{P_{x2,x_3}(k)}{|P_{x2,x_3}(k)|}\right] \tag{57}$$

Substituting (57) into Equation (39) provides an explicit expression for $P_{x1,z}(k)$:

$$P_{x_1,z}(k) = \\ F^*_{x_1}(k)[F_{x_2}(k) + iF_{x_2}(k)] = F^*_{x_1}(k)F_{x_2}(k) + iF^*_{x_1}(k)F_{x_3}(k) = \\ P_{x_1,x_2}(k) + i\left(\sqrt[+]{P_{x_1,x_1}(k)P_{x_3,x_3}(k)}\left[\frac{P_{x_1,x_2}(k)}{|P_{x1,x_2}(k)|}\right]\left[\frac{P_{x_2,x_3}(k)}{|P_{x2,x_3}(k)|}\right]\right) \tag{58}$$

□

Now we'll consider the second order statistics of the more general case, of multivariate mixture of real and complex-valued time series. To simplify our discussion, we made the assumption that the r lower indexes of the time-variates of $U(t)$ are real valued, and the rest are complex valued. We justify our assumption by the following. Assume that $U'(t)$ is a obtained from $U(t)$, by a random permutation,p, of the variates of $U(t)$, such that $U'(t) = \{u_{p(1)}(t), u_{p(2)}(t), ..., u_{p(r+\frac{m-r}{2}-1)}(t)\}$. Then the second order statistics of $U(t)$ and $U'(t)$ are similarly described by the power spectra and cross-power spectra of $U(t)$, as by those of $U'(t)$. Clearly, the power spectra of the time variates of $U(t)$ and $U'(t)$ are similar, and therefore we need to consider only the cross-power spectra that change under the permutation of variates order. In the following we show that the cross-power spectra of $U(t)$ and $U'(t)$ carry similar information.

Proposition 2.

$$P_{u_j,u_{j+1}}(k), j = 1, ...(r + \tfrac{m-r}{2} - 1)$$

$$\Longleftrightarrow$$

$$P_{u_{p(j)},u_{p(j)+1}}(k), j = 1, ...(r + \tfrac{m-r}{2} - 1)$$

Proof. Our proof is based on the observation that given the power spectra $P_{u_j,u_j}(k)$ and cross-power spectra $P_{u_j,u_{j+1}}(k)$, we may express the cross-power spectra of any pair of variates of U. Given $P_{u_j,u_j}(k)$ and $P_{u_j,u_{j+1}}(k)$, we may extract the phase difference $\theta_{u_j,u_{j+1}}(k) = (\theta_{u_{j+1}}(k) - \theta_{u_j}(k))$ between $F_{u_{j+1}}(k)$ and $F_{u_j}(k)$:

$$e^{i\theta_{u_j,u_{j+1}}}(k) = \frac{P_{u_j,u_{j+1}}(k)}{|P_{u_j,u_{j+1}}(k)|} = \frac{P_{u_j,u_{j+1}}(k)}{P_{u_j,u_j}(k)P_{u_{j+1},u_{j+1}}(k)} \tag{59}$$

By transitivity, the phase difference between $F_{u_{j1}}(k)$ and $F_{u_{j2}}(k)$, $j1 < j2, j1, j2 <= m$:

$$e^{i\theta_{u_{j1},u_{j2}}(k)} = \prod_{r=j1}^{j2} e^{i\theta_{u_{j1},u_{j2}}(k)} \tag{60}$$

And the cross-power spectrum $P_{u_{j1},u_{j2}}(k)$, is expressed as:

$$\begin{gathered} P_{u_{j1},u_{j2}}(k) = P_{u_{j1},u_{j2}}(t)(k)e^{\theta_{u_{j1},u_{j2}}(k)} = \\ \sqrt[+]{P_{u_{j1},u_{j1}}P_{u_{j2},u_{j2}}(k)} \prod_{r=j1}^{j2} e^{\theta_{u_{j1},u_{j2}}(k)} \end{gathered} \tag{61}$$

The cross-power spectrum $P_{u_{j1},u_{j2}}$ (between $u_{j1}(t)$ and $u_{j2}(t)$) is related to the cross-power spectrum $P_{u_{j2},u_{j1}}$ (between $u_{j2}(t)$ and u_{j1}) as: $P_{u_{j2},u_{j1}} = P^*_{u_{j1},u_{j2}}$. Therefore, with (61), we may express the cross-power spectrum between any pair variates of $U(t)$ and therefore also between any pair of variates of $U'(t)$.

□

Proposition 3.

$$P_{u_j,u_{j+1}}(k), j = 1, ..., (r + \tfrac{m-r}{2} - 1)$$
$$\Longleftrightarrow$$
$$P_{u_1,u_j}(k), j = 2, ..., (r + \tfrac{m-r}{2})$$

Proof. The first direction, $P_{u_j,u_{j+1}}(t), j = 1, ..., (r + \frac{m-r}{2} - 1) \Rightarrow P_{u_1,u_j}(t), j =$

$2, ..., (r + \frac{m-r}{2})$ follows immediately from Proposition 2. For the second direction,with similar calculations as in (56) the phase difference $\theta_{j,j+1}$ between $F_j(k)$ and $F_{j+1}(k)$ may be computed from the phase differences $\theta_{1,j+1}$ and $\theta_{1,j}$ as:$\theta_{j,j+1} = \theta_{1,j+1} - \theta_{1,j}$ and we get:

$$e^{i\theta_{u_j,u_{j+1}}(k)} = e^{i(\theta_{u_1,u_{j+1}}(k) - \theta_{u_1,u_j}(k))} = \frac{e^{i\theta_{u_1,u_{j+1}}(k)}}{e^{i\theta_{u_1,u_j}(k)}} =$$
$$\Big(\frac{P_{u_1,u_{j+1}}(k)}{|P_{u_1,u_{j+1}}(k)|}\Big) / \Big(\frac{P_{u_1,u_j}(k)}{|P_{u_1,u_j}(k)|}\Big) = \frac{P_{u_1,u_{j+1}}(k)|P_{u_1,u_j}(k)|}{|P_{u_1,u_{j+1}}(k)|P_{u_1,u_j}(k)} =$$

$$\frac{P_{u_1,u_{j+1}}(k) \sqrt[+]{|P_{u_1,u_1}(k)||P_{u_j,u_j}(k)|}}{\sqrt[+]{|P_{u_1,u_1}(k)||P_{u_{j+1},u_{j+1}}(k)|}P_{u_1,u_j}(k)} = \frac{P_{u_1,u_{j+1}}(k) \sqrt[+]{|P_{u_j,u_j}(k)|}}{P_{u_1,u_j}(k) \sqrt[+]{|P_{u_{j+1},u_{j+1}}(k)|}} =$$

$$\frac{P_{u_1,u_{j+1}}(k)}{P_{u_1,u_j}(k)} \sqrt[+]{\frac{P_{u_j,u_j}(k)}{P_{u_{j+1},u_{j+1}}(k)}}$$

and the cross-power spectra $P_{u_j,u_{j+1}}(t), j = 1, ..., (r + \frac{m-r}{2} - 1)$, may be expressed as:

$$P_{u_j,u_{j+1}}(k) = |P_{u_j,u_{j+1}}(k)|e^{i\theta_{u_j,u_{j+1}}(k)} = \sqrt[+]{P_{u_j,u_j}(k)P_{u_{j+1},u_{j+1}}(k)}e^{i\theta_{u_j,u_{j+1}}(k)} =$$
$$\sqrt[+]{P_{u_j,u_j}(k)P_{u_{j+1},u_{j+1}}(k)}\frac{P_{u_1,u_{j+1}}(k)}{P_{u_1,u_j}(k)} \sqrt[+]{\frac{P_{u_j,u_j}(k)}{P_{u_{j+1},u_{j+1}}(k)}} =$$
$$P_{u_j,u_j}(k)\frac{P_{u_1,u_{j+1}}(k)}{P_{u_1,u_j}(k)}$$

□

By Proposition 3 , from cross-power spectra $P_{u_j,u_{j+1}}, j = 1, ...(r + \frac{m-r}{2} - 1)$, we may derive the cross-power spectra $P_{u_1,u_j}(k), j = 2, ...(r + \frac{m-r}{2})$, and in particular, the cross-power spectra $P_{u_1,u_w}(k), w = r + 1, ..., (r + \frac{m-r}{2})$, for all of the complex-valued variates $u_w(t)$ of U, where $u_w(t) = z_{w-r}(t) = x_{2(w-r)-1}(t) + ix_{2(w-r)}(t)$. Since $u_1(k) = x_1(t)$ is real valued, from with $P_{u_1,u_1}(k)$, $P_{u_1,u_w}(k)$ and $P_{u_w,u_w}(k)$, by lemma (1) we may may obtain expressions for the power spectra of $P_{2(w-r)-1,2(w-r)-1}(k)$, $P_{2(w-r),2(w-r)}(k)$ and cross power spectra $P_{1,2(w-r)-1}(k)$,$P_{1,2(w-r)}(k)$ of the real valued time series $x_1(t)$, $x_{2(w-r)-1}(t)$,and $x_{2(w-r)}(t)$. At this point we've shown, that from the power and cross-power spectra of the variates of U, we may compute the the power spectra $P_{j,j}(k), j = 1, ..., m$ of all the variates of the real valued representation R, and the cross-power spectra a $P_{1,j}(k)$ between its first variate $x_1(t)$and the rest. Again by Proposition 3, it follows that from $P_{1,j}(k), j = 2, ..., m$ we may also get explicit expression for a $P_{j,j+1}(k), j = 1, ..., m - 1$, which completes our proof of Theorem 3.

□

6 Summary

Equivalence relations for second order statistics of complex-valued and real valued representation of multivariate time series have been provided, with explicit expressions relating the two. Our construction begun by introducing the complementary discrete cross-correlation function for finite complex-valued time series and deriving its frequency domain equivalent, the complementary power spectrum (by the complementary cross correlation theorem in section 3). We have shown that in addition to the auto- and cross-correlation functions, additional single complementary auto-correlation function of one of the complex valued variates is sufficient for a complete second order statistical description of the complex valued multivariate time series. The choice of defining the complementary correlation function was motivated by previous works, in which the complementary covariance matrix was defined for complex valued variables.

We have also considered mixtures of real and complex-valued time series. Our motivation originated from the idea that in some disciplines, such mixed representation might be a most natural choice. For example, in studies involving the analysis of correlations between multiple types of measurements such and oceanic surface currents, atmospheric pressure and temperature [White and Peterson 1996] - surface currents might be represented as complex valued . In this case, of a time series which is a mixture of real valued and complex valued variates, the second order statistics are fully captured by their auto- and cross-correlation functions, similar to the case of real-valued time series.This result is potentially of practical significance, rendering standard statistical methods for the analysis of real-valued multivariate time series applicable here as well.

Finally, our results are readily extended to complex valued infinite time series and continues time signals. Our choice of presentation through finite time series was motivated primarily by practical considerations, as providing direct recipe for implementation as a computer program.

References

Brillinger, D. (1981). *Time Series: Data Analysis and Theory*. Holden-Day, San Francisco.

Chang, P., Ji, L., and Li, H. (1997). A decadal climate variation in the tropical atlantic ocean from thermodynamic air-sea interactions. *Nature*, 385:516–518.

Horel, J. D. (1984). Complex principal component analysis: theory and examples. *Journal of Climate and Applied Meteorology*, 23(12):1660–1673.

Kaihatu, J., Handler, R. A., Marmorino, G. O., and Shay, L. K. (1998). Empirical orthogonal function analysis of ocean surface currents using complex and real-vector methods. *Journal of Atmospheric and Oceanic Technology*, 15:927941.

Mandic, D., Javidi, S.and Goh, S., , K. A., and Aihara, K. (2009). Complex-valued prediction of wind profile using augmented complex statistics. *Renewable Energy*, 34:196–201.

Mokeichev, A., Okun, M., Barak, O., Katz, Y., Ben-Shahar, O., and Lampl, I. (2007). Stochastic emergence of repeating cortical motifs in spontaneous membrane potential fluctuations in vivo. *Neuron*, 53:413–425.

Mooers, C. (1973). A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants. *Deep Sea Research*, 20(12):1129–1141.

Neseer, F. and Massey, J. (1993). Proper complex random processes with applications to information theory. *IEEE Transactions on Information Theory*, 39(4):1293–1302.

Ollila, E. (2008). On the circularity of a complex random variable. *IEEE Signal Processing Letters*, 15:841–844.

Parzen, E. (1961). An approach to time series analysis. *Annals of Mathematical Statistics*, 32(4):951–989.

Parzen, E. (1964). An approach to empirical time series analysis. *Radio Science*, 68(9):937–951.

Picinbono, B. and Bondon, P. (1997). Second-order statistics of complex signals. *IEEE Transactions on Signal Processing*, 45(2):411–420.

Pillow, J. W., Shlens, J., L., P., Sher, A., Litke, A., Chichilnisky, E. J., and Simoncelli, E. P. (2007). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454:995–999.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., and Stanley, H. E. (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters*, 83(7):14711474.

Ramchand, L. and Susmel, R. (1998). Volatility and cross correlation across major stock markets. *Journal of Empirical Finance*, 5(4):397–416.

Reinsel, G. C. (1997). *Elements of Multivariate Time Series Analysis*. Springer-Verlag.

Schreier, P. and Scharf, L. (2003). Second-order analysis of improper complex random vectors and processes. *IEEE Transactions on Signal Processing*, 51(3):714–725.

Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307:1615–1618.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne, F. J. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D. Nonlinear Phenomena*, 58:77–94.

Wahlberg, P. and Schreier, P. (2008). Spectral relations for multidimensional complex improper stationary and (almost) cyclostationary processes. *IEEE Transactions on Information Theory*, 54(4):1670–1682.

White, W. B. and Peterson, R. G. (1996). An antarctic circumpolar wave in surface pressure, wind, temperature and sea-ice extent. *Nature*, 380:699–702.

Wiener, N. (1930). Generalized harmonic analysis. *Acta Mathematica*, 55:117–258.

Witt, A., Kurths, J., and Pikovsky, A. (1998). Testing stationarity in time series. *Physical Review E*, 58(4):18001810.