# 13.741J/6.455J
# Multidimensional Random Variables and Vectors Real and Complex

Christ D. Richmond

October 30, Fall 2000

- Time Domain Data
  - In practice one typically has a discretized sampled function $x(t_n) = x(n\Delta T)$ often referred to as a *time series* or sequence
    - Sampled data obtained from a sensor (transducer) whose output was sampled with an A/D converter
  - It is convenient and reasonable to treat these time samples collectively as single random quantity (vector):

$$\mathbf{x} = \begin{bmatrix} x(t_1) \\ x(t_2) \\ \vdots \\ x(t_L) \end{bmatrix}. \tag{1}$$

Often optimal processors are designed to process data. Such processors rely upon some statistical description of data; hence, need to treat samples collectively as single random vector.

- Temporal Frequency Domain Data
  - The frequency domain is often preferred for various reasons (filtering, analysis, spectral estimation, etc.). Samples are complex in general and of the form $X(f_m) = X(m\Delta f)$

$$x(t) \stackrel{\mathcal{F}}{\longleftrightarrow} X(f), \qquad \mathbf{X} = \begin{bmatrix} X(f_1) \\ X(f_2) \\ \vdots \\ X(f_M) \end{bmatrix} \tag{2}$$

- Space-time Domain Data
  - A system with multiple discrete sensors distributed in space sample a space-time random process $x(t, \mathbf{z})$
  - Sensor outputs produce a vector of time dependent spatial samples. A single sample across array is at time $t$ is called a *snapshot* and is given by

$$\begin{bmatrix} x(t, \mathbf{z}_1) \\ x(t, \mathbf{z}_2) \\ \vdots \\ x(t, \mathbf{z}_N) \end{bmatrix} \tag{3}$$

– when sampled at multiple points in time, one produces a *data matrix*

$$\begin{array}{c} \uparrow \\ \text{space} \\ \downarrow \end{array} \quad \overset{\leftarrow \text{ time } \rightarrow}{\begin{bmatrix} x(t_1, \mathbf{z}_1) & \cdots & x(t_L, \mathbf{z}_1) \\ x(t_1, \mathbf{z}_2) & \cdots & x(t_L, \mathbf{z}_2) \\ \vdots & \ddots & \vdots \\ x(t_1, \mathbf{z}_N) & \cdots & x(t_L, \mathbf{z}_N) \end{bmatrix}} \tag{4}$$

- Spatial frequency (wavenumber) - temporal frequency
  - Often convenient to Fourier transform data both temporally and spatially
  - Data can be filtered spatially (for example, pass signals coming from a specific direction, while suppressing all other directions) and temporally (for example, bandpass filtering)
  - Data matrix of form

$$\begin{array}{c} \uparrow \\ \text{spatial frequency} \\ \downarrow \end{array} \quad \overset{\leftarrow \text{ temporal frequency } \rightarrow}{\begin{bmatrix} X(f_1, \mathbf{k}_1) & \cdots & X(f_M, \mathbf{k}_1) \\ X(f_1, \mathbf{k}_2) & \cdots & X(f_M, \mathbf{k}_2) \\ \vdots & \ddots & \vdots \\ X(f_1, \mathbf{k}_K) & \cdots & X(f_M, \mathbf{k}_K) \end{bmatrix}} \tag{5}$$

# Real Random Vectors

- Definition: A *random variable* is a mapping from an event space to the real number line. The likelihood of event occurence is described by the cumulative distribution function (cdf) and its corresponding probability density function (pdf) defined respectively by

$$Pr\,(x \le x_0) = F_x(x_0), \qquad f_x(x_0) = \frac{dF_x(x_0)}{dx_0}. \tag{6}$$

- In the multivariate case the outcome of an experiment can consist of several events, $x_1, x_2, \ldots, x_N$. The likelihood of joint event occurence is described by a multivariate cdf:

$$Pr\,(x_1 \le a_1, x_2 \le a_2, \ldots, x_N \le a_N) = F_{x_1,\ldots,x_N}(a_1, a_2, \ldots, a_N) \triangleq F_{\mathbf{x}}(\mathbf{a}) \tag{7}$$

- The multivariate probability density function is defined in terms of the cdf

$$f_{x_1,\ldots,x_N}(a_1, a_2, \ldots, a_N) = \frac{\partial^N F_{x_1,\ldots,x_N}(a_1, a_2, \ldots, a_N)}{\partial a_1 \partial a_2 \cdots \partial a_N} \triangleq f_{\mathbf{x}}(\mathbf{a}) \tag{8}$$

- Partial statistical characterizations
  - The *Expectation* or average is defined:

$$E\left\{g(\mathbf{x})\right\} \triangleq \int_{\mathcal{R}^N} g(\mathbf{a}) f_{\mathbf{x}}(\mathbf{a}) d\mathbf{a} \tag{9}$$

  - First moment, known as Mean vector

$$\mathbf{m}_{\mathbf{x}} = E\{\mathbf{x}\} \triangleq \begin{bmatrix} \vdots \\ E\{x_i\} \\ \vdots \end{bmatrix}, \qquad E\{x_i\} = \int_{\mathcal{R}^N} a_i f_{\mathbf{x}}(\mathbf{a}) d\mathbf{a} = \int_{-\infty}^{\infty} a_i f_{x_i}(a_i) da_i \tag{10}$$

  - Second moment, known as Correlation matrix

$$\mathbf{R}_{\mathbf{x}} = E\left\{\mathbf{x}\mathbf{x}^T\right\} \triangleq \begin{bmatrix} & \vdots & \\ \cdots & E\left\{x_i x_j\right\} & \cdots \\ & \vdots & \end{bmatrix}, \tag{11}$$

$$E\left\{x_i x_j\right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a_i a_j f_{x_i,x_j}(a_i, a_j) da_i da_j$$

    where the superscript $T$ represents the matrix transpose.

  - Properties of Correlation matrix:
    * Symmetric, *i.e.* $\mathbf{R}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}^T$.
    * Positive semi-definite, $\mathbf{a}^T \mathbf{R}_{\mathbf{x}} \mathbf{a} \ge 0$, for all vectors $\mathbf{a} \in \mathcal{R}^N$.
    * Eigenvalues are real and positive, $\lambda_i(\mathbf{R}_{\mathbf{x}}) > 0$.
    * Eigenvectors are orthonormal, $\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij}$.

– Second order *central moment*, known as Covariance matrix

$$\mathbf{K_x} = E\left\{(\mathbf{x}-\mathbf{m_x})(\mathbf{x}-\mathbf{m_x})^T\right\} \triangleq \begin{bmatrix} & \vdots & \\ \cdots & E\left\{(x_i - m_{x_i})(x_j - m_{x_j})\right\} & \cdots \\ & \vdots & \end{bmatrix} \tag{12}$$
$$= \mathbf{R_x} - \mathbf{m_x}\mathbf{m_x}^T$$

Note that covariance matrix shares same properties as correlation matrix. In addition we note that

* $[\mathbf{K_x}]_{i,i} = \sigma_{x_i}^2$
* If $[\mathbf{K_x}]_{i,j} = [\mathbf{K_x}]_{|i-j|}$ then covariance matrix has a *Toeplitz* structure

$$\mathbf{K_x} = \begin{bmatrix} K_0 & K_1 & & \\ K_1 & \ddots & \ddots & \\ & \ddots & & K_1 \\ & & K_1 & K_0 \end{bmatrix} \tag{13}$$

– Coherence/normalized Correlation matrix

$$\rho_{\mathbf{x}} \triangleq \begin{bmatrix} & \vdots & \\ \cdots & E\left\{(x_i - m_{x_i})(x_j - m_{x_j})\right\} / \left(\sigma_{x_i}\sigma_{x_j}\right) & \cdots \\ & \vdots & \end{bmatrix} \tag{14}$$

* Can show via Schwartz Inequality that $|\rho_{\mathbf{x}i,j}| \leq 1$
* If $|\rho_{\mathbf{x}i,j}| = 1$, then $x_i$, and $x_j$ are said to be perfectly *correlated*. Provides a measure of linear dependence.
* if $|\rho_{\mathbf{x}i,j}| = 0$, then $x_i$, and $x_j$ are said to be *uncorrelated.*

– Useful higher central moments:

* skewness (measure of how much pdf is lumped to one side of the mean)

$$\Delta_i = \frac{E\{(x_i - m_{x_i})^3\}}{\sigma_{x_i}^3} \tag{15}$$

* kurtosis (measure of the peakedness of pdf about mean)

$$k_i = \frac{E\{(x_i - m_{x_i})^4\}}{\sigma_{x_i}^4} \tag{16}$$

- Multidimensional Characteristic Function

  – Recall that characteristic function (c.f.) of a scalar random variable is given by

$$M_x(jv) \triangleq E\{e^{jvx}\} = \int_{\mathcal{R}} e^{jva} f_x(a)da \tag{17}$$

– The c.f. of a random vector is defined as

$$M_{\mathbf{x}}(j\mathbf{v}) \triangleq E\{e^{j\mathbf{v}^T\mathbf{x}}\} = E\left\{\exp\left(j\sum_{i=1}^{N} v_i x_i\right)\right\} = \int_{\mathcal{R}^N} e^{j\mathbf{v}^T\mathbf{a}} f_{\mathbf{x}}(\mathbf{a}) d\mathbf{a} \tag{18}$$

- Cramér-Wold Theorem: The pdf of a random vector $\mathbf{x} \in \mathcal{R}^N$ is completely determined by that of $y = \mathbf{a}^T\mathbf{x}$ if the pdf of $y$ is known for all $\mathbf{a}$. Proof:

$$f_y \longleftrightarrow M_y(jv) = E\{e^{jvy}\} = E\{e^{jv\cdot\mathbf{a}^T\mathbf{x}}\}\Big|_{v=1} = M_{\mathbf{x}}(j\mathbf{a}) \longleftrightarrow f_{\mathbf{x}} \tag{19}$$

## The Multivariate Gaussian Distribution

- A very popular example of a model for random multivariate data is the multivariate Gaussian pdfs (sometimes referred as the Normal distribution). Recall that the pdf and c.f. for a Gaussian scalar random variable are given respectively by

$$f_x = \frac{1}{\sigma_x\sqrt{2\pi}} e^{\frac{-(x-m_x)^2}{2\sigma_x^2}} \longleftrightarrow e^{jvm_x - \frac{1}{2}v^2\sigma_x^2} = M_x(jv). \tag{20}$$

The multivariate extension for an $N \times 1$ vector $\mathbf{x}$ yields the following pdf

$$f_{\mathbf{x}} = (2\pi)^{-(N/2)}|\mathbf{K}_{\mathbf{x}}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m}_{\mathbf{x}})^T\mathbf{K}_{\mathbf{x}}^{-1}(\mathbf{x}-\mathbf{m}_{\mathbf{x}})\right] \tag{21}$$

and c.f.

$$M_{\mathbf{x}}(j\mathbf{v}) = \exp\left[j\mathbf{v}^T\mathbf{m}_{\mathbf{x}} - \frac{1}{2}\mathbf{v}^T\mathbf{K}_{\mathbf{x}}\mathbf{v}\right]. \tag{22}$$

where $|\mathbf{A}|$ denotes the determinant of the matrix $\mathbf{A}$.

- The Gaussian distribution is *completely* determined/specified by its mean $\mathbf{m}_{\mathbf{x}}$ and covariance $\mathbf{K}_{\mathbf{x}}$.
- If a random vector $\mathbf{x}$ is normally distributed, then one typically denotes this by writing $\mathbf{x} \sim N(\mathbf{m}_{\mathbf{x}}, \mathbf{K}_{\mathbf{x}})$.
- Linear Transformations:
  – Let $\mathbf{x} \sim N(\mathbf{m}_{\mathbf{x}}, \mathbf{K}_{\mathbf{x}})$ and consider the *affine* transformation $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$.
  – Note that the c.f. of the random vector $\mathbf{y}$ is given by

$$\begin{aligned} M_{\mathbf{y}}(j\mathbf{v}_y) &= E\{e^{j\mathbf{v}_y^T\mathbf{y}}\} = E\{e^{j\mathbf{v}_y^T\mathbf{b}} \cdot e^{j\mathbf{v}_y^T\mathbf{A}\mathbf{x}}\} = e^{j\mathbf{v}_y^T\mathbf{b}} \cdot M_{\mathbf{x}}(j\mathbf{v}_x = j\mathbf{A}^T\mathbf{v}_y) \\ &= e^{j\mathbf{v}_y^T(\mathbf{A}\mathbf{m}_{\mathbf{x}}+\mathbf{b})} \cdot \exp\left[-\frac{1}{2}\mathbf{v}_y^T\mathbf{A}\mathbf{K}_{\mathbf{x}}\mathbf{A}^T\mathbf{v}_y\right] \end{aligned} \tag{23}$$

  – Hence, $\mathbf{y} \sim N(\mathbf{A}\mathbf{m}_{\mathbf{x}} + \mathbf{b}, \mathbf{A}\mathbf{K}_{\mathbf{x}}\mathbf{A}^T)$, *i.e.* affine transformations of Gaussian random vectors result in Gaussian distributed random vectors.
- Standardized Multivariate Normal:

- Let $\mathbf{x} \sim N(\mathbf{m_x}, \mathbf{K_x})$ and consider the vector $\mathbf{z}$ given by

$$\mathbf{z} = \mathbf{K_x}^{-1/2}(\mathbf{x} - \mathbf{m_x}) = \mathbf{K_x}^{-1/2}\mathbf{x} - \mathbf{K_x}^{-1/2}\mathbf{m_x}. \tag{24}$$

- From previous result, we conclude that $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$. The distribution $N(\mathbf{0}, \mathbf{I})$ is referred as the standardized multivariate normal.
- The shift in mean centers the distribution about the origin, and the multiplication by $\mathbf{K_x}^{-1/2}$ is said to *whiten* the process.
- For the scalar case, $z = (x - m_x)/\sigma_x$, and has pdf given by

$$\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \tag{25}$$

- Let $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$, and consider $\mathbf{z}_q = \mathbf{Qz}$ where the matrix $\mathbf{Q}$ is orthogonal, *i.e.* $\mathbf{QQ}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$. Clearly, $\mathbf{z}_q \sim N(\mathbf{0}, \mathbf{I})$.
  * The vectors $\mathbf{z}$ and $\mathbf{z}_q$ have exactly the same pdf and c.f, *i.e.* they are equal in distribution. This is often denoted in literature by writing

$$\mathbf{z} \overset{d}{=} \mathbf{z}_q. \tag{26}$$

  * The pdf of $\mathbf{z}$ is of the form $f_{\mathbf{z}}(\mathbf{a}) = g(||\mathbf{a}||^2)$ and its c.f. is of the form $M_{\mathbf{z}}(j\mathbf{v}) = G(||\mathbf{v}||^2)$. Such distributions are said to be *spherically symmetric*, and as we have shown are invariant to orthogonal transformations.

• Stochastic Representation:

- Let $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$, and $\mathbf{x} \sim N(\mathbf{m_x}, \mathbf{K_x})$, and consider the vector

$$\mathbf{h} = \mathbf{m_x} + \mathbf{K_x}^{1/2}\mathbf{z}. \tag{27}$$

- Clearly, this affine transformation leads to the result $\mathbf{h} \sim N(\mathbf{m_x}, \mathbf{K_x})$, *i.e.* $\mathbf{h} \overset{d}{=} \mathbf{x}$.
- The transformation represented by $\mathbf{h}$ is very useful for computer simulation of a multivariate Gaussian of a specified mean and covariance matrix.

• Uncorrelated Gaussian random vectors:

- Let $\mathbf{x} \sim N(\mathbf{m_x}, \mathbf{K_x})$, and consider the partition

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \qquad \mathbf{K_x} = \begin{bmatrix} \mathbf{K}_{\mathbf{x}_1} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{\mathbf{x}_2} \end{bmatrix} \tag{28}$$

- If $\mathbf{K}_{12} = \mathbf{K}_{21}^T = \mathbf{0}$, then

$$\mathbf{K_x}^{-1} = \begin{bmatrix} \mathbf{K}_{\mathbf{x}_1}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{\mathbf{x}_2}^{-1} \end{bmatrix}, \tag{29}$$

and $|\mathbf{K_x}| = |\mathbf{K}_{\mathbf{x}_1}| \cdot |\mathbf{K}_{\mathbf{x}_2}|$. Consequently,

$$f_{\mathbf{x}} = f_{\mathbf{x}_1} \cdot f_{\mathbf{x}_2}. \tag{30}$$

  - Thus, if two Gaussian random vectors are uncorrelated, then they are likewise independent.

- A useful moment theorem for Gaussians:

$$E\{x_1 \cdot x_2 \cdots x_N\} = \begin{cases} 0, & N \text{ odd} \\ \sum\limits_{\substack{All \\ distinct \\ pairs}} \prod E\{x_i x_j\}, & N \text{ even} \end{cases} \tag{31}$$

For example, if $N = 4$

$$E\{x_1 x_2 x_3 x_4\} = E\{x_1 x_2\}E\{x_3 x_4\} + E\{x_1 x_3\}E\{x_2 x_4\} + E\{x_1 x_4\}E\{x_2 x_3\} \tag{32}$$

# Complex Random Scalars/Vectors

- Motivation for use of Complex Data Models

  - Perhaps the most common appearance of complex data in practice consists of the in-phase and quadrature components of demodulated data. The complex envelope represents the concatentation of the in-phase and quadrature components into a single complex quantity and is ubiquitiously present in radar, sonar, and communication systems.

    * complex representation simplifies analysis, is more intuitive, and often leads to complex dual of results known for real data.
    * can be compared to simplication of Fourier Series using complex exponentials as opposed to sines and cosines

  - The in-phase/quadrature components of demodulated data typically satisfy the circularity assumption made on the covariance structure of the corresponding real process (more said later)
  - Often processing is done in frequency domain. Fourier transform of data is complex in general, even if original time series is real.

    * Filtering and Fourier transforms involve linear transformation of data. Such transformations often satisfy assumptions of central limit theorem allowing data to be modeled as complex Gaussian.

  - Optimization simplifies significantly with complexified arithmetic (for example, complex gradient techniques)

- Goals of Complex Density:

  1. Express pdf, cdf and all related statistical measures as a function of complex variables.
  2. Obtain duality between known results for real random variables and those obtained using complex random variables.

- Concept of complex density was introduced by practicing engineers (see Wooding [3]). Rigorous theory ultimately developed by pure mathematicians (see Goodman [4], Miller [5]). A nice tutorial introduction on complex Gaussian pdf/cdf can be found in Kay [2]. First we consider a random complex scalar quantity.

- Let $\widetilde{x} = x_R + jx_I$, and $\widehat{\mathbf{x}} = [x_R, x_I]^T$. Can we write

$$f_{\widehat{\mathbf{x}}}(u, v) = f_{\widetilde{x}}(u + jv), \tag{33}$$

  where $f_{\widetilde{x}}(u + jv)$ is in general a real valued function of a complex variable?

- How would we define and interpret moments of a complex random variable? Here are the natural definitions for the first two moments:

  - First moment/mean:

$$E\{\widetilde{x}\} \triangleq E\{x_R\} + jE\{x_I\} = m_{\widetilde{x}}. \tag{34}$$

– Second moment/correlation:

$$E\{\tilde{x}\tilde{x}^*\} = E\{|\tilde{x}|^2\} = E\{x_R^2\} + E\{x_I^2\}. \tag{35}$$

– Second central moment/variance:

$$\begin{aligned} E\{(\tilde{x} - m_{\tilde{x}})(\tilde{x} - m_{\tilde{x}})^*\} &= E\{|\tilde{x}|^2\} - |E\{\tilde{x}\}|^2 \\ &= E\{x_R^2\} + E\{x_I^2\} - E^2\{x_R\} + E^2\{x_I\} \\ &= \sigma_{x_R}^2 + \sigma_{x_I}^2. \\ &\triangleq \sigma_{\tilde{x}}^2 \end{aligned} \tag{36}$$

- Note that for the second central moment, or variance of a complex scalar $\sigma_{\tilde{x}}^2$, we have a single number describing variation about the mean $m_{\tilde{x}}$ in the complex plane $x_R$, $jx_I$. This is to be contrasted with the $2 \times 2$ covariance matrix one would have for the corresponding real vector $\widehat{\mathbf{x}}$, which typically describe variation about the mean in the real 2-D plane $x_R$, $x_I$.

  – It only makes sense to constrain the variances such that $\sigma_{x_R}^2 = \sigma_{x_I}^2 = \sigma^2/2$. This leads to a pdf with circular symmetry about the mean $m_{\tilde{x}}$ in the complex plane. This is often referred to as the *circularity constraint.*

- Regarding Cross-Correlation:

$$\sigma_{\tilde{x}\tilde{y}} \triangleq E\{\tilde{x}\tilde{y}^*\} - m_{\tilde{x}} m_{\tilde{y}}^*. \tag{37}$$

  Note that when $\tilde{x} = \tilde{y}$ this reduces to the correct definition of variance.

- Concerning vector quantities, let $\tilde{\mathbf{x}} = \mathbf{x}_R + j\mathbf{x}_I$, and $\widehat{\mathbf{x}} = [\mathbf{x}_R^T, \mathbf{x}_I^T]^T$. Note that if $\tilde{\mathbf{x}}$ is an $N \times 1$ complex random vector, then $\widehat{\mathbf{x}}$ is a $2N \times 1$ real random vector. Again, we ask can we write

$$f_{\widehat{\mathbf{x}}}(\mathbf{u}, \mathbf{v}) = f_{\tilde{\mathbf{x}}}(\mathbf{u} + j\mathbf{v}), \tag{38}$$

  where $f_{\tilde{\mathbf{x}}}(\mathbf{u} + j\mathbf{v})$ is in general a real valued function of several complex variables?

- How would we define and interpret moments of a complex random vector? Here are the natural definitions for the first two moments:

  – First moment/Mean vector:

$$E\{\tilde{\mathbf{x}}\} \triangleq E\{\mathbf{x}_R\} + jE\{\mathbf{x}_I\} = \mathbf{m}_{\tilde{\mathbf{x}}}. \tag{39}$$

  – Second moment/Correlation Matrix:

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H\} = \mathbf{R}_{\tilde{\mathbf{x}}}, \tag{40}$$

  where the superscript $H$ represents the conjugate matrix transpose. Note that the correlation matrix is hermitian symmetric, $\mathbf{R}_{\tilde{\mathbf{x}}} = \mathbf{R}_{\tilde{\mathbf{x}}}^H$. The outer product can be expanded as

$$\begin{aligned} \tilde{\mathbf{x}}\tilde{\mathbf{x}}^H &= (\mathbf{x}_R + j\mathbf{x}_I)(\mathbf{x}_R^T - j\mathbf{x}_I^T) \\ &= \mathbf{x}_R\mathbf{x}_R^T + \mathbf{x}_I\mathbf{x}_I^T + j(\mathbf{x}_I\mathbf{x}_R^T - \mathbf{x}_R\mathbf{x}_I^T). \end{aligned} \tag{41}$$

  Thus, the correlation matrix consists of the terms

$$\mathbf{R}_{\tilde{\mathbf{x}}} = \mathbf{R}_{\mathbf{x}_R} + \mathbf{R}_{\mathbf{x}_I} + j(\mathbf{R}_{IR} - \mathbf{R}_{RI}). \tag{42}$$

- Second central moment/Covariance Matrix:

$$\begin{aligned}\mathbf{K}_{\tilde{\mathbf{x}}} &= E\{(\tilde{\mathbf{x}} - \mathbf{m}_{\tilde{\mathbf{x}}})(\tilde{\mathbf{x}} - \mathbf{m}_{\tilde{\mathbf{x}}})^H\} \\ &= \mathbf{R}_{\tilde{\mathbf{x}}} - \mathbf{m}_{\tilde{\mathbf{x}}}\mathbf{m}_{\tilde{\mathbf{x}}}^H \\ &= \mathbf{K}_{\mathbf{x}_R} + \mathbf{K}_{\mathbf{x}_I} + j(\mathbf{K}_{IR} - \mathbf{K}_{RI})\end{aligned} \tag{43}$$

- Properties of Correlation/Covariance matrix:
    * Hermitian symmetric, *i.e.* $\mathbf{K}_{\tilde{\mathbf{x}}} = \mathbf{K}_{\tilde{\mathbf{x}}}^H$.
    * Positive semi-definite, $\mathbf{a}^H\mathbf{K}_{\tilde{\mathbf{x}}}\mathbf{a} \geq 0$, for all complex vectors $\mathbf{a} \in \mathcal{C}^N$.
    * Eigenvalues are real and positive, $\lambda_i(\mathbf{K}_{\tilde{\mathbf{x}}}) > 0$.
    * Eigenvectors are orthogonal, $\mathbf{q}_i^H\mathbf{q}_j = \delta_{ij}$.

- Regarding Cross-Correlation:

$$\mathbf{R}_{\tilde{\mathbf{x}}\tilde{\mathbf{y}}} \triangleq E\{\tilde{\mathbf{x}}\tilde{\mathbf{y}}^H\} - \mathbf{m}_{\tilde{\mathbf{x}}}\mathbf{m}_{\tilde{\mathbf{y}}}^H. \tag{44}$$

Note that when $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$ this reduces to the definition the correlation matrix $\mathbf{R}_{\tilde{\mathbf{x}}}$.

- Concerning the equality of density functions and corresponding c.f. suggested by eq(38), lets consider pdfs possessing elliptical symmetry (for example, the multivariate Gaussian).
    - Let

$$\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix}, \quad \widehat{\mathbf{m}}_{\hat{\mathbf{x}}} = \begin{bmatrix} \mathbf{m}_{\mathbf{x}_R} \\ \mathbf{m}_{\mathbf{x}_I} \end{bmatrix}, \quad \hat{\mathbf{v}} = \begin{bmatrix} \mathbf{v}_R \\ \mathbf{v}_I \end{bmatrix}, \quad \mathbf{K}_{\hat{\mathbf{x}}} = \begin{bmatrix} \mathbf{K}_{\mathbf{x}_R} & \mathbf{K}_{IR} \\ \mathbf{K}_{RI} & \mathbf{K}_{\mathbf{x}_R} \end{bmatrix} \tag{45}$$

$$\tilde{\mathbf{x}} = \mathbf{x}_R + j\mathbf{x}_I, \quad \mathbf{m}_{\tilde{\mathbf{x}}} = \mathbf{m}_{\mathbf{x}_R} + j\mathbf{m}_{\mathbf{x}_I}, \quad \tilde{\mathbf{v}} = \mathbf{v}_R + j\mathbf{v}_I, \quad \mathbf{K}_{\tilde{\mathbf{x}}} = \mathbf{K}_R + j\mathbf{K}_I \tag{46}$$

- Let the pdfs be of the form

$$\begin{aligned}f_{\hat{\mathbf{x}}} &= g_1\left[(\hat{\mathbf{x}} - \mathbf{m}_{\hat{\mathbf{x}}})^T\mathbf{K}_{\hat{\mathbf{x}}}^{-1}(\hat{\mathbf{x}} - \mathbf{m}_{\hat{\mathbf{x}}}), |\mathbf{K}_{\hat{\mathbf{x}}}|\right] \\ f_{\tilde{\mathbf{x}}} &= g_2\left[(\tilde{\mathbf{x}} - \mathbf{m}_{\tilde{\mathbf{x}}})^H\mathbf{K}_{\tilde{\mathbf{x}}}^{-1}(\tilde{\mathbf{x}} - \mathbf{m}_{\tilde{\mathbf{x}}}), |\mathbf{K}_{\tilde{\mathbf{x}}}|\right]\end{aligned} \tag{47}$$

Random vectors with pdfs satisfying this property are known in the literature as elliptically contoured distributed. Often they are presumptuously referred as spherically invariant random vectors/processes (SIRVs or SIRPs)[1]. The concomitant c.f.'s can be shown to similarly be of the form

$$\begin{aligned}M_{\hat{\mathbf{x}}}(j\hat{\mathbf{v}}) &= \phi_1\left(\hat{\mathbf{v}}^T\mathbf{m}_{\hat{\mathbf{x}}}, \hat{\mathbf{v}}^T\mathbf{K}_{\hat{\mathbf{x}}}\hat{\mathbf{v}}\right) \\ M_{\tilde{\mathbf{x}}}(j\tilde{\mathbf{v}}) &= \phi_2\left(\tilde{\mathbf{v}}^H\mathbf{m}_{\tilde{\mathbf{x}}}, \tilde{\mathbf{v}}^H\mathbf{K}_{\tilde{\mathbf{x}}}\tilde{\mathbf{v}}\right)\end{aligned} \tag{48}$$

[1]The presumption comes from the fact that often in a statistician's initial analysis it is assumed that the mean and covariance of a random process is known. One can, therefore, always center and whiten the process to produce a spherically symmetric random vector. In practice, however, it is seldom that one knows these parameters a priori. Often the mean and covariance must be estimated from sample data. The estimation process often limits our ability to fully center and whiten a random process.

– It can be shown [2] that if the covariance $\mathbf{K}_{\widehat{\mathbf{x}}}$ of the original real $2N \times 1$ vector is constrained such that

$$\begin{aligned} \mathbf{K}_{\mathbf{x}_R} &= \mathbf{K}_{\mathbf{x}_I} = \mathbf{K}_R/2 \\ \mathbf{K}_{IR} &= -\mathbf{K}_{RI}^T = \mathbf{K}_I/2, \end{aligned} \tag{49}$$

(these relations represent the generalization of the circularity constraint) then we obtain the following equivalences

$$\begin{aligned} \frac{1}{2}(\widehat{\mathbf{x}} - \mathbf{m}_{\widehat{\mathbf{x}}})^T \mathbf{K}_{\widehat{\mathbf{x}}}^{-1} (\widehat{\mathbf{x}} - \mathbf{m}_{\widehat{\mathbf{x}}}) &= (\widetilde{\mathbf{x}} - \mathbf{m}_{\widetilde{\mathbf{x}}})^H \mathbf{K}_{\widetilde{\mathbf{x}}}^{-1} (\widetilde{\mathbf{x}} - \mathbf{m}_{\widetilde{\mathbf{x}}}) \\ 2^{-(2N)/2} |\mathbf{K}_{\widehat{\mathbf{x}}}|^{-1/2} &= |\mathbf{K}_{\widetilde{\mathbf{x}}}|^{-1} \\ \widehat{\mathbf{v}}^T \mathbf{m}_{\widehat{\mathbf{x}}} &= \text{Re}\left\{\widetilde{\mathbf{v}}^H \mathbf{m}_{\widetilde{\mathbf{x}}}\right\} \\ \widehat{\mathbf{v}}^T \mathbf{K}_{\widehat{\mathbf{x}}} \widehat{\mathbf{v}} &= \frac{1}{2} \widetilde{\mathbf{v}}^H \mathbf{K}_{\widetilde{\mathbf{x}}} \widetilde{\mathbf{v}}. \end{aligned} \tag{50}$$

– These basic relations allows an exact algebraic equivalence to be drawn between the real and complex densities, *i.e.*

$$\begin{aligned} f_{\widehat{\mathbf{x}}}(\mathbf{u}, \mathbf{v}) &= f_{\widetilde{\mathbf{x}}}(\mathbf{u} + j\mathbf{v}) \\ M_{\widehat{\mathbf{x}}}(j\widehat{\mathbf{v}}) &= M_{\widetilde{\mathbf{x}}}(j\widetilde{\mathbf{v}}) \end{aligned} \tag{51}$$

– Next, as an example, we consider the most frequently used complex distribution.

- **Complex Multivariate Gaussian Distribution**

– Applying the appropriate equivalences to the multivariate real Gaussian distributed $2N \times 1$ random vector $\widehat{\mathbf{x}}$ one obtains the following multivariate complex Gaussian pdf for the $N \times 1$ complex random vector $\widetilde{\mathbf{x}}$:

$$f_{\widetilde{\mathbf{x}}} = \pi^{-N} |\mathbf{K}_{\widetilde{\mathbf{x}}}|^{-1} \exp\left[-(\widetilde{\mathbf{x}} - \mathbf{m}_{\widetilde{\mathbf{x}}})^H \mathbf{K}_{\widetilde{\mathbf{x}}}^{-1} (\widetilde{\mathbf{x}} - \mathbf{m}_{\widetilde{\mathbf{x}}})\right] \tag{52}$$

$$M_{\widetilde{\mathbf{x}}}(j\widetilde{\mathbf{v}}) \triangleq E\left\{\exp\left[j\text{Re}(\widetilde{\mathbf{v}}^H \widetilde{\mathbf{x}})\right]\right\} = \exp\left[j\text{Re}(\widetilde{\mathbf{v}}^H \mathbf{m}_{\widetilde{\mathbf{x}}}) - \frac{1}{4}\mathbf{v}^H \mathbf{K}_{\widetilde{\mathbf{x}}} \widetilde{\mathbf{v}}\right]. \tag{53}$$

– This is denoted by writing $\widetilde{\mathbf{x}} \sim CN(\mathbf{m}_{\widetilde{\mathbf{x}}}, \mathbf{K}_{\widetilde{\mathbf{x}}})$.

– When $N = 1$ one obtains

$$f_{\widetilde{x}} = \frac{1}{\pi \sigma_{\widetilde{x}}^2} \exp\left[-\frac{|\widetilde{x} - m_{\widetilde{x}}|^2}{\sigma_{\widetilde{x}}^2}\right] \tag{54}$$

$$M_{\widetilde{x}}(jv) = \exp\left[j\text{Re}(\widetilde{v}^* m_{\widetilde{x}}) - \frac{1}{4}|\widetilde{v}|^2 \sigma_{\widetilde{x}}^2\right]. \tag{55}$$

– All the properties of real multivariate Gaussian random vectors persist in an equivalent form for the complex multivariate Gaussian (for example, linear combinations of complex Gaussians produce complex Gaussians).

# Stochastic Processes

- Recall that a random variable is defined as a mapping from an event space $\Omega$ to the real number line, *i.e.* for $\omega \in \Omega$ one obtains $x = x(\omega) \in \mathcal{R}$. A stochastic process is similar, except a time-function is assigned to every event.
- A Stochastic Process (SP) is a rule for assigning to every event $\zeta_i \in Z$ a time function $x(t, \zeta_i)$. This function represents a single realization, or event. Sometimes called a sample function.
- The collection of all possible realizations of the stochastic process is called the ensemble. This family of time functions depends on both event $\zeta$ and time $t$.
- In practice, one samples a waveform typically modeled as a stochastic process. These samples collectively represent a multidimensional random vector $[x(t_1), x(t_2), \ldots, x(t_N)]^T$.
  - For any fixed $t_0$ ($\zeta$ varies) a sample of the process $x(t_0)$ is a random variable
  - For any fixed $\zeta_0$ ($t$ varies) one obtains a sample time function.
- Characterization of stochastic processes:
  - First order distribution (cdf/pdf):

$$F_{x(t_1)}(x_1) = Pr\left[x(t_1) \leq x_1\right], \qquad f_{x(t_1)}(x_1) = \frac{dF_{x(t_1)}(x_1)}{dx_1} \tag{56}$$

  - Second order cdf/pdf:

$$\begin{aligned} F_{x(t_1),x(t_2)}(x_1, x_2) &= Pr\left[x(t_1) \leq x_1, x(t_2) \leq x_2\right], \\ f_{x(t_1),x(t_2)}(x_1, x_2) &= \frac{dF_{x(t_1),x(t_2)}(x_1, x_2)}{\partial x_1 \partial x_2} \end{aligned} \tag{57}$$

  - N-th order cdf/pdf

$$F_{x(t_1),,x(t_2),\ldots,x(t_N)}\left(x_1, x_2, \ldots, x_N\right) = Pr\left[x(t_1) \leq x_1, x(t_2) \leq x_2, \ldots, x(t_N) \leq x_N\right] \tag{58}$$

$$f_{x(t_1),,x(t_2),\ldots,x(t_N)}\left(x_1, x_2, \ldots, x_N\right) = \frac{\partial^N F_{x(t_1),,x(t_2),\ldots,x(t_N)}\left(x_1, x_2, \ldots, x_N\right)}{\partial x_1 \cdots \partial x_N} \tag{59}$$

  - A *complete characterization* of a stochastic processes requires knowledge of this $N$-th order distribution function for all possible $t_i$, and for all $N > 0$.
  - If cdfs/pdfs are only known for $N$-tuples of $t_i$'s up to a fixed value of $N$, then we say that the stochastic process is only partially characterized.
- Statistics of Stochastic Processes
  - Mean (1st moment)

$$\eta_x(t) \stackrel{\triangle}{=} E\left\{x(t)\right\} = \int a f_{x(t)}(a) da \tag{60}$$

- Autocorrelation (2nd moments)

$$R_x(t_1,t_2) = E\left\{x(t_1)x^*(t_2)\right\} = \int\int a_1 a_2^* f_{x(t_1),x(t_2)}(a_1,a_2)da_1da_2 \tag{61}$$

  Note that $R_x(t,t) = E\{|x(t)|^2\}$, the average power. Symmetric, $R_x(t_1,t_2) = R_x^*(t_2,t_1)$.
- Auto-covariance (central moments)

$$K_x(t_1,t_2) = R_x(t_1,t_2) - \eta_x(t_1)\eta_x^*(t_2) \tag{62}$$

  Note that $K_x(t,t) = E\{|x(t) - \eta_x(t)|^2\}$, variance of $x(t)$.
- Cross-correlation/covariance

$$\begin{aligned} R_{xy}(t_1,t_2) &= E\{x(t_1)y^*(t_2)\} \\ K_{xy}(t_1,t_2) &= R_{xy}(t_1,t_2) - \eta_x(t_1)\eta_y^*(t_2) \end{aligned} \tag{63}$$

  Symmetric, $R_{xy}(t_1,t_2) = R_{yx}^*(t_2,t_1)$, and $K_{xy}(t_1,t_2) = K_{yx}^*(t_2,t_1)$.
  * Two processes $x(t)$ and $y(t)$ are said to be *orthogonal* if $R_{xy}(t_1,t_2) = 0$ for all $t_1,t_2$.
  * Two processes $x(t)$ and $y(t)$ are said to be *uncorrelated* if $K_{xy}(t_1,t_2) = 0$ for all $t_1,t_2$.

• Uncorrelated increment / Independent increment

- Let $t_1 < t_2 \le t_3 < t_4$, and $a = x(t_2) - x(t_1)$, $b = x(t_4) - x(t_3)$.
- If $a$ and $b$ are uncorrelated for any $t_i$, $i = 1,2,3,4$, then we say that the process has uncorrelated increments.
- If $a$ and $b$ are independent for any $t_i$, $i = 1,2,3,4$, then we say that the process has independent increments.
- For Gaussian processes uncorrelated increments implies independent increments.

• Joint Processes

- The complete characterization of joint processes $x(t)$ and $y(t)$ requires knowledge of the joint cdfs/pdfs

$$F_{x(t_1),,x(t_2),\ldots,x(t_N),y(t_1),y(t_2),\ldots,y(t_M)}\left(x_1,x_2,\ldots,x_N,y_1,y_2,\ldots,y_M\right) \tag{64}$$

$$f_{x(t_1),\ldots,x(t_N),y(t_1),\ldots,y(t_M)}\left(x_1,\ldots,x_N,y_1,\ldots,y_M\right) = \tag{65}$$

$$\frac{\partial^{(N+M)}F_{x(t_1),\ldots,x(t_N),y(t_1),\ldots,y(t_M)}\left(x_1,\ldots,x_N,y_1,\ldots,y_M\right)}{\partial x_1\cdots\partial x_N,\partial y_1\cdots\partial y_M}$$

  for all $t_n,t_m$, $n = 1,\ldots,N$, $m = 1,\ldots,M$, and all $N > 0$, and $M > 0$.

• Stationarity

- Strict (Strong) Sense Stationary (SSS): A stochastic process is said to be stationary in the strict sense if its statistical characteristics/properties are invariant to a shift of the time origin.
  * One consequence of this is that the all $N$-th order cdfs/pdfs are invariant to a uniform shift of $t_i$'s:

$$f_{x(t_1),\ldots,x(t_N)} = f_{x(t_1+\tau),\ldots,x(t_N+\tau)}. \tag{66}$$

  * Consider first order descriptions. SSS requires that $f_{x(t_1+\tau)} = f_{x(t_1)}$ for all $t_1$ and all $\tau$. If we choose $\tau = -t_1 + t$, then clearly $f_{x(t_1)}$ must be independent of time for SSS to hold. Thus, mean is constant

$$\eta_x(t) = E\{x(t)\} = \int a f_{x(t)}(a) da = \eta \quad \text{constant} \tag{67}$$

  * Consider second order descriptions: SSS requires that $f_{x(t_1+\tau),x(t_1)} = g(\tau)$.
- Wide (Weak) Sense Stationary (WSS): A stochastic process is said to be stationary in the weak sense if
  1. The mean of the process is constant, *i.e.* $E\{x(t)\} = \eta_x$.
  2. Autocorrelation function $R(t_1, t_2)$ depends only on the time difference $\tau = t_1 - t_2$. Thus,

$$\begin{aligned} R_x(t+\tau, t) &= E\{x(t+\tau)x^*(t)\} = R_x(\tau) \\ K_x(t+\tau, t) &= R_x(\tau) - |\eta_x|^2 = K_x(\tau) \end{aligned} \tag{68}$$

  * Symmetries become $R_x(\tau) = R_x^*(-\tau)$, and $K_x(\tau) = K_x^*(-\tau)$.
  * Note that $R_x(0) = E\{|x(t)|^2\}$, average power.
  * If a Gaussian process is WSS, then it is also SSS.
- Nearly WSS
  * If process $z(t)$ has mean $E\{z(t)\} = \eta_z(t)$ and autocorrelation $R_z(\tau)$, then we can create a WSS process $x(t)$ by letting $x(t) = z(t) - \eta_z(t)$. Note that $\eta_x(t) = 0$ and $R_x(\tau) = R_z(\tau)$.
- Jointly stationary: Two processes $x(t)$ and $y(t)$ are said to be jointly stationary if the joint statistics of $x(t)$ and $y(t)$ are the same as $x(t+\tau)$ and $y(t+\tau)$ for any $\tau$.
  * A complex process $\tilde{z}(t) = x(t) + jy(t)$ is said to be stationary if the processes $x(t)$ and $y(t)$ are jointly stationary.
- Processes $x(t)$ and $y(t)$ are said to be jointly WSS if
  1. $x(t)$ is WSS, and $y(t)$ is WSS.
  2. Cross-correlation function depends on time difference,

$$\begin{aligned} R_{xy}(t+\tau, t) &= E\{x(t+\tau)y^*(t)\} = R_{xy}(\tau) \\ K_{xy}(t+\tau, t) &= R_{xy}(\tau) - \eta_x\eta_y^* = K_{xy}(\tau) \end{aligned} \tag{69}$$

– Note by the Schwartz Inequality that

$$\begin{aligned}
|E\{x(t+\tau)y^*(t)\}|^2 &= \left|\int\int a_x a_y^* f_{x(t+\tau),y(t)}(a_x,a_y)da_x da_y\right|^2 \\
&= \left|\int\int a_x\sqrt{f_{x(t+\tau),y(t)}(a_x,a_y)}a_y^*\sqrt{f_{x(t+\tau),y(t)}(a_x,a_y)}da_x da_y\right|^2 \quad (70)\\
&\leq \int |a_x|^2 f_{x(t)}(a_x)da_x \int |a_y|^2 f_{y(t)}(a_y)da_y
\end{aligned}$$

Hence,

$$|R_{xy}(\tau)|^2 \leq R_x(0)\cdot R_y(0). \quad (71)$$

When $x(t)=y(t)$ we obtain the property that $|R_x(\tau)| \leq R_x(0)$.

- Ergodicity: A stochastic process is said to be ergodic if its ensemble averages equal appropriate time averages. Thus, any statistic of process $x(t)$ can be determined from a single sample function $x(t,\zeta)$.

  – For example, given process $x(t)$ with constant mean $E\{x(t)\}=\eta_x$, one could consider estimating this mean from a single sample function with the time average:

$$\widehat{\eta}_x = \frac{1}{T}\int_{-T/2}^{T/2} x(t,\zeta)dt. \quad (72)$$

# Power Spectral Density

- The power spectral density (PSD) is defined for WSS processes as the Fourier transform of the auto-correlation function:

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau. \tag{73}$$

Clearly,

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(f) e^{j2\pi f\tau} df. \tag{74}$$

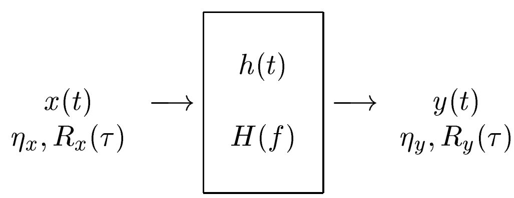
Note that

$$R_x(0) = E\{|x(t)|^2\} = \int_{-\infty}^{\infty} S_x(f) df. \tag{75}$$

$S_x(f)$ represents the distribution of average signal power over frequency. Hence, the name PSD.

# WSS Stochastic Processes Thru LTI Systems

- Consider the a linear time-invariant (LTI) system with impulse response $h(t)$ and system function $H(f)$ that is excited by a WSS stochastic process $x(t)$:




(76)

where the output and input are related through convolution:

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(a) h(t-a) da = \int_{-\infty}^{\infty} x(t-a) h(a) da. \tag{77}$$

- How are the statistical properties of the output process $y(t)$ related to those of input process $x(t)$?

  - Means:

$$E\{y(t)\} = \int_{-\infty}^{\infty} E\{x(t-a)\} h(a) da = \eta_x \cdot \int_{-\infty}^{\infty} h(a) da \tag{78}$$

Thus,

$$\boxed{E\{y(t)\} = \eta_x \cdot H(0)} \tag{79}$$

– Cross-correlation: Note that

$$x(t+\tau)y^*(t) = \int_{-\infty}^{\infty} x(t+\tau)x^*(t-b)h^*(b)db. \quad (80)$$

Thus, taking expectation one obtains

$$\boxed{R_{xy}(\tau) = \int_{-\infty}^{\infty} R_x(\tau+b)h^*(b)db = R_x(\tau) * h^*(-\tau).} \quad (81)$$

PSD relations obtained in frequency domain:

$$\boxed{S_{xy}(f) = S_x(f)H^*(f)} \quad (82)$$

– Autocorrelation: Note that

$$y(t+\tau)y^*(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t+\tau-a)x^*(t-b)h(a)h^*(b)dadb. \quad (83)$$

Thus, taking expectation one obtains

$$\begin{aligned} R_y(\tau) &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} R_x(\tau+b-a)h^*(b)db\right]h(a)da \\ &= R_{xy}(\tau) * h(\tau). \end{aligned} \quad (84)$$

Thus,

$$\boxed{R_y(\tau) = R_x(\tau) * h^*(-\tau) * h(\tau).} \quad (85)$$

PSD relations obtained in frequency domain:

$$\boxed{S_y(f) = S_x(f)H^*(f)H(f) = S_x(f)|H(f)|^2.} \quad (86)$$

- Note that if we choose the following bandpass filter

$$H(f) = \begin{cases} 1/\sqrt{\Delta f}, & |f-f_0| \le \Delta f \\ 0, & \text{Otherwise} \end{cases} \quad (87)$$

it follows that

$$0 \le E\{|y(t)|^2\} = R_y(0) = \int_{-\infty}^{\infty} S_y(f)df = \int_{f_0-\frac{\Delta f}{2}}^{f_0+\frac{\Delta f}{2}} \frac{S_x(f_0)}{\Delta f}df \approx S_x(f_0). \quad (88)$$

Since $f_0$ can be chosen arbitrarily we have the property that a PSD is real and non-negative:

$$\boxed{S_x(f) \ge 0 \quad \text{for all} \quad f.} \quad (89)$$

# References


[1] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Inc., 1985, Chapters 9-10.

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, PTR Prentice-Hall, Inc., 1993, Chapter 15.

[3] R. A. Wooding, "The Multivariate Distribution of Complex Normal Variables," *Biometrika*, Vol. 43, pp. 212–215 (1956).

[4] N. R. Goodman, "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," *Ann. Math. Stat.*, Vol. 34, 152–177 (1963).

[5] K. S. Miller, *Multidimensional Gaussian Distributions*, John Wiley and Sons, Inc., New York, 1964.