Published in final edited form as:

J Opt Soc Am A Opt Image Sci Vis. 2016 August 01; 33(8): 1464–1475.

Characteristic functionals in imaging and image-quality assessment: tutorial

Eric Clarkson* and Harrison H. Barrett

Department of Medical Imaging, College of Optical Sciences, Applied Mathematics Program,
University of Arizona, Tucson, Arizona 85721, USA

Abstract

Characteristic functionals are one of the main analytical tools used to quantify the statistical properties of random fields and generalized random fields. The viewpoint taken here is that a random field is the correct model for the ensemble of objects being imaged by a given imaging system. In modern digital imaging systems, random fields are not used to model the reconstructed images themselves since these are necessarily finite dimensional. After a brief introduction to the general theory of characteristic functionals, many examples relevant to imaging applications are presented. The propagation of characteristic functionals through both a binned and list-mode imaging system is also discussed. Methods for using characteristic functionals and image data to estimate population parameters and classify populations of objects are given. These methods are based on maximum likelihood and maximum *a posteriori* techniques in spaces generated by sampling the relevant characteristic functionals through the imaging operator. It is also shown how to calculate a Fisher information matrix in this space. These estimators and classifiers, and the Fisher information matrix, can then be used for image quality assessment of imaging systems.

1. INTRODUCTION

If we are imaging an ensemble of objects, then it is natural to describe this ensemble as a random field on some support region in $\mathbb{R}^q$. (Other terms used for this concept are random function and stochastic process.) For example, in medical imaging, the object is described by a function of three variables that varies unpredictably from one patient to the next. In x-ray computed tomography, this function gives the x-ray attenuation at each point, while in position emission tomography and single photon emission computed tomography imaging, this function gives the activity distribution inside the body. While there are anatomical structures that are common to all patients, there are also unpredictable variations in these structures throughout the patient population. As another example, for the x-ray imaging of luggage, there is an almost infinite variety of objects that can be packed in a bag and also infinite variation in the arrangement of these objects in the bag. The corresponding attenuation maps then vary in an unpredictable way from one bag to the next. In each of these examples, any given attenuation map or activity distribution can be regarded as a realization of a random field defined on a region of space. In medical imaging, this support

*Corresponding author: clarkson@radiology.arizona.edu.

OCIS codes: (110.0110) Imaging systems; (110.2960) Image analysis; (110.3000) Image quality assessment.

region can be thought of as a cylinder large enough to contain all of the patients being imaged. In a luggage scanner, the support set may be a box shape large enough to contain all of the bags being scanned. In all imaging situations, there will be unpredictable variations in the function being imaged. If there were no such variations, then the object function would be known already and there would be no need to create an image. Therefore, we will consider the object functions being imaged to be realizations of a random field defined on a support set.

Describing the statistical properties of a random field can be difficult since we are dealing with essentially a random vector in an infinite-dimensional space. The concept of a probability density function (PDF), which is used to describe the statistics of continuous finite-dimensional random vectors, is not defined in general for a random field. For finite-dimensional random vectors, an equivalent description of the statistics is given by the characteristic function, which is the Fourier transform of the PDF. Indeed, many concepts in probability theory and multivariate statistics are easier to understand if we use the characteristic function as our description of the random vector. Fortunately, the concept of the characteristic function generalizes easily to random fields and is referred to as the characteristic functional in this context.

After a brief discussion of probability models and random variables in Section 2 to establish notation, we introduce the idea of a random field in Section 3. In Section 4, we define the characteristic functional for a random field and discuss some of its properties. Sections 5 and 6 introduce generalized random fields and their characteristic functionals; these ideas are needed to describe point processes, for example. In Section 7, we show how certain mathematical operations with random fields translate to relations among the corresponding characteristic functionals. Section 8 consists of a myriad of examples of random fields and their characteristic functionals that are relevant to imaging applications. We then show how characteristic functionals are propagated through noisy imaging systems in Section 9. This discussion includes both binned and list-mode imaging systems. Finally, in Section 10, we show how characteristic functionals and image data can be used to estimate population parameters, classify populations of objects, and compute a Fisher information matrix for the estimation of population parameters. In all of these applications, an analytic expression for the characteristic functional is essential to carrying out the task. Fortunately, in Section 8, we have a long list of such analytic expressions. We hope that these results will convince readers that modeling object populations as random fields and describing their statistical properties with characteristic functionals is both useful and important for the analysis of imaging systems and the data generated by them.

Since this paper is intended as a tutorial on the use of characteristic functionals in image science, most of the results have been derived elsewhere. Exceptions to this rule are Subsection 8.L on dynamically evolving random fields and Sections 10 and 11 on using characteristic functionals to generate mathematical observers and figures of merit for imaging systems.

2. PROBABILITY MODELS AND RANDOM VARIABLES

To define the characteristic functional of a random field, we must first discuss the definition of a random field. This definition in turn necessitates a brief discussion of probability spaces and random variables.

To define a probability space [1–3] we start with a set Ω , sometimes called the sample space, the space of outcomes, or the set of elementary events. This set may be finite. For example, if N photons strike a detector and n of them are detected, then $\Omega = \{0, 1, 2, \dots, N\}$ because these are the possible outcomes from this experiment. The outcome space may be infinite but discrete. An example of this situation is operating a photon-counting detector being illuminated by a source for a fixed amount of time. In this case, it is standard practice to take $\Omega = \{0, 1, 2, \dots\}$. If an imaging system performs M real-valued measurements of an object, then $\Omega = \mathbb{R}^M$. In the theory of random fields, we have to allow for the possibility that Ω is an infinite-dimensional space. The set of all real-valued functions of q variables is an example of such a space. In general, we will use the symbol ω to represent a point in Ω . Thus, ω may represent an integer, a real number, a finite-dimensional vector, a function, or some other more abstract way of representing outcomes.

The next step in defining a probability model is to define the collection $\mathcal{A}$ of subsets of Ω that will be assigned a probability. These subsets are usually called events. When Ω is a discrete set, $\mathcal{A}$ is usually simply the collection of all possible subsets of Ω . Thus, in this case, we will assign a probability to any subset of outcomes. When Ω is not a discrete set, then $\mathcal{A}$ normally does not consist of all possible subsets of Ω . In general, all that we require of the collection $\mathcal{A}$ is three conditions. The first condition is that Ω itself is one of the subsets in the collection. The second condition is that, for any subset A in the collection, the complement of A , which is the subset of points in Ω that are not in A , is also in the collection. The third condition is that, if the subsets A_n are in the collection for $n = 1, 2, \dots$, then the subset consisting of the union of these subsets,

$$A = \bigcup_{n=1}^{\infty} A_n, \quad (1)$$

is also in the collection. A collection of subsets $\mathcal{A}$ satisfying these properties is called a σ -algebra.

The final ingredient for a probability model is a probability measure $Pr(A)$ that assigns a probability to each event A in $\mathcal{A}$. We always have $0 \leq Pr(A) \leq 1$. To state the only other condition on the probability measure, we begin with a sequence of subsets A_n that are in $\mathcal{A}$ for $n = 1, 2, \dots$. We assume that the events are mutually disjoint, i.e., $A_n \cap A_m$ is empty for $n \neq m$. Then, if A is the union event as just described, the probability measure must satisfy

$$Pr(A) = \sum_{n=1}^{\infty} Pr(A_n). \quad (2)$$

This is just the usual condition that probabilities add when the events are disjoint, and is called countable additivity.

A probability model is specified by the triple $(\Omega, \mathcal{F}, Pr)$ consisting of a sample space, a σ -algebra of subsets of this space, and a probability measure on this collection of subsets. The concept of a probability model forms the foundation of probability theory, which in turn is essential for any discussion of random fields and characteristic functionals. For those not familiar with σ -algebras and probability measures, we will discuss a few examples.

In the first example, we will consider the photon-counting system where $\Omega = \{0, 1, 2, \dots\}$ and an event E will be any subset of Ω . Since the one-element set $\{n\}$ is an event, it has an associated probability $P_n = Pr(\{n\})$. For any event, we then have

$$Pr(E) = \sum_{n \in E} P_n. \quad (3)$$

This is just the usual way to assign a probability to a subset of a discrete outcome space.

In the second example, we will take Ω to be the set $\mathbb{R}$ of all real numbers. If the outcome of an experiment is a measurement of a continuous quantity, then this is the appropriate outcome space. If the real-valued function $pr(x)$ on $\mathbb{R}$ satisfies $pr(x) \geq 0$ for all x and

$$\int_{\mathbb{R}} pr(x) dx = 1, \quad (4)$$

then we say that this function is a PDF. The corresponding probability measure is defined by

$$Pr(E) = \int_E pr(x) dx, \quad (5)$$

when the integral exists. This last caveat is important, since there are subsets of $\mathbb{R}$ for which this integral does not exist. The existence of these non-measurable sets has been known for some time and was a motivation for restricting the subsets to which we want to assign a probability, the events, to all be members of a σ -algebra. For the interested reader it is not difficult to find descriptions of such sets, but they are extremely unlikely to arise in applications. We say that a σ -algebra is generated by a collection of subsets of Ω if it is the smallest σ -algebra containing all of those subsets. When Ω is the real line $\mathbb{R}$, then we normally use the σ -algebra generated by the intervals, the Borel σ -algebra, to define a

probability measure. For events in the Borel σ -algebra, the integral in Eq. (5) is always defined.

As a final example of a situation where we need to restrict the subsets to which we want to assign a probability, consider Ω to be the space of all real-valued functions on $\mathbb{R}$. We may think of an experiment where the outcome is an oscilloscope trace. In a typical measurement, we want to assign a probability to an event of the form $E = \{f: a_n \leq f(x_n) \leq b_n \text{ for } n = 1, 2, \dots, N\}$. This event corresponds to a finite number of measurements of the function $f(x)$ at the points x_n , with the results falling into the given intervals. We can then take $\mathcal{F}$ to be the σ -algebra generated by these subsets for our probability model. In this case, however, the subset $C = \{f: a \leq f(x) \leq b \text{ for } c < x < d\}$ is not a member of $\mathcal{F}$ and cannot be assigned a probability since it requires an uncountably infinite number of measurements to verify that a given function is in C . We are usually not bothered by this, since we cannot make an uncountably infinite number of measurements anyway. However, this example does show that the natural σ -algebra for a probability model may be very far from containing all of the subsets of Ω . In any case, we will only use the general concept of a probability model as just outlined to define what a random field is and what the characteristic functional of a random field is. Once that is accomplished, we will not have to worry very much about the details of the probability model used to define any given random field.

A real random variable is a real-valued function $X = f(\omega)$ on the sample space Ω that satisfies the following measurability condition. For any real number a , the subset of Ω defined by $E_X(a) = \{\omega: f(\omega) \leq a\}$ is in the collection $\mathcal{F}$ of events, and may thus be assigned a probability. A complex random variable $Z = f(\omega)$ is defined by insisting that $\text{Re}Z$ and $\text{Im}Z$ are both real random variables. We can also have real random vectors $\mathbf{X} = \mathbf{f}(\omega)$ of any dimension by insisting that each component X_n of the vector is a real random variable. This definition is also valid for complex random vectors.

For a real random variable X , the PDF $pr_X(x)$ is defined by the condition

$$\int_{-\infty}^a pr_X(x) dx = Pr[E_X(a)] \quad (6)$$

for any real number a . We will use the angle-bracket notation for expectations of function of a random variable:

$$\langle F(X) \rangle = \int_{-\infty}^{\infty} F(x) pr_X(x) dx. \quad (7)$$

The condition for this integral to exist is that $F(X)$ is itself a random variable. This notation will also be used for expectations of functions of finite-dimensional random vectors, in which case there is a multivariate PDF $pr_{\mathbf{X}}(\mathbf{x})$ that can be used to compute them.

3. RANDOM FIELDS

Let S be some q -dimensional support set for all functions in this section. This may be a subset of $\mathbb{R}^q$ or a q -dimensional manifold such as a sphere. The support set S may even be a discrete set of points in $\mathbb{R}^q$, such as a regular grid in this space. A random field [1–3] is a function $f(\mathbf{r}, \omega)$ of points $\mathbf{r}$ in S and ω in Ω such that $X_{\mathbf{r}} = f(\mathbf{r}, \omega)$ is a random variable for each $\mathbf{r}$. We may also describe a random field as a collection of random variables $X_{\mathbf{r}}$ for each $\mathbf{r}$ in S . For a fixed sample point ω , the function $f_{\omega}(\mathbf{r}) = f(\mathbf{r}, \omega)$ is called a realization of the random field. Sometimes further constraints are placed on the random field, such as assuming that all of the realizations are continuous functions of $\mathbf{r}$. In general, however, we cannot expect the realizations to be continuous. We will use the common notation $f(\mathbf{r}) = X_{\mathbf{r}}$ for a random field. This notation has the potential for ambiguity, but it will be clear from the context whether we are referring to a random field or an ordinary function of $\mathbf{r}$.

Given an ordered set of points $\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_N]$ in S , the vector $\mathbf{f} = [f(\mathbf{r}_1), \dots, f(\mathbf{r}_N)]$ is a random vector with a corresponding multivariate PDF $p_{\mathbf{R}}(x_1, \dots, x_N)$ on $\mathbb{R}^N$ or $\mathbb{C}^N$. The Kolmogorov extension theorem [3] provides two sets of consistency conditions that this collection of PDFs has to satisfy. The two sets of conditions arise from permutations of the points in $\mathbf{R}$ and marginalization over any of the variables x_n . Conversely, if the conditions in the theorem are satisfied, then there is a random field corresponding to the collection of PDFs. This theorem is often used to derive the existence of a random field from a suitable collection of PDFs.

For any realization of a random field, we will need to consider the inner product defined by

$$(\varphi, f_{\omega}) = \int_S \varphi^*(\mathbf{r}) f_{\omega}(\mathbf{r}) d^q r. \quad (8)$$

For a discrete support set S , the integral is replaced with a sum here and anywhere below where an integral over S appears. Any linear combination of random variables is also a random variable. Therefore, if we think of the integral in this expression as a limit of Riemann sums,

$$(\varphi, f_{\omega}) = \lim_{\Delta V \rightarrow 0} \sum_{k=1}^K \varphi^*(\mathbf{r}_k) f(\mathbf{r}_k, \omega) \Delta V_k, \quad (9)$$

where $V = \max\{\Delta V_k\}$, then it is a limit of a sequence of random variables. As long as this limit exists for all ω , the integral defines a random variable X_{φ} , which will also be written as (φ, f) . We will also write

$$X_{\varphi} = (\varphi, f) = \int_S \varphi^*(\mathbf{r}) f(\mathbf{r}) d^q r. \quad (10)$$

If we have two functions $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ that satisfy this condition and $\varphi(\mathbf{r}) = c_1\varphi_1(\mathbf{r}) + c_2\varphi_2(\mathbf{r})$, then (φ, f) is also a random variable. Thus, the set of functions $\varphi(\mathbf{r})$ such that (φ, f) is a well-defined random variable forms a vector space. The details of the problem of determining which functions are in this space depend on the random field itself, and we will not be discussing that issue in this work. This issue is discussed in detail elsewhere and involves a considerable amount of mathematics [4,5]. In all of the equations that follow, we will simply assume that $\varphi(\mathbf{r})$ is a function in this space.

4. CHARACTERISTIC FUNCTIONALS OF RANDOM FIELDS

Another way to think of a random field is as a map that assigns the function $f_\omega(\mathbf{r})$ to each ω in the sample space. Since the set of functions on $\mathcal{S}$ forms a vector space, this is analogous to a random vector in an infinite-dimensional space. However, there is no easy way to generalize the concept of a PDF for a finite-dimensional random vector $\mathbf{X}$ to a PDF for an infinite-dimensional random vector. The main reason for this is that there is no way to easily generalize an integral of the form

$$\int_{\mathbb{R}^N} f(\mathbf{x}) pr_{\mathbf{x}}(\mathbf{x}) d^N x \quad (11)$$

to the case where $\mathbb{R}^N$ is replaced with an infinite-dimensional function space. Thus, the concept of a PDF is not very useful for random fields. On the other hand, the characteristic function for a random vector, which is given by

$$\psi_{\mathbf{x}}(\boldsymbol{\xi}) = \langle \exp[-2\pi i \text{Re}(\boldsymbol{\xi}^\dagger \mathbf{x})] \rangle, \quad (12)$$

is also very useful for problems involving random vectors and is easily generalized to random fields. In this equation, the dagger superscript indicates the conjugate transpose. Of course, for random vectors, the characteristic function is just the Fourier transform of the PDF $pr_{\mathbf{x}}(\mathbf{x})$ and is therefore an equivalent description of the statistics of the random variable. The same is true for the characteristic functional of a random field even though there may be no PDF.

The characteristic functional for a random field $f(\mathbf{r})$ is defined by

$$\Psi_f(\varphi) = \langle \exp[-2\pi i \text{Re}(\varphi, f)] \rangle. \quad (13)$$

As noted previously, we always assume that the function φ in the argument of the characteristic functional satisfies the condition that the scalar $X_\varphi = (\varphi, f)$ is a well-defined random variable. Thus, the expectation in the definition of the characteristic functional can be computed from the one-dimensional PDF $pr_\varphi(x)$ for X_φ . There are certain conditions that a characteristic functional must satisfy. The normalization condition is $\Phi_f(0) = 1$. The

Hermitian symmetry condition is $\Psi_f^*(\varphi) = \Psi_f(-\varphi)$. The magnitude of the characteristic functional always satisfies $|\Phi_f(\varphi)| = \langle \exp[-2\pi i \operatorname{Re}(\varphi, f)] \rangle = 1$. The last condition is called non-negative definiteness,

$$\sum_{j,k=1}^J \Psi_f(\varphi_j - \varphi_k) a_j a_k^* = \left\langle \left| \sum_{j=1}^J a_j \exp[-2\pi i \operatorname{Re}(\varphi_j, f)] \right|^2 \right\rangle \geq 0, \quad (14)$$

which must be true for any J , any set of functions $\{\varphi_1, \dots, \varphi_J\}$ for which $\Phi_f(\varphi_j)$ is defined for all j , and any complex vector $\mathbf{a} = [a_1, \dots, a_J]$. The condition in this equation is equivalent to constraining the matrix $\mathbf{A}$ given by $A_{jk} = \Phi_f(\varphi_j - \varphi_k)$ to be non-negative definite. A generalization of Bochner's theorem [3] to this setting shows that, if we add in a continuity condition, then these constraints on a functional ensure that it is the characteristic functional of a random field.

Suppose that $f(\mathbf{r})$ is a real random field. Since $X_{\mathbf{r}} = f(\mathbf{r})$ is a random variable, we may use its PDF to compute a mean value: $\bar{f}(\mathbf{r}) = \langle f(\mathbf{r}) \rangle$. Similarly, $[f(\mathbf{r}), f(\mathbf{r}')]^T$ is a random vector and its PDF may be used to compute the correlation function $R(\mathbf{r}, \mathbf{r}') = \langle f(\mathbf{r}) f(\mathbf{r}') \rangle$. The covariance function for the random field is then given by $K(\mathbf{r}, \mathbf{r}') = R(\mathbf{r}, \mathbf{r}') - \bar{f}(\mathbf{r}) \bar{f}(\mathbf{r}')$. For a real random field, we have $\Phi_f(\varphi) = \langle \exp[-2\pi i(\varphi, f)] \rangle$ when $\varphi(\mathbf{r})$ is a real-valued function, which will be the assumption made in this paragraph. For a parameter τ , we then have

$$\frac{d}{d\tau} \Psi_f(\tau\varphi) \Big|_{\tau=0} = -2\pi i(\varphi, \bar{f}). \quad (15)$$

So the mean of the random field is determined by the characteristic functional. Taking another derivative, we get

$$\frac{d^2}{d\tau^2} \Psi_f(\tau\varphi) \Big|_{\tau=0} = -4\pi^2 \int_S \int_S \varphi(\mathbf{r}) R(\mathbf{r}, \mathbf{r}') \varphi(\mathbf{r}') d^q r' d^q r. \quad (16)$$

We will write this equation in the form

$$\frac{d^2}{d\tau^2} \Psi_f(\tau\varphi) \Big|_{\tau=0} = -4\pi^2(\varphi, \mathcal{R}\varphi), \quad (17)$$

where the integral operator $\mathcal{R}$ is given by the inner integral in the previous equation. The characteristic functional therefore determines the correlation function and hence the covariance function also. There are similar expressions for higher-order moments of the random field. There is also an integral operator $\mathcal{K}$ whose kernel is the covariance function

$K(\mathbf{r}, \mathbf{r}')$. Spectral theory for random fields is the study of the spectrum of this operator, which can also be derived from the characteristic functional in the obvious way.

For complex random fields, the correlation function is usually defined by $R(\mathbf{r}, \mathbf{r}') = \langle f^*(\mathbf{r}) f(\mathbf{r}') \rangle$, with the corresponding covariance function $K(\mathbf{r}, \mathbf{r}') = R(\mathbf{r}, \mathbf{r}') - \bar{f}(\mathbf{r}) f(\mathbf{r}')$. If we introduce the complex parameter $\tau = \tau_R + i\tau_I$, then this correlation function is given by

$$\left(\frac{\partial^2}{\partial \tau_R^2} - \frac{\partial^2}{\partial \tau_I^2} \right) \Psi_f(\tau\varphi) \Big|_{\tau=0} = -4\pi^2(\varphi, \mathcal{R}\varphi). \quad (18)$$

Since $\mathcal{R}$ is a Hermitian operator, the quantity $(\varphi, \mathcal{R}\varphi)$ is real. To complete the description of the second-order statistics for a complex random process, we also need another correlation function defined by $\tilde{R}(\mathbf{r}, \mathbf{r}') = \langle f(\mathbf{r}) f(\mathbf{r}') \rangle$, with the corresponding covariance function $\tilde{K}(\mathbf{r}, \mathbf{r}') = \tilde{R}(\mathbf{r}, \mathbf{r}') - \tilde{f}(\mathbf{r}) \tilde{f}(\mathbf{r}')$. In terms of the characteristic functional, we then have

$$\left(\frac{\partial^2}{\partial \tau_R^2} + \frac{\partial^2}{\partial \tau_I^2} \right) \Psi_f(\tau\varphi) \Big|_{\tau=0} = -4\pi^2 \text{Re}(\varphi, \tilde{\mathcal{R}}\varphi^*) \quad (19)$$

and

$$2 \frac{\partial^2}{\partial \tau_R \partial \tau_I} \Psi_f(\tau\varphi) \Big|_{\tau=0} = -4\pi^2 \text{Im}(\varphi, \tilde{\mathcal{R}}\varphi^*). \quad (20)$$

Note that $\tilde{\mathcal{R}}$ is, in general, not a Hermitian operator.

5. GENERALIZED RANDOM FIELDS

Some physically important random fields, such as Poisson point processes, are actually generalized random fields. For a generalized random field we do not insist that $X_{\mathbf{r}} = f(\mathbf{r})$ is a random variable since, for any given realization, $f_{\omega}(\mathbf{r})$ may not be well defined as an ordinary function. Instead, we restrict $\varphi(\mathbf{r})$ to be a test function and define

$$X_{\varphi} = \int_S \varphi^*(\mathbf{r}) f(\mathbf{r}, \omega) d^q r. \quad (21)$$

Here, $f_{\omega}(\mathbf{r}) = f(\mathbf{r}, \omega)$ is taken to be a generalized function [1,6] for all sample points ω . The requirement for a generalized random field is that X_{φ} be a well-defined random variable for all test functions $\varphi(\mathbf{r})$. We will also write $X_{\varphi} = (\varphi, f) = f(\varphi)$ and think of $f(\varphi)$ for a particular test function $\varphi(\mathbf{r})$ as the output of a random distribution on the space of test functions defined on S . This notation emphasizes the fact that the space of test functions on S plays the

same role for a generalized random process as the set S itself does for an ordinary random process. Thus, we can also write $X_\varphi = f(\varphi, \omega)$ as defining the random variable X_φ , and $f_\omega(\varphi) = f(\varphi, \omega)$ as defining a realization of the generalized random field. If we have a vector of test functions $\varphi = [\varphi_1, \dots, \varphi_N]$, then we have a random vector $\mathbf{f} = [f(\varphi_1), \dots, f(\varphi_N)]$ with corresponding PDF $pr_\varphi(x_1, \dots, x_N)$.

6. CHARACTERISTIC FUNCTIONALS OF GENERALIZED RANDOM FIELDS

Just as with ordinary random fields, extending the concept of a PDF to a generalized random field is difficult. We would need the PDF to be defined on the space of distributions on S , which is an infinite-dimensional space that includes, for example, the space of locally integrable functions on S . However, the characteristic functional for a generalized random field is easily defined by $\Phi_f(\varphi) = \langle \exp[-2\pi i \text{Re}(\varphi, f)] \rangle$, where the argument $\varphi(\mathbf{r})$ of the functional is taken to be a test function. The expectation in this expression is computed using the PDF $pr_\varphi(x)$ for the random variable $X_\varphi = (\varphi, f)$. All of the statistical information about the generalized random field is contained in this functional.

A generalized random field is real-valued if $f(\varphi)$ is real whenever $\varphi(\mathbf{r})$ is a real-valued test function. In this case, we may define the mean of the generalized random field as a distribution by using $\bar{f}(\varphi) = \langle f(\varphi) \rangle$. Since, for any given test function $\varphi(\mathbf{r})$, $f(\varphi)$ is a random variable, the expectation in this definition may be computed by using the PDF for that random variable. We may also write this definition as $(\varphi, \bar{f}) = \langle (\varphi, f) \rangle$. Similarly, for a pair of test functions $[\varphi, \varphi']$, the random vector $[f(\varphi)f(\varphi')]$ has a PDF that can be used to compute the correlation operator via $\langle f(\varphi)f(\varphi') \rangle = (\varphi, \mathcal{R}\varphi')$. Notice that we do not attempt to define a correlation function, which in many cases would be a generalized function itself. The covariance operator is then defined by $(\varphi, \mathcal{R}\varphi') - \bar{f}(\varphi)\bar{f}(\varphi') = (\varphi, \mathcal{K}\varphi')$.

For complex generalized random fields, we have the Hermitian correlation operator $\mathcal{R}$ defined by $\langle f^*(\varphi)f(\varphi') \rangle = (\varphi, \mathcal{R}\varphi')$, and the corresponding covariance operator $\mathcal{K}$ given by $(\varphi, \mathcal{R}\varphi') - \bar{f}^*(\varphi)\bar{f}(\varphi') = (\varphi, \mathcal{K}\varphi')$. As with ordinary random fields, to complete the description of the second-order statistics of a complex generalized random field, we need the operator $\tilde{\mathcal{R}}$ defined by $\langle f(\varphi)\bar{f}(\varphi') \rangle = (\varphi, \tilde{\mathcal{R}}\varphi')$. The corresponding covariance operator $\tilde{\mathcal{K}}$ is then given by $(\varphi, \tilde{\mathcal{R}}\varphi') - \bar{f}(\varphi)\bar{f}(\varphi') = (\varphi, \tilde{\mathcal{K}}\varphi')$. The relations between the characteristic functional and these first- and second-order statistics are the same as those for ordinary random fields given previously.

7. OPERATIONS WITH RANDOM FIELDS

Various mathematical operations with random fields or generalized random fields give rise to relations between characteristic functionals. If $f_1(\mathbf{r}, \omega_1)$ and $f_2(\mathbf{r}, \omega_2)$ are two random fields on S with corresponding probability models $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$, then there is a probability model on $\Omega = \Omega_1 \times \Omega_2$ with a probability measure that satisfies $Pr(A_1 \times A_2) = P_1(A_1)P_2(A_2)$. We can then define a random field on S by $f(\mathbf{r}, (\omega_1, \omega_2)) = f_1(\mathbf{r}, \omega_1) + f_2(\mathbf{r}, \omega_2)$. We usually write $f(\mathbf{r}) = f_1(\mathbf{r}) + f_2(\mathbf{r})$ in this case and say that the two random fields $f_1(\mathbf{r})$ and $f_2(\mathbf{r})$ are statistically independent. The characteristic functionals for these random fields are related by the equation $\Phi_f(\varphi) = \Phi_{f_1}(\varphi)\Phi_{f_2}(\varphi)$. On the other hand, we may also

take this equation as defining when two random fields are statistically independent when they share a probability model $(\Omega, \mathcal{F}, Pr)$. The obvious condition for this equation to be valid is that it must be permissible to use $\phi(\mathbf{r})$ in the argument of all of the characteristic functionals involved.

In contrast to the technical issues for summing two random fields, multiplying a random field by a non-random function $\alpha(\mathbf{r})$ is straightforward and leads to the relation $\Phi_{\alpha f} = \Phi_f \alpha^*$. The one constraint here is that the function $\alpha^*(\mathbf{r}) \phi(\mathbf{r})$ must be in the space of functions that can be used in the argument of the functional Φ_f . For example, if $f(\mathbf{r})$ is a generalized random field and $\alpha^*(\mathbf{r})$ is infinitely differentiable, then $\alpha(\mathbf{r})f(\mathbf{r})$ is also a generalized random field. If $\phi(\mathbf{r})$ is now a test function, then $\alpha^*(\mathbf{r}) \phi(\mathbf{r})$ is also a test function and the equality is valid.

Now suppose that $\mathcal{H}$ is a linear integral operator given by

$$g(\mathbf{r}) = \int_{S'} h(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') d^q r' = \mathcal{H} f(\mathbf{r}). \quad (22)$$

If we can show that the integral defining $g_\omega(\mathbf{r}) = \mathcal{H} f_\omega(\mathbf{r})$ converges for all $\mathbf{r}$ in S and all sample points ω , then this equation defines a random field $g(\mathbf{r}, \omega) = g_\omega(\mathbf{r})$ on S . In this case, we usually write $g = \mathcal{H} f$ as a relation between two random fields. The characteristic functionals of these two random fields are related by $\Phi_g(\phi) = \Phi_f(\mathcal{H}^\dagger \phi)$, where the adjoint operator $\mathcal{H}^\dagger$ is defined by $(\phi, \mathcal{H} f) = (\mathcal{H}^\dagger \phi, f)$. In the relation just given between characteristic functionals, $\phi(\mathbf{r})$ must be in the space of functions for which both sides of the equation are well defined. Note that this relation between characteristic functionals makes sense even if $\mathcal{H}$ is a differential operator. In this case we can, if necessary, restrict $\phi(\mathbf{r})$ to be a test function so that $g(\mathbf{r})$ can be treated as a generalized random field defined by the relation $(\phi, g) = (\mathcal{H}^\dagger \phi, f)$.

Now suppose that we have a set of random fields $f_n(\mathbf{r}, \omega)$ on S such that they are all defined with the same probability model $(\Omega, \mathcal{F}, Pr)$ for $n = 1, \dots, N$. Also assume that we have a set of probabilities Pr_n . We can create a new probability model $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{Pr})$ such that the samples are pairs $\tilde{\omega} = (\omega, n)$ and $\tilde{Pr}(E, n) = Pr(E) Pr_n$. The mixture random field for this situation is defined by $f(\mathbf{r}, \tilde{\omega}) = f_n(\mathbf{r}, \omega)$. We can think of the realizations of this random field as being equal to $f_n(\mathbf{r}, \omega)$ with probability Pr_n . We then have that the characteristic functional for the mixture random field is given by the convex combination

$$\Psi_f(\phi) = \sum_{n=1}^N Pr_n \Psi_{f_n}(\phi) \quad (23)$$

of the characteristic functionals of the component random fields. Again, the condition for this equation to be valid is that it must be permissible to use $\phi(\mathbf{r})$ in the argument of all of

the characteristic functionals involved. With this constraint, we can see that any convex combination of characteristic functionals is also a characteristic functional.

8. EXAMPLES

We now look at some examples of characteristic functionals that are relevant for imaging. In most cases, we do not present a derivation of the characteristic functionals since they are readily available elsewhere.

A. Real Gaussian

A random field on S is Gaussian or normal if its characteristic functional has the form $\Phi_f(\varphi) = \exp[-2\pi i(\varphi, \bar{f})] \exp[-2\pi^2(\varphi, \mathcal{K}\varphi)]$ [1 (page 410),2]. In this equation, $\bar{f}(\mathbf{r})$ is the mean of the random field and $\mathcal{K}$ is the covariance operator as described above. If the mean is zero and the covariance operator is a multiple of the identity operator, then we have a zero-mean white noise Gaussian random field. This noise field is often used in the description of Brownian motion. Note that a white noise Gaussian random field is actually a generalized random field since the kernel of the covariance operator is given by $K(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$, which is a generalized function of two variables. Due to the central limit theorem, Gaussian random fields often occur in applications as an approximation for a random field that is a sum of many independent random fields.

Suppose that we have a random field $p(\mathbf{r})$ that satisfies $p(\mathbf{r}) > 0$ for all realizations and all $\mathbf{r}$ in S . If the random field $f(\mathbf{r}) = \ln p(\mathbf{r})$ is a Gaussian random field, then we say that $p(\mathbf{r})$ is a log-normal random field [1 (page 1462),7]. There is no analytic form for the characteristic functional of a log-normal random field. However, in this case, the functional

$$\begin{aligned} \mathcal{M}_p(\varphi) &= \langle \exp[-2\pi i(\varphi, \ln p)] \rangle \\ &= \exp[-2\pi i(\varphi, \bar{f})] \exp[-2\pi^2(\varphi, \mathcal{K}\varphi)] \end{aligned} \quad (24)$$

is well defined whenever $\varphi(\mathbf{r})$ is in the domain of the characteristic functional of the random field $f(\mathbf{r})$. Note that the operator $\mathcal{K}$ in this expression is the covariance operator for $f(\mathbf{r})$ and not for $p(\mathbf{r})$. If $p'(\mathbf{r})$ and $p''(\mathbf{r})$ are statistically independent positive random fields and $p(\mathbf{r}) = p'(\mathbf{r})p''(\mathbf{r})$, then $\mathcal{M}_p(\varphi) = \mathcal{M}_{p'}(\varphi)\mathcal{M}_{p''}(\varphi)$. Therefore, due to the central limit theorem, log-normal random fields often show up as an approximation for a random field that is a product of many independent positive random fields.

B. Circular Gaussian

A complex random field on S is a circular Gaussian if its characteristic functional is given by $\Phi_f(\varphi) = \exp[-\pi^2(\varphi, \mathcal{K}\varphi)]$, where the covariance operator $\mathcal{K}$ is a Hermitian operator. It is often assumed that the electric field in partially coherent light is a circular Gaussian random field. We can also have complex random fields that are Gaussian but not circular Gaussian. The characteristic functional in this case involves both covariance functions for the complex random field.

C. Poisson Random Field

A random point field is a generalized random field that satisfies

$$(\varphi, f) = \sum_{n=1}^N \varphi(\mathbf{r}_n), \quad (25)$$

where N is a random integer with a distribution $Pr(N)$, and, for a given N , the points $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ are also randomly selected [1 (page 649)]. A typical restriction we could place on $\varphi(\mathbf{r})$ for this expression to be well defined is that it be continuous.

For a Poisson random point field, we have the Poisson distribution for N :

$$Pr(N) = \frac{\bar{N}^N}{N!} \exp(-\bar{N}). \quad (26)$$

We also require that the points $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ be chosen independently from a PDF $p(\mathbf{r})$. The mean of this generalized random field satisfies $\bar{N}p(\mathbf{r}) = \bar{f}(\mathbf{r})$ and therefore

$$\bar{N} = \int_S \bar{f}(\mathbf{r}) d^q r. \quad (27)$$

If we define the function $\tilde{\varphi}(\mathbf{r})$ by $2\pi i \tilde{\varphi}(\mathbf{r}) = 1 - \exp[2\pi i \varphi(\mathbf{r})]$, then the characteristic functional of a Poisson random field is given by $\Phi_f(\varphi) = \exp[-2\pi i \langle \tilde{\varphi}, \bar{f} \rangle]$. Poisson random fields are often used to describe the distribution of photon locations on a photon-counting detector. In this case the mean of the field is called the photon fluence, and it determines the statistics of the field completely. Poisson random fields are also used to describe photon-emitting objects that are being imaged.

D. Random Fields Related to Poisson Random Fields

Suppose that the mean of a Poisson random field $f(\mathbf{r})$ is itself a random field. Then we have immediately that the characteristic functional of $f(\mathbf{r})$ is given by $\Phi_f(\varphi) = \Phi_{\tilde{f}}(\tilde{\varphi})$. In this case, we say that $f(\mathbf{r})$ is a doubly stochastic Poisson random field [1 (page 659), 8]. If we are imaging an ensemble of objects with a photon-counting detector, then a doubly stochastic random field is the appropriate model of the statistics of the emitted or detected photon locations.

If we convolve a Poisson random field with a continuous function $h(\mathbf{r})$, then we have a filtered Poisson random field [1 (page 662)]:

$$f(\mathbf{r}) = \sum_{n=1}^N h(\mathbf{r} - \mathbf{r}_n). \quad (28)$$

This type of field is sometimes used to model textures in objects being imaged. If we define the Hermitian conjugate of the filter function via $h^\dagger(\mathbf{r}) = h^*(-\mathbf{r})$ and define $\varphi_h(\mathbf{r}) = (h^\dagger * \varphi)(\mathbf{r})$, then we have $\Phi_f(\varphi) = \exp[-2\pi i(\tilde{\varphi}_h, \tilde{f})]$ for the characteristic functional of the filtered Poisson random field.

We can have a doubly stochastic filtered Poisson random field also. In this case, the characteristic functional is given by $\Phi_f(\varphi) = \Phi[\tilde{\chi} \tilde{\varphi}_h]$. Other generalizations are also possible. For example, the filter could have the more general form $h(\mathbf{r}, \mathbf{r}_n)$, i.e., a space-variant filter. It is straightforward to work out the characteristic functional in these cases also.

E. Partitioned Random Fields

We will say that the random field is partitioned if we can write the support set S as a disjoint union

$$S = \bigcup_{j=1}^J S_j, \quad (29)$$

and we have mutually independent random fields $f_j(\mathbf{r})$ on each set S_j such that

$$(\varphi, f) = \sum_{j=1}^J a_j (\varphi_j, f_j). \quad (30)$$

Here, $\varphi_j(\mathbf{r})$ is the restriction of $\varphi(\mathbf{r})$ to the set S_j [1 (page 428)]. The coefficients a_j in this sum are the components of a random vector $\mathbf{a}$ with PDF $pr_{\mathbf{a}}(\mathbf{a})$. We can think of the random fields $f_j(\mathbf{r})$ as generating random textures in each region S_j that are statistically independent of each other, but the random vector $\mathbf{a}$ introduces correlations in the amplitudes of these textures between the regions. This can be used, for example, to model the distribution of a drug or radiotracer in the organs of each member of an ensemble of subjects being imaged. If each $f_j(\mathbf{r})$ is a filtered Poisson random field with filter $h_j(\mathbf{r})$, then we define the functions $\tilde{\varphi}_{h_j}(\mathbf{r})$ by $2\pi i \tilde{\varphi}_{h_j}(\mathbf{r}) = 1 - \exp[2\pi i(h_j * \varphi_j)(\mathbf{r})]$ and the vector $\tilde{\varphi}_h$ by $[\tilde{\varphi}_h]_j = (\tilde{\varphi}_{h_j}, \tilde{f}_j)$. The characteristic functional for the partitioned random field is then given by $\Phi_f(\varphi) = \psi_{\mathbf{a}}(\tilde{\varphi}_h)$.

F. Square of a Real Gaussian Random Field

First we consider a finite-dimensional, zero-mean Gaussian random vector $\mathbf{f}$ with covariance matrix $\mathbf{K}$. Define a random vector $\mathbf{I}$ by $I_n = f_n^2$. If we use $\mathbf{D}(\boldsymbol{\xi})$ to denote the diagonal matrix with the components ξ_n along the main diagonal, then we may write the characteristic function of $\mathbf{I}$ as

$$\psi_{\mathbf{I}}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} [(2\pi)^N \det \mathbf{K}] \exp \left(-\frac{1}{2} \mathbf{f}^\dagger \mathbf{K}^{-1} \mathbf{f} \right) \exp[-2\pi i \mathbf{f}^\dagger \mathbf{D}(\boldsymbol{\xi}) \mathbf{f}] d^N f. \quad (31)$$

This integral can be performed analytically and the result is [1 (page 1255)]

$$\psi_I(\xi) = \{\det[\mathbf{I} + 2\pi i \mathbf{K} \mathbf{D}(\xi)]\}^{-\frac{1}{2}}. \quad (32)$$

Now consider a zero-mean Gaussian random field $f(\mathbf{r})$ with covariance operator $\mathcal{K}$. We define the corresponding intensity field $I(\mathbf{r})$ by $I(\mathbf{r}) = f^2(\mathbf{r})$. This is called the intensity field in analogy with optical applications where $f(\mathbf{r})$ would be the (real) electric field from a finite-bandwidth source, and the detector responds to the intensity field. If we define the diagonal operator $\mathcal{D}_\varphi$ for $\varphi(\mathbf{r})$ acting on a function $\beta(\mathbf{r})$ by the equation $\mathcal{D}_\varphi \beta(\mathbf{r}) = \varphi(\mathbf{r})\beta(\mathbf{r})$, then the obvious generalization of the finite-dimensional formula is

$$\Psi_I(\varphi) = \{\det[\mathcal{I} + 2\pi i \mathcal{K} \mathcal{D}_\varphi]\}^{-\frac{1}{2}}. \quad (33)$$

This determinant is well defined if $\mathcal{K}$ is a trace-class operator since, in this case, $2\pi i \mathcal{K} \mathcal{D}_\varphi$ is also a trace-class operator and the determinant can be defined as an infinite product of eigenvalues. However, just because this functional is well defined does not mean it is necessarily the characteristic functional of the random field $I(\mathbf{r})$.

To see why we might believe that the given expression for $\Phi_I(\varphi)$ is correct, we consider the following special cases:

$$\varphi(\mathbf{r}) = \sum_{n=1}^N \xi_n \delta(\mathbf{r} - \mathbf{r}_n). \quad (34)$$

We will assume that the Gaussian random field $f(\mathbf{r})$ is an ordinary, as opposed to a generalized, random field. In this case, these sums are allowable for $\varphi(\mathbf{r})$ because

$$(\varphi, f) = \sum_{n=1}^N \xi_n f(\mathbf{r}_n), \quad (35)$$

which is a well-defined random variable. We then have

$$\mathcal{K} \mathcal{D}_\varphi \beta(\mathbf{r}) = \sum_{n=1}^N \xi_n K(\mathbf{r}, \mathbf{r}_n) \beta(\mathbf{r}_n), \quad (36)$$

which shows that $\mathcal{K} \mathcal{D}_\varphi$ is a finite-rank operator with rank no greater than N . Therefore, there are at most N nonzero eigenvalues for this operator and we can find them by assuming that an eigenfunction has the form

$$\beta(\mathbf{r}) = \sum_{n'=1}^N v_{n'} K(\mathbf{r}, \mathbf{r}_{n'}). \quad (37)$$

After substituting this $\beta(\mathbf{r})$ into the eigenvalue equations, we arrive at

$$\sum_{n'=1}^N \xi_n K(\mathbf{r}_n, \mathbf{r}_{n'}) v_{n'} = \lambda v_n. \quad (38)$$

Thus, the nonzero eigenvalues of $\mathcal{K}\mathcal{D}_\varphi$ for this $\varphi(\mathbf{r})$ are the eigenvalues of $\mathbf{K}\mathbf{D}(\boldsymbol{\xi})$. Therefore we can say that, if there is a random field for which $\Phi_I(\varphi)$ (as defined previously) is the characteristic functional, then this random field will have the same finite-dimensional PDFs as $I(\mathbf{r})$. By the Kolmogorov theorem, we would then have that this random field is $I(\mathbf{r})$. By Bochner's theorem, $\Phi_I(\varphi)$ as just given is the characteristic functional of a random field if it satisfies the constraints given in Section 4. The constraint of non-negative definiteness is the hard one to check but we can argue as follows. Let the sampled function $\tilde{\varphi}(\mathbf{r})$ be defined by

$$\tilde{\varphi}(\mathbf{r}) = \sum_{n=1}^N \varphi(\mathbf{r}_n) \delta(\mathbf{r} - \mathbf{r}_n), \quad (39)$$

where the $\mathbf{r}_n$ form a regular grid in $\mathcal{S}$. Then the operator $\mathcal{K}\mathcal{D}_{\tilde{\varphi}}$ is a good approximation to $\mathcal{K}\mathcal{D}_\varphi$, and this approximation gets better as we refine the grid. Since $\mathcal{K}\mathcal{D}_{\tilde{\varphi}}$ is a finite-rank operator, we know that $\Psi_I(\varphi)$ satisfies all of the constraints if we restrict $\varphi(\mathbf{r})$ to sampled functions. By continuity, we then have that the constraints are satisfied for all $\varphi(\mathbf{r})$. This argument is really a sketch of a proof that we have the correct expression for $\Psi_I(\varphi)$ and needs work to make it rigorous. Filling this sketch out is, however, beyond the scope of this paper.

G. Square Magnitude of a Circular Gaussian

Now suppose that $k(\mathbf{r})$ is a circular Gaussian complex random field and $I(\mathbf{r}) = |k(\mathbf{r})|^2$. Using the results of the previous subsection, we have

$$\Psi_I(\varphi) = \{\det[\mathcal{I} + 2\pi i \mathcal{K}\mathcal{D}_\varphi]\}^{-1}. \quad (40)$$

This is the characteristic functional for a commonly used model of speckle in coherent imaging. The characteristic functions for the finite-dimensional PDFs are given by

$$\psi_1(\xi) = \{\det[\mathbf{I} + 2\pi i \mathbf{K} \mathbf{D}(\xi)]\}^{-1}. \quad (41)$$

For the $N = 1$ case, the inverse 1D Fourier transform integral can be computed analytically to give a chi-square PDF for the random variable $I(\mathbf{r})$ at a fixed $\mathbf{r}$. The $N = 2$ case can also be computed analytically in terms of the J_0 Bessel function [1 (page 1258),9].

H. Complex-Amplitude Poisson Random Field

An example of a complex-amplitude Poisson random field has the form

$$(\varphi, f) = \sum_{n=1}^N a_n \exp(i\phi_n) \varphi(\mathbf{r}_n), \quad (42)$$

where, for a given N , the amplitudes a_n , the phases ϕ_n and the locations $\mathbf{r}_n$ are all independent of each other. The number N is a Poisson random variable with mean $\bar{N}$. Conditioned on N , we have that the locations are i.i.d. with PDF $pr(\mathbf{r})$ and the amplitudes are i.i.d. with PDF $pr_a(a)$. A common assumption that we will follow is that the phases are i.i.d. and uniformly distributed on $[-\pi, \pi]$. This is a common model for a scattered radiation field from a random collection of point scatterers, and is also known as a Cox process [10].

We define the function $m(\mathbf{r}) = \bar{N}pr(\mathbf{r})$ and

$$\tilde{\varphi}(\mathbf{r}) = \frac{i}{2\pi} \int_0^\infty \{J_0[2\pi a|\varphi(\mathbf{r})] - 1\} pr_a(a) da, \quad (43)$$

where J_0 is the zero-order Bessel function. We then have that the characteristic functional of the scattered field is given by $\Psi_f(\varphi) = \exp[-2\pi i(\tilde{\varphi}, m)]$ [1 (page 1317)]. If we have an ensemble of objects, each consisting of a collection of point scatterers, then $m(\mathbf{r})$ is itself a random field and the characteristic functional of $f(\mathbf{r})$ is given by $\Psi_f(\varphi) = \Psi_m(\tilde{\varphi})$.

If the propagation of the field through space or an imaging system is defined by an operator $\mathcal{P}$, then $\Psi_{\mathcal{P}f}(\varphi) = \Psi_f(\mathcal{P}^\dagger \varphi)$. If the propagation is implemented by convolving with a point spread function $h(\mathbf{r})$ we define, as before, $\varphi_h = h * \varphi$ and have $\Psi_f(\varphi) = \exp[-2\pi i(\tilde{\varphi}_h, m)]$ for a fixed collection of scatterers and $\Psi_f(\varphi) = \Psi_m(\tilde{\varphi}_h)$ for an ensemble of objects.

I. Innovation Random Field

Sparsity of a random field is a property often assumed in compressed sensing that allows better object reconstructions. One model for a sparse random field $f(\mathbf{r})$ assumes that $f = \mathcal{L}w$ where $w(\mathbf{r})$ is a stationary white noise field, sometimes called an innovation field. We then have $\Psi_f(\varphi) = \Psi_w(\mathcal{L}^\dagger \varphi)$. To describe the statistical properties of an innovation field, we introduce the translation operators $\mathcal{T}_s \varphi(\mathbf{r}) = \varphi(\mathbf{r} - \mathbf{s})$ and the random variables $(\mathcal{T}_s \varphi, w)$. The stationarity assumption is that for any N and functions $\{\varphi_1(\mathbf{r}), \dots, \varphi_N(\mathbf{r})\}$, the N -

dimensional PDF for the N -dimensional random vector with components $(\mathcal{T}_s \varphi_n, w)$ is independent of the translation vector $\mathbf{s}$. For a stationary random field, we necessarily have $\mathcal{S} = \mathbb{R}^q$. The white noise assumption is that whenever $\varphi(\mathbf{r})\varphi'(\mathbf{r}) \equiv 0$ for all $\mathbf{r}$ in $\mathcal{S}$, the random variables X_φ and $X_{\varphi'}$ are statistically independent. In this case it can be shown that [4,5]

$$\Psi_w(\varphi) = \exp \left\{ \int_{\mathbb{R}^q} \Lambda[\varphi(\mathbf{r})] d^q r \right\}, \quad (44)$$

where, for some θ and σ ,

$$\Lambda(\omega) = -2\pi i \theta \omega - 2\pi^2 \sigma^2 \omega^2 + \int_{\mathbb{R}^*} [\exp(-2\pi i a \omega) - 1 + 2\pi i a \omega \text{step}(1 - |a|)] dV(a). \quad (45)$$

The differential $dV(a)$ is called the Levy measure and must satisfy the constraint

$$\int_{\mathbb{R}^*} \min(1, a^2) dV(a) < \infty. \quad (46)$$

Therefore any random field that is the result of a linear operator acting on an innovation field must have a characteristic functional of the form $\Psi_f(\varphi) = \Psi_w(\mathcal{L}^\dagger \varphi)$, where the functional Ψ_w is as just given. If the PDFs for the random variables $(\mathcal{L}^\dagger \varphi_n, w) = (\varphi_n, f)$ for some basis sequence $\varphi_1(\mathbf{r}), \varphi_2(\mathbf{r}), \dots$ are highly kurtotic with long tails, then this sequence can be thought of as a sparsity basis for the random field $f(\mathbf{r})$. This requirement on these PDFs can then be related to properties of the Levy measure for the innovation process.

J. Plane-Wave Random Field

Suppose that the random field $f(\mathbf{r})$ is a superposition of a random number N of plane waves with random amplitudes, phases, and spatial frequencies:

$$f(\mathbf{r}) = \sum_{n=1}^N A_n \exp(-2\pi i \mathbf{k}_n \cdot \mathbf{r}). \quad (47)$$

This would be the case, for example, if $f(\mathbf{r})$ was the inverse Fourier transform of a point process with complex amplitudes in frequency space. For a given N , we will assume that the amplitudes and spatial frequencies are independent of each other and are i.i.d. with PDFs $p_{r_A}(A_n)$ and $p_{r_k}(\mathbf{k}_n)$, respectively. If $\Phi(\mathbf{k})$ is the Fourier transform of $\varphi(\mathbf{r})$ and N is Poisson distributed with mean $\bar{N}$, then the characteristic functional of this random field is given by $\Psi_f(\varphi) = \exp\{\bar{N} \langle \psi_A[\Phi(\mathbf{k})] \rangle_{\mathbf{k}} - \bar{N}\}$. If we assume that the spatial frequencies follow a circular Gaussian distribution, with variance σ^2 , then

$$\Psi_f(\varphi) = \exp\{\overline{N}\langle \exp[-\pi^2 \sigma^2 |\Phi(\mathbf{k})|^2] \rangle_{\mathbf{k}} - \overline{N}\}. \quad (48)$$

The single-point characteristic function for the random variable $X = f(\mathbf{r}_0)$ is derived by setting $\varphi(\mathbf{r}) = \xi \delta(\mathbf{r} - \mathbf{r}_0)$, which results in

$$\psi_X(\xi) = \exp[\overline{N} \exp(-\pi^2 \sigma^2 |\xi|^2) - \overline{N}]. \quad (49)$$

There is no analytic form for the PDF corresponding to this characteristic function.

K. Infinite-Series Random Field

Suppose that the function $F(\mathbf{n}, \omega)$ describes a random field on a discrete set of integer vectors $\mathbf{n}$. A Markov random field [1 (page 428)] is an example of this situation. If the functions $\mu_{\mathbf{n}}(\mathbf{r})$ form a set of functions on S indexed by these integer vectors, then we may define a random field $f(\mathbf{r}, \omega)$ on S via

$$f(\mathbf{r}, \omega) = \sum_{\mathbf{n}} F(\mathbf{n}, \omega) \mu_{\mathbf{n}}(\mathbf{r}), \quad (50)$$

as long as the series converges for all ω in Ω . A wavelet series expansion, for example, would have this form. We then write

$$f(\mathbf{r}) = \sum_{\mathbf{n}} F(\mathbf{n}) \mu_{\mathbf{n}}(\mathbf{r}), \quad (51)$$

where $F(\mathbf{n})$ is understood to be a random field on the index set. We also have, for appropriate functions $\varphi(\mathbf{r})$, the series

$$(\varphi, f_{\omega}) = \sum_{\mathbf{n}} F_{\omega}(\mathbf{n}) (\varphi, \mu_{\mathbf{n}}), \quad (52)$$

which defines a random variable. We will write this equation as

$$(\varphi, f) = \sum_{\mathbf{n}} F(\mathbf{n}) (\varphi, \mu_{\mathbf{n}}). \quad (53)$$

If we define a vector $\boldsymbol{\varphi}$ on the index set via $[\boldsymbol{\varphi}]_{\mathbf{n}} = (\varphi, \mu_{\mathbf{n}})$, then the characteristic functional for the random field $f(\mathbf{r})$ is given by $\Psi_f(\varphi) = \Psi_F(\boldsymbol{\varphi})$. This type of model can be used for

compressed-sensing applications where the functions $\mu_{\mathbf{h}}(\mathbf{r})$ could form a wavelet basis or some similar basis or frame.

L. Dynamically Evolving Random Field

Suppose that we define a bounded spatial integral operator $\mathcal{A}$ on spatiotemporal functions via

$$\mathcal{A}f(\mathbf{r}, t) = \int_S A(\mathbf{r}, \mathbf{r}') f(\mathbf{r}', t) d^q r. \quad (54)$$

We may then consider the dynamic equation

$$\frac{\partial}{\partial t} f(\mathbf{r}, t) = \mathcal{A}f(\mathbf{r}, t) + I(\mathbf{r}, t), \quad (55)$$

with initial condition $f(\mathbf{r}, 0) = a(\mathbf{r})$. The solution to this equation can be written as [11]

$$f(\mathbf{r}, t) = \exp(t\mathcal{A})a(\mathbf{r}) + \int_0^t \exp[(t-t')\mathcal{A}]I(\mathbf{r}, t') dt', \quad (56)$$

where the exponential of an operator is defined using the power-series expansion. We will assume that the input function $I(\mathbf{r}, t)$ and the initial distribution $a(\mathbf{r})$ are independent random fields. The random fields $f(\mathbf{r}, t)$ and $I(\mathbf{r}, t)$ are then defined on $S \times [0, \infty)$. Now fix a time t and consider the random field $f_t(\mathbf{r}) = f(\mathbf{r}, t)$ on S . We will denote the characteristic functional for this field by $\Psi_t(\phi)$. For a function $\phi(\mathbf{r})$ in the domain of this functional, we define $\tilde{\phi}_t(\mathbf{r}, t') = \text{step}(t - t') \exp[(t - t')\mathcal{A}^\dagger]\phi(\mathbf{r})$ and $\phi_t(\mathbf{r}) = \exp(t\mathcal{A})\phi(\mathbf{r}) = \tilde{\phi}_t(\mathbf{r}, 0)$. Then we have the relation $\Psi_t(\phi) = \Psi_a(\phi_t)\Psi_I(\tilde{\phi}_t)$, assuming that all of the functional values in this equation are well defined. If the operator $\mathcal{A}$ is a differential operator, we may have a problem defining the exponential $\exp(t\mathcal{A})$. One case where this exponential is well defined is when $\mathcal{A}^\dagger = -\mathcal{A}$, i.e., $\mathcal{A}$ is anti-Hermitian. For example, if $\mathcal{A} = \mathbf{s} \cdot \nabla_{\mathbf{r}}$, then $\exp(t\mathcal{A})f(\mathbf{r}) = f(\mathbf{r} + \mathbf{s})$. This operator appears in the radiative transport equation [1].

In some applications, we may also want a vector-valued version of this type of dynamically evolving random field. A model for the distribution of various molecules in an organ or a tumor, for example, may take this form. In this case, we define a spatial integral operator $\mathcal{A}$ on spatiotemporal vector-valued functions via

$$\mathcal{A}\mathbf{f}(\mathbf{r}, t) = \int_S \mathbf{A}(\mathbf{r}, \mathbf{r}') \mathbf{f}(\mathbf{r}', t) d^q r, \quad (57)$$

where $\mathbf{A}(\mathbf{r}, \mathbf{r}')$ is now a matrix-valued kernel function. Then our dynamic equation is given by

$$\frac{\partial}{\partial t} \mathbf{f}(\mathbf{r}, t) = \mathcal{A} \mathbf{f}(\mathbf{r}, t) + \mathbf{I}(\mathbf{r}, t), \quad (58)$$

with initial condition $\mathbf{f}(\mathbf{r}, 0) = \mathbf{a}(\mathbf{r})$. The solution to the dynamic equation is given by [11]

$$\mathbf{f}(\mathbf{r}, t) = \exp(t\mathcal{A})\mathbf{a}(\mathbf{r}) + \int_0^t \exp[(t-t')\mathcal{A}]\mathbf{I}(\mathbf{r}, t') dt', \quad (59)$$

where the exponential of the operator can still be defined by using the power-series expansion. We again assume that the input function $\mathbf{I}(\mathbf{r}, t)$ and the initial distribution $\mathbf{a}(\mathbf{r})$ are independent vector-valued random fields. The vector-valued random fields $\mathbf{f}(\mathbf{r}, t)$ and $\mathbf{I}(\mathbf{r}, t)$ are then defined on $S \times [0, \infty)$. Now fix time t and consider the random field $\mathbf{f}_t(\mathbf{r}) = \mathbf{f}(\mathbf{r}, t)$ on S . We define inner products by

$$(\varphi, \mathbf{f}_t) = \int_S \varphi^\dagger(\mathbf{r}) \mathbf{f}(\mathbf{r}) d^q r, \quad (60)$$

and the characteristic functional for this field by $\Psi_t(\boldsymbol{\varphi}) = \langle \exp[-2\pi i(\boldsymbol{\varphi}, \mathbf{f}_t)] \rangle$. For a vector-valued function $\boldsymbol{\varphi}(\mathbf{r})$ in the domain of this functional, we define $\tilde{\boldsymbol{\varphi}}_t(\mathbf{r}, t') = \text{step}(t - t') \exp[(t - t')\mathcal{A}^\dagger] \boldsymbol{\varphi}(\mathbf{r})$ and $\boldsymbol{\varphi}_t(\mathbf{r}) = \exp(t\mathcal{A}) \boldsymbol{\varphi}(\mathbf{r}) = \tilde{\boldsymbol{\varphi}}_t(\mathbf{r}, 0)$. We then have the relation $\Psi_t(\boldsymbol{\varphi}) = \Psi_a(\boldsymbol{\varphi}_t) \Psi_I(\tilde{\boldsymbol{\varphi}}_t)$, assuming that all of the functional values in this equation are well defined. This model, perhaps in combination with Example 8.E, may be useful for describing pharmacokinetic processes in medical imaging applications.

M. Cascaded Poisson Process

For a cascaded Poisson process, we have

$$(\varphi, f) = \sum_{n=1}^N \sum_{k=1}^{K_n} \varphi(\mathbf{r}_n + \Delta \mathbf{r}_{nk}). \quad (61)$$

The idea here is that there are primary interactions that take place at positions $\mathbf{r}_n$, but that what is measured are the secondary interactions at positions $\mathbf{r}_n + \mathbf{r}_{nk}$. This is a model for scintillation detectors, for example [1 (page 670)]. We will examine the simplest situation from the statistical point of view, but more complex cases can also be considered. The random variables and vectors are N , K_n , $\mathbf{r}_n$, and $\mathbf{r}_{nk}$. These integers and vectors are all assumed to be independent of each other. The $\mathbf{r}_{nk}$ are i.i.d. with PDF $pr_{\mathbf{r}}(\mathbf{r})$, and the $\mathbf{r}_n$ are i.i.d. with PDF $pr_{\mathbf{r}}(\mathbf{r})$. The K_n are i.i.d. Poisson random variables with mean $\bar{K}$, and N is Poisson with mean $\bar{N}$. We define the functions $b(\mathbf{r}) = \bar{K} pr_{\mathbf{r}}(\mathbf{r})$ and $a(\mathbf{r}) = \bar{N} pr_{\mathbf{r}}(\mathbf{r})$.

In order to present the characteristic functional for this process, we start with the functions $\tilde{\varphi}_{\mathbf{r}}(\mathbf{r})$ defined by $2\pi i \tilde{\varphi}_{\mathbf{r}}(\mathbf{r}) = 1 - \exp[2\pi i \varphi(\mathbf{r} + \mathbf{r})]$. We then compute the function $\tilde{\varphi}(\mathbf{r})$ by

using $2\pi i\tilde{\varphi}(\mathbf{r}) = 1 - \exp[-2\pi i(\tilde{\varphi}_r, b)]$. Finally, we have $\Psi_f(\varphi) = \exp\{-2\pi i(\tilde{\varphi}, a)\}$. If $a(\mathbf{r})$ is itself a random field, as would happen if we were imaging an ensemble of objects using a scintillation camera, then we have $\Psi_f(\varphi) = \Psi_a(\tilde{\varphi})$. Note that the cascading of the Poisson process is reflected in the recursive chain that leads to the characteristic functional.

9. IMAGING SYSTEMS

For our purposes, an imaging system collects a finite amount of image data about an unknown object function $f(\mathbf{r})$. We will assume that the object function is a realization from a random field with characteristic functional $\Psi_f(\varphi)$. We will consider two types of systems corresponding to binned data and list-mode data.

A. Binned Data

For a linear imaging system with a binned M -dimensional data vector $\mathbf{g}$, we write $\mathbf{g} = \mathcal{H}f + \mathbf{n}$, where $\mathcal{H}f$ is given by

$$[\mathcal{H}f]_m = \int_S h_m(\mathbf{r}) f(\mathbf{r}) d^q r = (h_m^*, f), \quad (62)$$

and $\mathbf{n}$ is a zero-mean random vector that describes the measurement noise. In general, we do not make any further assumptions about the measurement noise, so $\mathbf{n}$ is not necessarily statistically independent of the object $f(\mathbf{r})$. If we define the mean data vector by $\bar{\mathbf{g}} = \mathcal{H}f$, then we have $\Psi_{\bar{\mathbf{g}}}(\xi) = \Psi_f(\mathcal{H}^\dagger \xi)$. By the defining equation of the imaging system, the function $\mathcal{H}^\dagger \xi$ is in the domain of the characteristic functional $\Psi_f(\varphi)$.

For additive measurement noise, where $\mathbf{n}$ is statistically independent of the object $f(\mathbf{r})$, we have the simple relation $\Psi_{\mathbf{g}}(\xi) = \Psi_f(\mathcal{H}^\dagger \xi) \Psi_{\mathbf{n}}(\xi)$. For Poisson measurement noise, we define

$$\tilde{\xi}_m = \frac{1}{2\pi i} [\exp(2\pi i \xi_m) - 1], \quad (63)$$

and we then have $\Psi_{\mathbf{g}}(\xi) = \Psi_f(\mathcal{H}^\dagger \tilde{\xi})$ [1 (page 666)]. If the measurement noise has a Poisson component plus an additive component, then we have $\Psi_{\mathbf{g}}(\xi) = \Psi_f(\mathcal{H}^\dagger \tilde{\xi}) \Psi_{\mathbf{n}}(\xi)$, where $\Psi_{\mathbf{n}}(\xi)$ is the characteristic function for the additive component. The data in all of these cases is one high-dimensional sample from $pr_{\mathbf{g}}(\mathbf{g})$ for each object.

B. List-Mode Data

The data for a list-mode imaging system can be viewed as a Poisson random field $g(\mathbf{v})$ on a p -dimensional space V of attribute vectors. The mean of this random process is given by

$$\bar{g}(\mathbf{v}) = \int_S L(\mathbf{v}, \mathbf{r}) f(\mathbf{r}) d^q r, \quad (64)$$

which we write as $\bar{g} = \mathcal{L}f$. The attributes assigned to each photon can include position, time, and energy. Such systems are also called photon-processing systems [12]. From our previous discussion, we know that the characteristic functional for the generalized random field $g(\mathbf{v})$ with a fixed object function $f(\mathbf{r})$ is given by $\Psi_g(\varphi) = \exp[-2\pi i \text{Re}(\varphi, \mathcal{L}f)]$. When the object $f(\mathbf{r})$ is a random field, we then have $\Psi_g(\varphi) = \Psi_f(\mathcal{L}^\dagger \tilde{\varphi})$. If we define the mean number of photons as

$$\bar{N} = \int_v \bar{g}(\mathbf{v}) d^p v, \quad (65)$$

and the attribute space PDF as $pr_{\mathbf{v}}(\mathbf{v}) = \bar{g}(\mathbf{v})/\bar{N}$, then the data for a given object is a set of N low-dimensional independent samples from this PDF, where N is a Poisson random number with mean $\bar{N}$.

10. MATHEMATICAL OBSERVERS FOR POPULATIONS OF OBJECTS

In this section, we will discuss how characteristic functionals can be used to create mathematical observers for estimation and classification tasks of populations of objects. For estimation, the task will be to estimate parameters of a model that generated a sample of object fields that are being imaged. Therefore, we are trying to estimate population parameters as opposed to parameters that describe a particular object function. Similarly, for classification, we are trying to use image data to classify a sample of object fields as belonging to one of a finite set of object models. In both cases, the strategy is to use the sample of data vectors to estimate the characteristic function for the data, and then compare this estimate to the analytical form for this characteristic function derived from the characteristic functional of the object model and the model for the measurement noise. We will consider only binned data, but a similar procedure can be implemented for list-mode data also.

A. Estimation Tasks

In this subsection, we assume that the random field is specified by a parameter vector θ , although more general parameter spaces are certainly possible. We will use the notation $f(\mathbf{r}|\theta) = f_\theta(\mathbf{r})$ for the random field and emphasize again that θ is a population parameter and does not determine any particular realization of the random field. Our goal is to produce an estimate of this parameter from the data generated by a binned-data imaging system. Define the imaging system operator by $[g]_m = g_m = (h_m^*, f_\theta)$ for $m = 1, \dots, M$. We will also write $\mathbf{g} = \mathcal{H}f_\theta$ and note that we are considering noise-free measurements in this section for simplicity. Generalizing these results to additive or Poisson noise is straightforward. If the parameterized characteristic functional is given by $\Psi_f(\varphi|\theta) = \langle \exp[-2\pi i \text{Re}(\varphi, f_\theta)] \rangle$, then we have the parameterized characteristic function of the data: $\psi_g(\xi|\theta) = \Psi_f(\mathcal{H}^\dagger \xi|\theta)$.

Consider a sample $\{f_1(\mathbf{r}|\theta), \dots, f_J(\mathbf{r}|\theta)\}$ of independent realizations of the object field and the corresponding data matrix $\mathbf{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_J\}$. We can form an estimate $\hat{\psi}_g(\xi|\theta)$ of $\psi_g(\xi|\theta)$ by using a sample mean:

$$\hat{\psi}_{\mathbf{g}}(\boldsymbol{\xi}|\boldsymbol{\theta}) = \frac{1}{J} \sum_{j=1}^J \exp[-2\pi i \text{Re}(\boldsymbol{\xi}, \mathbf{g}_j)]. \quad (66)$$

Since the terms in the sum are independent complex random variables of unit magnitude, the central-limit theorem applies and we may assume that the joint distribution of the random variables $Z_l = \hat{\psi}_{\mathbf{g}}(\boldsymbol{\xi}_l|\boldsymbol{\theta})$ for $l = 1, \dots, L$ is normal. The mean of this distribution is given by $\bar{Z}_l(\boldsymbol{\theta}) = \langle Z_l \rangle = \psi_{\mathbf{g}}(\boldsymbol{\xi}_l|\boldsymbol{\theta})$. Our goal is to use Maximum Likelihood (ML) or Maximum *a posteriori* (MAP) estimation using the vector $\mathbf{Z} = (Z_1, \dots, Z_L)$ as our data. Note that we may choose the number of frequencies L to be any positive integer. To perform the estimation, we need the covariance matrix for $\mathbf{Z}$. To compute this covariance matrix, we start with the quadratic moments

$$\langle Z_l^* Z_{l'} \rangle = \frac{J-1}{J} \psi_{\mathbf{g}}^*(\boldsymbol{\xi}_l|\boldsymbol{\theta}) \psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'}|\boldsymbol{\theta}) + \frac{1}{J} \psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'} - \boldsymbol{\xi}_l|\boldsymbol{\theta}) \quad (67)$$

and

$$\langle Z_l Z_{l'} \rangle = \frac{J-1}{J} \psi_{\mathbf{g}}(\boldsymbol{\xi}_l|\boldsymbol{\theta}) \psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'}|\boldsymbol{\theta}) + \frac{1}{J} \psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'} + \boldsymbol{\xi}_l|\boldsymbol{\theta}). \quad (68)$$

Now if we define $Z_l = Z_l - \bar{Z}_l$, there are two covariance matrices for this complex vector:

$$\begin{aligned} [\mathbf{K}_{\mathbf{Z}}(\boldsymbol{\theta})]_{ll'} &= \langle \Delta Z_l^* \Delta Z_{l'} \rangle \\ &= \frac{1}{J} [\psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'} - \boldsymbol{\xi}_l|\boldsymbol{\theta}) - \psi_{\mathbf{g}}^*(\boldsymbol{\xi}_l|\boldsymbol{\theta}) \psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'}|\boldsymbol{\theta})] \end{aligned} \quad (69)$$

and

$$\begin{aligned} [\mathbf{C}_{\mathbf{Z}}(\boldsymbol{\theta})]_{ll'} &= \langle \Delta Z_l \Delta Z_{l'} \rangle \\ &= \frac{1}{J} [\psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'} + \boldsymbol{\xi}_l|\boldsymbol{\theta}) - \psi_{\mathbf{g}}(\boldsymbol{\xi}_l|\boldsymbol{\theta}) \psi_{\mathbf{g}}(\boldsymbol{\xi}_{l'}|\boldsymbol{\theta})]. \end{aligned} \quad (70)$$

Now we are ready to formulate an estimator for the population parameter vector $\boldsymbol{\theta}$. From the data matrix $\mathbf{G}$, we follow the steps just mentioned to create the $2L$ -dimensional vector $\mathbf{W}$:

$$\mathbf{W}_{\mathbf{Z}} = \begin{bmatrix} \mathbf{Z} \\ \mathbf{Z}^* \end{bmatrix}. \quad (71)$$

From the characteristic functional of the object model, the system operator, and the noise model for noisy data, we compute the mean vector

$$\overline{\mathbf{W}}_Z(\boldsymbol{\theta}) = \begin{bmatrix} \overline{Z}(\boldsymbol{\theta}) \\ \overline{Z}^*(\boldsymbol{\theta}) \end{bmatrix} \quad (72)$$

and the Hermitian matrix

$$\mathbf{Q}_Z(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{K}_Z(\boldsymbol{\theta}) & \mathbf{C}_Z(\boldsymbol{\theta}) \\ \mathbf{C}_Z^\dagger(\boldsymbol{\theta}) & \mathbf{K}_Z^*(\boldsymbol{\theta}) \end{bmatrix}. \quad (73)$$

From the central-limit theorem we then have, to a good approximation for large J ,

$$pr_Z(\mathbf{Z}|\boldsymbol{\theta}) = \frac{l}{\pi^L \sqrt{\det \mathbf{Q}_Z(\boldsymbol{\theta})}} \exp \left\{ -\frac{1}{2} [\mathbf{W}_Z - \overline{\mathbf{W}}_Z(\boldsymbol{\theta})]^\dagger \mathbf{Q}_Z^{-1}(\boldsymbol{\theta}) [\mathbf{W}_Z - \overline{\mathbf{W}}_Z(\boldsymbol{\theta})] \right\}. \quad (74)$$

We may then compute the MAP estimate of $\boldsymbol{\theta}$ via

$$\boldsymbol{\theta}_{\text{MAP}}(\mathbf{Z}) = \arg \max_{\boldsymbol{\theta}} [pr_Z(\mathbf{Z}|\boldsymbol{\theta}) pr_{\boldsymbol{\theta}}(\boldsymbol{\theta})]. \quad (75)$$

Alternatively, we can take a logarithm and write $\hat{\boldsymbol{\theta}}_{\text{MAP}}(\mathbf{Z})$ as

$$\arg \min_{\boldsymbol{\theta}} \{ [\mathbf{W}_Z - \overline{\mathbf{W}}_Z(\boldsymbol{\theta})]^\dagger \mathbf{Q}_Z^{-1}(\boldsymbol{\theta}) [\mathbf{W}_Z - \overline{\mathbf{W}}_Z(\boldsymbol{\theta})] + \ln \det \mathbf{Q}_Z(\boldsymbol{\theta}) - 2 \ln [pr_{\boldsymbol{\theta}}(\boldsymbol{\theta})] \}. \quad (76)$$

A similar formula can be used for the ML estimate if we have no prior on $\boldsymbol{\theta}$. See also [13] for another approach to estimating population parameters from characteristic functionals.

B. Classification Tasks

We may use a similar approach for classification tasks as we did for estimation tasks. We suppose that the data is generated from one of a finite set $f_1(\mathbf{r}), \dots, f_N(\mathbf{r})$ of random fields. If we define $f(\mathbf{r}|n) = f_n(\mathbf{r})$, then the classification task comes down to estimating the number n from the data. If the prior probability for class n is Pr_n , then we can follow the same processing steps as we described for the estimation task and get a MAP classifier

$$\hat{n}_{\text{MAP}}(\mathbf{Z}) = \arg \max_n [pr_Z(\mathbf{Z}|n) Pr_n],$$

where the maximum is taken over the set $\{1, \dots, N\}$. With the obvious notation, this can also be written as

$$\arg \min_n \{ [\mathbf{W}_z - \overline{\mathbf{W}}_z(n)]^\dagger \mathbf{Q}_z^{-1}(n) [\mathbf{W}_z - \overline{\mathbf{W}}_z(n)] + \ln \det \mathbf{Q}_z(n) - 2 \ln P r_n \}. \quad (77)$$

When the central-limit theorem is applicable, this classifier minimizes the probability of error among all classifiers that use the processed data vector $\mathbf{Z}$.

11. FISHER INFORMATION MATRIX

The Fisher information matrix (FIM) is often used as a figure of merit for an imaging system. The matrix inverse of the FIM is related to estimation performance via the well known Cramer–Rao lower bound on the variance of an unbiased estimator. The FIM is also directly related to the performance, as measured by the area under the receiver operating characteristic (ROC) curve, of the ideal observer in detecting a small change in the population parameter vector [14,15]. In our case, we can compute the FIM of the processed data vector $\mathbf{Z}$ using

$$\mathbf{F}(\boldsymbol{\theta}) = \langle \nabla_{\boldsymbol{\theta}} \ln [pr_z(\mathbf{Z}|\boldsymbol{\theta})] \nabla_{\boldsymbol{\theta}}^\dagger \ln [pr_z(\mathbf{Z}|\boldsymbol{\theta})] \rangle_{\mathbf{Z}|\boldsymbol{\theta}}. \quad (78)$$

Assuming that the central-limit theorem is valid for $\mathbf{Z}$, we have

$$[\mathbf{F}(\boldsymbol{\theta})]_{pp'} = \left[\frac{\partial}{\partial \theta_p} \overline{\mathbf{W}}_z(\boldsymbol{\theta}) \right]^\dagger \mathbf{Q}_z^{-1}(\boldsymbol{\theta}) \left[\frac{\partial}{\partial \theta_{p'}} \overline{\mathbf{W}}_z(\boldsymbol{\theta}) \right] + \frac{1}{2} \text{tr} \left\{ \mathbf{Q}_z^{-1}(\boldsymbol{\theta}) \left[\frac{\partial}{\partial \theta_p} \mathbf{Q}_z(\boldsymbol{\theta}) \right] \mathbf{Q}_z^{-1}(\boldsymbol{\theta}) \left[\frac{\partial}{\partial \theta_{p'}} \mathbf{Q}_z(\boldsymbol{\theta}) \right] \right\}. \quad (79)$$

Therefore, this matrix can be computed from the characteristic functional of the object random field, the system operator, and the characteristic function of the noise.

12. CONCLUSIONS

We have provided a brief introduction to the concept of a characteristic functional for a random field and described some of its properties. Our primary purpose for doing this in an imaging context is that characteristic functionals provide a way to describe the statistical properties of object models that represent the ensemble of objects being imaged. We have provided examples of random fields and their characteristic functionals that are useful in imaging, and shown how they are propagated through a noisy imaging system. Finally, we

have described methods for using characteristic functionals for estimation and classification tasks, and for computing a Fisher information matrix for an imaging system.

We have seen that there are analytic expressions for many random field models that are useful in imaging. Furthermore, many of these expressions are not very complicated. This is in contrast to the other method of describing the statistics of a random field, the finite-dimensional PDFs for the random vectors consisting of samples of the field at a finite number of points. These PDFs are often complicated or unknown for the examples we have discussed. Since the characteristic functional contains all of the statistical information about the field and can be used for estimating population parameters and classifying populations, it is often a more useful description of a random object model.

Acknowledgments

Funding. National Institutes of Health (NIH) (P41-EB002035, R01-EB000803); U.S. Department of Homeland Security (DHS) (HSHQDC-14-C-BOOIO).

References

1. Barrett, HH., Myers, KJ. Foundations of Image Science. Wiley; 2004.
2. Parzen, E. Stochastic Processes. SIAM; 1999.
3. Shiryaev, AN. Probability. Springer; 2000.
4. Amini A, Unser M. Sparsity and infinite divisibility. IEEE Trans Inf Theory. 2014; 60:2346–2358.
5. Fageot J, Amini A, Unser M. On the continuity of characteristic functionals and sparse stochastic modeling. J Fourier Anal Appl. 2014; 20:1179–1211.
6. Rudin, W. Functional Analysis. McGraw-Hill; 1973.
7. Moller J, Syversveen AR, Waagepetersen RP. Log Gaussian Cox processes. Scand J Stat. 1998; 25:451–482.
8. Bouzas PR, Valderamma MJ, Aguilaera AM. On the characteristic functional of a doubly stochastic Poisson process: application to a narrow-band process. Appl Math Model. 2006; 30:1021–1032.
9. Goodman, JW. Statistical Optics. Wiley; 2015.
10. Bouzas PR, Ruiz-Fuentes N, Ocana FM. Functional approach to the random mean of a compound Cox process. Comput Statist. 2007; 22:467–479.
11. Hirsch, MW., Smale, S. Differential Equations, Dynamical Systems, and Linear Algebra. Academic; 1974.
12. Caucci L, Barrett HH. Objective assessment of image quality V: photon counting detectors and list-mode data. J Opt Soc Am A. 2012; 29:1003–1016.
13. Kupinski MA, Clarkson E, Hoppin J, Barrett HH. Experimental determination of object statistics from noisy images. J Opt Soc Am A. 2003; 20:421–429.
14. Clarkson E, Shen F. Fisher information and surrogate figures of merit for the task-based assessment of image quality. J Opt Soc Am A. 2010; 27:2313–2326.
15. Shen F, Clarkson E. Using Fisher information to approximate ideal observer performance on detection tasks for lumpy background images. J Opt Soc Am. 2006; A 23:2406–2414.