

S2 Appendix: Example Random Process Models

In this supplement, we offer a small collection of random processes models that are useful in modeling physiological and imaging systems, providing a functional form for the characteristic functional whenever possible. We display some numerical realizations, though we do not discuss simulation methods in depth; see e.g. the extensive recent book [1]. Refer also to the introductory paper [2].

Recall that the spatial characteristic functional is defined as

$$\Psi_{\mathbf{f}}[\phi, t] = \left\langle \exp [-2\pi i(\phi, \mathbf{f})] \right\rangle_{\mathbf{f}}. \quad (1)$$

By explicit functional form, we mean an expression which does not involve the expectation angle brackets in (1). As mentioned in S1, one can modify (1) to admit spatiotemporal test functions $\phi \equiv \phi(\mathbf{r}, t)$ by simply modifying the definition of the scalar product (ϕ, \mathbf{f}) to incorporate time.

Gaussian Processes

The class of Gaussian processes is the most commonly known. Its complete statistical characterization relies on only the mean function $\bar{\mathbf{f}} = \bar{f}(\mathbf{r}, t)$ and the covariance function $\mathbf{k} = k(\mathbf{r}_1, \mathbf{r}_2, t)$ (see S1). First, note that the covariance function defines a *covariance operator* \mathcal{K}_t , which acts on a test function ϕ via

$$(\mathcal{K}_t \phi)(\mathbf{r}_1) = \int_V d^3 r_2 k(\mathbf{r}_1, \mathbf{r}_2, t) \phi(\mathbf{r}_2)$$

Given $\bar{\mathbf{f}}$ and a valid \mathbf{k} , $\mathbf{f} \sim \mathcal{GP}(\bar{\mathbf{f}}, \mathbf{k})$ has characteristic functional [2, 3]

$$\Psi_{\mathbf{f}}[\phi(\mathbf{r}), t] = \exp [-2\pi i(\phi, \bar{\mathbf{f}})] \exp [-2\pi^2(\mathcal{K}_t \phi, \phi)] \quad (2)$$

where

$$(\phi, \bar{\mathbf{f}}) = \int_V d^3 r \phi(\mathbf{r}) \bar{f}(\mathbf{r}, t), \quad (\mathcal{K}_t \phi, \phi) = \int_V d^3 r_1 \int_V d^3 r_2 k(\mathbf{r}_1, \mathbf{r}_2, t) \phi(\mathbf{r}_2) \phi(\mathbf{r}_1)$$

One of the key features of a Gaussian process is that any finite dimensional sample is multivariate Gaussian. Indeed, let $\mathbf{f} \sim \mathcal{GP}(\bar{\mathbf{f}}, \mathbf{k})$, $n \geq 1$, $\mathbf{r}_1, \dots, \mathbf{r}_n \in V$, and $t \geq 0$. Then the random vector $\mathbf{X} = [f(\mathbf{r}_1, t), \dots, f(\mathbf{r}_n, t)]$ has characteristic function

$$\begin{aligned} \psi_{\mathbf{X}}(\boldsymbol{\xi}) &= \Psi_{\mathbf{f}} \left[\sum_{j=1}^n \xi_j \delta(\mathbf{r} - \mathbf{r}_j), t \right] \\ &= \exp \left[-2\pi i \sum_{j=1}^n \xi_j (\delta(\mathbf{r} - \mathbf{r}_j), \bar{\mathbf{f}}) \right] \\ &\quad \times \exp \left[-2\pi^2 \sum_{j,k=1}^n \xi_j \xi_k (\mathcal{K}_t \delta(\mathbf{r} - \mathbf{r}_j), \delta(\mathbf{r} - \mathbf{r}_k)) \right] \\ &= \exp [-2\pi i \boldsymbol{\xi}^\top \bar{\mathbf{X}}] \exp [-2\pi^2 \boldsymbol{\xi}^\top \mathbf{C} \boldsymbol{\xi}] \end{aligned} \quad (3)$$

where \top indicates the transpose and

$$\overline{\mathbf{X}}(i) = \bar{f}(\mathbf{r}_i, t), \quad C(i, j) = k(\mathbf{r}_i, \mathbf{r}_j, t). \quad (4)$$

Applying the inverse Fourier transform to (3), we obtain the PDF of $\mathbf{X}_{\mathbf{r}_1:\mathbf{r}_n,t}$:

$$p_{\mathbf{X}}(\mathbf{x}) = \frac{(2\pi)^{-n/2}}{\sqrt{\det(\mathbf{C})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \overline{\mathbf{X}})^\top \mathbf{C}^{-1}(\mathbf{x} - \overline{\mathbf{X}})\right) \quad (5)$$

The finite dimensional distributions (5) provide a simple method to draw samples from $\mathcal{GP}(\bar{\mathbf{f}}, \mathbf{k})$: simply choose t and sample locations $\mathbf{r}_1, \dots, \mathbf{r}_n \in V$, form $\overline{\mathbf{X}}$ and \mathbf{C} using (4), then use an existing multivariate Gaussian sampler (based on e.g. Cholesky factorization) to draw from $\mathcal{N}(\overline{\mathbf{X}}, \mathbf{C})$. However, this can be very inefficient if n is large, for instance sampling a fine grid of points in 3D: Cholesky factorization is $O(n^3)$ and matrix multiplication is $O(n^2)$, so to draw K samples on a grid with m points per dimension, this method would be $O(m^9)$ for the first sample, then $O(Km^6)$ thereafter. Faster specialized methods are available if for instance \mathbf{k} is shift-invariant; see e.g. [1].

In addition to the sample-based finite dimensional distributions (5), it is also useful to have the density of scalar random variables of the sort $X_{\phi,t} = (\phi, \mathbf{f}) = \int_V d^3r \phi(\mathbf{r})f(\mathbf{r}, t)$, i.e. a scalar product with a fixed test function. Such a random variable has (univariate) PDF $p_{\phi,t}$ given by

$$p_{\phi,t}(x) = \frac{1}{\sqrt{2\pi\sigma_{\phi,t}^2}} \exp\left(-\frac{(x - m_{\phi,t})^2}{2\sigma_{\phi,t}^2}\right) \quad (6)$$

where $\sigma_{\phi,t} = (\mathcal{K}_t \phi, \phi)$ and $m_{\phi,t} = (\phi, \bar{\mathbf{f}})$, with \mathcal{K} and $\bar{\mathbf{f}}$ respectively the covariance operator and mean function of the process \mathbf{f} . Note that as the mean and covariance functions may depend on time, the PDF (6) also evolves in time.

One application of (2) is to provide a rigorous description of the *Gaussian white noise* process \mathbf{w} . While somewhat challenging to describe using standard probabilistic definitions, the characteristic functional provides a direct method to define a process which has mean zero and ‘independent standard normal values at every point’. Intuitively, one would like the covariance function to be zero for $\mathbf{r}_1 \neq \mathbf{r}_2$, but nonzero for $\mathbf{r}_1 = \mathbf{r}_2$; the logical choice is a Dirac delta covariance, that is, $k(\mathbf{r}_1, \mathbf{r}_2, t) = \delta(\mathbf{r}_1 - \mathbf{r}_2)$. This leads to a covariance operator \mathcal{K} which is the identity operator, and hence the characteristic functional [4]

$$\Psi_{\mathbf{w}}[\phi] = \exp(-2\pi^2(\phi, \phi)) = \exp\left(-2\pi^2 \int_V d^3r |\phi(\mathbf{r})|^2\right)$$

A fairly wide variety of processes can be obtained under the Gaussian model by choosing different mean and covariance functions. The only requirement on $\bar{\mathbf{f}}$ and \mathbf{k} is that \mathbf{k} be *non-negative-type*, which means that any matrix of the form $C(i, j) = k(\mathbf{r}_i, \mathbf{r}_j, t)$ must be symmetric and non-negative definite i.e. a valid covariance matrix. Several common examples of covariance functions are provided in Table 1. In Figure 1 we display realizations for several choices of \mathbf{k} .

#	Name	Covariance function \mathbf{k}
(1)	Gaussian	$k(\mathbf{r}_1, \mathbf{r}_2; \sigma, \mathbf{A}) = \sigma^2 \exp\left(-\frac{1}{2}(\mathbf{r}_1 - \mathbf{r}_2)^t \mathbf{A}(\mathbf{r}_1 - \mathbf{r}_2)\right)$
(2)	Exponential	$k(\mathbf{r}_1, \mathbf{r}_2; \sigma, \ell) = \sigma^2 \exp(-\ \mathbf{r}_1 - \mathbf{r}_2\ /\ell)$
(3)	Matérn-Whittle	$k(\mathbf{r}_1, \mathbf{r}_2; \sigma, \nu, \ell) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} \left(\frac{\ \mathbf{r}_1 - \mathbf{r}_2\ }{\ell}\right)^\nu K_\nu\left(\frac{\ \mathbf{r}_1 - \mathbf{r}_2\ }{\ell}\right)$
(4)	Bessel	$k(\mathbf{r}_1, \mathbf{r}_2; \sigma, \nu) = \sigma^2 \Gamma(d/2) \frac{J_\nu(\ \mathbf{r}_1 - \mathbf{r}_2\)}{(\ \mathbf{r}_1 - \mathbf{r}_2\ /2)^\nu}$
(5)	White noise	$k(\mathbf{r}_1, \mathbf{r}_2) = \delta(\mathbf{r}_1 - \mathbf{r}_2)$

Table 1: A short list of common covariance functions used in spatial random process modeling, found in refs. [1, 5, 4, 6, 7]. The special functions $\Gamma(\cdot)$, K_ν , J_ν are respectively the Gamma function, modified Bessel function of the second kind, and Bessel function of the first kind. The matrix A in (1) must be symmetric nonnegative definite. Covariances (2)-(4) are their isotropic versions; anisotropic variants can also be defined. The covariance (5) must be interpreted as the generalized kernel of an integral operator.

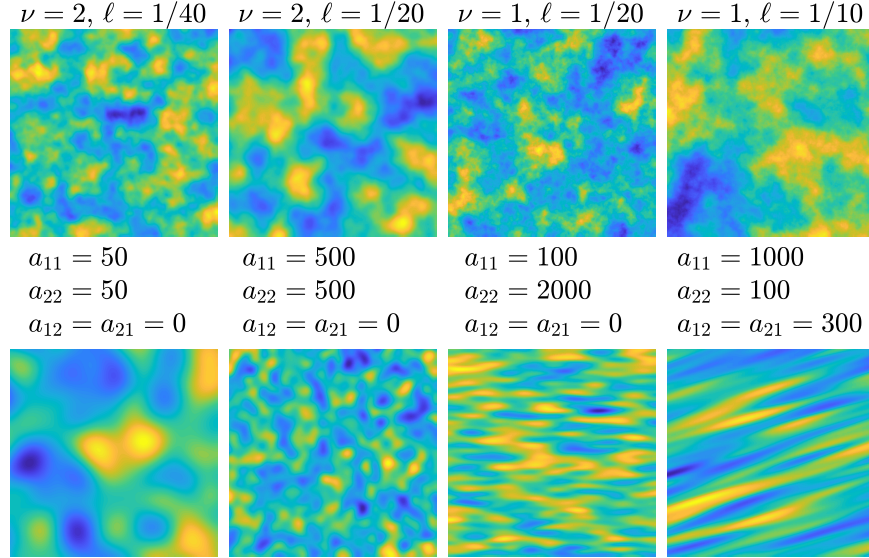


Figure 1: Simulated Gaussian random processes $f(\mathbf{r}, t)$ on $V = [0, 1]^2$. Top row uses the isotropic Matérn-Whittle covariance function (function (3) in Table 1) with parameters indicated above, while the bottom row uses the Gaussian covariance (function (1) in Table 1) with $A = (a_{ij})$ indicated above. All samples use variance $\sigma^2 = 1$ and were computed using Algorithm 7.6 in [1].

Lognormal Processes

One undesirable feature of the Gaussian model for physiology applications is that realizations are not guaranteed to be non-negative, whereas many physiological processes necessarily are. Given $\mathbf{z} \sim \mathcal{GP}(\bar{\mathbf{f}}, \mathbf{k})$, one possibility to easily generate other random processes from \mathbf{z} is to consider point transformations of \mathbf{z} , i.e. we generate a process \mathbf{f} by simply applying a deterministic function Υ to the realizations of \mathbf{z} :

$$f(\mathbf{r}, t, \omega) = \Upsilon(z(\mathbf{r}, t, \omega))$$

For the particular choice of $\Upsilon(z) = \exp(z)$, we form the class of *lognormal* processes, denoted $\mathcal{LN}(\bar{\mathbf{f}}, \mathbf{k})$, via

$$f(\mathbf{r}, t, \omega) = \exp(z(\mathbf{r}, t, \omega)), \quad \mathbf{z} \sim \mathcal{GP}(\bar{\mathbf{f}}, \mathbf{k})$$

It is well-known that the characteristic function of a lognormal random variable does not admit a closed-form representation [8], so similarly we cannot expect that a lognormal random process will admit a closed-form representation for the characteristic functional. In this work (namely S3), we were able to successfully manipulate $\mathbf{z} = \log \mathbf{f}$ using (2), so this was a non-issue.

Poisson Point Processes

A very useful class of non-Gaussian random process models are the various *point processes*. Realizations of a point process take the form

$$f(s) = \sum_{j=1}^N \delta(s - S_j) \quad (7)$$

where S_1, \dots, S_n is some collection of random states or sample points and N is a possibly random number of samples. We will consider only spatiotemporal *Poisson* point processes where $S_j = (\mathbf{R}_j, T_j)$ denotes a random spatiotemporal vector and the number of points N is a Poisson random variable. More general point processes are also possible [9, 3, 10].

To define a spatiotemporal Poisson Point Process (PPP), we require only an *intensity function* $\lambda(\mathbf{r}, t)$, which is a non-negative, integrable function on $V \times [0, \infty)$. We then define the mean number of points as

$$\bar{N} = \int_0^\infty \int_V \lambda(\mathbf{r}, t) d^3r dt < \infty$$

To form a point process of the sort (7), we first sample $N \sim \text{Poi}(\bar{N})$, where $\text{Poi}(\bar{N})$ is the Poisson distribution with rate \bar{N} . Then, we consider random variables $(\mathbf{R}_1, T_1), \dots, (\mathbf{R}_N, T_N)$, independent and identically distributed with probability density

$$p(\mathbf{r}, t) = \frac{\lambda(\mathbf{r}, t)}{\bar{N}}$$

Given a realization $N = n$ and samples $(\mathbf{R}_1, T_1) = (\mathbf{r}_1, t_1), \dots, (\mathbf{R}_n, T_n) = (\mathbf{r}_n, t_n)$, we form the point process as in (7):

$$f(\mathbf{r}, t) = \sum_{j=1}^n \delta(\mathbf{r} - \mathbf{r}_j) \delta(t - t_j) \quad (8)$$

Note that as (8) is a generalized function, one must take care in defining its statistical properties. For instance, the finite dimensional point sample vectors $\mathbf{X} = [f(\mathbf{r}_1, t), \dots, f(\mathbf{r}_n, t)]$ become meaningless. However, the finite dimensional vectors $\mathbf{X} = [(\phi_1, \mathbf{f}), \dots, (\phi_n, \mathbf{f})]$ and thus the characteristic functional is still well-defined, because of their definition in terms of test functions. To derive $\Psi_{\mathbf{f}}[\phi]$ (which is now a spatiotemporal characteristic functional; see S1), we let $\phi \equiv \phi(\mathbf{r}, t)$ be a (Schwartz class, say) test function. Then,

$$(\phi, \mathbf{f}) = \sum_{j=1}^N \phi(\mathbf{R}_j, T_j)$$

where $N, \{\mathbf{R}_j\}$ and $\{T_j\}$ are the random quantities discussed above. Thus

$$\exp[-2\pi i(\phi, \mathbf{f})] = \prod_{j=1}^N \exp[-2\pi i\phi(\mathbf{R}_j, T_j)]$$

We then perform the expected value in (1) using the law of total probability and the definition of the random vector (\mathbf{R}_j, T_j) :

$$\begin{aligned} \Psi_{\mathbf{f}}[\phi] &= \left\langle \left\langle \exp[-2\pi i(\phi, \mathbf{f})] \right\rangle_{\{(\mathbf{R}_j, T_j)\}_{|N}} \right\rangle_N \\ &= \left\langle \left(\int_0^\infty dt \int_V d^3r \exp(-2\pi i\phi(\mathbf{r}, t)) \frac{\lambda(\mathbf{r}, t)}{\bar{N}} \right)^N \right\rangle_N \\ &= \sum_{n=0}^\infty \frac{q_\phi^n \exp(-\bar{N}) \bar{N}^n}{\bar{N}^n n!} \\ &= \exp(-\bar{N}) \exp(q_\phi) \end{aligned}$$

where we have defined $q_\phi \equiv \int_0^\infty dt \int_V d^3r \exp(-2\pi i\phi(\mathbf{r}, t)) \lambda(\mathbf{r}, t)$. Recalling the definition of \bar{N} , we can simplify once more to obtain

$$\Psi_{\mathbf{f}}[\phi] = \exp \left[\int_0^\infty dt \int_V d^3r (\exp(-2\pi i\phi(\mathbf{r}, t)) - 1) \lambda(\mathbf{r}, t) \right]$$

We briefly discuss some extensions of the PPP model (8). First, we can consider a *filtered* version of (8), whereby we compute $g(\mathbf{r}, t) = (\mathcal{A}f)(\mathbf{r}, t)$ where \mathcal{A} is a linear operator (defined for generalized functions). For instance, suppose

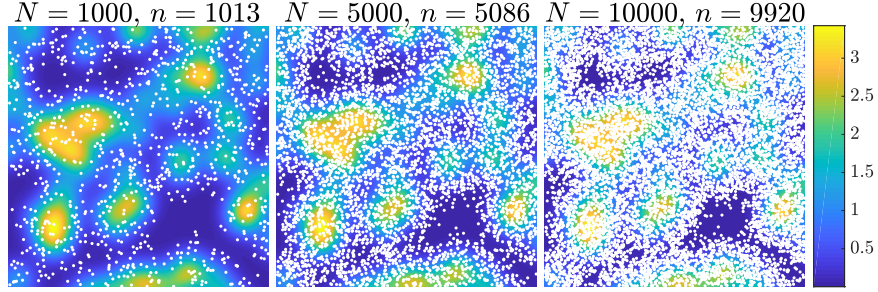


Figure 2: Simulation of three Poisson point processes with fixed PDF $p(\mathbf{r}, t) = \lambda(\mathbf{r}, t)/\bar{N}$ and varying \bar{N} . In each panel, a single realization of $f(\mathbf{r}, t)$ is displayed (white dots) over the color intensity plot of $p(\mathbf{r}, t)$. The mean number of points \bar{N} and the realized number of points n are displayed above each panel. The PDF $p(\mathbf{r}, t)$ is a realization of a lumpy background process.

$h(\mathbf{r}, t)$ is a real, continuous, compactly supported function. Then, define $\mathbf{g} = \mathcal{A}\mathbf{f}$ to be the convolution of \mathbf{f} with $h(\mathbf{r}, t)$:

$$\begin{aligned} g(\mathbf{r}, t) &= (h * f)(\mathbf{r}, t) \\ &= \int_0^\infty dt' \int_{\mathbb{R}^3} d^3r' h(\mathbf{r} - \mathbf{r}', t - t') \sum_{j=1}^N \delta(\mathbf{r}' - \mathbf{R}_j) \delta(t' - T_j) \\ &= \sum_{j=1}^N h(\mathbf{r} - \mathbf{R}_j, t - T_j) \end{aligned} \quad (9)$$

The characteristic functional of (9) then follows by the general relation for linearly transformed random processes (discussed in S1), namely that $\Psi_{\mathbf{g}}[\phi] = \Psi_{\mathbf{f}}[\mathcal{A}^\dagger \phi]$. With $\mathcal{A} = h*$, we have $\mathcal{A}^\dagger = \tilde{h}*$ where $\tilde{h}(\mathbf{r}, t) = h(-\mathbf{r}, -t)$, and so

$$\Psi_{\mathbf{g}}[\phi] = \exp \left[\int_0^\infty dt \int_V d^3r \left(\exp(-2\pi i(\tilde{h} * \phi)(\mathbf{r}, t)) - 1 \right) \lambda(\mathbf{r}, t) \right]$$

A particular example of (9) can be used to simulate extravascular drug diffusion component c^{diff} (see S4 and the main text). Let $h(\mathbf{r}, t)$ be the fundamental solution (Green's function) of a constant-coefficient diffusion equation, that is,

$$h(\mathbf{r}, t) = \frac{1}{(4\pi D_0 t)^{d/2}} \exp \left(-\frac{\|\mathbf{r}\|^2}{4D_0 t} \right)$$

then a solution to the diffusion equation with Poisson point process source $s(\mathbf{r}, t)$ results in a generalized Lumpy background process of the type (9):

$$c^{diff}(\mathbf{r}, t) = \sum_{j=1}^n h(\mathbf{r} - \mathbf{r}_j, t - t_j) = \sum_{j=1}^N \frac{1}{(4\pi D_0(t - t_j))^{d/2}} \exp \left(-\frac{\|\mathbf{r} - \mathbf{r}_j\|^2}{4D_0(t - t_j)} \right)$$

Taking (9) as a starting point, one can construct a wide array of possible random processes by choosing a different kernel functions \mathbf{h} , or even allowing multiple kernel functions $h_j(\mathbf{r}, t)$ in the sum (9). In general, we call a random process of the type

$$f(\mathbf{r}, t) = \sum_{j=1}^N h_j(\mathbf{r}, t; \boldsymbol{\theta}_j) \quad (10)$$

a *generalized lumpy background process*. The random variables that generate (10) are N and $(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N)$. The functions $h_j(\mathbf{r}, t; \boldsymbol{\theta})$ need not be the same type of function nor do they need to form a basis for any function space. Functions similar to (10) have been called *kernel density estimators* and *mixture models* [11], *shot noise* [10, 12], *lumpy backgrounds* [13, 14] and *texton noise* [15]. The model (10) can be evaluated in parallel at very high speed, and \mathbf{h}_j can be chosen to match observed texture statistics [15, 16] or display desired regularity, nonnegativeness, boundedness, or any other desired sample function property. In particular, (10) can be made highly non-stationary and non-Gaussian, making it a particularly appealing class of models for physiological processes.

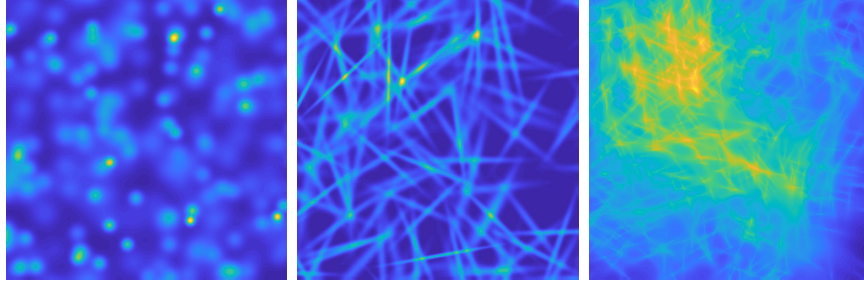


Figure 3: Three realizations of generalized lumpy background processes of the form (10) on the unit square $V = [0, 1]^2$. On the left, each lump takes the form $h_j(\mathbf{r}) = A_j \exp(-\|\mathbf{r}\|^2/2\sigma_j^2)$ with σ_j^2 drawn randomly from a lognormal distribution. In the middle, each lump takes the form $h_j(\mathbf{r}) = A \exp(-\mathbf{r}^t B_j \mathbf{r}/2)$ where B_j is a 2×2 positive definite matrix formed by rotating a matrix B_0 randomly. On the right is a *clustered* lumpy background [14] where each lump is of the form $h_j(\mathbf{r}) = \sum_{k=1}^{n_j} h(\mathbf{r} - \mathbf{r}_{jk})$.

The second direction we can generalize either (7), (9) or (10) is to allow the PPP intensity function $\boldsymbol{\lambda} \equiv \lambda(\mathbf{r}, t)$ to be a realization of a *secondary* random process, that is, we suppose that $\boldsymbol{\lambda}$ is a random process whose realizations are (almost surely) nonnegative and integrable. Then, if for each realization of $\boldsymbol{\lambda}$ we form a Poisson point process, we call the resulting random process a *doubly stochastic* or *Cox* Poisson point process [17, 9, 10], after which the filtered forms (9) or (10) can be formed.

References

- [1] Lord GJ, Powell CE, Shardlow T. Introduction to Computational Stochastic PDEs. Cambridge; 2014.
- [2] Clarkson E, Barrett HH. Characteristic functionals in imaging and image-quality assessment: tutorial. *JOSA A*. 2016;33(8):1464–1475.
- [3] Barrett H, Myers K. Foundations of image science. John Wiley & Sons; 2004.
- [4] Gel’fand IM, Vilenkin NY. Generalized functions, Vol. 4: applications of harmonic analysis. Academic Press; 1964.
- [5] Christakos G. Random field models in earth sciences. Courier Corporation; 2012.
- [6] Stein ML. Interpolation of spatial data: some theory for kriging. Springer Science & Business Media; 2012.
- [7] Matérn B. Spatial variation. vol. 36. Springer Science & Business Media; 2013.
- [8] Lukacs E. Characteristic functions. 1970;.
- [9] Reiss RD. A course on point processes. Springer Science & Business Media; 2012.
- [10] Snyder DL, Miller MI. Random point processes in time and space. Springer Science & Business Media; 2012.
- [11] Silverman BW. Density estimation for statistics and data analysis. vol. 26. CRC press; 1986.
- [12] Davenport WB, Root WL, et al. An introduction to the theory of random signals and noise. vol. 159. McGraw-Hill New York; 1958.
- [13] Rolland JPY. Factors influencing lesion detection in medical imaging. The University of Arizona.; 1990.
- [14] Bochud F, Abbey C, Eckstein M. Statistical texture synthesis of mammographic images with clustered lumpy backgrounds. *Optics Express*. 1999;4(1):33–42.
- [15] Galerne B, Leclaire A, Moisan L. Texton noise. In: Computer Graphics Forum. Wiley Online Library; 2017.
- [16] Kupinski MA, Clarkson E, Hoppin JW, Chen L, Barrett HH. Experimental determination of object statistics from noisy images. *JOSA A*. 2003;20(3):421–429.
- [17] Cox DR. Some statistical methods connected with series of events. *Journal of the Royal Statistical Society Series B (Methodological)*. 1955; p. 129–164.