On sampling spatially-correlated random fields for complex geometries

Simone Pezzuto[1][0000−0002−7432−0424], Alessio Quaglino[1,2], and Mark Potse[3,4,5][0000−0003−4166−2687]

[1] Center for Computational Medicine in Cardiology,
Institute for Computational Science,
Università della Svizzera italiana, Lugano, Switzerland
`simone.pezzuto@usi.ch`
[2] NNAISENSE SA, Lugano, Switzerland,
`alessio@nnaisense.com`
[3] Univ. Bordeaux, IMB, UMR 5251, Talence, France
[4] CARMEN Research Team, Inria Bordeaux Sud-Ouest, Talence, France
[5] IHU Liryc, fondation Bordeaux Université, Pessac, France
`mark@potse.nl`

Abstract. Extracting spatial heterogeneities from patient-specific data is challenging. In most cases, it is unfeasible to achieve an arbitrary level of detail and accuracy. This lack of perfect knowledge can be treated as an uncertainty associated with the estimated parameters and thus be modeled as a spatially-correlated random field superimposed to them. In order to quantify the effect of this uncertainty on the simulation outputs, it is necessary to generate several realizations of these random fields. This task is far from trivial, particularly in the case of complex geometries. Here, we present two different approaches to achieve this. In the first method, we use a stochastic partial differential equation, yielding a method which is general and fast, but whose underlying correlation function is not readily available. In the second method, we propose a geodesic-based modification of correlation kernels used in the truncated Karhunen-Loève expansion with pivoted Cholesky factorization, which renders the method efficient even for complex geometries, provided that the correlation length is not too small. Both methods are tested on a few examples and cardiac applications.

Keywords: Random fields · Geodesic distance · Stochastic PDE · Fibrosis · Heterogeneity

1 Introduction

The ubiquitous presence of tissue heterogeneities affects the electrophysiological and mechanical function of the heart. An example of the severity of these effects is seen in the atria, whose tissue is a characterized by a complex structure of several fibre bundles [15]. At the micro-structural level, atria are often affected by fibrosis [1].

Despite their relevance in cardiac modeling, heterogeneities in the cardiac tissue are often neglected in patient-specific studies because they cannot be extracted from clinical imaging. An alternative is to include such heterogeneities as *random* variables, reflecting the inability to accurately describe the spatial distribution of tissue properties. It would be desirable to explicitly quantify this lack of knowledge in the output of cardiac simulations, providing confidence intervals of the quantities of interest [14].

Spatial heterogeneities in the parameters of the model typically exhibit a certain degree of spatial correlation. Generating correlated, stationary, and isotropic random fields for simple geometries, e.g. a box domain, is relatively straightforward. Efficient methods based on circulant embedding can be exploited to quickly sample random fields for a given correlation function [6]. Problems arise, however, in non-convex domains such as the heart muscle. If the shape of the domain were just ignored and the field were sampled in the bounding box of the original domain, anatomically close but functionally distant regions may be more correlated than expected.

A solution is offered by stochastic partial differential equations (SPDEs) [13]. Each sample is the solution of a PDE. The correlation function is encoded implicitly in the equation, as the solution of the second-moment equation.

Alternatively, random fields are commonly sampled using the truncated Karhunen-Loève (KL) expansion, a linear combination with random coefficients of eigenfunctions of the Hilbert-Schmidt operator associated with the correlation function [11]. Obviously, we rely on the assumption that the correlation kernel can be evaluated easily, e.g. when the domain is geometrically simple and the kernel is expressed in terms of the Euclidean distance. However, for more complicated domains this is not possible.

Here we analyze these two approaches to generate random, spatially-correlated heterogeneities on complex geometries. In Section 2, we recall the SPDE approach, including non-stationary and anisotropic random fields. In Section 3 we propose generalized correlation kernels, obtained by replacing the Euclidean distance with the *geodesic* one. We conclude with a comparison between the two methods (in Section 4) and an application to atrial fibrosis (Section 5).

2   Random fields via SPDE

A simple, yet powerful method to sample random fields on arbitrarily complex geometries is based on SPDEs. We consider the following one:

$$\begin{cases} (\kappa^2 - \nabla \cdot \mathbf{D}\nabla)^{\frac{\alpha}{2}} u = \mathcal{W}, & x \in \Omega \subset \mathbb{R}^d, \\ \mathbf{D}\nabla(\kappa^2 - \nabla \cdot \mathbf{D}\nabla)^j \cdot \mathbf{n} = 0, & x \in \partial\Omega, \ j = 0, \ldots, \left\lfloor \frac{\alpha-1}{2} \right\rfloor. \end{cases} \tag{1}$$

where $\kappa > 0$, $\alpha = \nu + d/2$ with $\nu > 0$, and $\mathbf{D}(x)$ is a uniformly elliptic tensor field. On the right hand side we have a Gaussian white noise $\mathcal{W}$ [13, Def. 6, sec. B.2]. The linear fractional SPDE (1) was analyzed in detail by Lindgren *et al.* [13]. In summary, the correlation function of u is related to the Matérn kernel. The

parameter ν, therefore, is a measure of the smoothness of the random field, while κ is inversely proportional to the correlation length.

The solution for general α is discussed by Bolin *et al.* [2]. A simpler setting occurs when $\alpha/2 = K \in \mathbb{N}$, which corresponds to the choice $\nu = 2K - d/2$. Denoting by u_K the solution of (1) with this assumption, the following iterative scheme applies:

$$u_0 = \mathcal{W}, \quad \begin{cases} (\kappa^2 - \nabla \cdot \mathbf{D}\nabla)u_k = u_{k-1}, & x \in \Omega, \, k = 1, \ldots, K, \\ \mathbf{D}\nabla u_k \cdot \mathbf{n} = 0, & x \in \partial\Omega, \, k = 1, \ldots, K. \end{cases} \tag{2}$$

The Galerkin discretization of (2) follows by selecting a finite-dimensional V_h subspace of $\mathrm{H}^1(\Omega)$, for instance a finite-element basis. Denoting by $\{\phi_i\}_{i=1}^N$ a basis for V_h, of dimension N, we define the matrices:

$$[\mathbf{K}]_{ij} = \kappa^2 \langle \phi_j, \phi_i \rangle + \langle \nabla\phi_j, \nabla\phi_i \rangle, \quad [\mathbf{M}]_{ij} = \langle \phi_j, \phi_i \rangle.$$

Then the algorithm (2) reduces to:

$$\begin{cases} \mathbf{K}\mathbf{u}_1 = \mathbf{w}, \\ \mathbf{K}\mathbf{u}_k = \mathbf{M}\mathbf{u}_{k-1}, & k = 2, \ldots, K, \end{cases}$$

where $\mathbf{w}$ is the discrete Gaussian white noise, that is an N-dimensional Gaussian sample with zero mean and covariance $\mathbf{M}$. Efficient methods to sample the discretized white noise $\mathbf{w}$ have been proposed. Mass lumping diagonalizes the mass matrix using reduced quadrature, rendering the sampling procedure trivial. Mass lumping is easy for linear finite elements, but not for higher-order polynomials. Croci *et al.* computed the Cholesky decomposition of the mass matrix element-wise, making the evaluation of white noise very efficient and parallelizable [5]. Alternatively, one can select a Galerkin space V_h with an orthonormal basis, as for instance in spectral methods [3].

Figure 1 provides an example of a random field sampled with the SPDE approach for increasingly values of ν. The simulation also shows the effect of non-constant anisotropic $\mathbf{D}(x)$ for sampling non-stationary random fields.

3   Random fields via KL with geodesic distance

An alternative method to sample random fields is based on the KL expansion, for which we have to choose an appropriate correlation function. Given a square-integrable correlation function $r(x, y)$, i.e. r is in $\mathrm{L}^2(\Omega \times \Omega)$, the random field $u(x, \omega)$ with mean $\bar{u}(x)$ and covariance $r(x, y)$ reads as follows:

$$u(x, \omega) = \bar{u}(x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \psi_i(x) Z_i(\omega),$$

where Z_i are jointly Gaussian random variables with zero mean and unit variance and $\{\lambda_i\}$ and $\{\psi_i\}$ are respectively eigenvalues and eigenvectors of the Hilbert-Schmidt operator associated with $r(x, y)$. In the discrete setting, as above, the
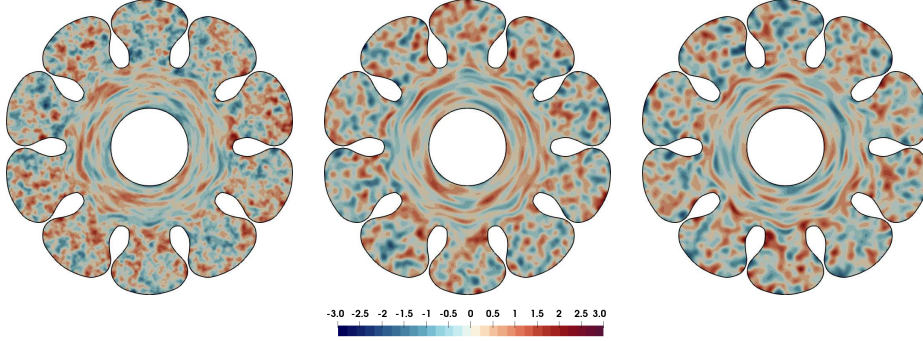
Fig. 1. Example of a random field sampled with the SPDE approach on a complex domain. From left to right we increased the value of ν, using $K = 1$, $K = 5$ and $K = 20$. The corresponding values for ν, given that $d = 2$, are respectively 1, 9 and 39. In the central area, we imposed an anisotropy factor 100 in the angular direction. The anisotropy ratio is reduced to isotropy towards the "leaves." The correlation length is 5 % of the domain size.

eigenvalue problem reduces to:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{M}\mathbf{v}, \qquad [\mathbf{A}]_{ij} = \int_{\Omega}\int_{\Omega} r(x,y)\phi_j(y)\phi_i(x)\,\mathrm{d}x\mathrm{d}y.$$

To compute the eigendecomposition efficiently we apply the low-rank pivoted Cholesky decomposition to replace the (large) matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ by the low-rank matrix $\mathbf{A}_m := \mathbf{L}_m\mathbf{L}_m^{\mathrm{T}}$, with $\mathbf{L}_m \in \mathbb{R}^{N \times m}$ and such that $\|\mathbf{A} - \mathbf{A}_m\| < \varepsilon$ [10]. In general, the performance of the method strongly depends on the decay rate of the spectrum of $\mathbf{A}$. Otherwise, methods based on $\mathcal{H}$-matrices for directly computing the square root of the covariance matrix show better performance [7].

A major limitation of this approach is that, for kernels depending on the distance function, that is $r(x,y) = h(\|x - y\|)$, the Euclidean distance does not always account for the geometry of the domain. It would be preferable to adopt the *geodesic distance*. In this way, geometrically-close but topologically distinct regions correctly show small correlation. There are two major problems to be addressed: 1) how to efficiently evaluate the distance function, and 2) how to ensure positive-definiteness of the corresponding Hilbert-Schmidt operator.

When relying on low-rank pivoted Cholesky decomposition, it is not necessary to assemble the full matrix $\mathbf{A}$: the algorithm only needs the diagonal entries and the function returning the i-th row of the matrix. Therefore, we approximate such functions with

$$[\mathbf{A}]_{i,j} \approx \int_{\Omega}\int_{\Omega} h(\delta(x_i,y))\phi_j(y)\phi_i(x)\,\mathrm{d}x\mathrm{d}y.$$

where $\delta(x_i, y)$ is the solution of the eikonal equation:

$$\begin{cases} \|\nabla_y\delta(x,y)\| = 1, & x \in \Omega \setminus \{x_i\}, \\ \delta(x_i, x_i) = 0. \end{cases} \tag{3}$$
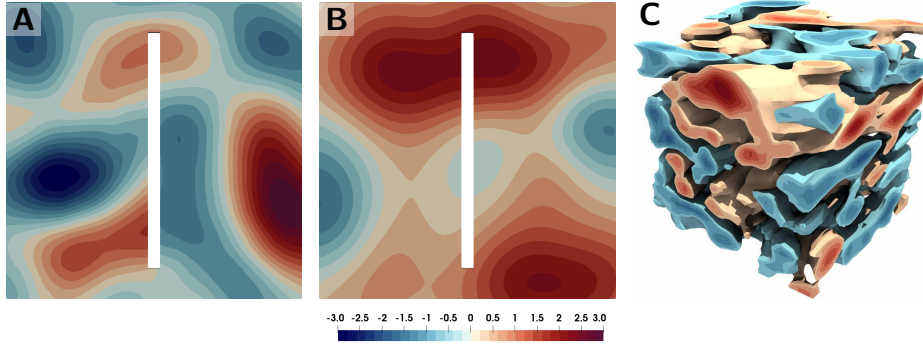
Fig. 2. Random fields generated with geodesic distance. A: a domain with a hole in the central area. Geometrically close regions around the cut are uncorrelated, as expected. B: when adopting the Euclidean distance, a correlation around the cut is present. C: a more complex example with non-constant and anisotropic metric for the distance function. The correlation length is $20\,\%$ of the domain size in all the examples.

The number of required eikonal evaluations matches the rank of the low-rank approximation of $\mathbf{A}$. In Figure 2, we compare random fields generated with geodesic distance (A) and with Euclidean distance (B).

The second issue, regarding the positive-definiteness of the Hilbert-Schmidt operator, is more subtle. From our experience, the pivoted Cholesky procedure often ends prematurely because of a negative pivot. It also occurs on simple domains with Euclidean distance approximated by the eikonal solution, suggesting that the numerical error may play a role. Nonetheless, the resulting eigenfunctions are sufficiently accurate to sample random fields via KL expansion. There are examples of correlation functions, e.g. the square-exponential, which are not positive definite even with exact geodesic distance: for instance, the sphere with great circle distance [8].

The geodesic distance can also be generalized to include varying velocity and anisotropy, substituting the norm in (3) with $\sqrt{\mathbf{G}(x)\mathbf{p}\cdot\mathbf{p}}$, with $\mathbf{G}$ symmetric positive definite. An example is provided in Figure 2C, where $\mathbf{G} = \sigma_{\mathrm{f}}\mathbf{f}\otimes\mathbf{f} + \sigma_{\mathrm{t}}(\mathbf{I}-\mathbf{f}\otimes\mathbf{f})$, $\sigma_{\mathrm{f}} = 1 = 10\sigma_{\mathrm{t}}$, $\mathbf{f} = \cos\alpha(z)\mathbf{e}_1 + \sin\alpha(z)\mathbf{e}_2$ and $\alpha(z) = \frac{\pi}{3}(2z-1)$.

4   Comparison of the two methods

In this section we compare the SPDE and the geodesic-based KL (geoKL) approaches in terms of quality of the samples and computational cost. In the test, we drew $10\,000$ samples from a random field defined on a hollow square domain, using both methods (see Figure 3). The domain was discretized with a uniform mesh of size $h = 1/100$. A squared-exponential kernel $h(d) = e^{-d^2/\rho^2}$ with fixed correlation length of $\rho = 0.2$ was used in geoKL. Correspondingly, we set the number of iterations K for the SPDE method to $K = 5$ and $\kappa = 2\sqrt{\nu}/\rho$.
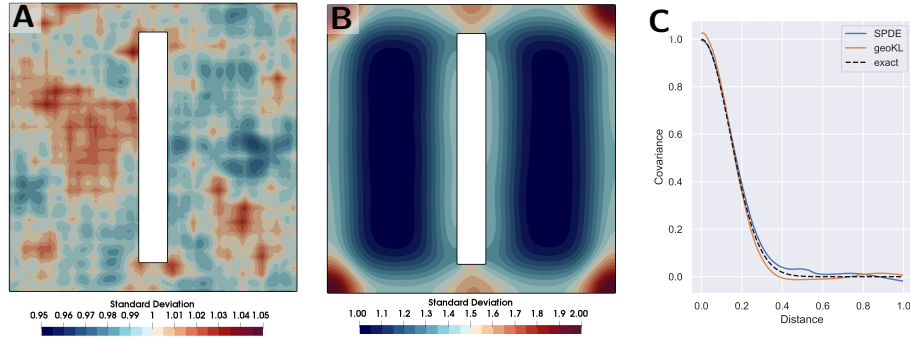
Fig. 3. Comparison between geoKL and SPDE methods for sampling of random fields. A: estimated standard deviation for geoKL method, B: estimated standard deviation for SPDE method, C: estimated correlation as function of distance along geodesic connecting points $(0.25, 0.25)$ and $(0.75, 0.25)$. A total of $10\,000$ samples were drawn.

Estimated mean and variance for each method are reported in Figure 3. The variance with the geoKL method was uniform across the domain, with some numerical artifacts (Figure 3A). Such oscillations also present in the trace of $\mathbf{A}_m$ and are likely due to numerical error in the geodesic distance. For problem (3), the singularity at the origin $y = x_i$ is indeed responsible for severe degradation of the convergence rate. The variance with the SPDE method (Figure 3B) showed instead a strong boundary effect, being significantly larger close to the boundary. This is a known effect of Neumann boundary conditions [12]. Finally, we computed the correlation with respect to the geodesic distance from point $(0.25, 0.25)$ towards point $(0.75, 0.25)$: see Figure 3C, reporting excellent agreement.

Method geoKL took $26\,$s to approximate 87 eigenfunctions (tolerance was 10^{-8}). Sampling via KL expansion took a fraction of a second. At shorter correlation lengths, the geoKL method took significantly longer time: respectively $87\,$s with $\rho = 0.1$ and $270\,$s with $\rho = 0.05$. In contrast, the computational cost of drawing one sample with the SPDE method was constant, regardless of the value of ρ. In this case the total time was $50\,$min ($3\,$s per samplex). As usual, timings are purely indicative, as they are greatly affected by implementation. Moreover, the SPDE-based samples were not computed in parallel.[6]

5   Application to atrial fibrosis

Initiation and perpetuation of atrial fibrillation (AF) is highly associated to heterogeneities in the substrate [9]. Hence, computational AF models for clinical applications should account for uncertainty in the anatomy and heterogeneity in the parameters. In this section we provide two examples of spatially-correlated heterogeneities for the atria.

---

[6] Python code is available at the address `https://github.com/pezzus/fimh2019`.
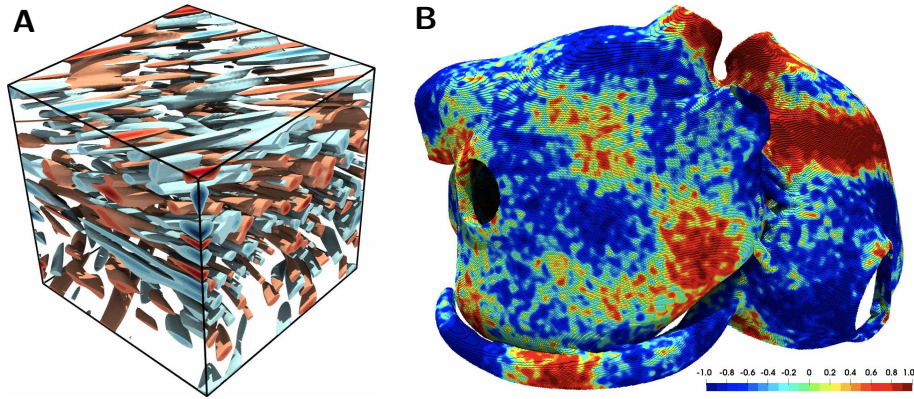
Fig. 4. Examples of randomly-generated fibrosis patterns on 3D tissues. A: the same configuration described in the last example of Figure 2, but with shorter correlation and stronger anisotropy. B: a fibrosis pattern on a realistic atrial geometry.

Cardiac myocytes form a structure of branching and merging fibers. Structural damage and fibrosis have a tendency to follow the main fibre orientation. Most commonly, these fibrosis patterns have dimensions in the order of the size of a few myocytes, i.e. ten to a few hundred micrometer. Such small structures cannot be assessed by clinical imaging methods; at best, techniques such as late gadolinium-enhanced magnetic resonance imaging provide the *density* of fibrosis averaged over a volume of several cubic millimetres [16]. Several modelling studies of cardiac electrophysiology have therefore used rule-based methods to generate anisotropic correlated random patterns to implement fibrosis in models [4,16].

Figure 4A shows a simulation of a random field with strong and non-stationary anisotropy in the correlation length. The interstitial space in the figure may represent fibrosis at microscale, for instance. We used the SPDE approach with $N = 10$, $\rho = 0.05$ (unit-length), and diffusion coefficient in the fibre direction 100 times larger than in the cross-fibre direction.

Figure 4B involves a more complex geometry. We simulated a fibrosis pattern on a patient-specific atrial anatomy using two random fields, one with short correlation length ($4\,\mathrm{mm}$), representing tissue-scale fibrosis, and another with longer correlation ($2\,\mathrm{cm}$), for organ-scale "patchiness." The two random fields were eventually averaged. The synthetic fibrosis pattern in Figure 4 resembles patterns observed with late gadolinium enhanced MRI [1].

6   Final remarks

We presented two alternative approaches to simulate spatially-correlated random fields for complex geometries. The SPDE approach is fast and sufficiently general, is easy to implement within any finite-element code, and handles general geometries. However, it does not provide much flexibility in the choice of the

correlation function. In contrast, sampling the random fields with KL expansion provides the freedom to select the kernel, including the geodesic distance, but it could be computationally demanding in some cases.

A potential application of such techniques is the automatic generation of realistic fibrosis patterns for the atria. This is important, for instance, for quantifying the uncertainty in cardiac electrophysiology, an emerging trend in the community and fundamental in our opinion in view of clinical applications.

References

1. Benito, E.M., Carlosena-Remirez, A., Guasch, E., Prat-González, S., Perea, R.J., Figueras, R., Borràs, R., Andreu, D., Arbelo, E., Tolosana, J.M., Bisbal, F., Brugada, J., Berruezo, A., Mont, L.: Left atrial fibrosis quantification by late gadolinium-enhanced magnetic resonance: A new method to standardize the thresholds for reproducibility. Europace **19**(8), 1272–1279 (2017). https://doi.org/10.1093/europace/euw219

2. Bolin, D., Kirchner, K., Kovács, M.: Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. IMA Journal of Numerical Analysis pp. 1–23 (2018). https://doi.org/10.1093/imanum/dry091

3. Canuto, C., Hussaini, M.Y., Quarteroni, A., Zang, T.A.: Spectral methods. Springer, Berlin, Heidelberg (2006)

4. Clayton, R.H.: Dispersion of Recovery and Vulnerability to Re-entry in a Model of Human Atrial Tissue With Simulated Diffuse and Focal Patterns of Fibrosis. Frontiers in Physiology **9**(AUG), 1–16 (2018). https://doi.org/10.3389/fphys.2018.01052

5. Croci, M., Giles, M.B., Rognes, M.E., Farrell, P.E.: Efficient White Noise Sampling and Coupling for Multilevel Monte Carlo with Nonnested Meshes. SIAM/ASA Journal on Uncertainty Quantification **6**(4), 1630–1655 (2018). https://doi.org/10.1137/18M1175239

6. Dietrich, C.R., Newsam, G.N.: Fast and Exact Simulation of Stationary Gaussian Processes through Circulant Embedding of the Covariance Matrix. SIAM Journal on Scientific Computing **18**(4), 1088–1107 (1997). https://doi.org/10.1137/S1064827592240555

7. Feischl, M., Kuo, F.Y., Sloan, I.H.: Fast random field generation with H-matrices. Numerische Mathematik **140**(3), 639–676 (2018). https://doi.org/10.1007/s00211-018-0974-2

8. Gneiting, T.: Strictly and non-strictly positive definite functions on spheres. Bernoulli **19**(4), 1327–1349 (2013). https://doi.org/10.3150/12-BEJSP06

9. Haissaguerre, M., Shah, A.J., Cochet, H., Hocini, M., Dubois, R., Efimov, I., Vigmond, E., Bernus, O., Trayanova, N.: Intermittent drivers anchoring to structural heterogeneities as a major pathophysiological mechanism of human

persistent atrial fibrillation. Journal of Physiology **594**(9), 2387–2398 (2016). https://doi.org/10.1113/JP270617

10. Harbrecht, H., Peters, M., Schneider, R.: On the low-rank approximation by the pivoted Cholesky decomposition. Applied Numerical Mathematics **62**(4), 428–440 (2012). https://doi.org/10.1016/j.apnum.2011.10.001

11. Harbrecht, H., Peters, M., Siebenmorgen, M.: Efficient approximation of random fields for numerical applications. Numerical Linear Algebra with Applications **22**(4), 596–617 (2015). https://doi.org/10.1002/nla.1976

12. Khristenko, U., Scarabosio, L., Swierczynski, P., Ullmann, E., Wohlmuth, B.: Analysis of boundary effects on PDE-based sampling of Whittle-Matérn random fields. arXiv e-prints arXiv:1809.07570 (2018)

13. Lindgren, F., Rue, H., Lindström, J.: An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **73**(4), 423–498 (2011). https://doi.org/10.1111/j.1467-9868.2011.00777.x

14. Pathmanathan, P., Gray, R.A.: Validation and trustworthiness of multiscale models of cardiac electrophysiology. Frontiers in Physiology **9**(FEB), 1–19 (2018). https://doi.org/10.3389/fphys.2018.00106

15. Schotten, U., Verheule, S., Kirchhof, P., Goette, A.: Pathophysiological Mechanisms of Atrial Fibrillation: A Translational Appraisal. Physiological Reviews **91**(1), 265–325 (2011). https://doi.org/10.1152/physrev.00031.2009

16. Vigmond, E., Pashaei, A., Amraoui, S., Cochet, H., Haïssaguerre, M.: Percolation as a mechanism to explain atrial fractionated electrograms and reentry in a fibrosis model based on imaging data. Heart Rhythm **13**, 1536–1543 (2016)