Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Extensions of Multiscale Gaussian Random Field Simulation Algorithms

Peter Kramer[1], Orazgeldi Kurbanmuradov[2], and Karl Sabelfeld[3] [4]

[1] Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, NY 12180 USA
E-mail: kramerp@rpi.edu

[2] Center for Phys. Math. Research
Turkmenian State University
Saparmyrat Turkmenbashy av. 31
744000 Ashgabad, Turkmenistan
E-mail: kurbanmuradov@yandex.ru

[3] Weierstrass Institute for Applied
Analysis and Stochastics
Mohrenstra"se 39
D – 10117 Berlin
Germany
E-Mail: sabelfeld@wias-berlin.de

[4] Institute of Computational Mathematics
and Mathematical Geophysics
Russian Acad. Sci.
Lavrentieva str., 6
630090 Novosibirsk
Russia

No. 1040
Berlin 2005

1991 *Mathematics Subject Classification.* 65C05, 65C20, 65T60.

Key words and phrases. Randomization method, Fourier-wavelet representation, multiscale random fields.

This work is partly supported by Grants: German DFG Grant 436 TUK 17/1/05, Russian RFBR Grant N 03-01-00914, NATO Linkage Grant N 978912, and CLG 981426. PRK is supported by an NSF grant DMS-0207242. For the kind hospitality the second author wishes to thank WIAS.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We analyze and compare the efficiency and accuracy of two simulation methods for homogeneous random fields with multiscale resolution. We consider in particular the Fourier-wavelet method and three variants of the Randomization method: (A) without any stratified sampling of wavenumber space, (B) with stratified sampling of wavenumbers with equal energy subdivision, (C) stratified sampling with a logarithmically uniform subdivision. We focus on fractal Gaussian random fields with Kolmogorov-type spectra. As noted in previous work by [3, 6], variants (A) and (B) of the Randomization method are only able to generate a self-similar structure function over three to four decades with reasonable computational effort. By contrast, variant (C), suggested by [34, 22], along with the Fourier-wavelet method developed by [6], is able to reproduce accurate self-similar scaling of the structure function over a number of decades increasing linearly with computational effort (for our examples we will show that nine decades can be reproduced). We provide some conceptual and numerical comparison of the various cost contributions to each random field simulation method (overhead, cost per realization, cost per evaluation).

When evaluating ensemble averaged quantities like the correlation and structure functions, as well as some multi-point statistical characteristics, the Randomization method can provide good accuracy with considerably less cost than the Fourier-wavelet method. The Fourier-wavelet method, however, has better ergodic properties, and hence becomes more efficient for the computation of spatial (rather than ensemble) averages which may be important in simulating the solutions to partial differential equations with random field coefficients.

1 Introduction

Random functions (generally referred to as random fields) provide a useful mathematical framework for representing disordered heterogeneous media in theoretical and computational studies. One example is in turbulent transport, where the velocity field representing the turbulent flow is modeled as a random field $\mathbf{v}(\mathbf{x}, t)$ with statistics encoding important empirical features, and the temporal dynamics of the position $\mathbf{X}(t)$ and velocity $\mathbf{V}(t) = \frac{d\mathbf{X}}{dt}$ of immersed particles is then governed by equations involving this random field such as

$$m\,\mathrm{d}\mathbf{V}(t) = -\gamma\Big(\mathbf{V}(t) - \mathbf{v}(\mathbf{X}(t), t)\Big)\,\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T\gamma}\,\mathrm{d}\mathbf{W}(t), \tag{1.1}$$

where m is particle mass, γ is its friction coefficient, k_{B} is Boltzmann's constant, T is the absolute temperature, and $\mathbf{W}(t)$ is a random Wiener process representing molecular collisions. Another example is in transport through porous media, such as groundwater aquifers, in which the hydraulic conductivity $\mathsf{K}(\mathbf{x})$ is modeled as random field reflecting

the empirical variability of the porous medium. The Darcy flow rate $\mathbf{q}(\mathbf{x})$ in response to pressure applied at the boundary is governed by the Darcy equation

$$\begin{aligned} \mathbf{q}(\mathbf{x}) &= -\mathsf{K}(\mathbf{x}) \operatorname{grad} \phi(\mathbf{x}), \\ \operatorname{div} \mathbf{q} &= 0, \end{aligned} \tag{1.2}$$

in which the random hydraulic conductivity function appears as a coefficient, and the applied pressure is represented in the boundary conditions for the internal pressure head ϕ. Our concern is with the computational simulation of random fields for applications such as these.

Interesting insights into the dynamics of transport in disordered media can be achieved already through relatively simple random models for the velocity field, such a finite superposition of Fourier modes, with each amplitude independently evolving according to an Ornstein-Uhlenbeck process [4, 37]. Here efficient and accurate numerical simulations of the flow can be achieved through application of the well-developed literature on simulating stochastic ordinary differential equations [18]. We will focus instead on the question of simulating random fields which involve challenging multiscale structures such as those relevant to porous media and turbulent flow simulations. Many questions remain open for the case of Gaussian multiscale random fields, so we confine our attention to this class.

We shall consider general real-valued Gaussian homogenous random fields $u(\mathbf{x})$ defined on multi-dimensional Euclidean space $\mathbb{R}^d$. (The extension to the vector-valued case is discussed in Appendices A and B.) Most of our main points can be made in the context of real-valued scalar random fields on the real line, and our numerical examples are all done within this simpler context, but we simply wish to indicate how some issues such as cost considerations extend to multiple dimensions.

Under quite general conditions, a real-valued Gaussian homogenous random field $u(\mathbf{x})$ can be represented through a stochastic Fourier integral [29]

$$u(\mathbf{x}) = \int_{R^d} \mathrm{e}^{2\pi \mathrm{i} \mathbf{k} \cdot \mathbf{x}} E^{1/2}(\mathbf{k}) \tilde{W}(\mathrm{d}\mathbf{k}) \tag{1.3}$$

where $\tilde{W}(\mathrm{d}\mathbf{k})$ is a complex-valued white noise random measure on $\mathbb{R}^d$, with $\tilde{W}(B) = \overline{\tilde{W}(-B)}$, $\langle \tilde{W}(B) \rangle = 0$, and $\langle \tilde{W}(B) \overline{\tilde{W}(B')} \rangle = \mu(B \cap B')$ for Lebesgue measure μ and all Lebesgue-measurable sets B, B'. The spectral density $E(\mathbf{k})$ is a nonnegative even function representing the strength (energy) of the random field associated to the wavenumber $\mathbf{k}$, meaning the length scale $1/|\mathbf{k}|$ and direction $\mathbf{k}/|\mathbf{k}|$.

Multiscale random fields will have a multiscale spectral density, meaning that $E(\mathbf{k})$ will have substantial contributions over a wide range of wavenumbers $k_{min} \ll |\mathbf{k}| \ll k_{max}$, with $k_{max}/k_{min} \gg 1$. This poses a challenge for efficient simulation.

Several approaches are based on various discretizations of (1.3) which give rise to finite sums of functions with independent Gaussian random coefficients. A Riemann sum discretization of the stochastic integral is easy to implement [36, 39, 40, 30, 14], and following [7, 17, 25], we shall refer to it as the standard Fourier method. As documented in [7], this method can suffer from false periodicity artifacts of the discretization, particularly if the wavenumbers are chosen with uniform spacing.

An alternative randomization method has been developed which evaluates the stochastic integral (1.3) through a finite set of randomly chosen wavenumbers generated by Monte

Carlo methods [19, 28, 35]. Yet another method designed for multiscale random field simulation is the Fourier-wavelet method [6], which arises from a wavelet decomposition of the white noise measure in (1.3). We stress that neither of these methods require that the multiscale structure be self-similar, though they are well-suited for this special case. The Fourier-wavelet and Randomization methods have been previously studied and compared, particularly in the context of (massless) turbulent diffusion problems [3, 6], with the general conclusion that the Randomization method performs well for simulating random fields with a self-similar multiscale scaling structure extending over a small number of decades, while the Fourier-wavelet method becomes more efficient for random fields with a large number of decades of self-similar scaling. The methods also have some theoretical differences. The Randomization Method reproduces the correlation function (or structure function) with only sampling error and no bias, while the Fourier-wavelet method incurs some bias from truncation of the sums in the associated random field representation. On the other hand, the Fourier-wavelet method simulates a truly Gaussian random field, while the statistics of quantities involving two or more evaluation points are generally simulated as non-Gaussian by the Randomization Method. However, as more wavenumbers are included in the random field representation, central limit theorem arguments indicate that the statistics simulated by the Randomization Method should approach Gaussian values [20]. In particular, the simulations of fractal random fields in [6] indicate that the kurtosis (normalized fourth order moment) of spatial increments in the random field was close to its Gaussian value of 3 over a range of scales which was one or two decades fewer than the range over which the second order moments were accurately simulated.

In the comparisons [3, 6], a particular version of the Randomization method was used; namely, the random wavenumbers were chosen according to a subdivision of wavenumber space into sampling bins of equal energy. In the studies [34, 22], a logarithmically stratified subdivision of wavenumber space was found to be significantly more efficient in representing self-similar power-law spectra such as those corresponding to Kolmogorov turbulence. This implementation of the Randomization method [22], a similar implementation of the standard Fourier method with wavenumber discretized uniformly and deterministically in logarithmic space [38], and a multiscale wavelet method [10] have all been employed to simulate dispersion of pair particles in isotropic Gaussian frozen pseudoturbulence with a Kolmogorov spectrum extending over several decades, in some cases with a constant mean sweep. Of particular interest in these works is whether the classical Richardson's cubic law can be observed in numerically generated pseudoturbulence [22, 38, 10].

We investigate the efficiency and accuracy of the Fourier-wavelet and Randomization methods, including alternative strategies for stratified sampling, along the following directions:

- The previous studies of which we are aware [3, 6] focus on the cost of simulating the value of the random field at a particular point x on demand, as is appropriate in the turbulent diffusion problem (1.1). By contrast, the solution of the porous medium problem requires the generation of the random field over the whole computational domain at once. We study the computational cost and accuracy of the Randomization and Fourier-wavelet methods for the construction and evaluation of a random field over a whole finite difference grid. Both the overhead cost and the cost of simulating each new realization of the random field is considered.

- We study finer statistical features of the random field beyond the second order structure function

$$S(r) = \langle (u(x+r) - u(x))^2 \rangle$$

and the kurtosis. In particular, with the logarithmically stratified sampling strategy for the Randomization Method, the random field is approximately represented with a considerably smaller number of random variables than the Fourier-wavelet method. We examine what types of statistics of the random field are still well represented by this compressed representation.

- We study the ergodicity properties of the simulation methods, which we expect to be important in the accurate simulation of statistics of the solutions of partial differential equations with random field coefficients such as the Darcy equation (1.2).

We begin in Section 2 by identifying some physical and numerical parameters which will play a key role in the development of the simulation algorithms and in quantifying their costs. We then present a brief but self-contained description of the Randomization Method (Section 3) and the Fourier-wavelet Method (Section 4), framing the discussion primarily in terms of one-dimensional random functions for notational simplicity. Some details of the extensions of the random field simulation algorithms to multiple dimensions can be found in Appendix A. We begin our examination of the numerical methods with a theoretical discussion in Section 5 of how their costs should scale with respect to various physical and numerical implementation parameters. We then revisit the question studied in [3], [9] concerning the comparative ability of the methods to generate random fields with self-similar fractal scaling of the second order structure functions over a large number of decades. Our contribution here is to consider variations of the Randomization Method which significantly improve its performance. We then turn to comparisons of the ergodic properties (Section 6) and the quality of the multi-point statistics of the random fields (Section 7) simulated by the Randomization Method and Fourier-wavelet Method. Our findings are summarized in Section 8.

2 General Simulation Framework

To discuss the implementation and the costs of the simulation methods, we first delimit the questions to be asked about the random field $u(\mathbf{x})$ to be simulated. We suppose here that $u(\mathbf{x})$ is a well-defined scalar-valued homogenous, Gaussian random field with given spectral density $E(\mathbf{k})$ (which is just the Fourier transform of the correlation function; see Appendix A). We quantify first in Subsection 2.1 some fundamental length scales of the random field which play a key role in choosing simulation parameters, then discuss in Subsection 2.2 some further length scales determined by the context of the problem or the choice of numerical implementation.

2.1 Length Scales of the Random Field

One of the most fundamental quantitative properties of a random field is its *correlation length*, which we define as:

$$\ell_c = \left(V_d^{-1} \frac{\sup_{\mathbf{k} \in R^d} E(\mathbf{k})}{\int_{R^d} E(\mathbf{k}) \, \mathrm{d}\mathbf{k}} \right)^{1/d}, \tag{2.1}$$

where $V_d = 2\pi^{d/2}/(d\Gamma(d/2))$ is the volume of the unit ball in d dimensions ($V_1 = 2$, $V_2 = \pi$, $V_3 = 4\pi/3$). The usual definition [27] has simply $E(\mathbf{0})$ in the numerator, but our definition generalizes meaningfully to random fields with the spectral density vanishing at the origin. Indeed $E(\mathbf{0})$ is the integral of the trace of the correlation function, which under many conditions gives the product of the random field variance and the correlation volume. The denominator precisely cancels out the random field variance $\langle u^2 \rangle$, and the remaining operations convert the correlation volume to a correlation length. If however the random field has oscillations, the integral of the correlation function may underrepresent the actual correlation volume (including the extent of negative correlations). This is why we have simply modified the definition to involve the value of the spectral density at its peak wavenumber; it coincides of course with the standard convention in the case of random fields with spectral density peaked at the origin (as is often the case in the absence of strong negative correlations in the random field).

We similarly define a *smoothness microscale* for the random field:

$$\ell_s = \left(V_d^{-1} \frac{\sup_{\mathbf{k} \in R} |\mathbf{k}|^2 E(\mathbf{k})}{\int_{R^d} |\mathbf{k}|^2 E(\mathbf{k}) \, \mathrm{d}\mathbf{k}} \right)^{1/d}, \tag{2.2}$$

which is really an analogous correlation length for the random field gradient $\boldsymbol{\nabla} u$. The two length scales ℓ_c and ℓ_s generalize the notion of integral length scale and Kolmogorov dissipation length scale in turbulent spectra to general homogenous random fields. The correlation length can be thought of as the largest length scale on which the random field $u(\mathbf{x})$ exhibits a nontrivial correlation structure. That is, for $|\mathbf{x} - \mathbf{x}'| \gg \ell_c$, the values of $u(\mathbf{x})$ and $u(\mathbf{x}')$ are independent to a good approximation. The smoothness microscale, conversely, describes the smallest length scale on which the random field has nontrivial correlation structure. On smaller scales, the random field appears smooth. More precisely, for $|\mathbf{x} - \mathbf{x}'| \ll \ell_s$, the random field over the line segment connecting $\mathbf{x}$ and $\mathbf{x}'$ can be well approximated by a linear interpolation between $u(\mathbf{x})$ and $u(\mathbf{x}')$ (with *relative* error $o((|\mathbf{x} - \mathbf{x}'|/\ell_s)^2)$).

We will contemplate only random fields with spectral density behaving well enough at small and large wavenumbers to be integrable, so that the correlation length ℓ_c in (2.1) is well-defined as a finite nonzero value. We admit random fields for which the integral in the denominator of (2.2) converges or diverges; in the latter case, we define $\ell_s = 0$. Idealized fractal random fields [26, 11], such as those associated with the Kolmogorov inertial range theory of turbulence $E(\mathbf{k}) \propto |\mathbf{k}|^{-5/3}$ can be placed within the present framework if we agree from the outset that the fractal scaling is smoothed out at a pre-defined large length scale. This will be appropriate for any physical application, and even from a purely mathematical point of view, we can think of this length scale cutoff as defining a concrete goal for the simulation of a fractal random field with finite effort. Any numerical

simulation or physical application will also necessarily have a positive lower limit on the length scale of fractal scaling, but we do not need to enforce this within our mathematical framework.

For some random fields, the smoothness length scale is comparable to the correlation length. This is true in particular if the random field depends only on one physical length scale. For such "single-scale" Gaussian homogenous random fields, a wide variety of simulation techniques beyond the Fourier-wavelet and Randomization methods may well be adequate [30]. Our concern is with simulating multiscale random fields, meaning that $\ell_s \ll \ell_c$. One situation where this can arise is in two-scale random fields (such as those contemplated in homogenization theory [5], where $E(\mathbf{k})$ is peaked near widely separated wavenumbers $|\mathbf{k}| \sim \ell_c^{-1}$ and $|\mathbf{k}| \sim \ell_s^{-1}$ but rapidly decaying away from these values. In this case, the random field could be simply simulated by expressing it as the superposition of two independent single-scale random fields, one varying on the large scale, and one on the small scale. But in applications such as turbulence and porous media flow, and in any context involving fractal random models, the spectral density has nontrivial contributions over a great range of wavenumbers within the wide interval $\ell_c^{-1} \ll |\mathbf{k}| \ll \ell_s^{-1}$, and the random field is not well-approximated by a superposition of a few single-scale random fields. It is in these situations that the power of the Fourier-wavelet and Randomization methods is indicated.

2.2 Length Scales Introduced in Computation

In some applications, such as the simulation of a particle moving through a prescribed turbulent field (1.1), one may wish to be able to evaluate the random field $u(\mathbf{x})$ at an arbitrary point on demand (not known in advance of the generation of the random field). Both the Randomization method and Fourier-wavelet method are able to do this in an efficient way, as shown in previous works [35], [9].

In other applications, such as in the simulation of the flow through a porous medium [16, 33], we may instead require the random field to be simulated over an entire pre-defined region. For this to require finite computational work, we must agree upon a domain length scale L and a sampling length scale h. The domain length scale describes the linear extent of the region over which the random field is evaluated, and h denotes the distance between points at which the random field is calculated. In the simplest case (which is sufficient for our illustrative purposes), the random field is to be simulated on a Cartesian grid with linear extent L in each direction, with grid spacing h. We stress however that both the Randomization Method and the Fourier-wavelet Method are perfectly capable of generating random field samples over an irregularly arranged collection of points.

Finally, for any application, the ideal random field to be simulated has nontrivial structure on length scales ranging from ℓ_s to ℓ_c. A numerical multiscale representation of this random field will be associated with certain finite minimum and maximum length scales ℓ_{min} and ℓ_{max}, outside of which the method cannot be expected to accurately represent the structure of the ideal random field. The length scales ℓ_{max} and ℓ_{min} can be related to more fundamental parameters of the Monte Carlo simulation methods, as we describe in subsequent sections. Generally speaking, ℓ_{max} should be chosen to be at least as large as the correlation length ℓ_c (but need not be as large as the domain length L). On the

other hand, $\ell_{\min}$ is either set equal or somewhat smaller than $\min(h, \ell_s)$ or at a larger value determined by computational cost constraints. In the latter case, the numerically simulated random field can only be expected to be an accurate representation of the desired random field when viewed on length scales larger than $\ell_{\min}$.

3 Randomization methods

In the main text, we will present the numerical algorithms for the case of a real-valued homogenous Gaussian scalar random field $u(x)$ defined on the one-dimensional real line. The generalization to the multi-dimensional case is given in Appendices A and B.

The simplest form of the Randomization method, which we shall refer to as *variant A*, reads [35]

$$u^{(\mathrm{R})}(x) = \frac{\sigma}{\sqrt{n_0}} \sum_{j=1}^{n_0} \Big[\xi_j \, \cos(2\pi \, k_j \, x) + \eta_j \, \sin(2\pi \, k_j \, x) \Big] \, , \tag{3.1}$$

where ξ_j, η_j, $j = 1, \ldots, n_0$ are mutually independent standard Gaussian random variables (mean zero and unit variance), and $\sigma^2 = \int E(k) \, \mathrm{d}k = 2 \int_0^\infty E(k) \, \mathrm{d}k$. The wave numbers k_j, $j = 1, \ldots, n_0$ are chosen as independent random variables in $[0, \infty)$ according to the probability density function (pdf) $p(k) = 2E(k)/\sigma^2$, and are also independent of the ξ_j and η_j. This variant A of the Randomization Method may be thought of as the most straightforward way to approximate the Fourier stochastic integral (1.3) through a Monte Carlo integration approach, using the complex conjugacy between the simulated random variables associated to wavenumbers $\pm k$.

While the Randomization Method always produces random field approximations with the correct mean and correlation function *when averaged over a theoretically complete ensemble of realizations*, the practical concern is how well one or a finite number of samples of the simulated random field replicate the statistics of the true random field which is to be simulated. The randomization of the choice of wavenumbers creates some additional variability in the simulated random field (such as realizations where, say, the low wavenumbers happen to be undersampled). A common practice in improving Monte Carlo calculations is the employment of "variance reduction" techniques which constrain the random choices somewhat to mitigate the problem of generating an artifically large number of strongly deviant samples. An extreme remedy would be to prescribe the wavenumbers deterministically, as in the standard Fourier method discussed in Section 1, but this has its own artifacts [7].

A compromise which seeks to avoid the problems of both purely deterministic and purely random choices of random wavenumbers is to partition wavenumber space into bins, and a prescribed number of wavenumbers are chosen at random locations within each bin. This Monte Carlo variance reduction technique is an example of "stratified sampling" [32]. It ensures a certain coverage of wavenumber space, but still takes advantage of Monte Carlo integration techniques.

The mathematical framework for stratified sampling in the Randomization Method is given as follows. We take Δ as the total space from which the wavenumbers are to

be sampled (it can in general be chosen as $\Delta = [0, \infty)$ but can also be chosen as the possibly smaller support of the spectrum E on the nonnegative real axis. We then choose a partition of Δ into a union of smaller non-overlapping intervals $\Delta = \cup_{j=1}^{n} \Delta_j$. Within each interval Δ_j, we sample n_0 independent random wavenumbers k_{jl}, $l = 1, \ldots, n_0$ according to the probability distribution function

$$p_j(k) \equiv \begin{cases} \frac{2E(k)}{\sigma_j^2} & \text{for } k \in \Delta_j, \\ 0 & \text{for } k \notin \Delta_j, \end{cases}$$

where

$$\sigma_j^2 = 2 \int_{\Delta_j} E(k)\, \mathrm{d}k \ .$$

The simulation formula then reads

$$u^{(\mathrm{R,s})}(x) = \sum_{j=1}^{n} \frac{\sigma_j}{\sqrt{n_0}} \sum_{l=1}^{n_0} \Big[\xi_{jl} \cos(2\pi\, k_{jl}\, x) + \eta_{jl} \sin(2\pi\, k_{jl}\, x) \Big] \ . \tag{3.2}$$

The amplitudes ξ_{jl}, η_{jl}, $j = 1, \ldots, n$; $l = 1, \ldots, n_0$ are again standard Gaussian random variables which are mutually independent and independent of the choice of wavenumbers k_{jl}. One natural choice of stratified sampling for variance reduction is to choose a number of sampling bins n and then choose the sampling intervals Δ_j so that each of them contains an equal amount of "energy" (integral of the spectral density): $\sigma_j^2 = \frac{\sigma^2}{n}$ for $1 \le j \le n$. We refer to this stratified sampling strategy as *variant B* of the Randomization Method.

We will also explore an alternative stratified sampling strategy in which the sampling bins are simply assigned to be equally spaced with respect to $\ln k$. That is, the wave number intervals $\Delta_j = (\hat{k}_j, \hat{k}_{j+1}]$ for $j = 1, \ldots, n$ are defined according to a geometric distribution with ratio parameter q: $\hat{k}_{j+1} = q\hat{k}_j$, $j = 2, \ldots, n-1$. If the spectral density of the random field is confined to a bounded domain of wavenumbers $k \le k_{max}$, then we can choose $\hat{k}_n = k_{max}$; otherwise we take $k_{max} = \infty$. Similarly, if the minimal wavenumber k_{min} in the support of the spectral density is positive, we can choose $\hat{k}_1 = k_{min}$. Otherwise, we would choose $\hat{k}_1$ in some other way suggested by the spectrum, perhaps as the wavenumber at which the spectral density is maximal, and then adjoin a sampling bin $\Delta_0 = (0, \hat{k}_1)$. We call this stratified sampling strategy based on a logarithmically uniform subdivision *variant C* of the Randomization Method.

The motivation for this externally imposed subdivision scheme is that a multiscale random field, particularly one with a self-similar fractal property, might be well represented in a hierarchical manner with a certain number of computational elements at each important "length scale," with geometrically distributed length scales. Indeed, this is precisely what the Fourier-wavelet simulation method (in fact any wavelet method) does, and appears to be one of the essential elements behind its demonstrated efficiency in simulating random fields with multiscale structure over many decades [6, 9, 8]. We are led to consider, therefore, how incorporating a similar distribution of wavenumbers within the Randomization approach would compare with the Fourier-wavelet method. Other externally imposed subdivision strategies may be chosen based on the spectral density of the random field to be simulated as well as the type of statistics which are sought in the application.

All variants of the Randomization method provide unbiased estimators of the correlation function of the simulated random field, meaning that these statistics can in principle be

recovered with arbitrary precision through a sufficiently large sample size (though with the usual relatively slow convergence of Monte Carlo sampling), even with fixed finite values of the discretization parameters. In particular, there need not be a rigid maximal and minimal length scale, $\ell_{\max}$ and $\ell_{\min}$ within which the Randomization Method confines its effort, because the wavenumber sampling bins can extend to $k = 0$ or $k = \infty$. However, in practice, the finite number n_0 of wavenumbers sampled in these bins does impose effective maximum and minimum length scales over which the random field structure can be expected to be adequately represented. We discuss this in the context of a particular example in Subsubsection 5.1.2. The quality of higher order and multi-point statistics simulated by the Randomization Method is less clear; in particular the randomization of the wavenumbers makes the simulated field non-Gaussian. Central limit theorem concepts, however, suggest that with a sufficiently rich sampling of wavenumbers, the simulated field should have some approximately Gaussian properties. (See Appendix A.1 for a summary of some rigorous results along these lines). We investigate these questions in detail in Sections 5–7.

4 Fourier-wavelet simulation method

We present here some details for the one-dimensional case, and discuss multi-dimensional generalizations in Appendix B. A homogeneous Gaussian random field $u(x)$ can be represented using the Fourier-wavelet representation [6]:

$$u(x) = u_0 \sum_{m=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \gamma_{mj}\, f_m(2^m\,(x/\ell) - j)\ , \tag{4.1}$$

where u_0 is a dimensional constant having the same dimensions as the field variable u, ℓ is an arbitrary length scale, γ_{mj} is a family of mutually independent standard Gaussian random variables, and

$$f_m(\xi) = \int_{-\infty}^{\infty} e^{-2\pi \mathrm{i}\,\tilde{k}\,\xi}\, 2^{m/2}\, \tilde{E}^{1/2}(2^m\tilde{k})\hat{\phi}(\tilde{k})\, d\tilde{k}\ , \tag{4.2}$$

where $\tilde{E}(\tilde{k})$ is a dimensionless spectral density defined through

$$\tilde{E}(\tilde{k}) = \frac{1}{l u_0^2} E(\tilde{k}/l)\ .$$

Here, in order to ensure an efficient wavelet representation of the random field, $\hat{\phi}(k)$ is chosen as a compactly supported function which is the Fourier transform of the Meyer mother wavelet function based on a pth order perfect B-spline [6]:

$$\hat{\phi}(k) = -\mathrm{i}\,\mathrm{sign}(k)\, e^{\mathrm{i}\pi k}\, b(|k|), \tag{4.3}$$

where

$$b(k) = \begin{cases} \sin(\frac{\pi}{2}\nu_p(3k-1)), & k \in \left(\frac{1}{3}, \frac{2}{3}\right] \\ \cos(\frac{\pi}{2}\nu_p(\frac{3}{2}k-1)), & k \in \left(\frac{2}{3}, \frac{4}{3}\right] \\ 0, & \text{else.} \end{cases}, \tag{4.4}$$

and the function $\nu_p(x)$ is defined by

$$\nu_p(x) = (-1)^p \frac{4^{p-1}}{p} \left\{ [x - x_0]_+^p + [x - x_p]_+^p + 2 \sum_{j=1}^{p-1} (-1)^j [x - x_j]_+^p \right\} , \tag{4.5}$$

where $x_j = (1/2)[\cos(((p-j)/p)\pi) + 1]$, and $[a]_+ = \max(a, 0)$. The positive integer parameter p is chosen in [6] equal to 2.

The representation (4.1) expresses the random field $u(x)$ as a hierarchical random superposition of real, deterministic functions f_m and their translates. The function f_m can be thought of as encoding the structure of $u(x)$ on the length scale $2^{-m}\,\ell$. This can best be seen by the dual relation between spatial lengths and Fourier wavenumbers, since f_m is completely determined by the contributions of the spectrum $E(k)$ over the interval $\frac{1}{3\ell} 2^m \le k \le \frac{4}{3\ell} 2^m$.

The implementation of the Fourier-wavelet method of course requires that the sums over m and j in (4.1) be truncated to finite sums. This is done through consideration of the length scales over which the random field is to be sampled.

First, we choose $\ell = \ell_{\max}$ as some convenient length scale that represents the largest length scale of the random field which we wish to resolve in our simulation. If the random field is to be simulated over a grid with spacing h, it will be convenient to choose ℓ so that $\ell/h = 2^m$ for some nonnegative integer m. By setting $\ell = \ell_{\max}$, it is now convenient to truncate the sum over m to run over $0 \le m \le M-1$, thereby formally representing the random field down to length scale $\ell_{\min} = 2^{1-M} \ell_{\max}$. Note that this truncated Fourier-wavelet random field representation will only incorporate information from the energy spectrum $E(k)$ over the wavenumber range $\frac{1}{3}\ell_{\max}^{-1} \le k \le \frac{4}{3}\ell_{\min}^{-1}$, so one should be careful that the energy outside this range can be safely neglected for the application.

We turn now to the truncation of the sum over the translation index j. It is shown in [6], that if the spectrum $E(k)$ is smooth enough then the functions $f_m(\xi)$ decay like $|\xi|^{-p}$ where p is the order of the spline used to construct the Meyer wavelet. So long as $p \ge 2$, then, we can choose a "bandwidth" cutoff b so that the total mean-square contribution from terms with $|j - 2^m x/\ell| \ge b$ to the random field value at x is as small as desired. For illustation, we show in Fig. 1 typical curves $f_m(\xi)$ for $m = 0, 1, 2$ for the spectral density defined in (5.1), which we will study in Section 5. We therefore choose an appropriate value for b which meets our accuracy needs, and then, when evaluating the random field $u(x)$ at a desired point x, only incorporating the $2b+1$ terms satisfying $|2^m(x/\ell) - j| \le b$, where the notation $\lfloor y \rfloor$ denotes the greatest integer not exceeding the real value y.

Hence the finitely truncated Fourier-wavelet representation for the value of the random field at any location x can be written as follows

$$u^{(\mathrm{FW})}(x) = u_0 \sum_{m=0}^{M-1} \sum_{j'=-b}^{b} \gamma_{m,\bar{n}_m(x)+j'} \, f_m(2^m \, (x/\ell) - \bar{n}_m(x) - j') \, , \tag{4.6}$$

where $\bar{n}_m(x) \equiv \lfloor 2^m (x/\ell) \rfloor$. One must be careful when evaluating the random field at various locations x to be sure that the random variables γ_{mj} used for each evaluation are the same (and not independent!) when the same indices m and j are involved.

Detailed analysis of the errors of interpolation, discretization, and aliasing in the evaluation of the Fourier transform (4.2) can be found in [6]. We shall simply make use of these results in choosing suitable numerical parameters p, b, and $\Delta\xi$ in our simulations.

We will focus our attention on the cost of the Fourier-wavelet method (Section 5) and the quality of the random field statistics which it generates (Section 6 and 7), particularly in comparison to the Randomization Method. The Fourier-wavelet method is considerably more complicated than the Randomization Method, and it does incur a statisical bias through truncation of the sums in (4.6) and the need to approximate the functions f_m through interpolation from a finite set of data points. We will therefore be particularly interested to examine the circumstances in which the extra complexity of the Fourier-wavelet method make it worthwhile relative to the Randomization Method.

5 General Considerations of Cost

We begin our studies of the Randomization and Fourier-wavelet Methods by revisiting the question of how much computational effort is required by these methods to generate fractal self-similarity over a desired number of decades [6, 3]. We consider the computational cost in two settings in turn: those in which the random field is to be evaluated at points on demand (Subsection 5.1), or on a pre-defined computational domain (Subsection 5.2). Our theoretical considerations are intended to apply rather broadly to multiscale random fields in multiple dimensions with characteristic parameters defined in Section 2, but for illustrative purposes, we will refer in our discussion to some numerical results for a one-dimensional random field example $u(x)$ with spectral density

$$E(k) = \begin{cases} C_E|k|^{-\alpha}, & |k| \geq k_0; \\ 0, & |k| < k_0. \end{cases} \tag{5.1}$$

This random field has correlation length $\ell_c = (\alpha - 1)/(2k_0)$ and $\ell_s = 0$. In numerical calculations we choose specifically $\alpha = 5/3$ (corresponding to the Kolmogorov spectrum for the inertial range of a turbulent flow [13, 31, 24], $k_0 = 1$, and $C_E = 1$.

To assess the basic quality of the simulated fractal random fields, we shall use the correlation function $B(\rho) = \langle u(x+\rho)u(x)\rangle$ and the second order structure function $D(\rho) = \langle [u(x+\rho) - u(x)]^2\rangle$, where the angle brackets denote an average over an ensemble of independent random field realizations which are related to the spectrum by the formulae

$$B(\rho) = \int_0^\infty 2\,E(k)\,\cos(2\pi\,k\rho)\,dk, \quad D(\rho) = \int_0^\infty 4\,E(k)\,[1 - \cos(2\pi\,k\rho)]\,dk\ .$$

Note that the main contribution to the structure function $D(\rho)$ at a given value of ρ comes from the wavenumbers which are of order of $1/\rho$. The power law structure of the energy spectrum implies, by Fourier duality, that the structure function should exhibit a self-similar power law scaling on scales small compared to the cutoff length scale k_0^{-1}:

$$D(\rho) \sim J_\alpha \rho^{\alpha-1} \text{ for } \rho \ll k_0^{-1}$$

where

$$J_\alpha = 4C_E \int_0^\infty k^{-\alpha}\,[1-\cos(2\pi\,k)]\,dk = \begin{cases} -\frac{2^{1+\alpha}}{\pi^{1-\alpha}}\,\Gamma(1-\alpha)\sin(\alpha\,\pi/2), & 1<\alpha<3, \alpha\neq 2 \\ 4\pi^2, & \alpha = 2 \end{cases} \tag{5.2}$$

and $\Gamma(x)$ for $x < 0$ is obtained by analytical continuation of the standard definition [15, 23]. This induced power-law scaling in the structure function is most clearly seen by rewriting the structure function in the form

$$D(\rho) = \left[J_\alpha - 4\int_0^{\rho k_0} (1-\cos(2\pi\,k'))k'^{-\alpha}\,\mathrm{d}k'\right]\rho^{\alpha-1}. \tag{5.3}$$

To more clearly display the accuracy of the simulation methods in replicating the correct scaling (5.3) of the structure function, we will look at a rescaled form [6]

$$G_2(\rho) = D(\rho)/(J_\alpha \rho^{\alpha-1}),$$

which should satisfy $G_2(\rho) \sim 1$ for $\rho k_0 \ll 1$. We can therefore analyze rather rigorously how well the Monte Carlo methods are simulating the second order statistics of the fractal random field described by spectral density (5.1) by observing over how many decades the function $G_2(\rho)$ remains near the constant value 1.

5.1 Random Field Simulations with On Demand Evaluations

We begin by revisiting the cost of simulating a multiscale random field which is to be evaluated at certain points on demand (not specified in advance). The primary three components of either simulation method may be categorized as follows:

- the *preprocessing cost* of taking the desired energy spectrum and making the deterministic calculations needed for the random field representation,
- the *cost in simulating a new realization* of the random field, which involves the generation of a new set of random numbers, and
- the *cost of evaluation* of the current realization of the random field at a location specified on demand.

Certain aspects of these cost contributions have been considered previously [6, 3], but we wish to re-evaluate the cost analysis of the Randomization Method in light of new stratified sampling strategies. The preprocessing cost can probably be treated as less important than the other costs if many realizations of a random field with a single energy spectrum is needed [3], but it could play a more important role in dynamical simulations where the energy spectrum evolves in time. We therefore include a brief consideration of the preprocessing cost.

We quickly consider the costs of a standard Gaussian simulation scheme (Subsubsection 5.1.1), then consider the Randomization Method (Subsubsection 5.1.2) and the Fourier-wavelet Method (Subsubsection 5.1.3) in turn.

5.1.1 Standard Simulation Method

One generic approach to simulating a Gaussian random field at a set of points specified on demand is to recognize that the values of the random field at these points form a mean zero, jointly Gaussian collection of random variables [41]. Consequently, the value of the random field can be simulated progressively at each new point specified on demand by the standard regression formulas that describe the conditional expectation and variance of a Gaussian random variable from the covariance and realized values of the other variables to which the new random variable is correlated [12]. Implementing this in a straightforward way would yield a negligible preprocessing cost, and a total cost proportional to N_e^4 to simulate one realization of the values at N_e points specified on demand. (With this iterated regression approach, the division of the simulation cost into random field realization and cost of evaluation per point is not meaningful).

When the points at which the random field is to be evaluated becomes sufficiently dense, it may be more efficient to simulate the random field on a regular grid (with resolution h) and then evaluate the random field at the points to be specified on demand through interpolation. Then the preprocessing cost is determined by the calculation and Cholesky decomposition of the covariance matrix of the random field values on the grid; this cost scales as $(L/h)^d(\ell_c/h)^{2d}$. A realization of the random field entails the calculation of the random field values on the grid, and involves the multiplication of a vector of independent standard Gaussian random variables by the square root of the covariance matrix associated to the grid locations, which is obtained from the Cholesky decomposition. This operation has cost scaling as $(L/h)^d(\ell_c/h)^d$ per realization of the random field. Each evaluation of the random field through interpolation has order unity cost.

The essential point is that these standard Gaussian random field simulation approaches grow as a superlinear power law in N_e, $(L/h)^d$ and $(\ell_c/h)^d$. Whenever all these quantities are large, as is the case for a simulation at which a multiscale random field is to be evaluated at a large set of points, then the cost of this straightforward approach can be expected to be quite large. We will see that the Randomization and Fourier-wavelet Methods have costs which scale much less rapidly with respect to these parameters.

5.1.2 Randomization Method

Choice of Numerical Parameters As discussed in Section 2, we choose parameters in a Monte Carlo simulation method based on the maximum and minimum length scales $(\ell_{\text{max}}, \ell_{\text{min}})$ of the random field structure which we aim to capture. The crucial determinant of cost is the ratio of these length scales, which we will express in terms of the number of decades separating them:

$$N_{\text{dec}} = \log_{10}(\ell_{\text{max}}/\ell_{\text{min}}).$$

In our numerical example (5.1), we will choose $\ell_{\text{max}} = k_0^{-1}$ (in general it should be comparable to the correlation length ℓ_c). The remaining parameters to be determined are the number of sampling bins, n, and the number of wavenumbers chosen per sampling bin, n_0. These parameters are really set by N_{dec}, the number of decades of accurate simulation desired, as well as our choice of bin widths. We elaborate for each of the variants of the stratified sampling strategy.

Variant A Without stratified sampling, $n = 1$. Following [3], we can estimate the smallest length scale $\ell_{\min}$ which will be reasonably approximated by the Randomization Method using n_0 random wavenumbers as that length scale for which the average number of wave number samples lying in the interval $(1/\ell_{\min}, \infty)$ exceeds some critical numerical value c. The larger we choose c, the more stringently we are interpreting the phrase "accurate representation of the random field down to length scales $\ell_{\min}$."

For our example (5.1), the average number of wave numbers in the interval $(1/\ell_{\min}, \infty)$ is equal to $n_0(k_0\ell_{\min})^{\alpha-1}$, then from $n_0(k_0\ell_{\min})^{\alpha-1} = c$ we get

$$\ell_{\min} = \frac{1}{k_0}\left(\frac{n_0}{c}\right)^{-\frac{1}{\alpha-1}} \quad . \tag{5.4}$$

So with n_0 wavenumber samples, the number of decades accurately described is $N_{\text{dec}} = \frac{1}{\alpha-1}\log_{10}(n_0/c)$. Equivalently, the number of wavenumber samples required grows exponentially with the number of decades desired: $n_0 = c10^{N_{\text{dec}}(\alpha-1)}$.

We can see, for example, in Figures 2 and 3 that increasing the number of simulated wavenumbers from $n_0 = 160$ to $n_0 = 1000$ extends the domain of accurate self-similar scaling by less than a decade. If we wished to simulate 9 decades of scaling, then for $\alpha = 5/3$ we have to sample 10^6 wavenumbers, which is practically unrealistic. Increasing the number of decades of scaling with variant A of the Randomization Method therefore requires a very large extra investment of computational effort.

Variant B Here we select n sampling bins, each with equal energy, and sample n_0 wavenumbers from each bin. Applying similar arguments as in our analysis of variant A, and assuming that the accuracy parameter c is smaller than n_0 (so that $\ell_{\min}$ is assumed to fall in the wavenumber bin with the highest wavenumbers), we obtain

$$c = \frac{n_0 \int_{1/\ell_{\min}}^{\infty} k^{-\alpha}\,\mathrm{d}\,k}{\frac{1}{n}\int_{k_0}^{\infty} k^{-\alpha}\,\mathrm{d}k}$$

and consequently $\ell_{\min} = \frac{1}{k_0}(nn_0/c)^{-\frac{1}{\alpha-1}}$. The number of decades of accurate scaling is therefore related to our sampling effort as:

$$N_{\text{dec}} = \frac{1}{\alpha-1}\log_{10}(nn_0/c).$$

So the stratification of the sampling into bins of equal energy seems to lead to no improvement in efficiency; the number of decades is again logarithmically related to the total number of wavenumbers nn_0 sampled. The quality of the simulation is also not markedly improved by the equal energy stratified sampling, as seen in Figures 4 and 5.

Variant C We finally consider how the cost of the logarithmically stratified sampling strategy is related to the range of scales over which one wishes to simulate a multiscale random field accurately. Applying the same criterion as in the previous variants for

determining the smallest scale $\ell_{\min}$ which is simulated accurately given the number of sampling bins n, the ratio q between the bin boundaries, and the number of samples n_0 per bin, we obtain $c = n_0(\ell_{\min}\hat{k}_n)^{\alpha-1}$, where $\hat{k}_n = k_0 q^{n-1}$ is the left endpoint of the highest wavenumber sampling bin. Solving for $\ell_{\min}$, we obtain $\ell_{\min} = k_0^{-1} q^{1-n}(n_0/c)^{-1/(\alpha-1)}$, and so the number of decades of accuracy can be estimated as:

$$N_{\text{dec}} = \frac{1}{\alpha - 1}\log_{10} n_0 + (n-1)\log_{10} q. \tag{5.5}$$

Note that in contrast to Variants A and B of the Randomization Method, the number of decades resolved in this case scales linearly with the number of bins n. So if we fix the bin ratio q and the number of samples per bin n_0 at some reasonable values, our theoretical estimate suggests that we can simulate a number of decades proportional to our computational cost by simply increasing the number of sampling bins. This is well illustrated by numerical results presented in the left panel of Figures 6 and 7. We see that with the same effort as in the previous variants of the Randomization Method, we are able to simulate 9 decades of self-similar scaling accurately. The structure function for larger values of ρ is also well calculated by the Randomization method with the logarithmic wavenumber subdivision (Figure 8).

To emphasize the quality of the second order structure function simulated by the Randomization Method, we compare it against a direct Monte Carlo simulation of the collection of random variables $\delta u_i = \frac{u(\rho_i)-u(0)}{(J_\alpha \rho_i^{\alpha-1})^{1/2}}$, $i = 1, \ldots, n_p$, with $\rho_1 = l_{min}$ and $\rho_i = q\rho_{i-1}$ for $i \geq 1$. These random variables are Gaussian with zero mean and covariance

$$\langle \delta u_i \delta u_j \rangle = \frac{1}{2\, J_\alpha \rho_i^{\frac{\alpha-1}{2}} \rho_j^{\frac{\alpha-1}{2}}}\Big[D(\rho_i) + D(\rho_j) - D(\rho_i - \rho_j)\Big]$$

but in our simulations we approximate the right hand side by its limiting value for $\rho_i k_0 \ll 1$ where $D(\rho)$ is replaced by $J_\alpha \rho^{\alpha-1}$. The results of this direct simulation, displayed in the right panel of Figure 7, represent a Monte Carlo estimate $\hat{G}_2(\rho_m) = \langle (\delta u_m)^2 \rangle$ of the structure function which only exhibits sampling error. This direct simulation approach is of course impractical for actually simulating the values of a multiscale random field over a large number of points, as discussed in Subsubsection 5.1.1. Comparison of the panels in Figure 7 shows that the structure function simulated by the Randomization Method is of almost as good quality over 9 decades as the direct simulation with only sampling error.

The same verification was made for the kurtosis

$$G_4(\rho) = \frac{\langle (u(\rho) - u(0))^4 \rangle}{\langle (u(\rho) - u(0))^2 \rangle^2}\ ;$$

see the curves in Figure 9.

From our exploration of the multiscale random field with spectral density (5.1), we have found that the Randomization Method can be made much more efficient by using more general stratified sampling schemes other than subdivision into sampling bins of equal energy. We have attempted a logarithmic subdivision strategy because of its natural association with self-similar fractal random fields, using essentially an equal level of resolution at each length scale within the range of the simulation. The Fourier-wavelet method

(indeed any wavelet method) employs a similar representation. We do not claim that the logarithmic subdivision strategy is optimal, only that it appears to greatly improve the efficiency of the Randomization Method relative to an equal-energy subdivision. Nor do we take the relation (5.5) too seriously by, for example, optimizing it with respect to the numerical parameters. This would lead to silly strategies because the formula (5.5) does not take into account the need to adequately sample wavenumbers *throughout* the range of scales from $\ell_{\min}$ to $\ell_{\max}$. Our rough theoretical considerations are only meant to suggest what to expect with a reasonable choice of parameters to ensure a decent level of accuracy. The main point is that the estimate (5.5), along with the numerical results in Figures 6–9 suggests that the Randomization Method with a logarithmic subdivision strategy should be able to simulate a multiscale random field with the number of computational elements growing linearly with the number of decades of random field structure simulated, at least insofar as producing an accurate simulated structure function. We expect this cost scaling to apply to more general multiscale random fields as well.

That is, to simulate a random field for the purposes of evaluation at points on demand, we expect that a Randomization Method with logarithmic subdivision strategy could be adequate with some fixed reasonable bin ratio (such as $q = 2$), some fixed reasonable number of wavenumbers sampled per bin (such as $n_0 \sim 4 - 10$), and the number of sampling bins n chosen in proportion to the number of decades of random field structure to be simulated (so proportional to $\log_{10}(\ell_c/\ell_s)$ if the full structure of the random field is to be represented). We emphasize that these choices of parameters are expected to be adequate only insofar as accurate simulation of two-point statistics (such as the structure function) evaluated at arbitrary points suffices for the application. In Sections 6 and 7, we examine how well the Randomization Method is able to recover multi-point statistical properties.

Multidimensional Simulation The above discussion focused on random fields defined over one dimension. We consider briefly how the choice of parameters can be expected to change in higher dimensions. For concreteness, we will consider the isotropic case (for which the Randomization formulas are presented in Appendices B), though similar conclusions can be expected to hold for the anisotropic case as well (Appendix A), particularly if the subdivision of wavenumber space is arranged into radially symmetric shells. We expect that a logarithmic subdivision strategy along the radial direction to again yield an efficient simulation (at least for the purpose of simulating statistics such as the structure function), so that the number of bins n (radial shells) should scale with $\log(\ell_{\max}/\ell_{\min})$. It is not so clear, on the other hand, how the number of wavenumbers n_0 per bin needed for an accurate simulation should depend on the length scales of the random field to be simulated. The answer likely depends on the type of statistics in which one is interested. If it suffices to simulate the second order structure function accurately and for the second order correlation function to appear approximately isotropic, then a fixed number n_0 wavenumbers per bin, independent of the length scales of the random field (but presumably depending on the number of dimensions), is likely to be adequate. There may be more complex statistics involving correlations of the random field along different directions which may require n_0 to be chosen to increase with ℓ_c/ℓ_s. However, since we will not examine multi-dimensional random field statistics in much detail, we will not dwell much on this point but rather think of n_0 as needing to depend somewhat

on dimension but not on the length scales of the random field.

Preprocessing Cost The Randomization Method has a preprocessing cost proportional to the number of stratified sampling bins n.

For subdivision strategies which are determined without detailed computation involving the energy spectrum (such as variant C), one needs to prepare a transform or rejection method in each bin to convert a standard uniform random number to the correct probability distribution of wavenumbers within each samping bin. For bins set by equal energy distribution (variant B), one must also compute where the bin divisions lie. We will not concern ourselves with quantifying this additional cost because the equal energy distribution strategy does not seem to have an advantage (compared to, say, variant C) justifying the extra computation.

Cost per Realization A new random field is simulated by choosing n_0 wavenumbers randomly within each of the n bins, and then generating a Gaussian random amplitude for each of these wavenumbers. The cost is proportional to a small multiple of $n_0 n$.

Cost per Evaluation The evaluation of the random field at a point specified on demand is done by straightforward evaluation of the Fourier series approximation (3.2), with a cost proportional to nn_0, the number of terms in the sum.

Summary of Cost Considerations The preprocessing cost is proportional to the number of sampling bins n, while the costs per realization and evaluation are proportional to the total number of wavenumbers sampled, $n_0 n$. Each of these costs are expected to scale linearly with the number of decades of random field structure to be simulated, at least if accuracy of the simulated second-order statistics is all that is required.

5.1.3 Fourier-wavelet Method

Choice of Numerical Parameters As with the Randomization Method, one must choose the maximal and minimal length scales, $\ell_{\max}$ and $\ell_{\min}$, to be resolved by the random field simulation. The maximal length scale $\ell_{\max}$ is generally taken to be comparable to the correlation length of the random field; in our example (5.1), we choose $\ell_{\max} = k_0 = 1$. The ratio between the minimal and maximal length scales is set by the choice of the number of scales M in the truncated random field representation (4.6); namely, $(\ell_{\max}/\ell_{\min}) = 2^{M-1}$. The number of decades which one is attempting to capture is

$$N_{\text{dec}} = \log_{10}(\ell_{\max}/\ell_{\min}) = (M-1)\log_{10} 2.$$

Good statistical quality will generally be somewhat less than this ideal figure, but we should expect the number of decades for which the random field will be accurately simulated to scale linearly with $M - 1$.

One must additionally choose the truncation parameter b to be large enough that the functions $f_m(\xi)$ derived from the wavelets can be considered negligible for $|\xi| > b$. Generally speaking, $f_m(\xi)$ decays algebraically, with power law $|\xi|^p$ if the Meyer mother wavelet

is built out of a pth order perfect B-spline [6]. The values of p and b primarily affect the relative error of the statistics of the simulated random field (arising from the truncation of the sum over translates in (4.6)), and in general need not be adjusted when simulating random fields with various length scales, so long as the relative accuracy required remains fixed. Following [6], we choose $p = 2$ and $b = 10$.

Finally, we must choose a finite spacing $\Delta\xi$ between the points $\xi = \xi_j = -b + (j - 1)\Delta\xi, j = 1, \ldots, 2b/\Delta\xi + 1$ at which the functions $f_m(\xi)$ are numerically evaluated through their Fourier integrals (4.2). We will assume that $1/\Delta\xi$ is an integer. The choice of $\Delta\xi$ determines how accurately the functions f_m are approximated through interpolation from the computed values throughout the interval $|\xi| \leq b$ over which they may need to be evaluated in the representation (4.6). Like b and p, the numerical value $\Delta\xi$ is determined by the amount of bias due to numerical discretization which is tolerable in the statistics, and is insensitive to the length scales characterizing the random field to be simulated, since the functions $f_m(\xi)$ are each single-scale functions. The value $\Delta\xi = 0.01$ was used in our calculations.

An example of the correlation function and normalized structure function $G_2(\rho)$ for the energy spectrum (5.1) as simulated by the Fourier-wavelet method with $M = 40$ scales and $b = 10$ is shown in Figure 10. For multi-dimensional simulations, one must choose how many one-dimensional random fields N_{a} to use in the plane wave superposition (see Appendix B). This is determined both by the angular resolution desired and the number of plane waves required per angular direction. [9] show that a fixed number N_{a} of plane waves (depending on dimension but not on the length scales of the random field) is adequate to ensure a desired approximation to isotropy of the simulated random field. As discussed in Subsubsection 5.1.2, there may be more complex multi-dimensional statistics that require N_{a} to increase with ℓ_c/ℓ_s, but we will not investigate this possibility in the present work. We will rather think of N_{a} as independent of the length scales of the random field, as should be adequate at least for statistics involving a small number of evaluation points.

Preprocessing Cost Once the numerical parameters have been chosen, the functions f_m used to represent the random field on various length scales each need to be computed through evalulation of the Fourier transforms (4.2). Some details of how these values can be calculated through a fast Fourier transform are given in Appendix C. The cost of each integration is $b/\Delta\xi \log_2(b/\Delta\xi)$ and M functions f_m need to be computed. Since these numerical integrations dominate the preprocessing cost, we can estimate it as $Mb/\Delta\xi \log_2(b/\Delta\xi)$. The extra preprocessing cost in the extension to multiple dimensions through plane wave superposition is negligible because the same functions are involved.

Cost per Realization We first consider the case in which the points at which the random field are to be evaluated, though not specified in advance, are known to all lie within a one-dimensional domain with length scale L. Then, one has two options with the Fourier-wavelet method:

- Simulate at the beginning of the calculation the random field over the whole domain, then simply evaluate the random field at the desired locations.

- Simulate the random field only as needed to evaluate its values at the desired points of interest.

The execution and accounting is simpler for the first approach, which we now discuss. Later we will remark on how one may be able to reduce computational cost through the second approach, particularly if the number of points at which the random field is to be evaluated are sparsely distributed over the simulation domain.

The number of scales M in the Fourier-wavelet representation is specified through consideration of relevant length scales in the random field as described in Section 2.1. We now fix an index $0 \leq m \leq M$ (which fixes a length scale $\ell 2^{-m}$), and note that we must simulate and store γ_{mj} for $2^m L/\ell + 2b$ different indices j in (4.1), so that the sum (4.6) can be accurately evaluated at any value of x on demand in the one-dimensional domain of length L. This follows from counting the number of integers j such that $|\bar{n}_m(x) - j| \leq b$ for some x within an interval of length L. Consequently, the cost to simulate a realization of a one-dimensional random field comprehensively (to the specified level of accuracy) over a domain of length L is

$$\sum_{m=0}^{M} (2^m L/\ell + 2b) = 2^{M+1} L/\ell + 2bM$$

For a random field to be simulated over a multi-dimensional domain with length scale L, this cost is multiplied by the number N_{a} of one-dimensional random fields used in the Radon plane wave decomposition.

One can cut back on this cost if the points at which the random field is to be evaluated are rather sparsely distributed over the domain in which they are known to lie. In this case, the random field is only simulated as needed for evaluation [10, 9, 8], and the cost for simulating the random field is essentially proportional to the cost of evaluation (discussed next). In this approach, however, one must be careful with managing the random numbers γ_{mj} so that the same values are used when the same indices are referred to in random field evaluations at different locations x. One can either store all random numbers that have been generated and develop an efficient data handling routine to check whether a random variable γ_{mj} appearing in an evaluation needs to be generated or recalled from a previous generation. Alternatively, one can explicitly use the structure of a reversible pseudo-random number generator to simulate all random variables as needed, maintaining the idenities of random variables already realized, without actually storing them [8]. In short, one may be able to save on computation time by only simulating the random field as needed, but one must adopt a more sophisticated code to handle the random numbers γ_{mj}.

Finally, if the points at which the random field are to be evaluated cannot be constrained to lie in a pre-specified domain (as may be the case in turbulent diffusion simulations on an unbounded domain), then one must adopt something like the second strategy described above in which the random field must be simulated as needed.

Cost per Evaluation The evaluation of the random field at a given point involves the calculation of a sum of the form (4.6). This involves interpolation to evaluate the functions f_m at the indicated values, and possibly the generation or recollection of the

random variables γ_{mj} if they are only being simulated as needed. In any event, the cost is generally proportional to the number of terms in the sum, bM.

Summary of Cost Considerations The preprocessing cost is proportional to the quantity $Mb/\Delta\xi \log_2(b/\Delta\xi)$, the cost per evaluation is proportional to $2Mb$, while the cost per realization may be proportional to $2^{M+1}L/\ell + 2bM$, but possibly simply proportional to MbN_e, where N_e is the number of evaluations of the random field, for the case of sparse sampling of the random field and a sufficiently sophisticated algorithm for handling the random numbers. The preprocessing cost appears negligible relative to the other costs, and the total cost for simulating one realization of the random field at N_e locations should scale somewhere between MbN_e (for sparse sampling of points on demand) and $2^{M+1}L/\ell + 2bM + MbN_e$ (for dense sampling of points on demand which lie within a fixed bounded domain).

5.1.4 Comparison of Costs

For both algorithms, the total simulation cost is proportional to the number of realizations, so we discuss the cost of simulating one realization of the random field at a certain number N_e of points specified on demand. We only consider the most competitive Variant C, with logarithmically uniform subdivision, of the Randomization Method.

For this Randomization Method, the computational cost per realization is rather simply found to be proportional to the number of decades resolved in the random field and the number of points to be evaluated. The prefactor in the cost is not very large (say, order 10), being simply determined by the number of wavenumbers that should be simulated per decade to provide sufficient statistical accuracy. The cost of simulating a multiscale Gaussian random field with the Fourier-wavelet method appears to be usually greater. In the case of sparse sampling, the cost scales nominally with MbN_eN_{a}. This may be viewed as ostensibly comparable to the cost scaling in the Randomization Method, but one must recall that to achieve such cost scaling in the Fourier-wavelet method for $N_e > 1$, the code must involve a somewhat sophisticated handling of the random numbers γ_{mj} in the expansion (4.1), thereby increasing the amount of work per calculation.

If one wishes to avoid the need for a delicate management of random variables in the Fourier-wavelet method, and one can pre-specify a bounded domain in which the points to be evaluated must lie, then one can simulate the random field over the whole domain, before evaluation, in which case the cost will generally scale as $N_{\mathrm{a}}\, 2^{M+1}L/\ell + 2bMN_{\mathrm{a}} + MbN_eN_{\mathrm{a}}$. The first term has the potential for growing quite large for random fields with many scales, and has no counterpart in variant C of the Randomization Method. We can understand the difference in scaling by noting that the random field structure on length scale $2^{-m}\ell$ is represented in terms of n_0 random variables (the number of wavenumbers sampled at this length scale) in the Randomization Method, but by $2^m L/\ell + 2b$ random variables (γ_{mj}) in the Fourier-wavelet Method. The Fourier-wavelet Method is therefore using many more independent random numbers to represent the smaller scales of the random field than the Randomization Method.

Both the Randomization and Fourier-wavelet Methods are much less expensive than the standard simulation approach described in Subsubsection 5.1.1 when a multiscale random

field (with $\ell_c/\ell_{\min} \gg 1$ and $L/\ell_{\min} \gg 1$) is to be evaluated at a large number N_e of points. Indeed, the cost of the Randomization Method scales linearly in N_e, logarithmically with respect to $\ell_c/\ell_{\min}$, and is independent of $L/\ell_{\min}$. The Fourier-wavelet method has similar cost scaling with careful random variable management, but even with the simpler approach of simulation on a grid followed by interpolation, the cost of the Fourier-wavelet method scales as $N_{\mathrm{a}}L/\ell_{\min} + N_{\mathrm{a}}b\ln(\ell_c/\ell_{\min}) + \ln(\ell_c/\ell_{\min})bN_eN_{\mathrm{a}}$, which scales logarithmically with respect to $\ell_c/\ell_{\min}$ and linearly (and additively) with respect to $L/\ell_{\min}$ and N_e. The standard simulation method described in Subsubsection 5.1.1, by contrast, has cost scaling superlinearly with respect to these parameters.

We conclude that to simulate a given random field, the Fourier-wavelet Method is generally more expensive than the Randomization Method with logarithmically uniform subdivision. In particular, the computer time needed to achieve a 1%-accuracy of the simulated structure function over 9 decades (see Figure 6) was approximately 12 times less with the Randomization Method ($n = 40$, $n_0 = 4$, $q = 2$) than for the Fourier-wavelet method ($M = 40$, $b = 10$).

We will see through numerical examples in the next few sections that the Randomization Method can also simulate multi-point statistical characteristics of the random field more quickly than the Fourier-wavelet method if the number of points involved is not too high (e.g., less than 100). The Fourier-wavelet method, however has better ergodic properties and appears to simulate statistical characteristics involving many points accurately, without the need to increase the cost of the simulation beyond that associated with the parameters used to simulate the structure function accurately (Figure 6).

We remark that the other variants (A and B) of the Randomization Method would have costs growing faster than that of the Fourier-wavelet method for multiscale random fields with many decades; it is crucial to use the logarithmically uniform subdivision to render the Randomization Method competitive for such applications.

5.2 Random Field Simulations on Pre-Specified Domain

We now discuss the relative costs of the Randomization and Fourier-wavelet Methods when the random field is to be simulated on a uniform grid with spacing h and domain length L. We only point out the differences from the above discussion concerning simulation of the random field at points which can be specified on demand.

5.2.1 Standard Simulation Method

As discussed in Subsubsection 5.1.1, one can use Cholesky decomposition to simulate a Gaussian random field on a pre-specified set of points. The number of points at which the random field is to be evaluated scales with $(L/h)^d$, so the preprocessing cost would scales as $(L/h)^d(\ell_c/h)^{2d}$, and the cost to simulate one realization of the random field over the grid scales as $(L/h)^d(\ell_c/h)^d$.

5.2.2 Randomization Method

The cost considerations for the Randomization Method are essentially unchanged. The preprocessing cost scales with the number of sampling bins n_0, and then each realization of the random field over the $(L/h)^d$ lattice points requires a number of computations proportional to $nn_0(L/h)^d$. We recall for variant C, with logarithmically uniform stratified sampling, the number of decades simulated is proportional to n. There does not appear to be any cost savings available from the evaluation points falling on a regular lattice; the Fast Fourier Transform is not available due to the irregular spacing of the wavenumbers used in the Randomization Method.

5.2.3 Fourier-wavelet Method

To first approximation, we can relate the cost for simulating a random field over a regular lattice by simply viewing it as a special case of making $N_e = (L/h)^d$ evaluations of the random field over a domain with length scale L. In this case, it is clearly more efficient to directly simulate all the random variables γ_{mj} needed, so long as they can be stored. We therefore estimate a preprocessing cost proportional to $Mb/\Delta\xi \log_2(b/\Delta\xi)$, and a cost of simulating each realization of the random field over the prescribed lattice proportional to $N_{\mathrm{a}}2^{M+1}L/\ell + 2bMN_{\mathrm{a}}(L/h)^d$.It is natural to take the sampling distance h comparable to the smallest length scale $\ell_{\min} = \ell 2^{1-M}$ resolved in the Fourier-wavelet representation, in which case we can re-express the cost per realization of the random field on the lattice as proportional to $(L/h)N_{\mathrm{a}} + bMN_{\mathrm{a}}(L/h)^d \sim bMN_{\mathrm{a}}(L/h)^d$.

5.2.4 Comparison of Costs

With the dense sampling of the random field implicit in simulating the random field over a lattice, both the Randomization Method and Fourier-wavelet method scale similarly with respect to the length scales involved. Namely, they are both proportional to the number of lattice points $(L/h)^d$, the number of decades simulated (which is logarithmic in (ℓ_c/h)), and some numerical implementation parameters. The Fourier-wavelet parameter N_{a} is probably the largest in multiple dimensions; in [9], it was chosen as $N_{\mathrm{a}} = 32$ in two dimensions. For one-dimensional random fields, the simulation costs of the two methods appear comparable if the lattice is very dense, with $h = \ell 2^{1-M}$. Both the Randomization and Fourier-wavelet Methods are again much more efficient than the standard simulation approach based on Cholesky decomposition of the covariance matrix associated to the pre-specified grid for multiscale applications ($(L/h)^d \gg 1$ and $(\ell_c/h)^d \gg 1$), since the latter has cost scaling as a superlinear power law with respect to these large parameters.

6 Ergodic Properties of Simulated Random Fields

An important feature of numerically simulated statistically homogenous random fields is the quality of their ergodicity, by which is meant the convergence of spatial averages of quantities to their theoretical averages taken over a statistical ensemble. Ergodicity is a particularly useful feature of a simulated random field when each realization is expensive

to compute, because statistics can be extracted by processing spatial averages of one or a small number of realizations instead of by averaging over a large set of realizations. Simulation of a porous medium flow through the Darcy equation (1.2) for a given realization of the conductivity, for example, is a rather time-consuming computational procedure.

As a basic example, we consider the second order correlation function and structure function of the random field, which have, respectively, the theoretical ensemble-averaged definitions:

$$\begin{aligned} B(\rho) &= \langle u(\rho + x)u(x)\rangle, \\ D(\rho) &= \langle (u(\rho + x) - u(x))^2\rangle. \end{aligned} \tag{6.1}$$

Neither depend on x due to statistical homogeneity. Rather than considering the quality of ensemble averages (as we did in Section 5), we now study how well corresponding spatial averages of a *single realization* of the simulated random field:

$$\begin{aligned} B_{N_b}(\rho) &\equiv \frac{1}{N_b}\sum_{j=1}^{N_b} u(\rho + (j-1)\ell_b)\, u((j-1)\ell_b)\ , \\ D_{N_b}(\rho_i) &\equiv \frac{1}{N_b}\sum_{j=1}^{N_b} [u(\rho_i + (j-1)L) - u((j-1)L)]^2\ , \quad \rho_i = \Delta\rho(i-1), \quad i = 1, \ldots, n_\rho\ . \end{aligned} \tag{6.2}$$

converge to the ensemble-averaged expressions $B(\rho)$ and $D(\rho)$ as the number of spatial samples N_b is taken large. Here ℓ_b is a length scale describing the spatial translation between each spatial sample. To study the ergodicity properties of the simulated random fields, we compare simulated spatial averages for the correlation function and the normalized structure function $G_2(\rho) = D(\rho)/(J_\alpha \rho^{\alpha-1})$ against the exact results for the energy spectrum (5.1).

6.1 Randomization Method

In Figure 11, these comparisons are made for the Randomization Method. Note that even for $N_b = 16000$ spatial samples, $n = 25$ bins, and $n_0 = 10$ wavenumbers per bin, the agreement is not satisfactory. Increasing n_0, the number of wavenumbers per bin, improves the results (Figures 12 and 13).

Thus we see that the number of wavenumbers per bin must be drastically increased for the Randomization Method to exhibit good ergodic properties. This phenomenon actually arises also for single-scale random fields. To show this, we simulate a random field with the exponential correlation function

$$B(\rho) = C\exp(-\mu|\rho|), \quad E(k) = \frac{2C\mu}{\mu^2 + 4\pi^2 k^2}\ , \tag{6.3}$$

with $\mu = 1$, $C = 2\pi^2$. This random field has correlation length $\ell_c = 1/\mu = 1$, and we only consider spatial scales $10^{-3} \le \rho$, over which the random field cannot be considered fractal.

In Figure 14, we present spatially averaged statistics from a Randomization Method simulation of the random field with exponential correlation function (6.3). Here the bins are constructed by subdividing the wavenumber range $0 < k < 10^3$ into $n = 5$ logarithmically uniform subintervals. $n_0 = 20$ random wavenumbers are sampled in each bin, and even averages over $N_b = 16000$ spatial blocks exhibit large deviations from the true statistics. Increasing the number of wavenumbers per bin to $n_0 = 200$ improves the results (Figure 15).

We see that achieving good ergodic properties in the random fields simulated by the Randomization Method requires a substantial increase in the number of wavenumbers sampled per bin and therefore the expense of the simulation.

6.2 Fourier-wavelet Method

In Figure 16, the correlation function and the normalized structure function as estimated by spatial averages of random fields simulated by a single realization of the Fourier-wavelet method are compared against the exact result for the Kolmogorov spectrum (5.1). With only $N_b = 2000$ spatial blocks, the accuracy achieved was approximately the same as in the Randomization method with $N_b = 16000$ spatial blocks, $n = 25$, and $n_0 = 250$; compare Figures 13 and 16. The cost of the Fourier-wavelet method in this case is about ten times less than that of the Randomization method.

6.3 Comparison

We see that in order for computed spatial averages to approximate the desired correlation function or structure function, the Randomization Method requires a drastic increase in the number of wavenumbers per bin, n_0, as compared to the values of n_0 adequate for ensemble average calculations. This can be understood by noting that the statistical quality of ergodic averages over a large spatial domain is related to the number of effectively independent samples in the collection of spatial observations. The Randomization method with logarithmically uniform stratified sampling uses a relatively small number of independent random numbers to generate the random field, and spatial averages will fail to improve once they already involve a number of effectively independent samples comparable to the number of independent random numbers $3nn_0$ used in the construction of the random field. The Fourier-wavelet Method, by contrast, involves a sufficiently rich collection of random variables so that spatial averages exhibit good ergodic properties without the need to increase the expense of the simulation beyond that necessary for ensemble averages to approximate the second order correlation function and structure function adequately.

7 Multi-point statistical characteristics of simulated random fields

We now return to consideration of the quality of ensemble averages in multiscale random field simulations, but now examine the quality of statistics involving more than two points. One aim is to examine whether the non-Gaussianity of the Randomization Method may exhibit itself in a more pronounced manner in multi-point statistics as compared to two-point statistics (such as the second order correlation fucntion and structure function). Another objective is to determine whether the cost of the Randomization Method must be increased (as it was for spatial averages) in order to simulate multi-point statistics with comparable quality to that simulated by the theoretically Gaussian Fourier-wavelet method. We remark that the consideration of spatial averages (6.2) in Section 6 yields a statistic that involves more than two points, but we only considered single realizations rather than ensemble averages of these random variables.

We therefore consider the following sequence of normalized random increments

$$\Delta u_i = \frac{u(ih) - u((i-1)h)}{[J_\alpha h^{\alpha-1}]^{1/2}}, \quad \delta u_i = \frac{u(h\,q^{i-1}) - u(0)}{[J_\alpha (hq^{i-1})^{\alpha-1}]^{1/2}}, \quad i = 1, \ldots, n' \,. \tag{7.1}$$

We introduce the random variable

$$\zeta_{n'} = \max_{1 \le i \le n'} |\Delta u_i| \,,$$

and denote by $p_{\zeta,n'}$ the probability density function of $\zeta_{n'}$. All numerical simulations in this section refer to the random field with Kolmogorov spectrum (5.1).

We first simulate the histogram for $p_{\zeta,n'}$ using variant C of the Randomization method, and compare with direct Monte Carlo simulation. The procedure of direct Monte Carlo simulation is analogous to that described in Subsubsection 5.1.2 for the case of the normalized structure function, and its results differs from the exact results only through sampling error.

From Figure 17 it is seen that 25 wavenumbers ($n = 25$, $n_0 = 1$) do not give satisfactory accuracy (right panel) for the statistics of ζ_{10} when $h = 1000\,\ell_{\min}$, while 250 wavenumbers ($n = 25$, $n_0 = 10$) show good agreement.

Other details can be extracted from the statistical moments

$$\mu_m(k) = \langle \Big(\max_{1 \le i \le k} |\Delta u_i| \Big)^m \rangle, \quad k = 1, \ldots, n' \,. \tag{7.2}$$

In Figure 18, we show the first four moments simulated by variant C of the Randomization method for 10 values of k, with the same parameter choices as in Figure 17. It is seen from the right panel that 25 wavenumbers is too small to describe satisfactorily the moments $\mu_3(k)$, $\mu_4(k)$ for $k \gtrsim 4$.

In Figures 19 and 20, the statistics of ζ_{10} are seen to deteriorate as $h/\ell_{\min}$ is reduced. This can be understood from the observation that wavenumbers with $k \sim 1/h$ are the most influential on the increments Δu_i, and that as $h/\ell_{\min}$ is reduced, then number of

simulated wavenumbers with $k \sim 1/h$ becomes smaller and variability of the computed statistics increases. The multi-point statistic ζ_{10} appears to be adequately simulated by the Randomization Method for $h \gg 10\ell_{\min}$. Similar conclusions can be drawn from consideration of maximal statistics of collections of the normalized increments δu_i with geometric spacing (Figure 21).

We expect that the number of wavenumbers n_0 in each bin must be increased if we desire statistical characteristics involving a larger number of points. We illustrate this in Figure 22 for the moments $\mu_m(k)$ of the random variable ζ_{100}. It is seen that $n_0 = 10$ wavenumbers per bin is not sufficient for the Randomization Method to yield an accurate evaluation of $\mu_4(k)$ for $k \geq 50$ (left panel), but for $n_0 = 40$ wavenumbers per bin, the first four moments are each simulated with decent accuracy.

The Fourier-wavelet method generates theoretically Gaussian multipoint statistics, and therefore its accuracy in simulating multipoint statistics is completely determined by the number of Monte Carlo samples and the accuracy in simulating the correct second order correlation function . Of course the accurate evaluation of more complex statistics will generally require better accuracy of the correlation function.

8 Conclusions

1. The Randomization Method is generally easier to implement than the Fourier-wavelet method. For simulating random fields with accurate second order statistics, the Randomization method with logarithmically uniform spectral subdivision can often be considerably less expensive than that of the Fourier-wavelet method. The scenario with greatest relative advantage for the Randomization Method appears to be the sampling of a multiscale random field over a large but sparse set of points to be specified on demand. Then, as is generally case, the cost of the Randomization Method scales linearly with the number N_e of evaluation points and logarithmically in the range of scales $\ell_{\max}/\ell_{\min}$ to be represented in the random field. A relatively simple implementation of the Fourier-wavelet method (simulation on a grid followed by interpolation as needed) has cost scaling linearly (and additively) in both N_e and $\ell_{\max}/\ell_{\min}$. A more sophisticated management of random numbers in the code can reduce the cost scaling of the Fourier-wavelet method in this scenario to the same as that of the Randomization Method, but with a typically larger prefactor. Both the Randomization Method and the simple implementation of the Fourier-wavelet method have comparable scaling when the random field is to be simulated over a pre-specified mesh of points.

2. The cost of the Randomization Method increases substantially if statistics involving large numbers of points are to be simulated accurately.

3. In particular, when statistics are to be evaluated through spatial averages (and an appeal to ergodicity) rather than ensemble averages, the Fourier-wavelet method appears more efficient than the Randomization Method. Good ergodic properties are important in applications which involve the solution of partial differential equations with random coefficients, such as the Darcy equation with random hydraulic conductivity.

A Appendix: Randomization method for homogeneous vector random fields

Here we briefly present the Randomization technique for simulation of a Gaussian homogeneous vector random field, and give the conditions for the weak convergence of the method.

Let $\mathbf{u}(\mathbf{x}) = [u_1(\mathbf{x}), \ldots, u_l(\mathbf{x})]^T$, $\mathbf{x} \in I\!R^d$ be a Gaussian homogeneous vector-valued random field with prescribed correlation tensor $\mathsf{B}(\mathbf{r}) = \langle \mathbf{u}(\mathbf{x}+\mathbf{r}) \otimes \mathbf{u}(\mathbf{x}) \rangle$, or prescribed spectral density tensor $\mathsf{F}(\mathbf{k})$; the two are related through Fourier transforms:

$$\mathsf{B}(r) = \int_{R^d} \exp\{2\pi\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}\}\mathsf{F}(\mathbf{k})\,\mathrm{d}\mathbf{k}\ , \quad \mathsf{F}(\mathbf{k}) = \int_{R^d} \exp\{-2\pi\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}\}\mathsf{B}(\mathbf{r})\,\mathrm{d}\mathbf{r}\ . \tag{A.1}$$

The symbol $()^T$ denotes the transpose operation and $\otimes$ denotes a tensor (outer) product.

The spectral density function is defined

$$E(\mathbf{k}) = \sum_{j=1}^{l} \mathsf{F}_{jj}(\mathbf{k}),$$

and we assume that $\sigma^2 \equiv \int_{R^d} E(\mathbf{k})\,\mathrm{d}\mathbf{k} < \infty$.

We will use the Cholesky decomposition to factor the nonnegative definite tensorial function $\mathsf{F}(\mathbf{k})$ into square roots:

$$\mathsf{F}(\mathbf{k}) = E(\mathbf{k})\,\mathsf{Q}(\mathbf{k})\,\mathsf{Q}^*(\mathbf{k})\ , \tag{A.2}$$

where the matrix Q^* is defined as the Hermitian conjugate $Q^* = \overline{Q^T}$.

The spectral density function $E(\mathbf{k}) = \mathrm{Tr}\,\mathsf{F}(\mathbf{k})$ is a scalar, nonnegative function describing the overall strength of the random field fluctuations at wavenumber $\mathbf{k}$, while $\mathsf{Q}(\mathbf{k})$ is a matrix-valued function describing the anisotropy and correlation structure among and along different directions. Further details can be found in [29, 41].' In the purely scalar-valued case ($l = 1$), $\mathsf{F}(\mathbf{k}) = E(\mathbf{k})$ and $\mathsf{Q}(\mathbf{k}) = \mathsf{I}$, and all our main points can be understood in this simpler context.

We next decompose the tensor Q into real and imaginary parts: $\mathsf{Q}(\mathbf{k}) = \mathsf{Q}'(\mathbf{k}) + \mathrm{i}\,\mathsf{Q}''(\mathbf{k})$, and subdivide the support of the spectral density $\Delta = \mathrm{supp}\ (E)$ into a finite number of nonoverlapping sets: $\Delta = \cup_{i=1}^{n} \Delta_i$.

For each $1 \le i \le n$, we choose a collection of independent, identically distributed random wavenumbers $\mathbf{k}_{i1}, \ldots, \mathbf{k}_{in_0}$ within the set Δ_i according to the probability distribution function

$$p_i(\mathbf{k}) = \begin{cases} \frac{E(\mathbf{k})}{\sigma_i^2}, & \mathbf{k} \in \Delta_i, \\ 0, & \text{else}\ , \end{cases} \qquad \sigma_i^2 = \int_{\Delta_i} E(\mathbf{k})\,\mathrm{d}\mathbf{k}\ . \tag{A.3}$$

The random field as simulated by the Randomization method is finally expressed as

$$\mathbf{u}^{(\mathrm{R})}(\mathbf{x}) = \sum_{i=1}^{n} \frac{\sigma_i}{\sqrt{n_0}} \sum_{j=1}^{n_0} \Big\{ \boldsymbol{\xi}_{ij} \Big[Q'(\mathbf{k}_{ij}) \cos\,\theta_{ij} - Q''(\mathbf{k}_{ij}) \sin\,\theta_{ij} \Big] + \boldsymbol{\eta}_{ij} \Big[Q''(\mathbf{k}_{ij}) \cos\,\theta_{ij} + Q'(\mathbf{k}_{ij}) \sin\,\theta_{ij} \Big] \Big\} , \qquad \text{(A.4)}$$

where $\theta_{ij} = 2\pi\,\mathbf{k}_{ij} \cdot \mathbf{x}$, and $\boldsymbol{\xi}_{ij}, \boldsymbol{\eta}_{ij}$, $i = 1, \ldots, n$; $j = 1, \ldots, n_0$ are mutually independent standard Gaussian l-dimensional random row vectors (mean zero and unity covariance) which are moreover independent of the family of random wavenumbers $\mathbf{k}_{ij}$, $i = 1, \ldots, n$; $j = 1, \ldots, n_0$.

A.1 Weak Convergence of Randomization Method

The random field representation (A.4) used in the Randomization Method is an approximation to the desired Gaussian random field whose convergence is understood as follows.

We first define one notion of convergence of a sequence of scalar random fields $u_n(\mathbf{x})$ to a limiting random field $u(\mathbf{x})$, all functions being defined on a compact domain $K \subset I\!R^d$. We assume that the samples of u are continuous (i.e., belong to $C(K)$) with probability one. The weak convergence $u_n \to u$ as $n \to \infty$ in $C(K)$ is defined to mean that for any uniformly bounded continuous functional $f : C(K) \to I\!R$ we have $\langle f(u_n) \rangle \to \langle f(u) \rangle$ as $n \to \infty$. Note that from the weak convergence in $C(K)$, it follows that all the finite dimensional distributions of u_n converge to those of u [1]. A sequence of vector-valued random fields is said to converge weakly to a limit if each component converges weakly.

In [21], it is proved that the random field representation (A.4) used in the Randomization Method converges weeakly in $C(K)$ for a fixed spectral subdivision $n \geq 1$ as $n_0 \to \infty$, provided that

$$\int_{R^d} \log^{1+\varepsilon}(1 + |\mathbf{k}|)\, E(\mathbf{k})\, d\mathbf{k} < \infty \text{ for some } \varepsilon > 0. \qquad \text{(A.5)}$$

If we fix instead n_0 and let the spectral subdivision become finer ($n \to \infty$), then the convergence conditions are more complicated. Let

$$\underline{\rho}_i = \inf\{|\mathbf{k}|, \mathbf{k} \in \Delta_i\}, \quad \overline{\rho}_i = \sup\{|\mathbf{k}|, \mathbf{k} \in \Delta_i\} ,$$

and assume that $\Delta_n = \{\mathbf{k} \in I\!R^d : |\mathbf{k}| \geq r_n\}$ where r_n is a sequence of real positive numbers such that $\lim_{n\to\infty} r_n = \infty$.

We assume that positive constants C_0, R_0 and $\varepsilon_0 \in (0, 1)$ can be chosen so that either $\overline{\rho}_i \leq R_0$, or $R_0 \leq \overline{\rho}_i \leq C_0(\underline{\rho}_i)^{1+\varepsilon_0}$ for all $i = 1, \ldots, n-1$. Under these assumptions, together with (A.5), the random field representation (A.4) converges weakly $\mathbf{u}^{(\mathrm{R})} \to \mathbf{u}$ in $C(K)$ for each fixed $n_0 \geq 1$ as $n \to \infty$ [20].

We mention also that other types of functional convergence in $C(K)$ can be established; convergence in probability and L_p-convergence was studied in [2].

A.2 Generalizations of Randomization Method

One way to generalize the Randomization Method is to sample the wavenumbers $\mathbf{k}_{ij}$ in each Δ_i from quite arbitrary probability density functions $p_i(\mathbf{k})$ satisfying the consistency condition: $p_i(\mathbf{k}) \neq 0$ if $E(\mathbf{k}) \geq 0$. Secondly, rather than choosing the same number n_0 of wavenumbers from each sampling bin, one can choose a different number N_i of random wavenumbers from each set Δ_i, $i = 1, \ldots, n$.

The random field representation for the Randomization method with these generalizations has the following form:

$$\mathbf{u}^{(\mathrm{R})} = \sum_{i=1}^{n} \frac{\sigma_i}{\sqrt{N_i}} \sum_{j=1}^{N_i} \Big(\frac{E(\mathbf{k}_{ij})}{p_i(\mathbf{k}_{ij})}\Big)^{1/2} \Big\{ \boldsymbol{\xi}_{ij} \Big[\mathsf{Q}'(\mathbf{k}_{ij}) \cos\, \theta_{ij} - \mathsf{Q}''(\mathbf{k}_{ij}) \sin\, \theta_{ij} \Big] + \boldsymbol{\eta}_{ij} \Big[\mathsf{Q}''(\mathbf{k}_{ij}) \cos\, \theta_{ij} + \mathsf{Q}'(\mathbf{k}_{ij}) \sin\, \theta_{ij} \Big] \Big\} \ . \tag{A.6}$$

B Appendix: Isotropic vector-valued random fields

. A homogeneous d-dimensional vector-valued random field $\mathbf{u}(x)$, $\mathbf{x} \in I\!R^d$ is called isotropic if the random field $U^T\, \mathbf{u}(U\mathbf{x})$ has the same finite-dimensional distributions as those of the random field $\mathbf{u}(\mathbf{x})$ for any rotation matrix $U \in SO(d)$ [9, 29, 41]. The spectral density tensor of an isotropic random field has the following general structure [29, 41]:

$$\mathsf{F}(\mathbf{k}) = \frac{1}{2A_d\, k^{d-1}} \left\{ E_1(k) \mathsf{P}^{(1)}(\mathbf{k}) + E_2(k) \mathsf{P}^{(2)}(\mathbf{k}) \right\} \tag{B.1}$$

where $k = |\mathbf{k}|$, A_d is the area of the unit sphere in $I\!R^d$, E_1 and E_2 are the transverse and longitudinal radial spectra (scalar even nonnegative functions), and the projection tensors are defined componentwise as:

$$P_{ij}^{(1)}(\mathbf{k}) = \delta_{ij} - \frac{k_i k_j}{k^2}, \quad P_{ij}^{(2)}(\mathbf{k}) = \frac{k_i k_j}{k^2}, \quad i, j = 1, \ldots, d \ , \tag{B.2}$$

with δ_{ij} defined as the usual Kronecker delta symbol.

This representation of the random field can be used to simplify the implementation of the Randomization Method and has also been used to construct a multi-dimensional isotropic version of the Fourier-wavelet method. We describe each briefly in turn.

B.1 Randomization Method

The isotropic spectral representation (B.1) can be associated with the Helmholtz decomposition of the random field: $\mathbf{u}(\mathbf{x}) = \mathbf{u}^{(1)}(\mathbf{x}) + \mathbf{u}^{(2)}(\mathbf{x})$ where $\mathbf{u}^{(1)}$ and $\mathbf{u}^{(2)}$ are, respectively, the incompressible and potential parts of $\mathbf{u}$ with spectral density tensors

$$\mathsf{F}^{(1)}(\mathbf{k}) = \frac{1}{2A_d\, k^{d-1}} E_1(k) \mathsf{P}^{(1)}(\mathbf{k}) \ , \quad \mathsf{F}^{(2)}(\mathbf{k}) = \frac{1}{2A_d\, k^{d-1}} E_2(k)\, \mathsf{P}^{(2)}(\mathbf{k}) \ , \tag{B.3}$$

respectively.

Each of the random fields, $\mathbf{u}^{(1)}(\mathbf{x})$ and $\mathbf{u}^{(2)}(\mathbf{x})$, can be simulated as independent Gaussian random fields using the approach described in Appendix A. The Cholesky factorizations (A.2)

$$\mathsf{F}^{(i)}(\mathbf{k}) = p_i(\mathbf{k})\,\mathsf{Q}^{(i)}\,\mathsf{Q}^{(i)*}\,, \tag{B.4}$$

take the special form

$$p_1(\mathbf{k}) = \sum_{i=1}^{d} F_{ii}^{(1)}(\mathbf{k}) = \frac{(d-1)E_1(k)}{2A_d\,k^{d-1}}\,, \quad p_2(\mathbf{k}) = \sum_{i=1}^{d} F_{ii}^{(2)}(\mathbf{k}) = \frac{E_2(k)}{2A_d\,k^{d-1}}\,. \tag{B.5}$$

Note in particular that $p_i(\mathbf{k}) = p_i(k)$, which generally greatly simplifies the simulation of random wavenumbers according to the probability distributions p_i.

The matrices $\mathsf{Q}^{(1)}$ and $\mathsf{Q}^{(2)}$ are to be chosen in any way such that

$$\frac{1}{d-1}\,P^{(1)}(\mathbf{k}) = \mathsf{Q}^{(1)}(\mathbf{k})\mathsf{Q}^{(1)*}(\mathbf{k})\,, \quad P^{(2)}(\mathbf{k}) = \mathsf{Q}^{(2)}(\mathbf{k})\mathsf{Q}^{(2)*}(\mathbf{k})\,.$$

One convenient explicit choice in three dimensions is [35]

$$\mathsf{Q}^{(1)}(\mathbf{k}) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & \frac{k_3}{k} & -\frac{k_2}{k} \\ -\frac{k_3}{k} & 0 & \frac{k_1}{k} \\ \frac{k_2}{k} & -\frac{k_1}{k} & 0 \end{pmatrix}\,, \quad \mathsf{Q}^{(2)}(\mathbf{k}) = \begin{pmatrix} \frac{k_1}{k} & 0 & 0 \\ \frac{k_2}{k} & 0 & 0 \\ \frac{k_3}{k} & 0 & 0 \end{pmatrix}\,. \tag{B.6}$$

Because $p_i(\mathbf{k}) = p_i(k)$ in the isotropic case, it is natural to choose the spectral subdivision $\Delta = \sum_{i=1}^{n} \Delta_i$ to be radially symmetric: $\Delta_i = \{\mathbf{k} : a_i \le |\mathbf{k}| \le b_i\}$. Using the tensors (B.6), we obtain the following simulation formula for the incompressible part of an isotropic three-dimensional random vector field:

$$\mathbf{u}^{(1)}(\mathbf{x}) = \sum_{i=1}^{n} \frac{\sigma_i^{(1)}}{\sqrt{n_0}} \sum_{j=1}^{n_0} \left[\left(\mathbf{\Omega}_{ij}^{(1)} \times \boldsymbol{\xi}_{ij} \right) \cos(\theta_{ij}^{(1)}) + \left(\mathbf{\Omega}_{ij}^{(1)} \times \boldsymbol{\eta}_{ij} \right) \sin(\theta_{ij}^{(1)}) \right]$$

where $(\sigma_i^{(1)})^2 = \int_{\Delta_i} p_1(\mathbf{k})\,\mathrm{d}\mathbf{k} = \frac{1}{2}\int_{a_i}^{b_i} E_1(k)\,\mathrm{d}k$, $\mathbf{\Omega}_{ij}^{(1)}$, $i = 1,\ldots,n$; $j = 1,\ldots,n_0$ is a family of mutually independent random vectors distributed uniformly on the unit sphere in $I\!R^3$; $\boldsymbol{\xi}_{ij}$ and $\boldsymbol{\eta}_{ij}$, $i = 1,\ldots,n$; $j = 1,\ldots,n_0$ are mutually independent families of three-dimensional standard Gaussian random vectors; $\theta_{ij}^{(1)} = 2\pi k_{ij}^{(1)}\,(\mathbf{\Omega}_{ij}^{(1)} \cdot \mathbf{x})$; and for each $i = 1,\ldots,n$, the $k_{ij}^{(1)}$, $j = 1,\ldots,n_0$ is a sequence of independent random wavenumbers sampled from the interval (a_i, b_i) according to the probability density function proportional to $E_1(k)$.

The potental component $\mathbf{u}^{(2)}$ is simulated in three dimensions through the representation

$$\mathbf{u}^{(2)}(\mathbf{x}) = \sum_{i=1}^{n} \frac{\sigma_i^{(2)}}{\sqrt{n_0}} \sum_{j=1}^{n_0} \left[\xi_{ij}\,\mathbf{\Omega}_{ij}^{(2)}\,\cos(\theta_{ij}^{(2)}) + \eta_{ij}\,\mathbf{\Omega}_{ij}^{(2)}\,\sin(\theta_{ij}^{(2)}) \right]\,.$$

Here, unlike in the previous simulation formula, the ξ_{ij} and η_{ij}, $i = 1,\ldots,n$; $j = 1,\ldots,n_0$ are families of *scalar* standard Gaussian random variables, which are all mutually independent. The remaining inputs are constructed analogously: $(\sigma_i^{(2)})^2 = \int_{\Delta_i} p_2(\mathbf{k})\,d\mathbf{k} =$

$\frac{1}{2}\int_{a_i}^{b_i} E_2(k)\,dk$; $\Omega_{ij}^{(2)}$ is a family of mutually independent random vectors distributed uniformly on the unit sphere in $\mathbb{R}^3$; $\theta_{ij}^{(2)} = 2\pi k_{ij}^{(2)}\,(\Omega_{ij}^{(2)} \cdot \mathbf{x})$; and for each $i = 1, \ldots, n$, the $k_{ij}^{(2)}, j = 1, \ldots, n_0$ is a sequence of independent random wavenumbers sampled in the interval (a_i, b_i) from the probability density function which is proportional to the function $E_2(k)$.

B.2 Fourier-wavelet method

The Fourier-wavelet method can be extended to simulate multi-dimensional isotropic random fields by using the Helmholtz decomposition along with a further angular decomposition into plane waves varying only along one direction [9]. Discretization of the angular integral produces the following simulation formula:

$$\mathbf{u}_N(\mathbf{x}) = \Big(\frac{1}{N_\mathrm{a}}\Big)^{1/2} \sum_{j=1}^{N_\mathrm{a}} \Big[\mathsf{P}^{(1)}(\mathbf{\Omega}_j)\,\mathbf{v}_j^{(1)}(\mathbf{x}\cdot\mathbf{\Omega}_j) + \mathsf{P}^{(2)}(\mathbf{\Omega}_j)\,\mathbf{v}_j^{(2)}(\mathbf{x}\cdot\mathbf{\Omega}_j)\Big]\,,$$

where $\mathbf{\Omega}_j,\ j = 1, \ldots, N_\mathrm{a}$ is a collection of deterministic vectors which discretize the unit sphere in $\mathbb{R}^d$. The $\mathbf{v}_j^{(1)}(r)$, $\mathbf{v}_j^{(2)}(r)$, $j = 1, \ldots, N_\mathrm{a}$, $\quad r \in \mathbb{R}$ are mutually independent Gaussian d-dimensional random fields. Each component of $\mathbf{v}_j^{(1)}(r)$ (respectively $\mathbf{v}_j^{(2)}(r)$) is an independent Gaussian homogenous random field with spectral density $\frac{1}{4}E_1(k)$ (respectively $\frac{1}{4}E_2(k)$), and can be simulated using the one-dimensional Fourier-wavelet method presented in Section 4. Of course, in practice, one handles the projections by simulating only $d-1$ components of $\mathbf{v}_j^{(1)}$ and 1 component of $\mathbf{v}_j^{(2)}$ and rotating them according to the direction $\mathbf{\Omega}_j$ [9].

One might contemplate anisotropic versions of the Fourier-wavelet method using Helmholtz and more general plane wave decompositions, but we are not aware of any detailed elaboration of such a simulation scheme nor how difficult it would be to implement in practice.

C Appendix: Calculation of the functions f_m

Here we give some technical details on the calculation of the functions (4.2) which reads in our case

$$f_m(\xi) = \int_{-4/3}^{4/3} e^{-2\pi \mathrm{i}\, k\, \xi}\, g(k)\,dk \tag{C.1}$$

where $g(k) = 2^{m/2}\tilde{E}^{1/2}(2^m k)\hat{\phi}(k)$.

We calculate this function on the grid of points $\xi_j = -\frac{N}{2}\Delta\xi + (j-1)\Delta\xi$, $j = 1, \ldots, N$, where N is an even number, and $\Delta\xi \geq 0$ is the grid step. In order to evaluate the truncated sums appearing in the Fourier-wavelet representation (4.6), we must choose $N\Delta\xi/2 \geq b$.

We approximate the intergral (C.1) by a Riemann sum:

$$f_m(\xi_j) = \int_{-a}^{a} e^{-2\pi \mathrm{i}\, k\, \xi}\, g(k)\, dk \;\simeq \sum_{l=1}^{N} \Delta k\, e^{-2\pi \mathrm{i}\, k_l\, \xi_j}\, g(k_l) \tag{C.2}$$

where

$$k_l = -a + (l - 1/2)\Delta k, \quad l = 1, \ldots, N; \quad \Delta k = \frac{2a}{N}.$$

We use the same number of points $N = 2^r$ (where r is some positive integer) to discretize the integral as we use to represent $f_m(\xi)$ in physical space so that we can use the discrete fast Fourier transform. We also clearly need the cutoff on the integral in (C.2) to satisfy $a > 4/3$ (with $g(k)$ set to zero whenever evaluated for $|k| > 4/3$). Finally, the use of the fast Fourier transform requires the steps in physical and wavenumber space be related through $\Delta\xi\Delta k = 1/N$. Indeed, simple transformations then yield

$$\begin{aligned} \xi_j k_l &= \Big[-\frac{N}{2}\Delta\xi + (j-1)\Delta\xi \Big] [-a + (l - 1/2)\Delta k] \\ &= \frac{N-1}{4} - \frac{j-1}{2}\Big(1 - \frac{1}{N}\Big) - \frac{l-1}{2} + \frac{(j-1)(l-1)}{N}\,, \end{aligned} \tag{C.3}$$

hence

$$f_m(\xi_j) = \exp\Big\{\pi \mathrm{i}\,(j-1)\Big(1 - \frac{1}{N}\Big)\Big\} \sum_{l=1}^{N} G_l \exp\Big\{ -2\pi \mathrm{i}\, \frac{(j-1)(l-1)}{N} \Big\}\,, \tag{C.4}$$

where

$$G_l = \Delta k\, g(k_l) \exp\Big\{ -2\pi \mathrm{i} \Big[\frac{N-1}{4} - \frac{l-1}{2} \Big] \Big\}\,,$$

which is in the form of a discrete Fourier transform.

The constraints imposed on the discretization of the integral (C.2) to obtain an expression amenable to fast Fourier transform imply the following sequence of choosing parameters. First a bandwidth value b is chosen according to the desired accuracy in the Fourier-wavelet representation (4.6). Then a spatial resolution $\Delta\xi$ for the $f_m(\xi)$ is selected, either according to the grid spacing h on a prespecified set of evaluation points or such that $f_m(\xi)$ can be calculated accurately enough by interpolation from the computed values. (In any event, we must have $\Delta\xi < 3/8$). Next, a binary power $N = 2^r$ is chosen large enough so that $2b/N \le \Delta\xi$. Then we set $a = \frac{1}{2\Delta\xi}$, and discretize the integral (C.2) with step size $\Delta k = 2a/N = 1/(N\Delta\xi)$.

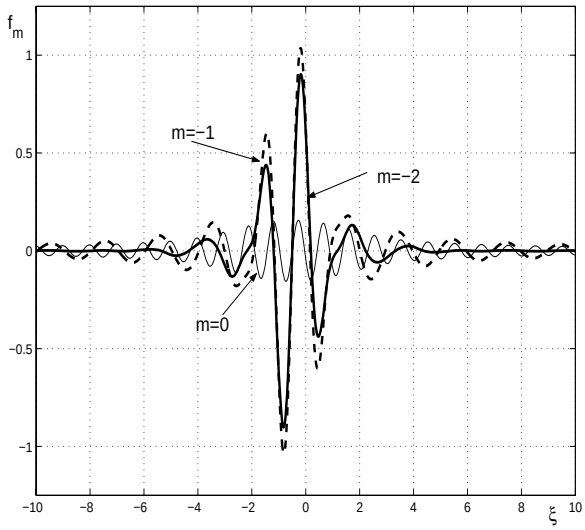

Figure 1: The functions $f_m(\xi)$, for $m = 0$ (solid line), $m = 1$ (dashed line), and $m = 2$ (the bold solid line) appearing in the Fourier-wavelet expansion computed from (4.2) for the energy spectrum (5.1) with $\alpha = 5/3$ and $k_0 = 1$. In fact, due to the self-similarity of the energy spectrum for $k \geq k_0$, the functions f_m for $m \geq 3$ are identical to f_2. For $m = 0$ and $m = 1$, the decay of f_m is somewhat slow because they feel the abrupt (nonsmooth) cutoff of the energy spectrum at $k = k_0 = 1$. But for $m \geq 2$, the function $f_m(\xi)$ is mainly supported within the interval $(-8, 8)$

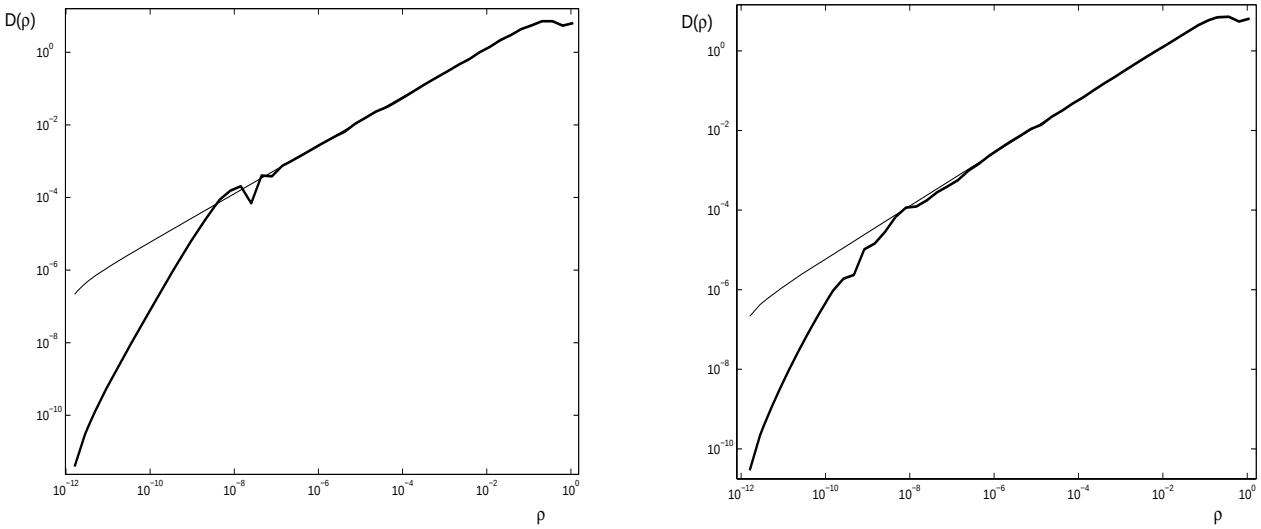

Figure 2: The structure function $D(\rho)$: exact formula (thin solid line) and calculated by averaging over an ensemble of $N_s = 2000$ Monte Carlo samples simulated by variant A of the Randomization method (bold solid line). Number of wavenumbers: $n_0 = 160$ (left panel) and $n_0 = 1000$ (right panel).

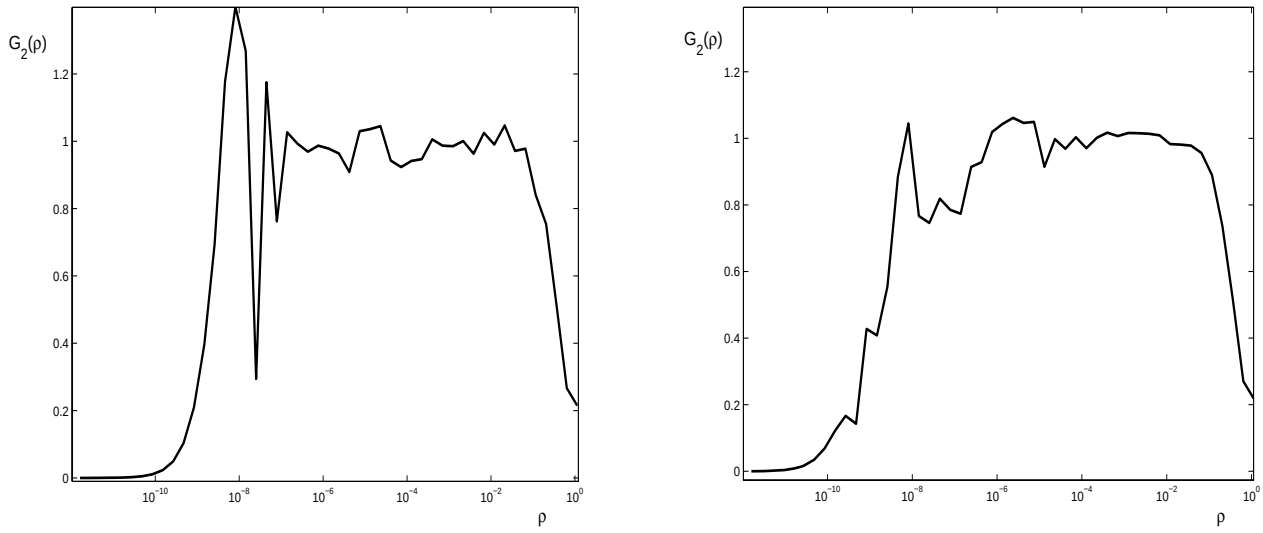

Figure 3: The normalized structure function $G_2(\rho) = \frac{D(\rho)}{J_{5/3}\,\rho^{2/3}}$ calculated by variant A of the Randomization method with $N_s = 2000$ samples. Number of wavenumbers: $n_0 = 160$ (left panel) and $n_0 = 1000$ (right panel).

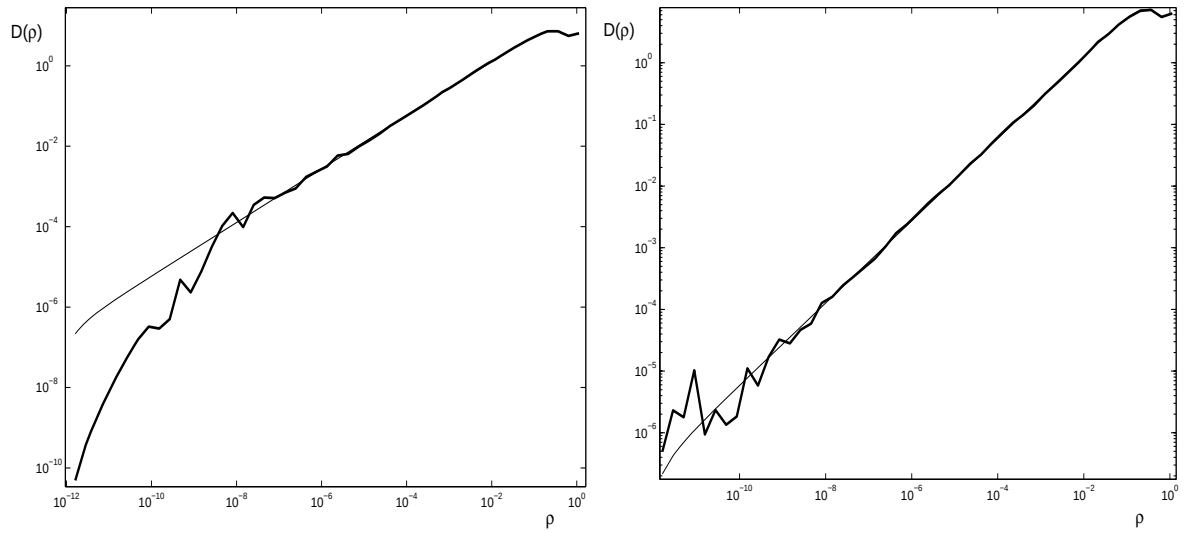

Figure 4: The structure function $D(\rho)$: exact formula (thin solid line) and calculated by averaging over an ensemble of $N_s = 2000$ Monte Carlo samples simulated by variant B of the Randomization method (bold solid line). Number of wavenumbers: $n = 160$ (left panel) and $n = 1000$ (right panel), $n_0 = 1$ in both cases.

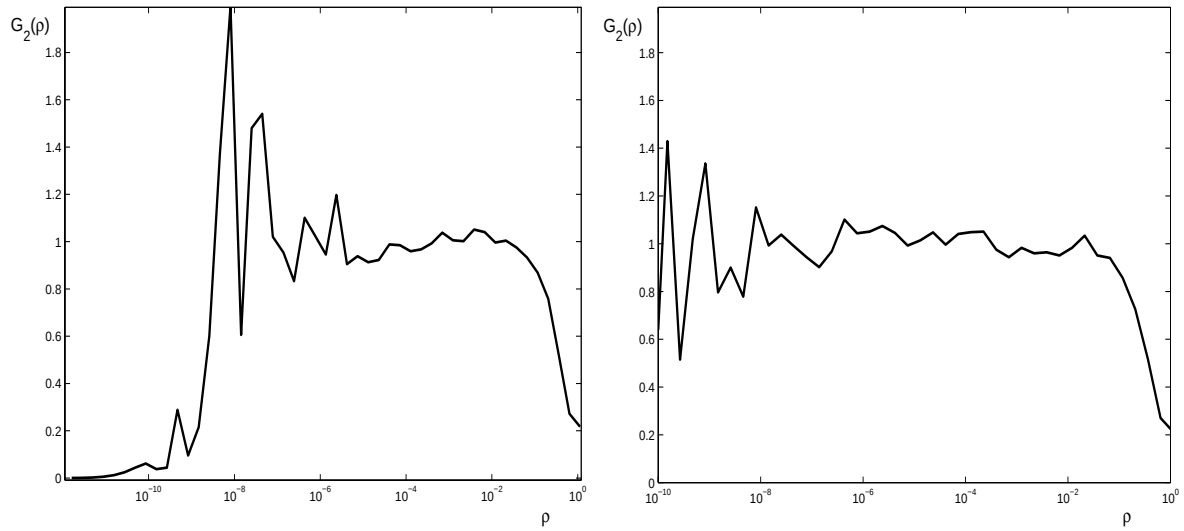

Figure 5: The normalized structure function $G_2(\rho) = \frac{D(\rho)}{J_{5/3}\,\rho^{2/3}}$ calculated by variant B of the Randomization method with $N_s = 2000$ samples. Number of wavenumbers: $n = 160$ (left panel) and $n = 1000$ (right panel).

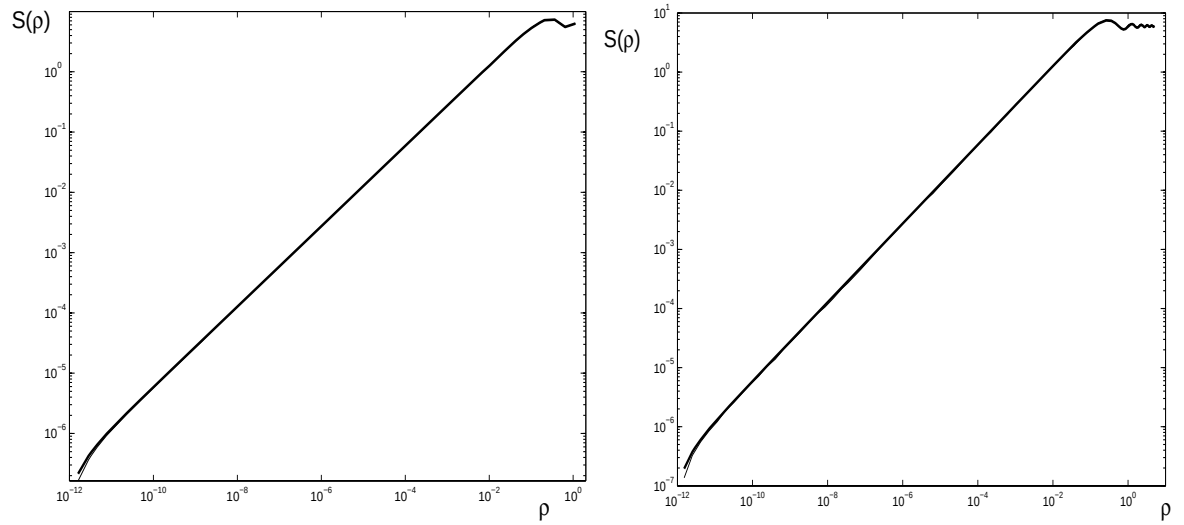

Figure 6: Comparison of structure function $D(\rho)$ as simulated by: variant C of the Randomization method (left panel) with 160 wavenumbers ($n_0 = 4$ samples from each of $n = 40$ bins, $q = 2$) and $N_s = 4000$ Monte Carlo samples; and by the Fourier-wavelet method (right panel) with $M = 40$, $b = 10$, and $N_s = 4000$ Monte Carlo samples.

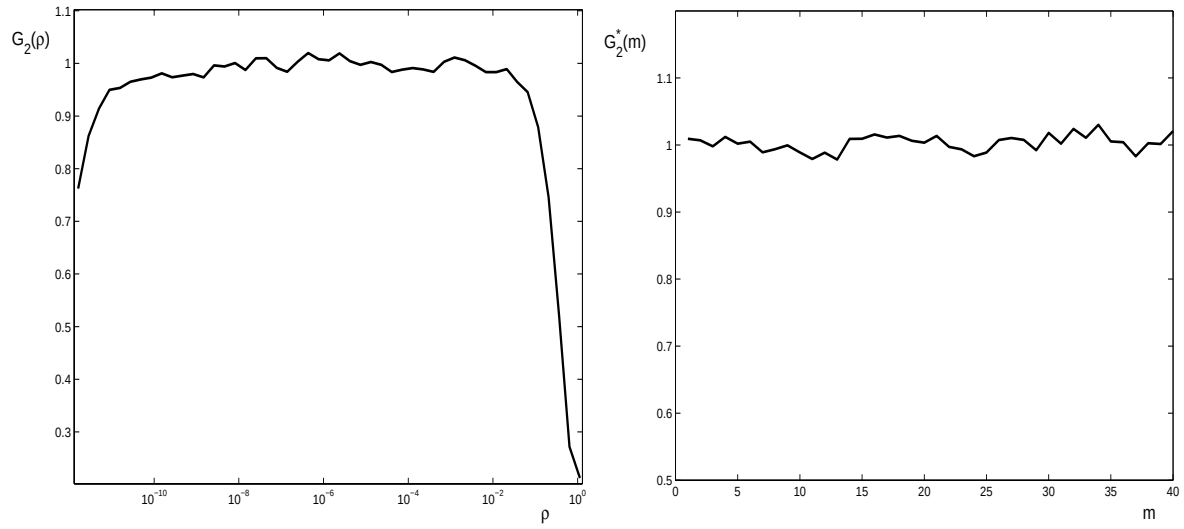

Figure 7: The normalized structure function $G_2(\rho) = \frac{D(\rho)}{J_{5/3}\,\rho^{2/3}}$ calculated by: variant C of the Randomization method (left panel) with 160 wavenumbers ($n_0 = 4$ samples from each of $n = 40$ bins), bin ratio $q = 2$, and $N_s = 20000$ Monte Carlo samples; and direct Monte Carlo simulation (right panel) with $N_s = 20000$ samples.

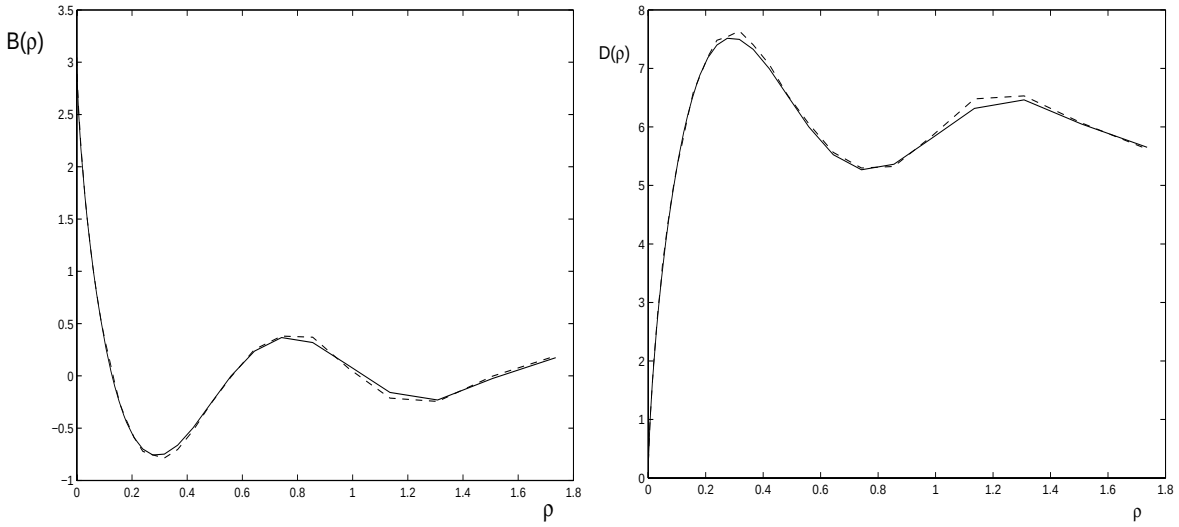

Figure 8: The correlation function (left panel) and structure function (right panel) simulated by variant C of the Randomization method (dashed line) with $n = 25$ bins, $q = 3.16$, $n_0 = 10$ wavenumbers per bin, and $N_s = 16000$ Monte Carlo samples. The solid line represents the exact formula.

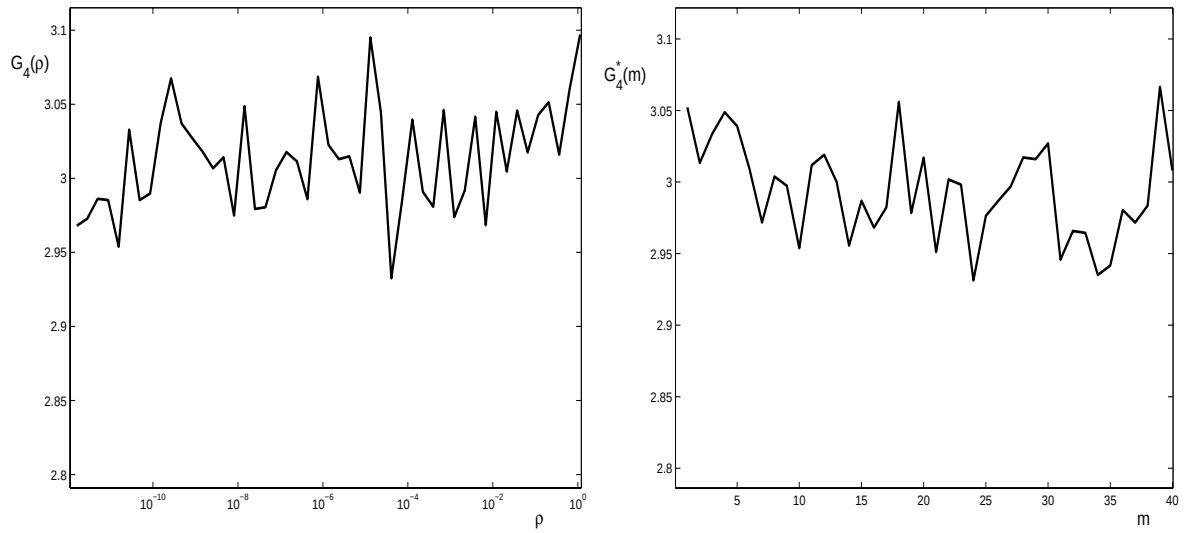

Figure 9: Kurtosis $G_4(\rho)$ calculated by: variant C of the Randomization method (left panel) with $n = 40$ bins, bin ratio $q = 2$, and $n_0 = 1$ wavenumbers per bin; and direct Monte Carlo simulation (right panel). $N_s = 20000$ samples are used in each case.

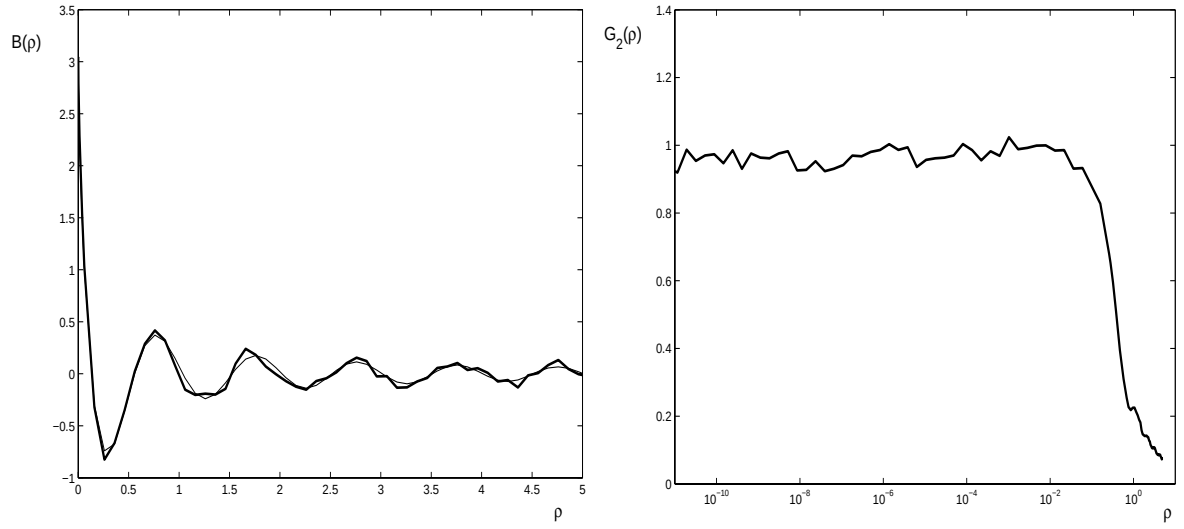

Figure 10: The correlation function (left panel) and normalized structure function $G_2(\rho)$ (right panel) for the spectrum (5.1). The bold line indicates the simulated results using a Fourier-wavelet method with $N_s = 4000$ Monte Carlo samples, $M = 40$ scales, and $b = 10$.

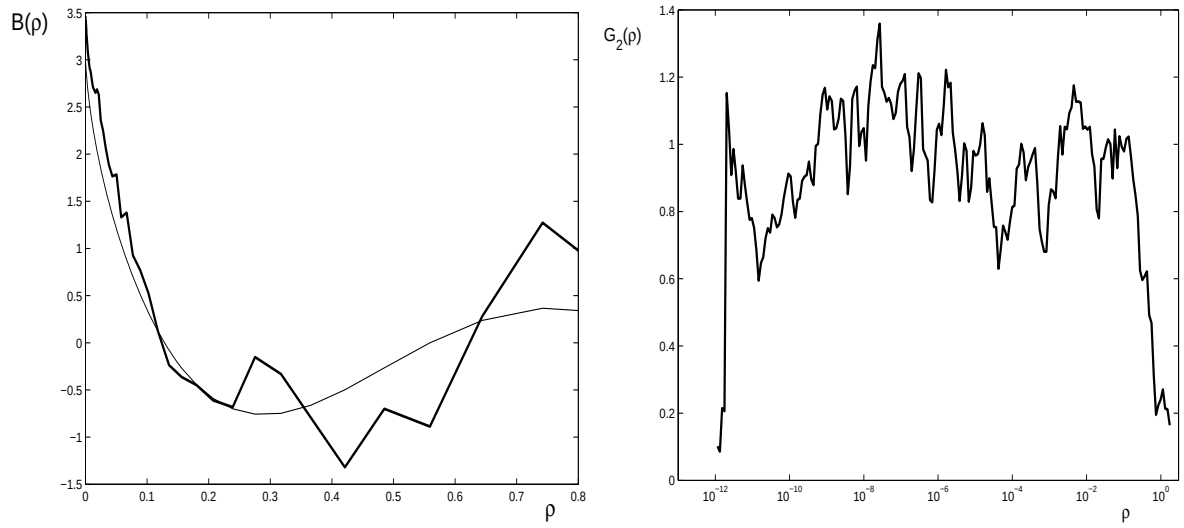

Figure 11: Ergodic averaging for the correlation function $B(\rho)$ (left panel) and normalized structure function $G_2(\rho)$ (right panel) for the Kolmogorov spectrum (5.1) over $N_b = 16000$ spatial blocks using variant C of the Randomization method with $n = 25$ sampling bins, $n_0 = 10$ wavenumbers per bin, bin ratio $q = 2$, and $\ell_b = 2\ell_{\mathbf{max}}$. In the left panel, the thin solid line denotes the exact ensemble average (6.1), while the bold solid line denotes the simulated average over spatial translations (6.2). In the right panel, the normalized structure function should be the constant value 1.

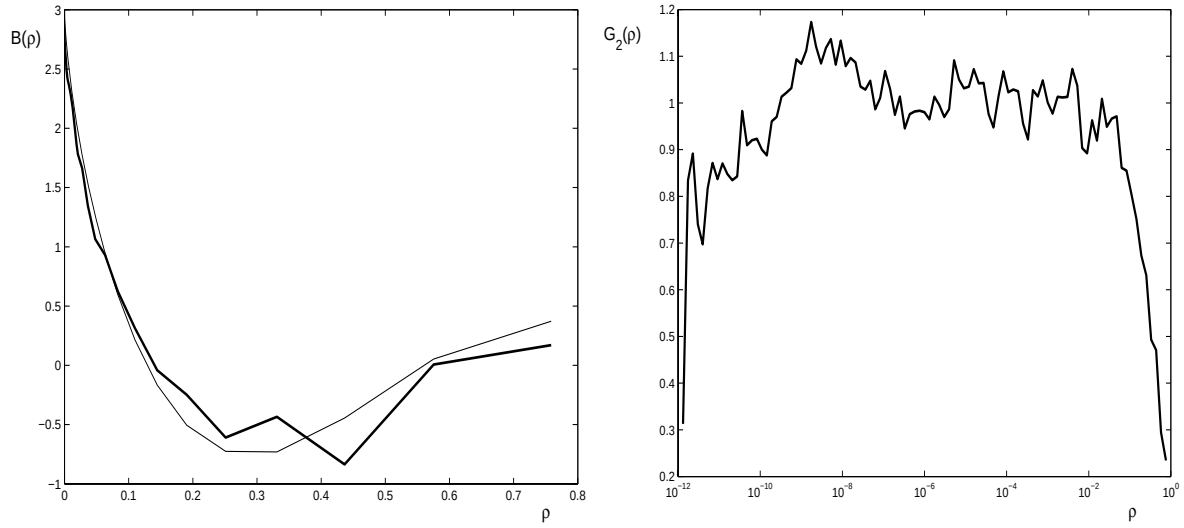

Figure 12: Same as in Figure 11, but the spatial averaging is now over $N_b = 16000$ blocks and $n_0 = 50$ wavenumbers are simulated per bin.

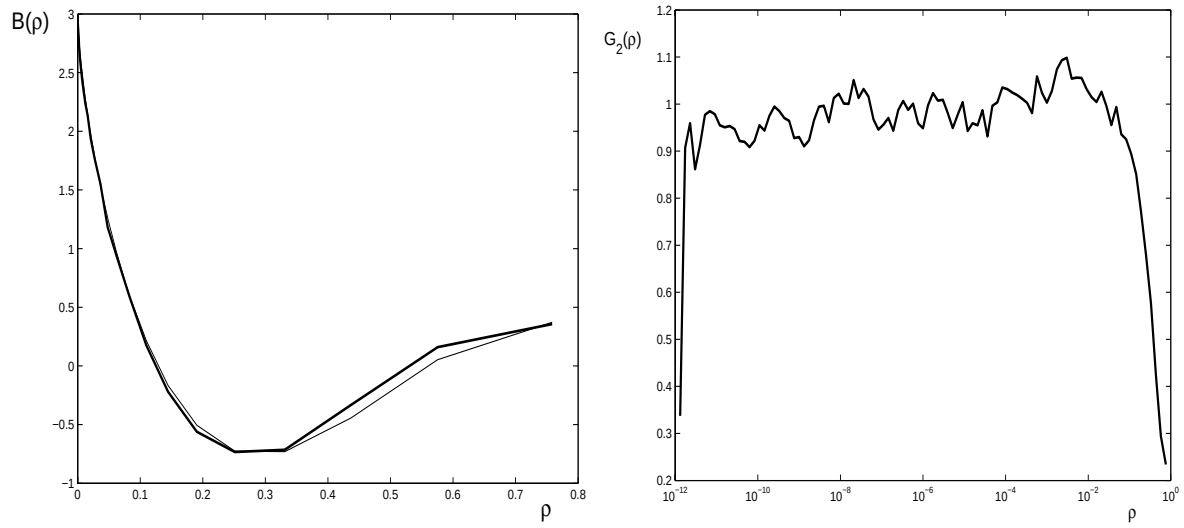

Figure 13: Same as in Figure 12, but with $n_0 = 250$ wavenumbers per bin.

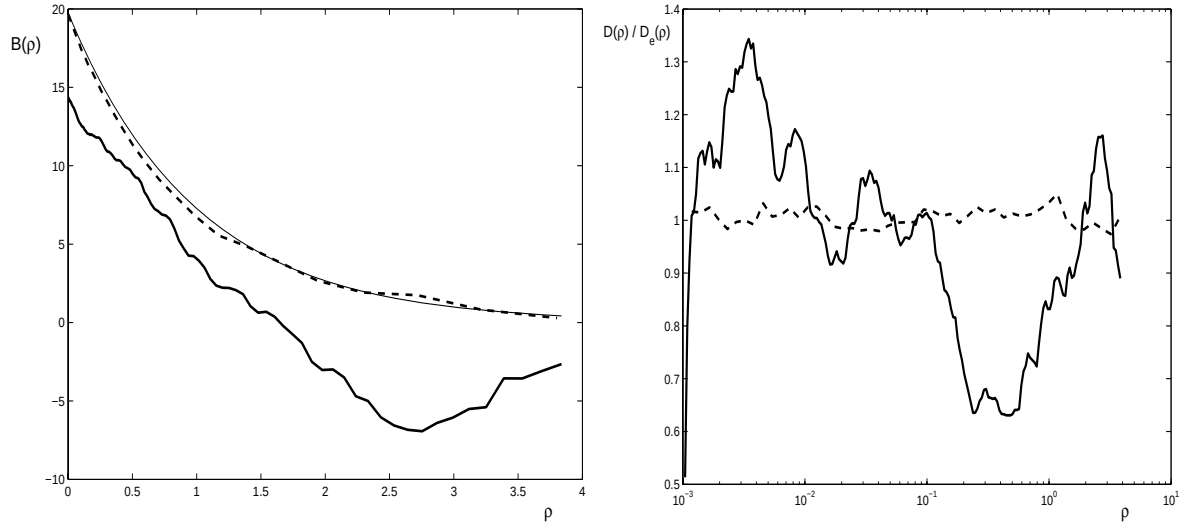

Figure 14: Ergodic averaging for the correlation function $B(\rho)$ (left panel) and ratio of spatially-averaged to exact structure function $D_{N_b}(\rho)/D(\rho)$ (right panel) for the exponential correlation function (6.3) using $N_b = 16000$ spatial blocks and variant C of the Randomization method with $n = 5$ sampling bins, $n_0 = 20$ wavenumbers per bin, bin ratio $q = 2$, and $\ell_b = 2\ell_{\mathbf{max}}$. In the left panel, the thin solid line denotes the exact result (6.3) for $B(\rho)$, while the bold solid line denotes the simulated average over spatial translations (6.2). In the right panel, the simulated ratio function should be the constant value 1. In both panels, the dashed line is obtained via ensemble average over 16000 samples.

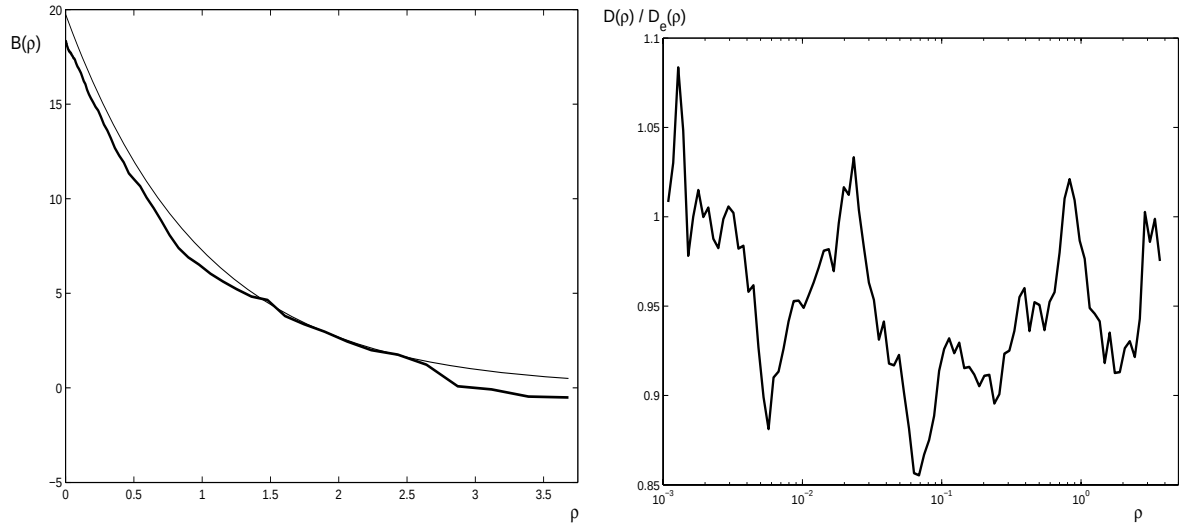

Figure 15: Same as in Figure 14, but the spatial avergaing is taken over 4000 blocks, and $n_0 = 200$ wavenumbers are sampled per bin.

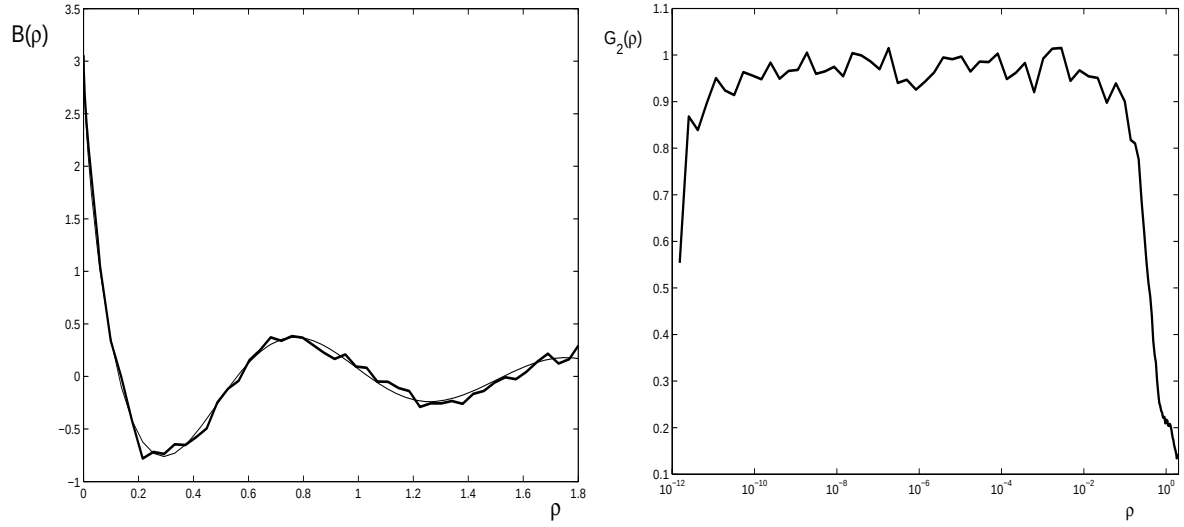

Figure 16: Ergodic averaging for the correlation function $B(\rho)$ (left panel) and normalized structure function $G_2(\rho)$ (right panel) for the Kolmogorov spectrum (5.1) over $N_b = 2000$ spatial blocks using Fourier-wavelet method with with $M = 40$ scales, $b = 10$, and $\ell_b = 2\ell_{\mathbf{max}}$. In the left panel, the thin solid line denotes the exact ensemble average (6.1), while the bold solid line denotes the simulated average over spatial translations (6.2). In the right panel, the normalized structure function should be the constant value 1.

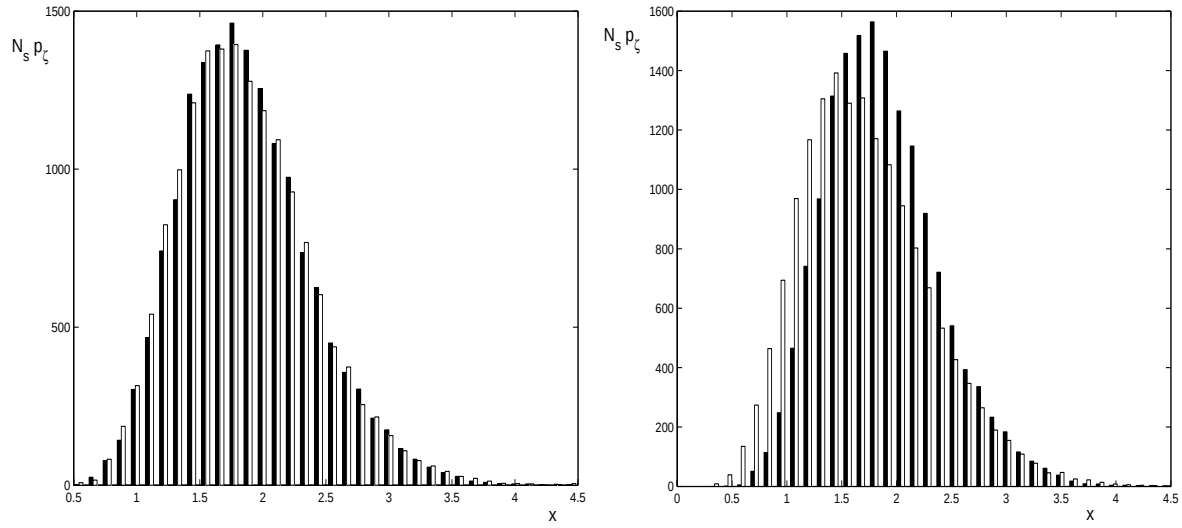

Figure 17: Histograms p_ζ for $\zeta_{10} = \max_{1 \le i \le 10} |\Delta u_i|$ multiplied by N_s, the number of samples, calculated by variant C of the Randomization method (filled bars), with $n = 25$ bins, $N_s = 16000$ samples, and $h = 1000\,\ell_{\min}$ and compared against the results of direct Monte Carlo simulations (empty bars) (also with 16000 samples). Left panel: $n_0 = 10$ wavenumbers per bin. Right panel: $n_0 = 1$ wavenumber per bin.

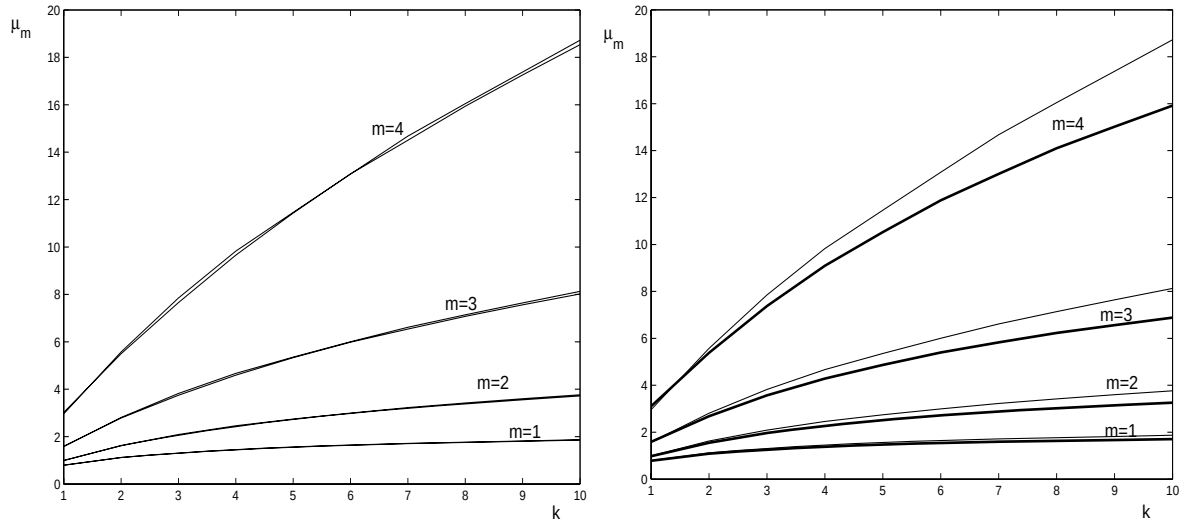

Figure 18: Statistical moments $\mu_m(k)$ as simulated from ensemble averages using variant C of the Randomization method (bold line) and compared with direct Monte Carlo simulation (thin line). The parameters are the same as in the corresponding panels of Figure 17.

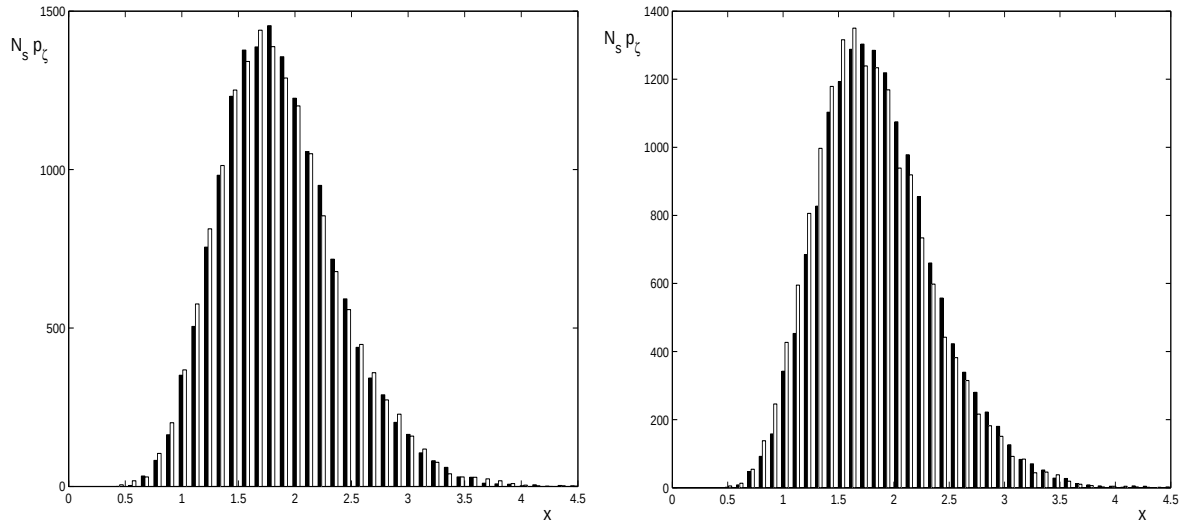

Figure 19: Histograms p_ζ for $\zeta_{10} = \max_{1 \le i \le 10} |\Delta u_i|$ multiplied by the number of samples, calculated by variant C of the Randomization method (filled bars), with $n = 25$ bins, $n_0 = 10$ wavenumbers per bin, and $N_s = 16000$ samples, and compared against the results of direct Monte Carlo simulations (empty bars) (also with 16000 samples). Left panel: $h = 100\,\ell_{\min}$. Right panel: $h = 10\,\ell_{\min}$.

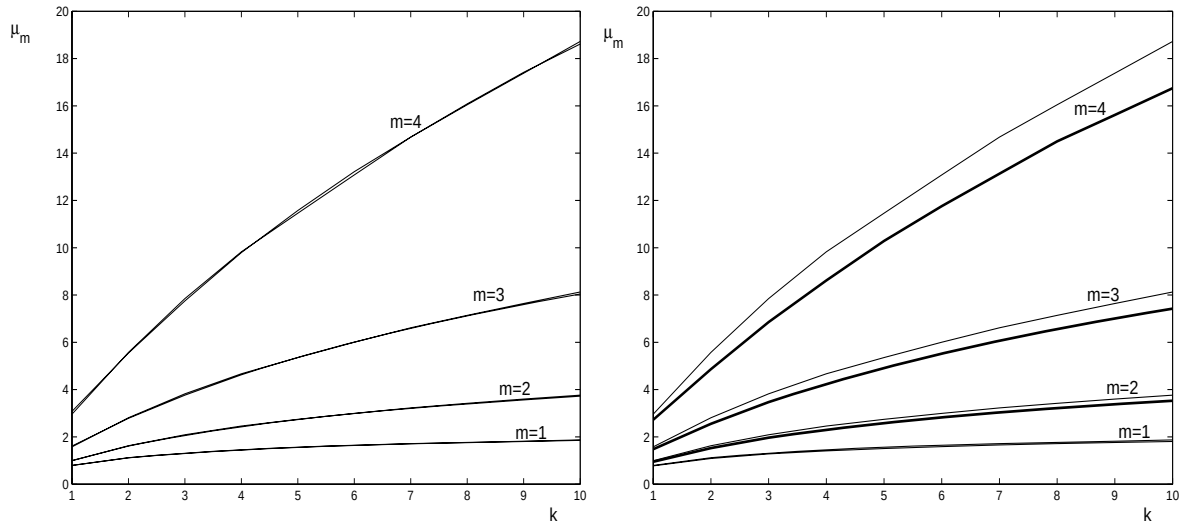

Figure 20: Statistical moments $\mu_m(k)$ as simulated from ensemble averages using variant C of the Randomization method (bold line) and compared with direct Monte Carlo simulation (thin line). The parameters are the same as in the corresponding panels of Figure 19.

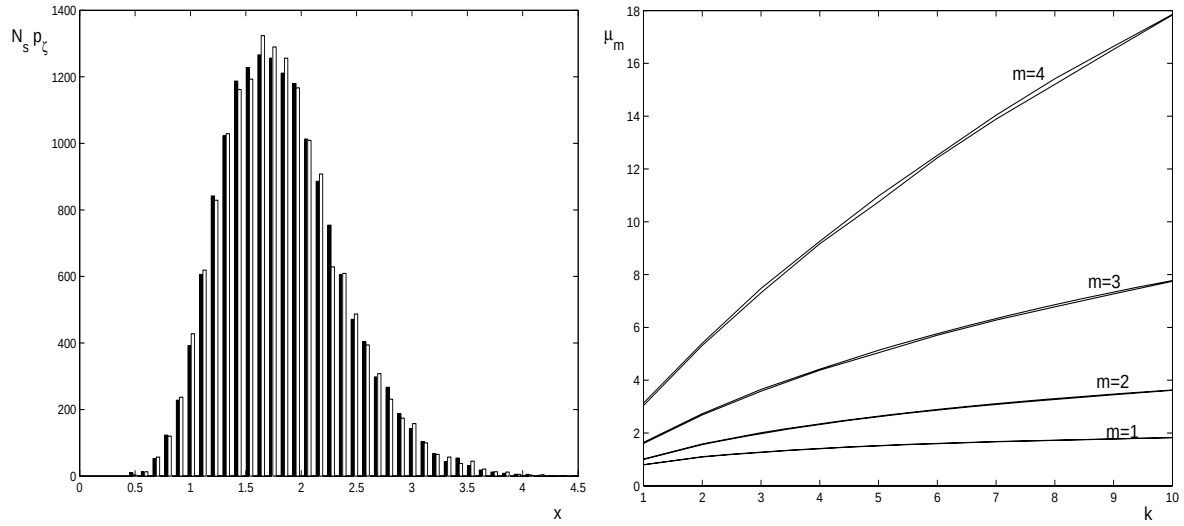

Figure 21: Left panel: Histogram p_ζ for $\max_{1\le i\le 10} |\delta u_i|$ multiplied by the number of samples, calculated by variant C of the Randomization method (filled bars), compared with the results of direct Monte Carlo simulation (empty bars).
Right panel: Statistical moments (defined as $\mu_m(k)$ in (7.2) except with δu_i replacing Δu_i) calculated by variant C of the Randomization method (bold line), compared against the results obtained by direct Monte Carlo simulation (thin line). Both calculations involve ensemble averages over $N_s = 16000$ Monte Carlo samples and $h = 1000\, l_{min}$. The Randomization Method uses a total of 250 wavenumbers ($n = 25$ bins with $n_0 = 10$ wavenumbers in each bin) and a bin ratio $q = 4$.

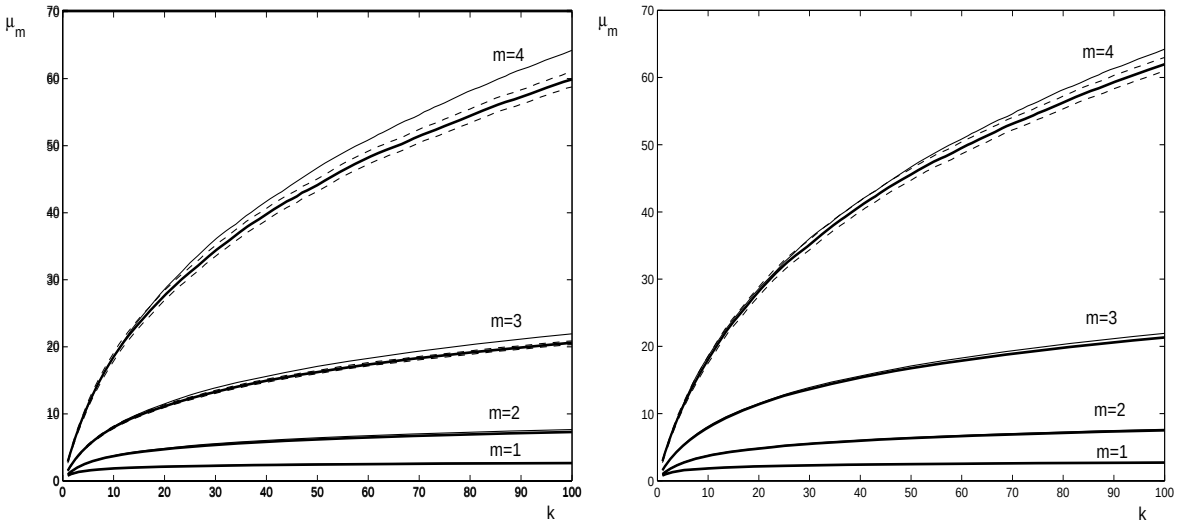

Figure 22: Statistical moments $\mu_m(k)$ of ζ_{100} with $h = 100\,\ell_{\min}$ calculated by variant C of the Randomization method (bold solid line) with $n = 25$ sampling bins, compared against the results obtained by direct Monte Carlo simulation (thin solid line). $N_s = 16000$ samples are used in both calculations. Left panel: $n_0 = 10$ wavenumbers per bin; right panel: $n_0 = 40$ wavenumbers per bin. The dashed lines show the Monte Carlo statistical sampling error for $m = 4$.

References

[1] P. Billingsley. Convergence of Probability Measures. John Willey, New York, 1968

[2] Buglanova N.A. and Kurbanmuradov O. Convergence of the Randomized Spectral Models of Homogeneous Gaussian Random Fields. *Monte Carlo Methods and Appl.*, vol. 1, No-3, 173-201, 1995.

[3] Chris Cameron. Relative efficiency of Gaussian stochastic process sampling procedures. *J. Comput. Phys.*, 192(2):546–569, 2003.

[4] René A. Carmona, Stanislav A. Grishin, and Stanislav A. Molchanov. Massively parallel simulations of motions in a Gaussian velocity field. In *Stochastic modelling in physical oceanography*, volume 39 of *Progr. Prob.*, pages 47–68. Birkhäuser Boston, Boston, 1996.

[5] D. Cioranescu and P. Donato. *An Introduction to Homogenization*. Oxford University Press, New York, 1999.

[6] Frank W. Elliott, Jr, David J. Horntrop, and Andrew J. Majda. A Fourier-wavelet Monte Carlo method for fractal random fields. *J. Comp. Phys*, 132(2):384–408, 1997.

[7] Frank W. Elliott, Jr, David J. Horntrop, and Andrew J. Majda. Monte Carlo methods for turbulent tracers with long range and fractal random velocity fields. *Chaos*, 7(1):39–48, 1997.

[8] Frank W. Elliott, Jr and Andrew J. Majda. A wavelet Monte Carlo method for turbulent diffusion with many spatial scales. *J. Comp. Phys*, 113(1):82–111, July 1994.

[9] Frank W. Elliott, Jr and Andrew J. Majda. A new algorithm with plane waves and wavelets for random velocity fields with many spatial scales. *J. Comp. Phys*, 117:146–162, 1995.

[10] Frank W. Elliott, Jr and Andrew J. Majda. Pair dispersion over an inertial range spanning many decades. *Phys. Fluids*, 8(4):1052–1060, 1996.

[11] Jens Feder. *Fractals*, chapter 9–14, pages 163–243. Physics of Solids and Liquids. Plenum Press, New York and London, 1988.

[12] William Feller. *An introduction to probability theory and its applications*, volume 2, section III.6, pages 47, 48. John Wiley & Sons, New York, London, Sydney, second edition, 1971.

[13] Uriel Frisch. *Turbulence*. Cambridge University Press, Cambridge, 1995. The legacy of A. N. Kolmogorov.

[14] J. C. H. Fung, J. C. R. Hunt, N. A. Malik, and R. J. Perkins. Kinematic simulation of homogenous turbulence by unsteady random Fourier modes. *J. Fluid Mech.*, 236:281–318, 1992.

[15] I. M. Gel'fand and G. E. Shilov. *Generalized functions. Properties and operations*, volume 1, section 3.3. Academic Press, New York, 1964.

[16] Lynn W. Gelhar. *Stochastic Subsurface Hydrology*. Prentice-Hall, Englewood Cliffs, N.J., 1993.

[17] D. Horntrop and A. Majda. An overview of Monte Carlo simulation techniques for the generation of random fields. In P. Muller and D. Henderson, editors, *Monte Carlo Simulations in Oceanography*, pages 67–79, 1997. Proceedings of the Ninth 'Aha Huliko'a Hawaiian Winter Workshop.

[18] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics: Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 1992.

[19] R.H. Kraichnan. Diffusion by a random velocity field. Phys.Fluids, **13** (1970), N1, 22-31.

[20] Kurbanmuradov O. Weak Convergens of Approximate Models of Random Fields. *Russian Journal of Numerical Analysis and Mathematical Modelling*, **10** (1995), N6, 500-517.

[21] Kurbanmuradov O. Weak Convergence of Randomized Spectral Models of Gaussian Random Vector Fields. *Bull.Novosibirsk Computing Center, Numerical Analysis*, issue 4, 19-25, 1993.

[22] O. Kurbanmuradov, K. Sabelfeld, and D. Koluhin. Stochastic Lagrangian models for two-particle motion in turbulent flows. Numerical results. *Monte Carlo Methods Appl.*, 3(3):199–223, 1997.

[23] N. N. Lebedev. *Special functions and their applications*, chapter 1, pages 1–15. Dover, New York, 1972.

[24] Marcel Lesieur. *Turbulence in fluids*. Number 1 in Fluid Mechanics and its Applications. Kluwer, Dordrecht, second revised edition, 1990.

[25] Andrew J. Majda and Peter R. Kramer. Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena. *Phys. Rep.*, 314(4-5):237–574, 1999.

[26] Benoit B. Mandelbrot. *The fractal geometry of nature*. W.H. Freeman, San Francisco, New York, updated and augmented edition, 1983.

[27] W. D. McComb. *The physics of fluid turbulence*, volume 25 of *Oxford Engineering Science Series*, chapter 2. Clarendon Press, New York, 1991.

[28] Mikhailov G.A. Approximate models of random processes and fields. Russian J. Comp. Mathem. and mathem. Physics, vol.23 (1983), N3, 558-566. (in Russian).

[29] A.S. Monin and A.M. Yaglom. *Statistical Fluid Mechanics: Mechanics of Turbulence*, Volume 2. The M.I.T. Press, 1981.

[30] F. Poirion and C. Soize. Numerical methods and mathematical aspects for simulation of homogenous and non homogenous Gaussian vector fields.

In Paul Kree and Walter Wedig, editors, *Probabilistic methods in applied physics*, volume 451 of *Lecture Notes in Physics*, pages 17–53. Springer-Verlag, Berlin, 1995.

[31] Stephen B. Pope. *Turbulent flows*. Cambridge University Press, Cambridge, 2000.

[32] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in FORTRAN*, section 7.8, pages xxvi+963. Cambridge University Press, Cambridge, second edition, 1992. The art of scientific computing, With a separately available computer disk.

[33] Yoram Rubin. *Applied Stochastic Hydrogeology*. Oxford University Press, New York, Oxford, 2003.

[34] K. K. Sabelfeld and O. Kurbanmuradov. Stochastic Lagrangian models for two-particle motion in turbulent flows. *Monte Carlo Methods Appl.*, 3(1):53–72, 1997.

[35] Karl K. Sabelfeld. *Monte Carlo methods in boundary value problems*, chapter 1,5, pages 31–47,228–238. Springer Series in Computational Physics. Springer-Verlag, Berlin, 1991.

[36] M. Shinozuka. Simulation of multivariate and multidimensional random processes. J. of Acoust. Soc. Am. **49** (1971), 357-368.

[37] H. Sigurgeirsson and A. M. Stuart. A model for preferential concentration. *Phys. Fluids*, 14(12):4352–4361, December 2002.

[38] D. J. Thomson and B.J. Devenish. Particle pair in kinematic simulations. Journal of Fluid Mechanics, **526** (2005), 277-302.

[39] J. A. Viecelli and E. H. Canfield, Jr. Functional representation of power-law random fields and time series. *J. Comp. Phys*, 95:29–39, 1991.

[40] Richard F. Voss. Random fractal forgeries. In Rae A. Earnshaw, editor, *Fundamental algorithms for computer graphics*, volume 17 of *NATO ASI Series F: Computer and System Sciences*, pages 805–835, Berlin, 1985. NATO Science Affairs Divison, Springer-Verlag.

[41] A. M. Yaglom. *Correlation theory of stationary and related random functions. Volume I: Basic results*. Springer-Verlag, Berlin, 1987.