

# Comparative Analysis of Multiscale Gaussian Random Field Simulation Algorithms <sup>★</sup>

Peter R. Kramer <sup>a,\*</sup> Orazgeldi Kurbanmuradov <sup>b</sup>  
Karl Sabelfeld <sup>c,d</sup>

<sup>a</sup>*Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy,  
NY 12180 USA*

<sup>b</sup>*Center for Phys. Math. Research, Turkmenian State University, Saparmyrat  
Turkmenbashy av. 31, 744000 Ashgabad, Turkmenistan*

<sup>c</sup>*Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstraße 39,  
D-10117 Berlin Germany*

<sup>d</sup>*Institute of Computational Mathematics and Mathematical Geophysics, Russian  
Acad. Sci., Lavrentieva str., 6, 630090 Novosibirsk, Russia*

---

## 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) with stratified sampling with a logarithmically uniform subdivision. We focus primarily on fractal Gaussian random fields with Kolmogorov-type spectra. Previous work has shown that 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), along with the Fourier-wavelet method, 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.

We find that 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 less cost than the Fourier-wavelet method. The cost of the Randomization method relative to the Fourier-wavelet method, however, appears to increase with the complexity of the random field statistics which are to be calculated accurately. Moreover, the Fourier-wavelet method 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 [30,11,45,46,21,17,12,42,27], where the velocity field representing the turbulent flow is modeled as a random field  $\vec{v}(\vec{x}, t)$  with statistics encoding important empirical features, and the temporal dynamics of the position  $\vec{X}(t)$  and velocity  $\vec{V}(t) = \frac{d\vec{X}}{dt}$  of immersed particles is then governed by equations involving this random field such as

$$m d\vec{V}(t) = -\gamma \left( \vec{V}(t) - \vec{v}(\vec{X}(t), t) \right) dt + \sqrt{2k_B T \gamma} d\vec{W}(t), \quad (1)$$

where  $m$  is particle mass,  $\gamma$  is its friction coefficient,  $k_B$  is Boltzmann's constant,  $T$  is the absolute temperature, and  $\vec{W}(t)$  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  $K(\vec{x})$  is modeled as random field reflecting the empirical variability of the porous medium [40,19,6,23]. The Darcy flow rate  $\vec{q}(\vec{x})$  in response to pressure applied at the boundary is governed by the Darcy equation

$$\begin{aligned} \vec{q}(\vec{x}) &= -K(\vec{x}) \text{grad } \phi(\vec{x}), \\ \text{div } \vec{q} &= 0, \end{aligned} \quad (2)$$

in which the random hydraulic conductivity function appears as a coefficient, and the applied pressure is represented in the boundary conditions for the

---

\* 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 CLG 981426. PRK is partly supported by an NSF grant DMS-0207242.

\* Corresponding author. Address: Tel.: +001 (518) 276-6896; fax: +001 (518) 276-4824.

*Email addresses:* [kramep@rpi.edu](mailto:kramep@rpi.edu) (Peter R. Kramer),  
[kurbanmuradov@yandex.ru](mailto:kurbanmuradov@yandex.ru) (Orazgeldi Kurbanmuradov),  
[sabelfel@wias-berlin.de](mailto:sabelfel@wias-berlin.de) (Karl Sabelfeld).

*URL:* <http://www.rpi.edu/~kramep> (Peter R. Kramer).

internal pressure head  $\phi$ . 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 [4,45]. 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 [22]. 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  $u(\vec{x})$  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  $u(\vec{x})$  can be represented through a stochastic Fourier integral [35]

$$u(\vec{x}) = \int_{\mathbb{R}^d} e^{-2\pi i \vec{k} \cdot \vec{x}} E^{1/2}(\vec{k}) \tilde{W}(d\vec{k}) \quad (3)$$

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

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

Several approaches are based on various discretizations of (3) which give rise to finite sums of functions with independent Gaussian random coefficients [39]. A Riemann sum discretization of the stochastic integral is easy to implement [44,47,48,36,16], and following [8,20,30], we shall refer to it as the stan-

standard Fourier method. As documented in [8], 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 (3) through a finite set of randomly chosen wavenumbers generated by Monte Carlo methods [24,34]. Yet another method designed for multiscale random field simulation is the Fourier-wavelet method [7], which arises from a wavelet decomposition of the white noise measure in (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 [3,7], 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 the truncation of the sums in the associated random field representations, which can however be significantly reduced through a good choice of wavelet functions [9,7]. 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 [25]. In particular, the simulations of fractal random fields in [7] 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 [3,7], a particular version of the Randomization method was used; namely, the random wavenumbers were chosen according to a stratified sampling strategy with a subdivision of wavenumber space into sampling bins of equal energy. In the studies [41,27], 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 [27], a similar implementation of the standard Fourier method with wavenumber discretized uniformly and deterministically in logarithmic space [46], and a multiscale wavelet method [11] have all been employed to simulate disper-

sion 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 [27,46,11].

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 [3,7] 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). By contrast, the solution of the porous medium problem requires the generation of the random field over the whole computational domain at once. We consider how the computational cost of the Randomization and Fourier-wavelet methods compare for various tasks in which the random field is to be evaluated at a set of points specified in advance or on demand, regularly or irregularly distributed within the computational domain. Both the overhead cost and the cost of simulating each new realization of the random field is considered.
- We study statistical features of the random field beyond two-point quantities such as the second order structure function

$$D(\rho) = \langle (u(x + \rho) - u(x))^2 \rangle \quad (4)$$

and associated kurtosis which have been the focus of previous work [7,3]. We point out that the Randomization method has the flexibility to simulate a random field using a smaller set of random variables than the Fourier-wavelet method, which permits for example the faster simulation of a random field with good accuracy of the simulated second order structure function. We examine how the number of random variables (and associated simulation cost) required by the Randomization method changes when more complex multi-point statistics are to be simulated accurately.

- 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 (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 Appendices A and B. 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 [3,7] 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.

The general conclusions are that the use of a logarithmically uniform stratified sampling strategy for the Randomization method greatly increases its competitiveness with the Fourier-wavelet method for simulating multiscale random fields. The Randomization method can simulate low order statistics such as the second order structure function (4) and associated kurtosis accurately using a smaller number of random variables and therefore lower cost than the Fourier-wavelet method. The cost of the Randomization method relative to the Fourier-wavelet method, however, increases with the number of observation points involved in the statistic to be simulated. In particular, to have good ergodic properties, meaning the accurate simulation of the statistics of spatial averages over large regions, the Randomization method incurs a cost considerably larger than the Fourier-wavelet method. The choice of which method to be used in the simulation of a multiscale random field in an application, therefore, should be guided by what types of statistics of the random field need to be simulated accurately. The Randomization method is simpler to implement and is faster for the accurate simulation of statistics involving a relatively small number of observation points. The Fourier-wavelet method is somewhat more complicated, but appears to be considerably more efficient in the accurate simulation of statistics involving large numbers of points, particularly ergodic averages.

## 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  $u(\vec{x})$  to be simulated. We suppose here that  $u(\vec{x})$  is a well-defined scalar-valued homogenous, Gaussian random field with given spectral density  $E(\vec{k})$  (which is just the Fourier transform of the correlation function; see Appendix A). The simulation of vector-valued multiscale Gaussian random fields can be conducted through standard extensions from the scalar-valued case. Some of the required formalism is reviewed briefly in Appendix A and B, but we only treat the scalar-valued case in the main text in order to keep the notation simple and

the discussion focused on the challenges of simulating a multiscale random field. 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:

$$\ell_c = \left( V_d^{-1} \frac{\sup_{\vec{k} \in \mathbb{R}^d} E(\vec{k})}{\int_{\mathbb{R}^d} E(\vec{k}) \, d\vec{k}} \right)^{1/d}, \quad (5)$$

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 [33] has simply  $E(\vec{0})$  in the numerator, but our definition generalizes meaningfully to random fields with the spectral density vanishing at the origin. Indeed  $E(\vec{0})$  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:

$$\ell_s = \left( V_d^{-1} \frac{\sup_{\vec{k} \in \mathbb{R}^d} |\vec{k}|^2 E(\vec{k})}{\int_{\mathbb{R}^d} |\vec{k}|^2 E(\vec{k}) \, d\vec{k}} \right)^{1/d}, \quad (6)$$

which is really an analogous correlation length for the random field gradient  $\vec{\nabla}u$ . The two length scales  $\ell_c$  and  $\ell_s$  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  $u(\vec{x})$  exhibits a nontrivial correlation structure. That is, for  $|\vec{x} - \vec{x}'| \gg \ell_c$ , the values of  $u(\vec{x})$  and  $u(\vec{x}')$  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  $|\vec{x} - \vec{x}'| \ll \ell_s$ , the random field over the line segment connecting  $\vec{x}$  and  $\vec{x}'$  can be well approximated by a linear interpolation between  $u(\vec{x})$  and  $u(\vec{x}')$  (with *relative* error  $o(|\vec{x} - \vec{x}'|/\ell_s)$ ).

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  $\ell_c$  in (5) is well-defined as a finite nonzero value. We admit random fields for which the integral in the denominator of (6) converges or diverges; in the latter case, we define  $\ell_s = 0$ . Idealized fractal random fields [32,13], such as those associated with the Kolmogorov inertial range theory of turbulence  $E(\vec{k}) \propto |\vec{k}|^{-5/3}$  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 [36]. 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  $E(\vec{k})$  is peaked near widely separated wavenumbers  $|\vec{k}| \sim \ell_c^{-1}$  and  $|\vec{k}| \sim \ell_s^{-1}$  but rapidly decaying away from these values. In this case, the random field could be simulated simply 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  $\ell_c^{-1} \ll |\vec{k}| \ll \ell_s^{-1}$ , 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, we may need to evaluate the random field at a pre-specified set of points in  $d$  dimensions. One example would be the simulation

of the flow through a porous medium [19,40], in which the random field would need to be simulated on a computational grid 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 the sampling length denotes the distance between points at which the random field is calculated. In the simplest case, the random field is to be simulated on a Cartesian grid with linear extent  $L$  in each direction, with grid spacing  $h$ . We can however consider the more general situation in which the prespecified points are irregularly arranged. Our computational cost considerations will still apply provided that the collection of points can still be described meaningfully by a characteristic length scale  $L$  of the overall diameter of the cluster of points and a length scale  $h$  for the typical separation between neighboring points.

In other applications, such as the simulation of a particle moving through a prescribed turbulent field (1), one may wish to be able to evaluate the random field  $u(\vec{x})$  at arbitrary points on demand (not known in advance of the generation of the random field). In situations in which the points to be specified on demand are expected to be rather densely distributed over a fixed computational domain, one might approach this simulation task by laying down a grid of points over the domain, pre-computing the random field over this grid, and then interpolating from this computed grid of random field values to evaluate the random field at points which are requested later. Since the interpolation procedure is separate from the type of algorithm used to simulate the multiscale Gaussian random field, it will not concern us. The cost and considerations involved in simulating the random field on the grid for interpolation falls within the category of random field simulations over a pre-specified set of points. On the other hand, if the points at which the random field is to be evaluated are not known to fall within a particular computational domain or if the evaluation points are expected to be rather sparsely distributed in the computational domain, then the strategy of pre-computation of the random field over a grid followed by interpolation may be impractical or inefficient. In this case, the computational procedure truly involves the evaluation of the simulated random field representation at an arbitrary collection of points not known in advance. This task has been the emphasis of previous studies of multiscale Gaussian random field simulation algorithms [43], [10].

Finally, for any application, the ideal random field to be simulated has non-trivial structure on length scales ranging from  $\ell_s$  to  $\ell_c$ . A numerical multiscale representation of this random field will be associated with certain finite minimum and maximum length scales  $\ell_{\min}$  and  $\ell_{\max}$ , outside of which the method cannot be expected to accurately represent the structure of the ideal random field. The length scales  $\ell_{\max}$  and  $\ell_{\min}$  can be related to more fundamental pa-

parameters of the Monte Carlo simulation methods, as we describe in subsequent sections. Generally speaking,  $\ell_{\max}$  should be chosen to be at least as large as the correlation length  $\ell_c$  (but need not be as large as the domain length  $L$ ). On the other hand,  $\ell_{\min}$  is set either equal to or somewhat smaller than  $\max(h, \ell_s)$  or at a larger value determined by computational cost constraints.

### 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 [43]

$$u^{(\text{R})}(x) = \frac{\sigma}{\sqrt{n_0}} \sum_{j=1}^{n_0} \left[ \xi_j \cos(2\pi k_j x) + \eta_j \sin(2\pi k_j x) \right], \quad (7)$$

where  $\xi_j, \eta_j, j = 1, \dots, n_0$  are mutually independent standard Gaussian random variables (mean zero and unit variance), and  $\sigma^2 = \int E(k) dk = 2 \int_0^\infty E(k) dk$ . The wave numbers  $k_j, j = 1, \dots, n_0$  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  $\xi_j$  and  $\eta_j$ . This variant A of the Randomization method may be thought of as the most straightforward way to approximate the Fourier stochastic integral (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 artificially 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 [8].

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 to choose a prescribed number of wavenumbers at random locations within each bin. This Monte Carlo variance reduction technique is an example of “stratified sampling” [38]. 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  $\Delta$  as the total space from which the wavenumbers are to be sampled (it can in general be chosen as  $\Delta = [0, \infty)$ ) but it can also be chosen as the possibly smaller support of the spectrum  $E$  on the nonnegative real axis. We then choose a partition of  $\Delta$  into a union of smaller non-overlapping intervals  $\Delta = \cup_{j=1}^n \Delta_j$ . Within each interval  $\Delta_j$ , we sample  $n_0$  independent random wavenumbers  $k_{jl}$ ,  $l = 1, \dots, n_0$  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

$$\sigma_j^2 = 2 \int_{\Delta_j} E(k) dk .$$

The simulation formula then reads

$$u^{(R,s)}(x) = \sum_{j=1}^n \frac{\sigma_j}{\sqrt{n_0}} \sum_{l=1}^{n_0} \left[ \xi_{jl} \cos(2\pi k_{jl} x) + \eta_{jl} \sin(2\pi k_{jl} x) \right] . \quad (8)$$

The amplitudes  $\xi_{jl}$ ,  $\eta_{jl}$ ,  $j = 1, \dots, n$ ;  $l = 1, \dots, n_0$  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  $\Delta_j$  so that each of them contains an equal amount of “energy” (integral of the spectral density):  $\sigma_j^2 = \frac{\sigma^2}{n}$  for  $1 \leq j \leq n$ . 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  $j = 1, \dots, n$  are defined according to a geometric distribution with ratio parameter  $q$ :  $\hat{k}_{j+1} = q\hat{k}_j$ ,  $j = 2, \dots, n-1$ . If the spectral density of the random field is confined to a bounded domain of wavenumbers  $k \leq k_{max}$ , then we can choose  $\hat{k}_{n+1} = k_{max}$ ; otherwise we take  $\hat{k}_{n+1} = \infty$ . 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 multi-scale 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 [7,10,9]. 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 (which can be rigorously formulated as in Appendix A.1), however, indicate that with a sufficiently rich sampling of wavenumbers, the simulated field should have some approximately Gaussian properties. We investigate these questions in some detail in Sections 5–7.

## 4 Fourier-wavelet simulation method

We present here the one-dimensional formulation of the Fourier-wavelet method, and discuss multi-dimensional generalizations in Appendix B. A homogeneous Gaussian random field  $u(x)$  can be represented using the Fourier-wavelet representation [7]:

$$u(x) = u_0 \sum_{m=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \gamma_{mj} f_m(2^m(x/\ell) - j), \quad (9)$$

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

$$f_m(\xi) = \int_{-\infty}^{\infty} e^{-2\pi i \tilde{k} \xi} 2^{m/2} \tilde{E}^{1/2}(2^m \tilde{k}) \hat{\phi}(\tilde{k}) d\tilde{k}, \quad (10)$$

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

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

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  $p$ th order perfect B-spline [7]:

$$\hat{\phi}(k) = -i \operatorname{sign}(k) e^{i\pi k} b(|k|), \quad (11)$$

where

$$b(k) = \begin{cases} \sin(\frac{\pi}{2} \nu_p(3k - 1)), & k \in \left(\frac{1}{3}, \frac{2}{3}\right], \\ \cos(\frac{\pi}{2} \nu_p(\frac{3}{2}k - 1)), & k \in \left(\frac{2}{3}, \frac{4}{3}\right], \\ 0, & \text{else.} \end{cases} \quad (12)$$

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

$$\nu_p(x) = \frac{4^{p-1}}{p} \left\{ [x - x_0]_+^p + [x - x_p]_+^p + 2 \sum_{j=1}^{p-1} (-1)^j [x - x_j]_+^p \right\}, \quad (13)$$

where  $x_j = (1/2)[\cos((p-j)/p)\pi + 1]$ , and  $[a]_+ = \max(a, 0)$ . The positive integer parameter  $p$  is chosen in [7] equal to 2.

The representation (9) 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  $\frac{1}{3\ell} 2^m \leq k \leq \frac{4}{3\ell} 2^m$ .

The implementation of the Fourier-wavelet method of course requires that the sums over  $m$  and  $j$  in (9) 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  $\ell$  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  $0 \leq m \leq M - 1$ , 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  $\frac{1}{3}\ell_{\max}^{-1} \leq k \leq \frac{4}{3}\ell_{\min}^{-1}$ , 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 [7], 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  $p \geq 2$ , then, we can choose an integer “bandwidth” cutoff  $b$  so that the total mean-square contribution from terms with  $|j - 2^m x/\ell| > b$  to the random field value at  $x$  is as small as desired. When evaluating the random field  $u(x)$  at a desired point  $x$ , then, only the  $2b$  terms satisfying  $|2^m(x/\ell) - j| \leq b$  are incorporated.

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

$$u^{(\text{FW})}(x) = u_0 \sum_{m=0}^{M-1} \sum_{j'=-b+1}^b \gamma_{m, \bar{n}_m(x)+j'} f_m(2^m(x/\ell) - \bar{n}_m(x) - j'), \quad (14)$$

where  $\bar{n}_m(x) \equiv \lfloor 2^m(x/\ell) \rfloor$ , and the notation  $\lfloor y \rfloor$  denotes the greatest integer not exceeding the real value  $y$ . One must be careful when evaluating the random field at various locations  $x$  to be sure that the random variables  $\gamma_{mj}$  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 (10) can be found in [7]. In our simulations, we shall simply use the numerical parameters  $p$ ,  $b$ , and  $\Delta\xi$  which were found to work well in [7].

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 somewhat more complicated than the Randomization method, and it does incur a (controllable) statistical bias through truncation of the sums in (14) 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 (as measured by the second order structure function (4)) over a desired number of decades [7,3]. We will also briefly discuss their costs relative to a direct Gaussian random field simulation approach which is not adapted for multiscale applications. In neither the Randomization method nor the Fourier-wavelet method does the random field construction make reference to the points at which the random field is to be eventually evaluated. When the evaluation points are irregularly distributed, therefore, the cost of random field construction and evaluation at the desired points does not depend on whether the evaluation points are specified in advance or later on demand. (This is not true of the direct Gaussian random field simulation approach, which we discuss briefly in Subsubsection 5.1.1). Some cost savings can be achieved by the Fourier-wavelet method if the evaluation points are pre-specified and distributed in a regular way, particularly as on a grid covering a specified computational domain. We therefore will begin by discussing the relative computational costs of the Gaussian random field simulation methods in Subsection 5.1 without any special assumptions on how the evaluation points are distributed, allowing them as well to be specified in advance or on demand. In Subsection 5.2, we specialize our discussion to the case in which the evaluation points are specified in advance with a regular distribution over a given computational domain. Also included in this special case is the situation where the random field is pre-computed over a regular computational grid, with evaluations at the desired points obtained by a subsequent interpolation step.

The costs of any of the Gaussian random field simulation methods 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 one new realization* of the random field *and evaluating* it at the points desired.

For the Randomization method, the second cost can be cleanly separated into the cost of building a realization of the simulated random field and an additive cost per evaluation. For the Fourier-wavelet method and the direct simulation method, such a decomposition is less clear. Certain aspects of computational cost of the multiscale Gaussian random field simulation algorithms have been considered previously [7,3], 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 are needed [3], but it could play a more important role in dynamical simulations where the energy spectrum evolves in time. We therefore do include a brief consideration of the preprocessing cost.

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

$$E(k) = \begin{cases} C_E |k|^{-\alpha}, & |k| \geq k_0; \\ 0, & |k| < k_0, \end{cases} \quad (15)$$

with  $1 < \alpha < 3$ . This random field has correlation length  $\ell_c = (\alpha - 1)/(4k_0)$  and  $\ell_s = 0$ ; the latter statement follows formally by introducing a high-wavenumber cutoff and noticing that the expression in (6) vanishes as the cutoff is removed. In numerical calculations we choose specifically  $\alpha = 5/3$  (corresponding to the Kolmogorov spectrum for the inertial range of a turbulent flow [15,37,29]),  $k_0 = 1$ , and  $C_E = 1$ .

To assess the basic quality of the simulated fractal random fields, we shall use 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. The second order structure function is related to the spectral density by the formula [49]

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

Note that the main contribution to the structure function  $D(\rho)$  at a given value of  $\rho$  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

$$J_\alpha = 4C_E \int_0^\infty k^{-\alpha} [1 - \cos(2\pi k)] dk = \begin{cases} -\frac{2^{1+\alpha}}{\pi^{1-\alpha}} \Gamma(1 - \alpha) \sin(\alpha \pi/2), & 1 < \alpha < 3, \alpha \neq 2 \\ 4\pi^2, & \alpha = 2 \end{cases} \quad (16)$$

and  $\Gamma(x)$  for  $x < 0$  is the Gamma function (extended by analytical continuation) [18,28]. This induced power-law scaling in the structure function is most clearly seen by rewriting the structure function in the form

$$D(\rho) = \left[ J_\alpha - 4 \int_0^{\rho k_0} (1 - \cos(2\pi k')) k'^{-\alpha} dk' \right] \rho^{\alpha-1}. \quad (17)$$

To display more clearly the accuracy of the simulation methods in replicating the correct scaling (17) of the structure function, we will look at a rescaled form [7]

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

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 (15) by observing over how many decades the function  $G_2(\rho)$  remains near the constant value 1.

As discussed in Section 2, we choose parameters in a random field simulation method based on the maximum and minimum length scales ( $\ell_{\max}$ ,  $\ell_{\min}$ ) 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:

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

### 5.1 Random Field Simulations with Evaluations at Irregular Locations

We begin by considering the cost of simulating a Gaussian random field at a collection of  $N_e$  points which may be specified in advance or on demand, and are not assumed to have any regular distribution. We briefly consider the costs of a direct 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 Direct Simulation method

One generic approach to simulating a Gaussian random field at a set of points is to recognize that the values of the random field at these points form a mean zero, jointly Gaussian collection of random variables [49]. The covariance matrix for these Gaussian random variables is obtained from evaluating the correlation function at the relative displacements between each pair of points. If the set of evaluation points is specified in advance, then the simulation task reduces to a preprocessing step of calculating the covariance matrix and its square root (through a Cholesky decomposition), followed by a multiplication of this matrix square root by a vector of standard independent Gaussian random variables for each realization. Performing this procedure for the simulation of the random field at  $N_e$  pre-specified points would incur a preprocessing cost which scales as  $N_e^3$  and a cost per realization scaling as  $N_e^2$ . The preprocessing step would have to be repeated each time the points of evaluation are changed.

Unlike the multiscale simulation procedures discussed in the present work, the direct simulation procedure must be fundamentally modified if the evaluation points are specified on demand. The value of the random field at each new evaluation point is then given by a Gaussian random variable with mean and variance conditioned upon the values of the random field at the previous evaluation points. (This statement remains true even if the location of the new evaluation point is a random variable depending on the previous evaluations of the Gaussian random field and possibly some additional independent random variables). The simulation of the new random field value is therefore equivalent to a standard regression with respect to the previously simulated data [14]. Implementing this regression 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. The cost is dominated by the calculation of the regression coefficients [14].

We therefore see that in the generic case, the cost of the direct Gaussian random field simulation approach is a superlinear power law in the number of evaluation points  $N_e$ . One may be able to reduce the costs of this direct simulation approach somewhat through exploiting information about the particular arrangement of the evaluation points; we consider in particular the case of a regular computational grid in Subsubsection 5.2.1.

### 5.1.2 Randomization Method

**5.1.2.1 Choice of Numerical Parameters** In our numerical example (15), we will choose  $\ell_{\max} = k_0^{-1}$  (in general it should be comparable to the correlation length  $\ell_c$ ). 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  $N_{\text{dec}}$ , 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  $\ell_{\text{min}}$  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  $(1/\ell_{\text{min}}, \infty)$  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  $\ell_{\text{min}}$ .”

For our example (15), the average number of wave numbers in the interval  $(1/\ell_{\text{min}}, \infty)$  is equal to  $n_0(k_0\ell_{\text{min}})^{\alpha-1}$ ; then from  $n_0(k_0\ell_{\text{min}})^{\alpha-1} = c$  we get

$$\ell_{\text{min}} = \frac{1}{k_0} \left( \frac{n_0}{c} \right)^{-\frac{1}{\alpha-1}}. \quad (18)$$

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 Figure 1 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  $\ell_{\text{min}}$  is assumed to fall in the wavenumber bin with the highest wavenumbers), we obtain

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

and consequently  $\ell_{\text{min}} = \frac{1}{k_0} (nn_0/c)^{-\frac{1}{\alpha-1}}$ . The number of decades of accurate

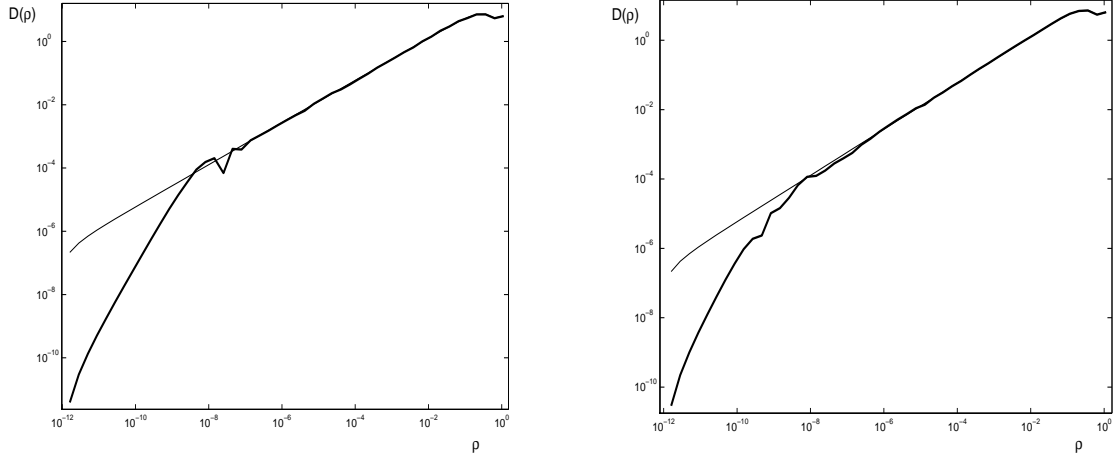


Fig. 1. The structure function  $D(\rho)$ : exact formula (thin solid line) and calculated by averaging over an ensemble of  $N_s = 2000$  Monte Carlo samples simulated by variant A of the Randomization method (bold solid line). Number of wavenumbers:  $n_0 = 160$  (left panel) and  $n_0 = 1000$  (right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

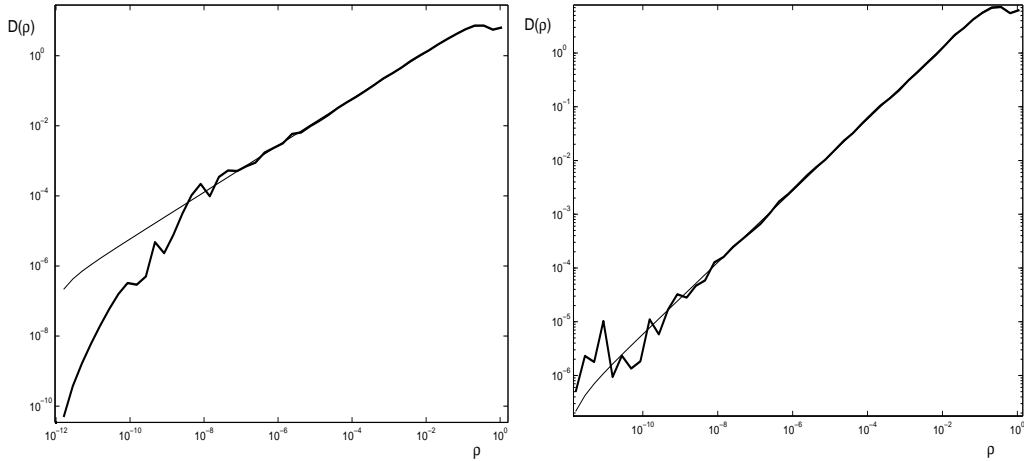


Fig. 2. The structure function  $D(\rho)$ : exact formula (thin solid line) and calculated by averaging over an ensemble of  $N_s = 2000$  Monte Carlo samples simulated by variant B of the Randomization method (bold solid line). Number of sampling bins:  $n = 160$  (left panel) and  $n = 1000$  (right panel);  $n_0 = 1$  wavenumber per bin in both cases.

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 2 and 3.

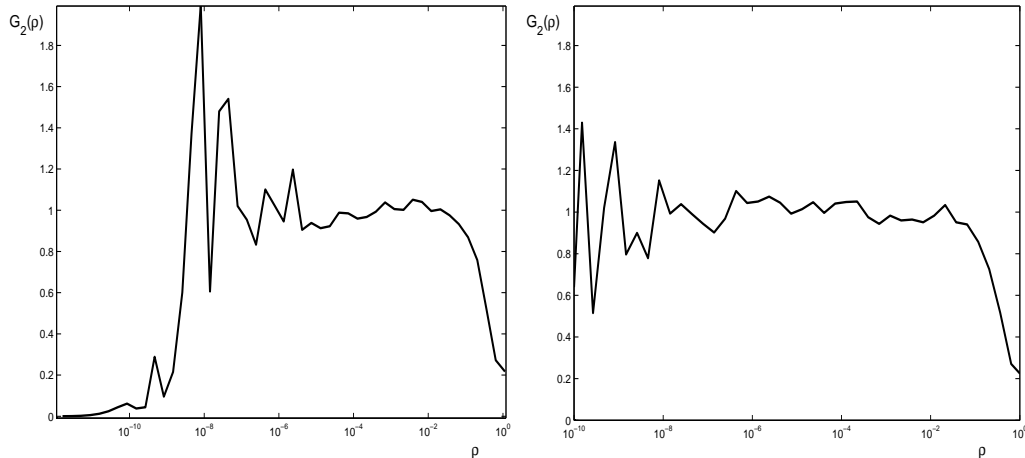


Fig. 3. The normalized structure function  $G_2(\rho) = \frac{D(\rho)}{J_{5/3} \rho^{2/3}}$  calculated by variant B of the Randomization method with  $N_s = 2000$  samples. Number of sampling bins:  $n = 160$  (left panel) and  $n = 1000$  (right panel);  $n_0 = 1$  wavenumber per bin in both cases.

**Variante 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:

$$N_{\text{dec}} = \frac{1}{\alpha - 1} \log_{10} \frac{n_0}{c} + (n - 1) \log_{10} q. \quad (19)$$

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 panels of Figures 4 and 5. 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  $\rho$  is also well calculated by the Randomization method with the logarithmic wavenumber subdivision (Figure 6).

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  $\delta u_i = \frac{u(\rho_i) - u(0)}{(J_\alpha \rho_i^{\alpha-1})^{1/2}}$ ,  $i = 1, \dots, n_p$ ,

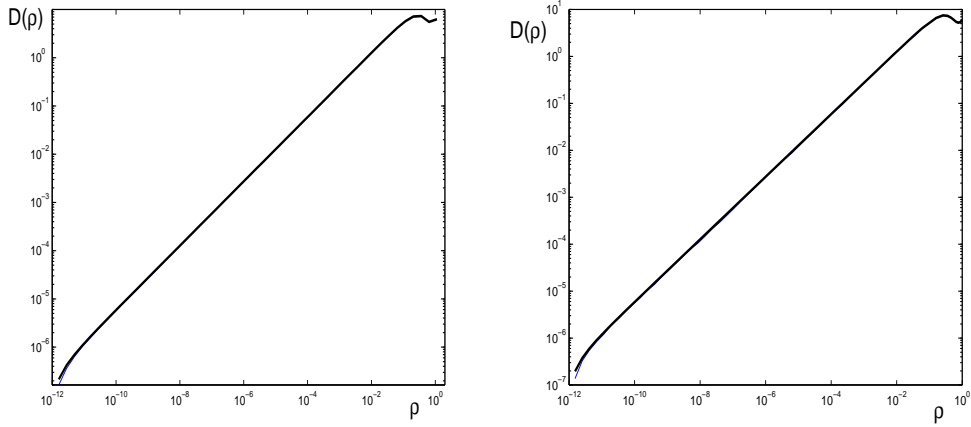


Fig. 4. Comparison of structure function  $D(\rho)$  as simulated by: variant C of the Randomization method (left panel) with 160 wavenumbers ( $n_0 = 4$  samples from each of  $n = 40$  bins,  $q = 2$ ) and  $N_s = 4000$  Monte Carlo samples; and by the Fourier-wavelet method (right panel) with  $M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

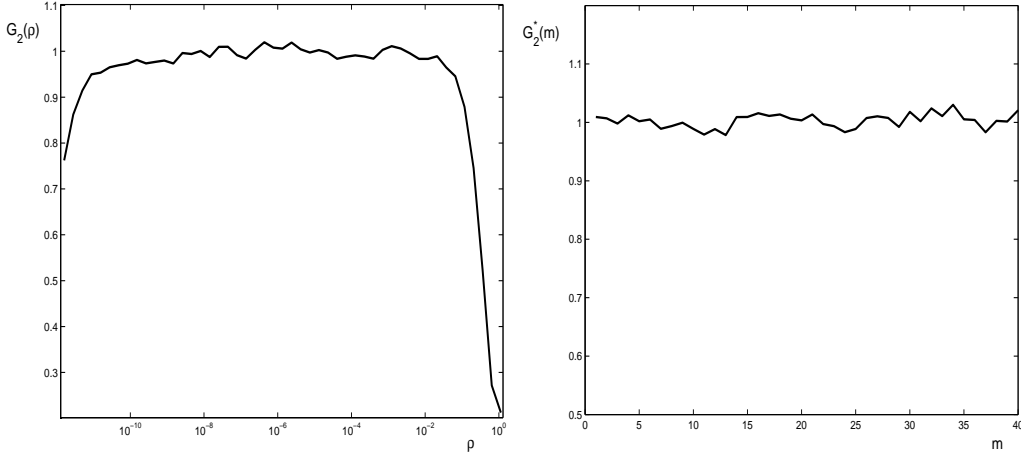


Fig. 5. The normalized structure function  $G_2(\rho) = \frac{D(\rho)}{J_{5/3} \rho^{2/3}}$  calculated by: variant C of the Randomization method (left panel) with 160 wavenumbers ( $n_0 = 4$  samples from each of  $n = 40$  bins), bin ratio  $q = 2$ , and  $N_s = 20000$  Monte Carlo samples; and direct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

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

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

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 5, represent a Monte Carlo estimate  $\hat{G}_2(\rho_m) = \langle (\delta u_m)^2 \rangle$  of the structure function which only ex-

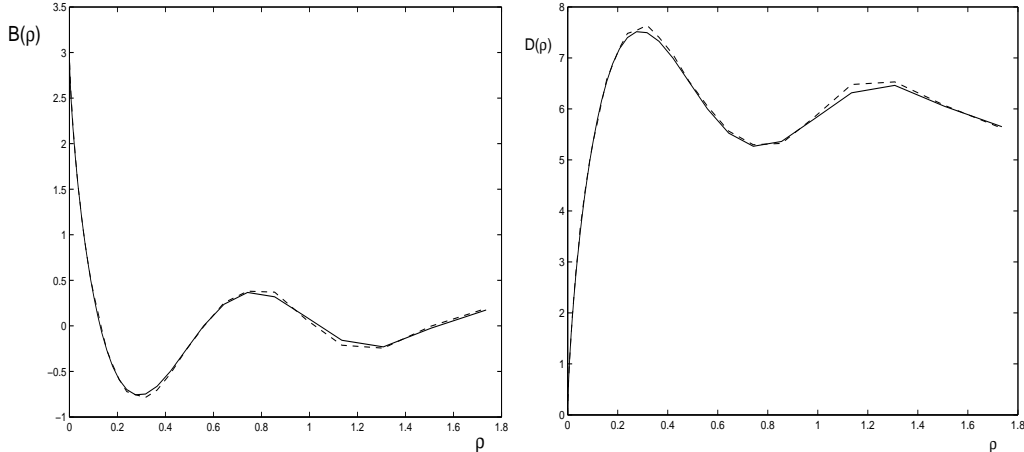


Fig. 6. The correlation function (left panel) and structure function (right panel) simulated by variant C of the Randomization method (dashed line) with  $n = 25$  bins,  $q = 3.16$ ,  $n_0 = 10$  wavenumbers per bin, and  $N_s = 16000$  Monte Carlo samples. The solid line represents the exact formula.

hibits sampling error (and the error of the asymptotic approximation in the previous sentence). 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 5 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

$$G_4(\rho) = \frac{\langle (u(\rho) - u(0))^4 \rangle}{\langle (u(\rho) - u(0))^2 \rangle^2}; \quad (20)$$

see the plots in Figure 7.

From our exploration of the multiscale random field with spectral density (15), we have found that the Randomization method can be made much more efficient by using 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, but only that it appears to improve greatly the efficiency of the Randomization method relative to an equal-energy subdivision. Nor do we take the relation (19) too seriously by, for example, optimizing it with respect to the numerical parameters. This would lead to silly strategies because the formula (19) does not take into account the need to adequately sample wavenumbers *throughout* the range of scales from  $\ell_{\min}$  to  $\ell_{\max}$ . Our

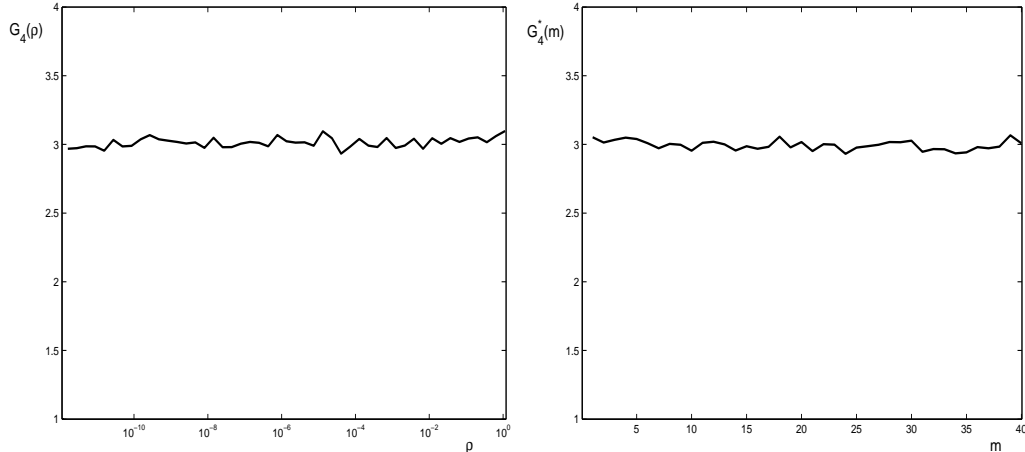


Fig. 7. Kurtosis  $G_4(\rho)$  calculated by: variant C of the Randomization method (left panel) with  $n = 40$  bins, bin ratio  $q = 2$ , and  $n_0 = 4$  wavenumbers per bin; and direct Monte Carlo simulation (right panel).  $N_s = 20000$  samples are used in each case.

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 (19), along with the numerical results in Figures 4–7 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 and kurtosis. 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 an irregularly situated set of points, 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, based on the results presented so far, we can only expect these choices of parameters to be adequate insofar as accurate simulation of two-point statistics (such as the structure function (4) and two-point kurtosis (20)) 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 pre-

sented 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  $\ell_c/\ell_s$ . 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 strongly on the length scales of the random field.

**5.1.2.2 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 sampling 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.

**5.1.2.3 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  $n_0n$ . The evaluation of the simulated random field at each desired point is accomplished by straightforward summation of the Fourier series approximation (8), with a cost proportional to  $nn_0$ , the number of terms in the sum. Therefore, the total cost in generating a realization of the random field at  $N_e$  irregularly spaced points scales as  $N_en_0n$ .

**5.1.2.4 Summary of Cost Considerations** The preprocessing cost is proportional to the number of sampling bins  $n$ , while the cost per realization is proportional to the total number of wavenumbers sampled and number of evaluations,  $n_0 n N_e$ . 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

**5.1.3.1 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 (15), 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 (14), 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  $p$ th order perfect B-spline [7]. 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 (14)), 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 [7], we choose  $p = 2$  and  $b = 10$ .

Finally, we must choose a finite spacing  $\Delta\xi$  between the points  $\xi = \xi_j = -b + (j - 1)\Delta\xi$ ,  $j = 1, \dots, 2b/\Delta\xi + 1$  at which the functions  $f_m(\xi)$  are numerically evaluated through their Fourier integrals (10). 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 (14). 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.

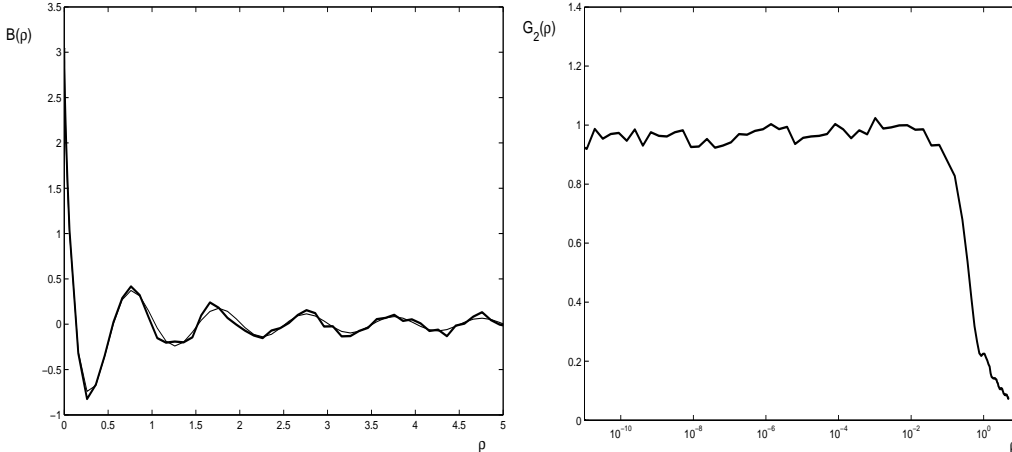


Fig. 8. The correlation function (left panel) and normalized structure function  $G_2(\rho)$  (right panel) for the spectrum (15). The bold line indicates the simulated results using a Fourier-wavelet method with  $N_s = 4000$  Monte Carlo samples,  $M = 40$  scales, and  $b = 10$ . The thin line in the left panel represents the exact formula.

An example of the correlation function and normalized structure function  $G_2(\rho)$  for the energy spectrum (15) as simulated by the Fourier-wavelet method with  $M = 40$  scales and  $b = 10$  is shown in Figure 8. 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. In [10], a fixed number  $N_a$  of plane waves (depending on dimension but not on the length scales of the random field) is found to be adequate to ensure a desired approximation to isotropy of the simulated random field. As discussed in Subsubsection 5.1.2.1, there may be more complex multi-dimensional statistics that require  $N_a$  to increase with  $\ell_c/\ell_s$ , 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 points.

**5.1.3.2 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 evaluation of the Fourier transforms (10). 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  $M(b/\Delta\xi) \log_2(b/\Delta\xi)$ . The extra preprocessing cost in the extension to multiple dimensions through plane wave superposition is negligible because the same functions are involved.

**5.1.3.3 Cost per Realization** We first consider the one-dimensional case, in which the random field is to be evaluated directly on an irregularly situated set of points.

One key observation that distinguishes the Fourier-wavelet method from the Randomization method as well as the direct simulation method discussed in Subsubsection 5.1.1 is that the evaluation of random field at a point does not involve a summation over all the computational elements and associated random numbers. Rather, because of the good localization properties of the wavelet basis, one need only sum at each scale over a fixed number ( $2b$ ) of wavelets which are situated closest to the point of evaluation. Consequently, one has two options with the Fourier-wavelet method:

- Simulate at the beginning of the calculation the complete random field representation over the whole computational domain, then evaluate this random field representation 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, though it requires the ability to store a large number of random variables. Later we will remark on how one may be able to reduce computational cost and memory requirements 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.

**Pre-computation of all random variables:** We can estimate the amount of work needed to pre-compute all coefficients needed for evaluations on a one-dimensional domain of length  $L$  by first fixing an index  $0 \leq m \leq M$  (which fixes a length scale  $\ell 2^{-m}$ ), and noting that we must simulate and store  $\gamma_{mj}$  for  $2^m L/\ell + 2b$  different indices  $j$  in (9), so that the sum (14) can be accurately evaluated at any value of  $x$  in the domain. 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 one realization of all coefficients of the one-dimensional random field representation comprehensively (to the specified level of accuracy) over a domain of length  $L$  is

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

The evaluation of the random field at a given point involves the calculation of a sum of the form (14). This involves interpolation to evaluate the functions  $f_m$  at the indicated values, and a summation over the indicated wavelets with their associated random numbers (already simulated). The cost of this evaluation step is proportional to  $bMN_e$ . For a random field to be simulated over a

multi-dimensional domain with length scale  $L$ , the costs are multiplied by the number  $N_a$  of one-dimensional random fields used in the Radon plane wave decomposition.

Consequently, the total cost for simulating one realization of the random field at a set of  $N_e$  irregularly situated points should scale as  $N_a (2^M L/\ell + bMN_e)$  if all random variables in the random field representation are calculated in advance of evaluation at the desired points.

**Evaluation of Random Variables as Needed** One can cut the memory and run-time costs if the random variables in the Fourier-wavelet random field representation are only simulated as needed for evaluation [11,10,9]. The cost savings would be most dramatic in a situation where the random field is to be evaluated over a sparsely distributed set of points (which may still be large in number). In particular, one can apply this approach without specifying the computational domain in advance. In this strategy, however, one must be careful with managing the random numbers  $\gamma_{mj}$  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  $\gamma_{mj}$  appearing in an evaluation needs to be generated or recalled from a previous generation. Alternatively, one can use explicitly the structure of a reversible pseudo-random number generator to simulate all random variables as needed, maintaining the identities of random variables already realized, without actually storing them [9]. 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  $\gamma_{mj}$ . The run-time cost of simulating the random field at an irregularly situated set of  $N_e$  points should then simply scale with the cost of evaluating the sums (14), which scales as  $N_a N_e bM$ .

**5.1.3.4 Summary of Cost Considerations** The preprocessing cost of the Fourier-wavelet method is proportional to the quantity  $M(b/\Delta\xi) \log_2(b/\Delta\xi)$ , while the cost per realization over an irregularly distributed set of  $N_e$  points is proportional to  $N_a(2^M L/\ell + bMN_e)$  if all random variables in the random field representation are computed in advance of evaluation, or simply proportional to  $N_a MbN_e$  if the random variables are simulated only as needed and managed by a sufficiently sophisticated algorithm. The preprocessing cost appears negligible relative to the cost per realization when  $N_e$  is large. We recall that the number of scales  $M$  in the Fourier-wavelet representation is related to the range of scales in the random field by  $\ell_{\max}/\ell_{\min} = 2^{M-1}$ .

#### 5.1.4 Comparison of Costs

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 found rather simply 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 determined by the number of wavenumbers that should be simulated per decade to provide sufficient statistical accuracy. To simulate statistics involving a small number of points accurately, this prefactor is on the order of 10.

The cost of simulating a multiscale Gaussian random field with the Fourier-wavelet method appears usually to be greater. If the random variables in the Fourier-wavelet representation are simulated only as needed for evaluation, the cost scales nominally with  $bN_e N_a \log_2(\ell_{\max}/\ell_{\min})$ . 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  $\gamma_{mj}$  in the expansion (9), 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  $N_a L/\ell_{\min} + bN_a N_e \log_2(\ell_{\max}/\ell_{\min})$ . 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.

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 pre-computing all random variables associated to the computational domain, the cost of the Fourier-wavelet method scales as  $N_a L/\ell_{\min} + N_a b N_e \log_2(\ell_c/\ell_{\min})$ , 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  $N_e$ .

We observe, then, that the Randomization method with logarithmically uniform subdivision can be expected to simulate a random field with accurate

two-point statistics with less expense than the Fourier-wavelet method. The reason for the reduced cost of the Randomization method is easily traced to its use of a smaller set of computational elements. To simulate the random field structure at each length scale  $2^{-m}\ell_{\max}$ , the Randomization method uses a fixed number  $n_0$  of wavenumbers, while the Fourier-wavelet method uses an increasing number  $2^m L/\ell_{\max} + 2b$  of wavelets at smaller scales (larger  $m$ ). The Randomization method has the flexibility in design in allowing the number of wavenumbers per sampling bin  $n_0$  and therefore the number  $n_0 n$  of computational elements to be chosen according to the statistical accuracy requirements. The Fourier-wavelet method, by contrast, really requires reference to a complete set of  $\sum_{m=0}^{M-1} (2^m L/\ell_{\max} + 2b) \approx 2^M L/\ell_{\max} + 2Mb = 2(L/\ell_{\min} + b \log_2(\ell_{\max}/\ell_{\min}))$  wavelets and associated random numbers to represent the random field meaningfully. The numerical parameters governing statistical accuracy in the Fourier-wavelet method relate to the number of terms used in evaluating the random field at a given location.

We remark that in one dimension (so that  $N_a = 1$ ), the standard approach described in Subsection 5.1 would require  $L/h$  random variables to represent the random field over a domain  $L$  with computational grid spacing  $h$ . The Fourier-wavelet method would use approximately the same number of random variables if  $h \geq \ell_s$ , so that the random field has structure all the way down to the grid scale and  $\ell_{\min} = h$ . If  $h < \frac{1}{2}\ell_s$ , then the Fourier-wavelet method would be using a smaller number of computational elements than the direct approach because the random field structure on length scales smaller than  $\ell_s$  can be obtained accurately by interpolation (without the need for additional random numbers) from the smoothness length scale  $\ell_s$  of the random field. The Fourier-wavelet method, therefore, has the number of computational elements (and associated random variables) set essentially by theoretical considerations of how many degrees of freedom of randomness are needed to represent effectively a random field over a computational domain.

The Randomization method, by contrast, has the number of its random variables and computational elements (wavenumbers) set by user specification. In particular, in an accurate simulation over nine decades of the second order structure function of the random field described by (17), the Randomization method used only 480 random variables to represent the random field while the Fourier-wavelet method used, in principle,  $2^{42}$  random variables. Apparently, the smaller number of random variables is sufficient for the random field simulated by the Randomization method to exhibit a good fidelity in the second order structure function as well as the kurtosis of the random field increment between two points (20). An interesting question is to what extent the Randomization method can exploit its flexibility in design and use of a relatively small number of computational elements and random variables to represent the random field over a computational domain when more complex statistics of the random field are important.

We will see through numerical examples in the next few sections that the Randomization method must increase the number of wavenumbers  $n_0$  simulated per scale as accuracy is desired in statistical characteristics involving a greater number of points. By contrast, the Fourier-wavelet method appears to simulate these multi-point statistics accurately with the same choice of parameters as was used to simulate the second order structure function accurately (Figure 4). We will find that the Randomization method remains competitive for the simulation of statistics involving tens of points, but that the Fourier-wavelet method becomes more efficient in simulating statistics involving a greater number of points. In particular, the Fourier-wavelet method is more efficient in simulating a random field with good ergodic properties (so that spatial averages over the whole domain approximate ensemble statistics).

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 Regular Grid over Pre-Specified Domain*

We now discuss the relative costs of the Randomization and Fourier-wavelet methods, as well as the direct simulation approach, when the random field is to be simulated on a pre-specified regular grid of points with spacing  $h$  and domain length  $L$ . One may think of a uniform cubic grid on a cubic domain, but our cost scaling considerations apply more generally to any domain which can be described by a single length scale  $L$  and grid configuration which covers the domain with spacing between neighbors characterized by a single length scale  $h$ . In this subsection, we only point out how the cost scalings change and compare in a relative sense due to the regular arrangement of evaluation points. This discussion is also intended to cover the case in which the random field is to be evaluated at a dense set of irregularly spaced points through pre-computation over a regular grid followed by interpolation. Our considerations of course apply to the pre-computation over the regular grid; the subsequent interpolation is independent of the random field simulation algorithm and is clearly linear in the number of evaluation points.

### 5.2.1 *Direct 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 on the computational

grid scales with  $(L/h)^d$ , so the preprocessing cost from a naive implementation would scale as  $(L/h)^{3d}$ , and the cost to simulate one realization of the random field over the grid would scale as  $(L/h)^{2d}$ .

However, when the domain length  $L$  is large compared with the correlation length  $\ell_c$ , one can easily reduce the cost by neglecting the correlations between points separated by a distance large compared with the correlation length  $\ell_c$ . This gives the covariance matrix a banded structure, reducing the computational linear algebra costs so that the preprocessing cost scales with  $(L/h)^d(\ell_c/h)^{2d}$  and the cost per realization 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 a 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 simulate directly all the random variables  $\gamma_{mj}$  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  $N_a 2^M L/\ell + bMN_a(L/h)^d$ . Fast wavelet transform methods (filter bank algorithms) [31] have the potential of reducing the cost even more, in analogy to fast Fourier transforms, but we are unaware of any actual implementations in the random field context. 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  $(L/h)N_a + bMN_a(L/h)^d \sim bMN_a(L/h)^d$ . The cost scaling would be the same if the random variables were computed only as needed, and the random variable management would be considerably easier for evaluation on a regular computational grid than in the case of irregularly distributed points (as discussed in

Subsubsection 5.1.3).

#### 5.2.4 Comparison of Costs

With the dense sampling of the random field implicit in simulating the random field over a regular 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  $(\ell_c/h)$ ), and some numerical implementation parameters. The numerical prefactors appear to be smaller for the Randomization method. In particular, a 1%-accuracy could be achieved by the Randomization method ( $n = 40$ ,  $n_0 = 2$ ,  $q = 2$ ) over 9 decades of the simulated structure function (see Figure 4) with approximately 12 times less computer time than the Fourier-wavelet method with parameters ( $M = 40$ ,  $b = 10$ ) specified as in [7]. The structure function was evaluated over a complete regular spatial grid by the dependent sampling technique. 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.

We conclude our theoretical cost considerations with a brief inference about the expected relative costs of the methods in partially regular situations, where the evaluation points are not distributed over a regular computational grid but do have some spatial structure that can be exploited to reduce computational costs from the generic irregularly distributed case discussed in Subsection 5.1. We interpolate our conclusions from the estimates for the irregularly and regularly distributed cases discussed above. The direct simulation approach should generally have cost scaling as a superlinear power law in  $N_e$ ,  $(L/h)^d$ , and  $(\ell_c/h)^d$ . The Randomization and Fourier-wavelet methods, by contrast, should have costs scaling linearly in  $N_e$  and logarithmically in  $\ell_c/h$ . This cost-scaling for the Fourier-wavelet method applies when the random variables in its random field representation are only computed as needed (and therefore managed in a careful way). A simpler implementation of the Fourier-wavelet method in which all the random variables are pre-computed and stored will have a cost scaling linearly with  $N_e$  and  $(L/h)^d$ , and logarithmically with  $\ell_{\max}/\ell_{\min}$ , the ratio of the maximal and minimal scales of the simulated random field. This would generally be considerably more expensive than the Randomization method except in situations where the number of evaluation points  $N_e$  is comparable to  $(L/h)^d$  (as for the case of a regular computational grid). Finally, both the Randomization method and Fourier-wavelet method are more efficient than the direct simulation method when simulating a multiscale random field ( $(\ell_c/h)^d \gg 1$ ) at a large number of points  $N_e \gg 1$ , because the

direct approach has cost scaling as a superlinear power law with respect to these parameters, while the multiscale methods scale as products of linear and logarithmic functions of these 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 (2) for a given realization of the conductivity, for example, is a rather time-consuming computational procedure. Computation of statistical flow properties through spatial rather than ensemble averages would therefore improve efficiency if they could be calculated accurately.

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:

$$\begin{aligned} B(\rho) &= \langle u(\rho + x)u(x) \rangle, \\ D(\rho) &= \langle (u(\rho + x) - u(x))^2 \rangle. \end{aligned} \tag{21}$$

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:

$$\begin{aligned} B_{N_b}(\rho) &\equiv \frac{1}{N_b} \sum_{j=1}^{N_b} u(\rho + (j-1)\ell_b) u((j-1)\ell_b) , \\ D_{N_b}(\rho) &\equiv \frac{1}{N_b} \sum_{j=1}^{N_b} [u(\rho + (j-1)\ell_b) - u((j-1)\ell_b)]^2 , \end{aligned} \tag{22}$$

converge to the ensemble-averaged expressions  $B(\rho)$  and  $D(\rho)$  as the number of spatial samples  $N_b$  is taken large. Here  $\ell_b$  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

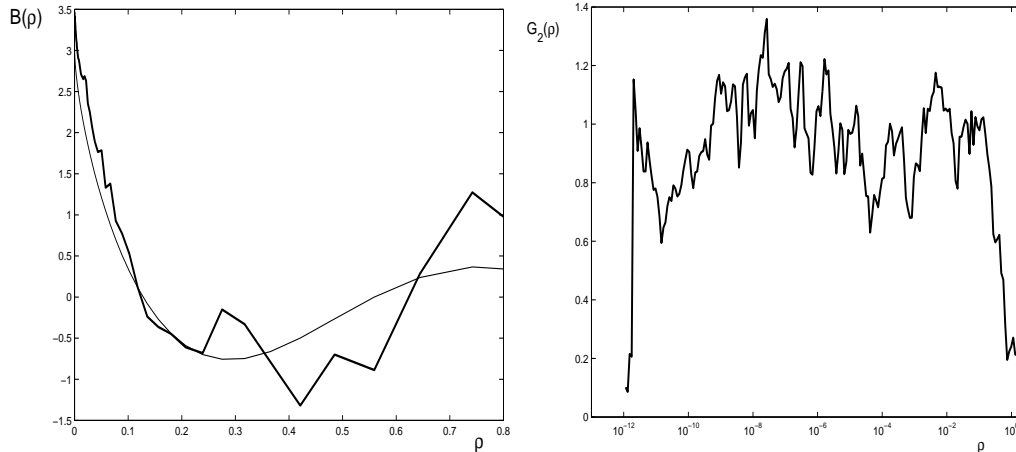


Fig. 9. Ergodic averaging for the correlation function  $B(\rho)$  (left panel) and normalized structure function  $G_2(\rho)$  (right panel) for the Kolmogorov spectrum (15) over  $N_b = 16000$  spatial blocks using variant C of the Randomization method with  $n = 25$  sampling bins,  $n_0 = 10$  wavenumbers per bin, bin ratio  $q = 3.16$ , and  $\ell_b = 2\ell_{\max}$ . In the left panel, the thin solid line denotes the exact ensemble average (21), while the bold solid line denotes the simulated average over spatial translations (22). In the right panel, the normalized structure function should be the constant value 1.

$G_2(\rho) = D(\rho)/(J_\alpha \rho^{\alpha-1})$  against the exact results for the energy spectrum (15).

### 6.1 Randomization method

In Figure 9, 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 10 and 11).

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 spectral function

$$E(k) = \begin{cases} C_E |k|^{-\alpha}, & k_0 \leq |k| \leq k_{\max}; \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

with  $\alpha = 5/3$ ,  $C_E = 1$ ,  $k_0 = 1$ , and  $k_{\max} = 8$ . We present spatially averaged statistics from a Randomization method simulation of the random field with this spectral function (23) in Figure 12. Here the bins are constructed by subdividing the wavenumber range  $k_0 < k < k_{\max}$  into  $n = 3$  logarithmically uniform subintervals. In the left panel,  $n_0 = 10$  random wavenumbers are

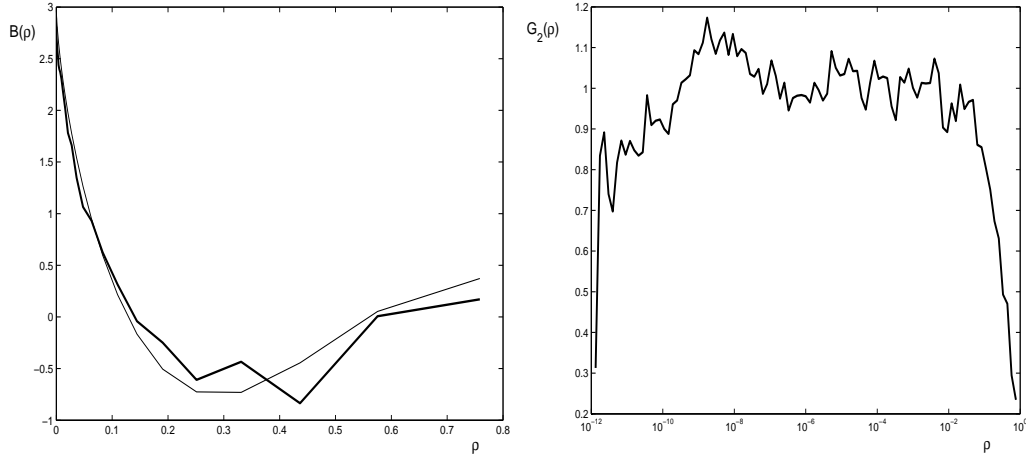


Fig. 10. Same as in Figure 9, but the spatial averaging is now over  $N_b = 16000$  blocks and  $n_0 = 50$  wavenumbers are simulated per bin.

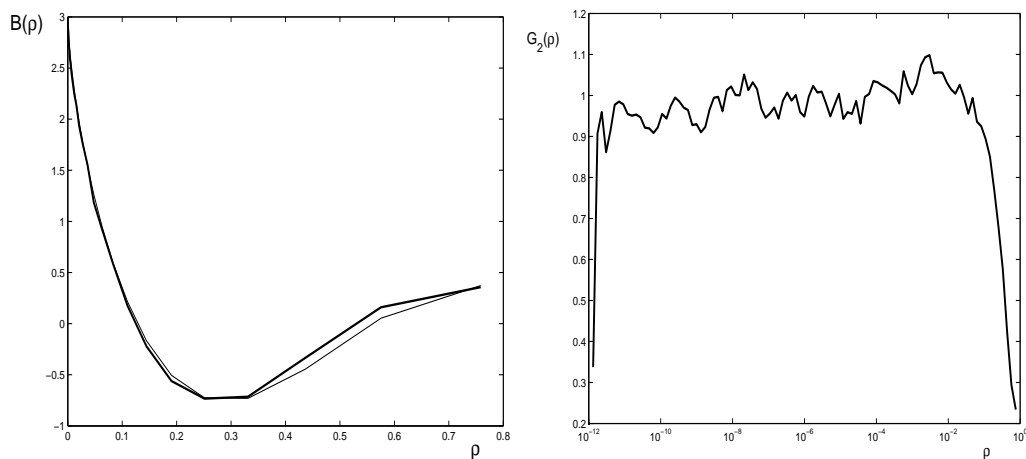


Fig. 11. Same as in Figure 10, but with  $n_0 = 250$  wavenumbers per bin.

sampled in each bin; note that even averages over  $N_b = 16000$  spatial blocks exhibit large deviations from the true statistics. Increasing the number of wavenumbers per bin to  $n_0 = 200$  (see the right panel) improves the results.

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 13, 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 (15). With only  $N_b = 2000$  spatial blocks, the accuracy

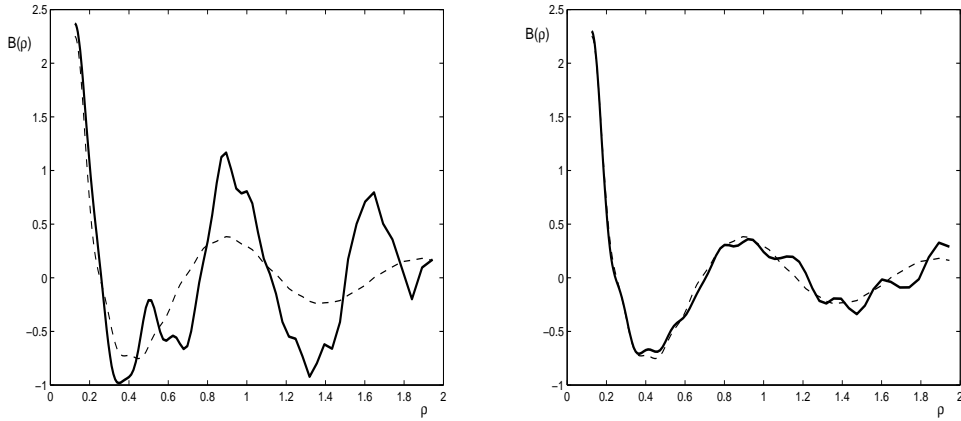


Fig. 12. Ergodic averaging for the correlation function  $B(\rho)$  with spectral function (23) using  $N_b = 16000$  spatial blocks and variant C of the Randomization method with  $n = 3$  sampling bins,  $k_0 = 1$ ,  $k_{max} = 8$ , bin ratio  $q = 2$ , and  $\ell_b = 2\ell_{max}$ . In the left panel,  $n_0 = 10$ , and in the right panel  $n_0 = 200$ . The dashed line denotes the exact correlation function corresponding to the spectral function (23), while the bold solid line denotes the simulated average over spatial translations (22).

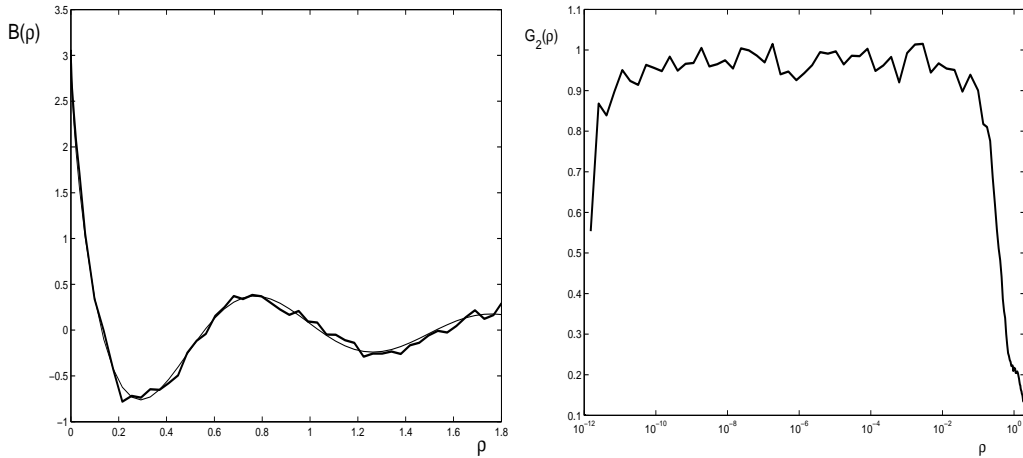


Fig. 13. Ergodic averaging for the correlation function  $B(\rho)$  (left panel) and normalized structure function  $G_2(\rho)$  (right panel) for the Kolmogorov spectrum (15) over  $N_b = 2000$  spatial blocks using Fourier-wavelet method with  $M = 40$  scales,  $b = 10$ , and  $\ell_b = 2\ell_{max}$ . In the left panel, the thin solid line denotes the exact ensemble average (21), while the bold solid line denotes the simulated average over spatial translations (22). In the right panel, the normalized structure function should be the constant value 1.

achieved was approximately the same as in the Randomization method with  $N_b = 16000$  spatial blocks,  $n = 25$ , and  $n_0 = 250$ ; compare Figures 11 and 13.

### 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. In our numerical example, the Randomization method required more than 14 times the computational time ( $n = 40$ ,  $n_0 = 200$ ) as the Fourier-wavelet method (with  $M = 40$  and  $b = 10$ ) to achieve comparable accuracy over 9 decades in the structure function when calculated using spatial averages of a single realization.

## 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 function 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 through ensemble averaging with comparable quality to that simulated by the theoretically Gaussian Fourier-wavelet method. We remark that the consideration of spatial averages (22) 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.

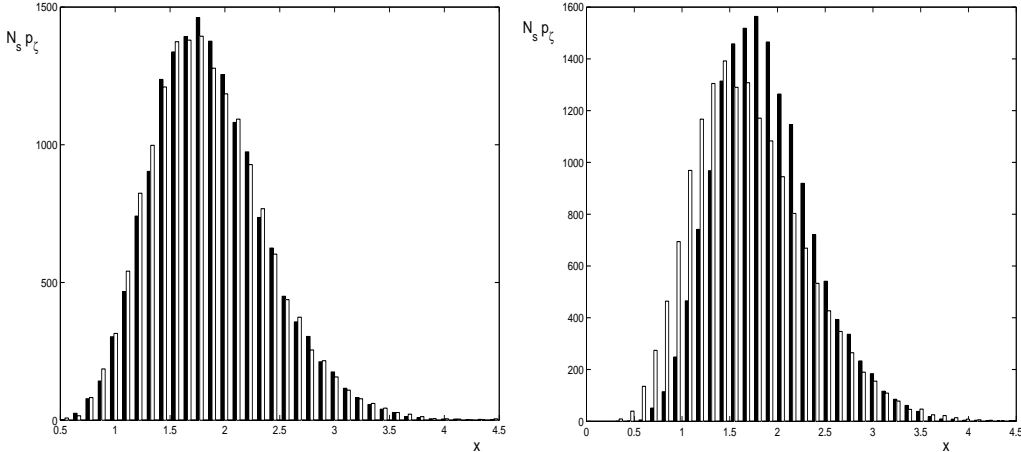


Fig. 14. Histograms  $p_\zeta$  for  $\zeta_{10} = \max_{1 \leq i \leq 10} |\Delta u_i|$  multiplied by  $N_s$ , the number of samples, calculated by variant C of the Randomization method (filled bars), with  $n = 25$  bins,  $N_s = 16000$  samples, and  $h = 1000 \ell_{\min}$  and compared against the 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.

We therefore consider the following sequence of normalized random increments

$$\Delta u_i = \frac{u(ih) - u((i-1)h)}{[J_\alpha h^{\alpha-1}]^{1/2}}, \quad \delta u_i = \frac{u(h q_0^{i-1}) - u(0)}{[J_\alpha (h q_0^{i-1})^{\alpha-1}]^{1/2}}, \quad i = 1, \dots, n'. \quad (24)$$

We introduce the random variable

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

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 (15).

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 5.1.2.1 for the case of the normalized structure function, and its results differs from the exact results only through sampling error.

From Figure 14 it is seen that 25 wavenumbers ( $n = 25$ ,  $n_0 = 1$ ) do not give satisfactory accuracy (right panel) for the statistics of  $\zeta_{10}$  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

$$\mu_m(k) = \left\langle \left( \max_{1 \leq i \leq k} |\Delta u_i| \right)^m \right\rangle, \quad k = 1, \dots, n'. \quad (25)$$

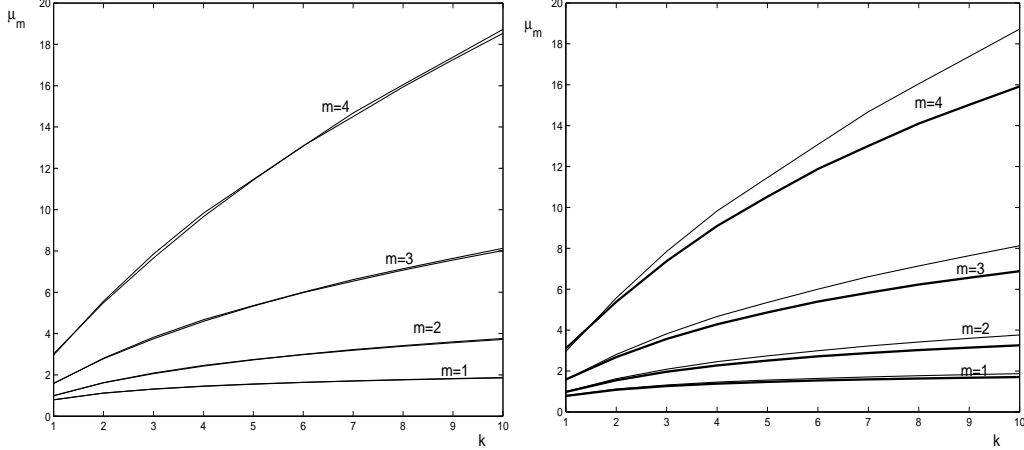


Fig. 15. Statistical moments  $\mu_m(k)$  as simulated from ensemble averages using variant C of the Randomization method (bold line) and compared with direct Monte Carlo simulation (thin line). The parameters are the same as in the corresponding panels of Figure 14.

In Figure 15, 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 14. 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 16 and 17, the statistics of  $\zeta_{10}$  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  $\Delta u_i$ , 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  $\zeta_{10}$  appears to be simulated adequately 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  $\delta u_i$  with geometric spacing (Figure 18).

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 19 for the moments  $\mu_m(k)$  of the random variable  $\zeta_{100}$ . 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. This is illustrated by the results presented in the left panel of Figure 19 where the dashed curves

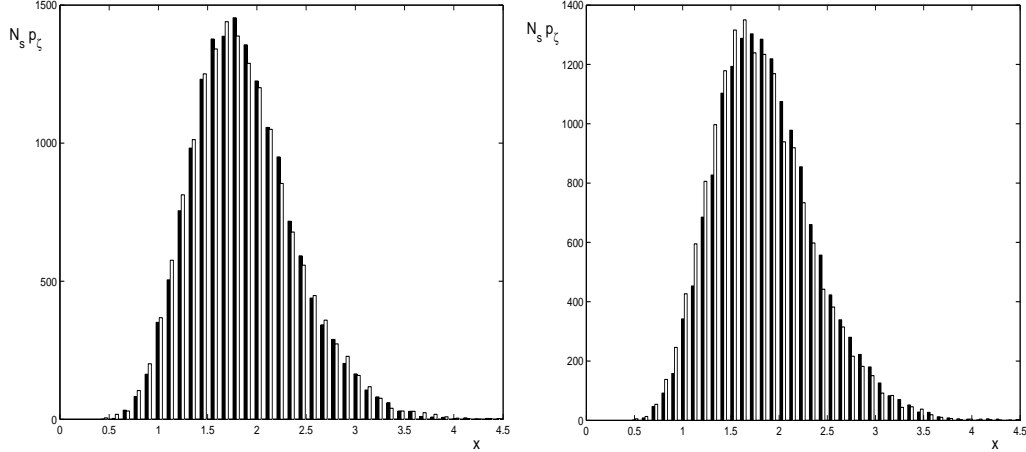


Fig. 16. Histograms  $p_\zeta$  for  $\zeta_{10} = \max_{1 \leq i \leq 10} |\Delta u_i|$  multiplied by the number of samples, calculated by variant C of the Randomization method (filled bars), with  $n = 25$  bins,  $n_0 = 10$  wavenumbers per bin, and  $N_s = 16000$  samples, and compared against the results of direct Monte Carlo simulations (empty bars) (also with 16000 samples). Left panel:  $h = 100 \ell_{\min}$ . Right panel:  $h = 10 \ell_{\min}$ .

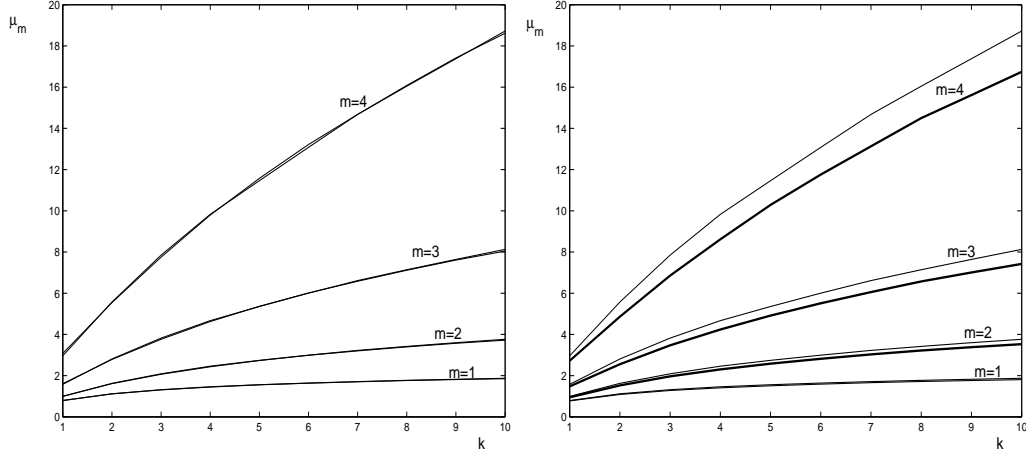


Fig. 17. Statistical moments  $\mu_m(k)$  as simulated from ensemble averages using variant C of the Randomization method (bold line) and compared with direct Monte Carlo simulation (thin line). The parameters are the same as in the corresponding panels of Figure 16.

obtained by the Fourier-wavelet method with  $M = 40$ ,  $b = 10$  practically coincide with the direct Monte Carlo simulation results. Of course the accurate evaluation of more complex statistics will generally require better accuracy of the correlation function.

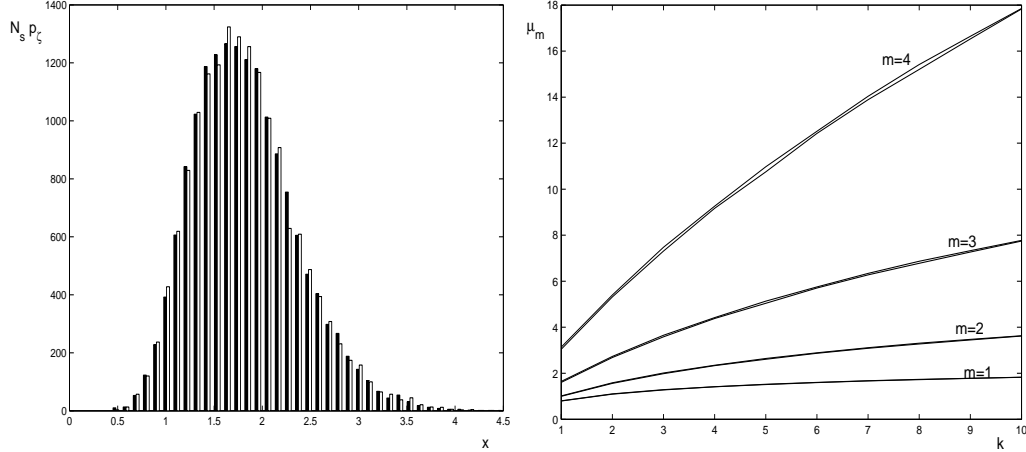


Fig. 18. Left panel: Histogram  $p_\zeta$  for  $\zeta_{10} = \max_{1 \leq i \leq 10} |\delta u_i|$  multiplied by the number of samples, calculated by variant C of the Randomization method (filled bars), compared with the results of direct Monte Carlo simulation (empty bars). Right panel: Statistical moments (defined as  $\mu_m(k)$  in (25) except with  $\delta u_i$  replacing  $\Delta u_i$ ) 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  $h = 1000 \ell_{min}$ . The Randomization method uses a total of 250 wavenumbers ( $n = 25$  bins with  $n_0 = 10$  wavenumbers in each bin) and  $q = 4$ .

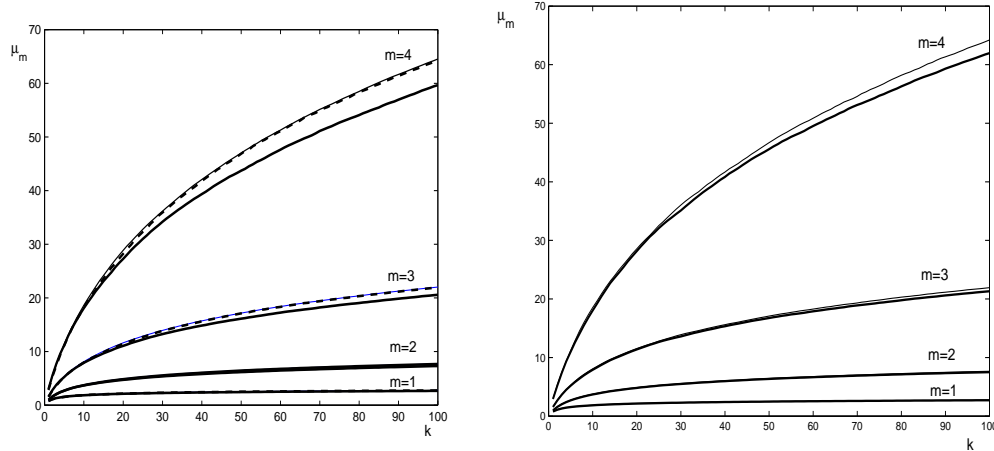


Fig. 19. Statistical moments  $\mu_m(k)$  of  $\zeta_{100}$  with  $h = 100 \ell_{min}$  calculated by variant C of the Randomization method (bold solid line) with  $n = 25$  sampling bins, compared against the results obtained by direct Monte Carlo simulation (thin solid line), and the Fourier-wavelet method ( $M = 40$ ,  $b = 10$ , dashed line in left panel);  $N_s = 16000$  samples are used in all calculations. Left panel:  $n_0 = 10$  wavenumbers per bin; right panel:  $n_0 = 40$  wavenumbers per bin for the Randomization method.

## 8 Conclusions

- (1) The Randomization method is generally easier to implement than the Fourier-wavelet method. For simulating random fields with accurate two-

point statistics, the Randomization method with logarithmically uniform spectral subdivision can often be 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. 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 (pre-computation of all random variables) has cost scaling linearly (and additively) in both the number of evaluations  $N_e$  and the ratio of the largest and smallest scales of the random field simulated,  $\ell_{\max}/\ell_{\min}$ . 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. The reason for the relative efficiency of the Randomization method in simulating two-point statistics accurately is that it appears adequate to use a fixed number  $n_0$  of random variables to represent the random field at each length scale. The Fourier-wavelet method uses, as part of its essential design,  $2^m(L/\ell_{\max})$  random variables to represent the random field structure over a computational domain of linear extent  $L$  at length scale  $2^{-m}\ell_{\max}$ . In particular, the Fourier-wavelet method uses many more random variables to represent the random field structure at smaller scales. 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 regular computational grid of points, because the number of evaluations are comparable to the number of random variables in the Fourier-wavelet representation.

- (2) The cost of the Randomization method increases substantially as statistics involving larger numbers of points are to be simulated accurately. The reason is that the number of computational elements (wavenumbers) per scale,  $n_0$  must be increased to achieve accuracy with these more complex statistics. The Fourier-wavelet method, on the other hand, from the start employs a rich set of computational elements and random variables, and does not require a substantial increase in computational effort to simulate statistics involving large numbers of points 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  $\vec{u}(\vec{x}) = [u_1(\vec{x}), \dots, u_l(\vec{x})]^T$ ,  $\vec{x} \in \mathbb{R}^d$  be a Gaussian homogeneous vector-valued random field with prescribed correlation tensor  $\mathbf{B}(\vec{r}) = \langle \vec{u}(\vec{x} + \vec{r}) \otimes \vec{u}(\vec{x}) \rangle$ , or prescribed spectral tensor  $\mathbf{F}(\vec{k})$ ; the two are related through Fourier transforms:

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

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

The spectral density function is defined

$$E(\vec{k}) = \sum_{j=1}^l F_{jj}(\vec{k}),$$

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

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

$$\mathbf{F}(\vec{k}) = E(\vec{k}) \mathbf{Q}(\vec{k}) \mathbf{Q}^*(\vec{k}), \quad (\text{A.2})$$

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

The spectral density function  $E(\vec{k})$  is a scalar, nonnegative function describing the overall strength of the random field fluctuations at wavenumber  $\vec{k}$ , while  $\mathbf{Q}(\vec{k})$  is a matrix-valued function describing the anisotropy and correlation structure among and along different directions. Further details can be found in [35,49]. In the purely scalar-valued case ( $l = 1$ ),  $\mathbf{F}(\vec{k}) = E(\vec{k})$  and  $\mathbf{Q}(\vec{k}) = 1$ .

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

For each  $1 \leq i \leq n$ , we choose a collection of independent, identically distributed random wavenumbers  $\vec{k}_{i1}, \dots, \vec{k}_{in_0}$  within the set  $\Delta_i$  according to the

probability distribution function

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

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

$$\begin{aligned} \vec{u}^{(R)}(\vec{x}) = \sum_{i=1}^n \frac{\sigma_i}{\sqrt{n_0}} \sum_{j=1}^{n_0} \left\{ \vec{\xi}_{ij} \left[ Q'(\vec{k}_{ij}) \cos \theta_{ij} - Q''(\vec{k}_{ij}) \sin \theta_{ij} \right] \right. \\ \left. + \vec{\eta}_{ij} \left[ Q''(\vec{k}_{ij}) \cos \theta_{ij} + Q'(\vec{k}_{ij}) \sin \theta_{ij} \right] \right\}, \quad (\text{A.4}) \end{aligned}$$

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

### 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  $u_n(\vec{x})$  to a limiting random field  $u(\vec{x})$ , all functions being defined on a compact domain  $K \subset \mathbb{R}^d$ . We assume that the samples of  $u$  are continuous (i.e., belong to  $C(K)$ ) with probability one. The weak convergence  $u_n \rightarrow u$  as  $n \rightarrow \infty$  in  $C(K)$  is defined to mean that for any uniformly bounded continuous functional  $f : C(K) \rightarrow \mathbb{R}$  we have  $\langle f(u_n) \rangle \rightarrow \langle f(u) \rangle$  as  $n \rightarrow \infty$ . 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 [26], it is proved that the random field representation (A.4) used in the Randomization method converges weakly in  $C(K)$  for a fixed spectral subdivision  $n \geq 1$  as  $n_0 \rightarrow \infty$ , provided that

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

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

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

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

We assume that positive constants  $C_0, R_0$  and  $\varepsilon_0 \in (0, 1)$  can be chosen so that either  $\bar{\rho}_i \leq R_0$ , or  $R_0 \leq \bar{\rho}_i \leq C_0(\underline{\rho}_i)^{1+\varepsilon_0}$  for all  $i = 1, \dots, n-1$ . Under these assumptions, together with (A.5), the random field representation (A.4) converges weakly  $\vec{u}^{(R)} \rightarrow \vec{u}$  in  $C(K)$  for each fixed  $n_0 \geq 1$  as  $n \rightarrow \infty$  [25].

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  $\vec{k}_{ij}$  in each  $\Delta_i$  from quite arbitrary probability density functions  $p_i(\vec{k})$  satisfying the consistency condition:  $p_i(\vec{k}) \neq 0$  if  $E(\vec{k}) \geq 0$ . 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  $\Delta_i$ ,  $i = 1, \dots, n$ .

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

$$\begin{aligned} \vec{u}^{(R)} = \sum_{i=1}^n \frac{1}{\sqrt{N_i}} \sum_{j=1}^{N_i} \left( \frac{E(\vec{k}_{ij})}{p_i(\vec{k}_{ij})} \right)^{1/2} \left\{ \vec{\xi}_{ij} \left[ Q'(\vec{k}_{ij}) \cos \theta_{ij} - Q''(\vec{k}_{ij}) \sin \theta_{ij} \right] \right. \\ \left. + \vec{\eta}_{ij} \left[ Q''(\vec{k}_{ij}) \cos \theta_{ij} + Q'(\vec{k}_{ij}) \sin \theta_{ij} \right] \right\}. \end{aligned} \quad (\text{A.6})$$

## B Appendix: Isotropic vector-valued random fields

A homogeneous  $d$ -dimensional vector-valued random field  $\vec{u}(x)$ ,  $\vec{x} \in \mathbb{R}^d$  is called isotropic if the random field  $\mathbf{U}^T \vec{u}(\mathbf{U}\vec{x})$  has the same finite-dimensional distributions as those of the random field  $\vec{u}(\vec{x})$  for any rotation matrix  $\mathbf{U} \in SO(d)$  [10,35,49]. The spectral density tensor of an isotropic random field has

the following general structure [35,49]:

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

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

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

with  $\delta_{ij}$  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:  $\vec{u}(\vec{x}) = \vec{u}^{(1)}(\vec{x}) + \vec{u}^{(2)}(\vec{x})$  where  $\vec{u}^{(1)}$  and  $\vec{u}^{(2)}$  are, respectively, the incompressible and potential parts of  $\vec{u}$  with spectral density tensors

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

respectively.

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

$$\mathbf{F}^{(i)}(\vec{k}) = p_i(\vec{k}) \mathbf{Q}^{(i)} \mathbf{Q}^{(i)*}, \quad (\text{B.4})$$

take the special form

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

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

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

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

One convenient explicit choice in three dimensions is [43]

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

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

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

where  $(\sigma_i^{(1)})^2 = \frac{1}{2} \int_{\Delta_i} p_1(\vec{k}) d\vec{k} = 2 \int_{a_i}^{b_i} E_1(k) dk$ ,  $\vec{\Omega}_{ij}^{(1)}$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, n_0$  is a family of mutually independent random vectors distributed uniformly on the unit sphere in  $\mathbb{R}^3$ ;  $\vec{\xi}_{ij}$  and  $\vec{\eta}_{ij}$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, n_0$  are mutually independent families of three-dimensional standard Gaussian random vectors;  $\theta_{ij}^{(1)} = 2\pi k_{ij}^{(1)} (\vec{\Omega}_{ij}^{(1)} \cdot \vec{x})$ ; and for each  $i = 1, \dots, n$ , the  $k_{ij}^{(1)}$ ,  $j = 1, \dots, n_0$  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 potential component  $\vec{u}^{(2)}$  is simulated in three dimensions through the representation

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

Here, unlike in the previous simulation formula, the  $\xi_{ij}$  and  $\eta_{ij}$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, n_0$  are families of *scalar* standard Gaussian random variables, which are all mutually independent. The remaining inputs are constructed analogously:  $(\sigma_i^{(2)})^2 = \int_{\Delta_i} p_2(\vec{k}) d\vec{k} = 2 \int_{a_i}^{b_i} E_2(k) dk$ ;  $\vec{\Omega}_{ij}^{(2)}$  is a family of mutually independent random vectors distributed uniformly on the unit sphere in  $\mathbb{R}^3$ ;  $\theta_{ij}^{(2)} = 2\pi k_{ij}^{(2)} (\vec{\Omega}_{ij}^{(2)} \cdot \vec{x})$ ; and for each  $i = 1, \dots, n$ , the  $k_{ij}^{(2)}$ ,  $j = 1, \dots, n_0$  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 [10]. Discretization of the angular integral produces the following simulation formula:

$$\vec{u}_N(\vec{x}) = \left(\frac{1}{N_a}\right)^{1/2} \sum_{j=1}^{N_a} \left[ \mathbf{P}^{(1)}(\vec{\Omega}_j) \vec{v}_j^{(1)}(\vec{x} \cdot \vec{\Omega}_j) + \mathbf{P}^{(2)}(\vec{\Omega}_j) \vec{v}_j^{(2)}(\vec{x} \cdot \vec{\Omega}_j) \right],$$

where  $\vec{\Omega}_j$ ,  $j = 1, \dots, N_a$  is a collection of deterministic vectors which discretize the unit sphere in  $\mathbb{R}^d$ . The  $\vec{v}_j^{(1)}(r)$ ,  $\vec{v}_j^{(2)}(r)$ ,  $j = 1, \dots, N_a$ ,  $r \in \mathbb{R}$  are mutually independent Gaussian  $d$ -dimensional random fields. Each component of  $\vec{v}_j^{(1)}(r)$  (respectively  $\vec{v}_j^{(2)}(r)$ ) is an independent Gaussian homogenous random field with spectral density  $E_1(k)$  (respectively  $E_2(k)$ ), 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  $\vec{v}_j^{(1)}$  and 1 component of  $\vec{v}_j^{(2)}$  and rotating them according to the direction  $\vec{\Omega}_j$  [10].

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 or 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 (10) which in our case reads

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

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  $\xi_j = -\frac{N}{2}\Delta\xi + (j-1)\Delta\xi$ ,  $j = 1, \dots, N$ , 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 (14), we must choose  $N\Delta\xi/2 \geq b$ .

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

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

where

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

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

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

hence

$$f_m(\xi_j) \simeq \exp \left\{ \pi i (j-1) \left( 1 - \frac{1}{N} \right) \right\} \sum_{l=1}^N G_l \exp \left\{ -2\pi i \frac{(j-1)(l-1)}{N} \right\}, \quad (\text{C.4})$$

where

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

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 (14). 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  $2b/N \leq \Delta\xi$ . Then we set  $a = \frac{1}{2\Delta\xi}$ , and discretize the integral (C.2) with step size  $\Delta k = 2a/N = 1/(N\Delta\xi)$ .

## References

- [1] P. Billingsley, *Convergence of Probability Measures*. John Wiley (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.* **1**(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, Progr. Prob.* (Birkhäuser Boston, 1996), Vol. **39**, pp. 47–68. .
- [5] D. Cioranescu and P. Donato. *An Introduction to Homogenization*, Oxford University Press (New York, 1999).
- [6] G. Dagan, A. Fiori, and I. Janković, Flow and transport in highly heterogeneous formations: 1. Conceptual framework and validity of first-order approximations, *Water Resour. Res.* **39**(9), 1268 (2003).
- [7] Frank W. Elliott, Jr, David J. Horntrop, and Andrew J. Majda, A Fourier-wavelet Monte Carlo method for fractal random fields, *J. Comput. Phys.* **132**(2), 384–408 (1997).
- [8] 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).
- [9] Frank W. Elliott, Jr and Andrew J. Majda, A wavelet Monte Carlo method for turbulent diffusion with many spatial scales, *J. Comput. Phys.* **113**(1), 82–111 (July 1994).
- [10] 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. Comput. Phys.* **117**, 146–162 (1995).
- [11] Frank W. Elliott, Jr and Andrew J. Majda, Pair dispersion over an inertial range spanning many decades, *Phys. Fluids* **8**(4), 1052–1060 (1996).
- [12] A. ElMaihy and F. Nicolleau, Investigation of the dispersion of heavy-particle pairs and Richardson’s law using kinematic simulation, *Phys. Rev. E* **71**(4), 046307 (April 2005).
- [13] Jens Feder, *Fractals*, Physics of Solids and Liquids, Plenum Press (New York and London, 1988), chapter 9–14, pp. 163–243.
- [14] William Feller, *An introduction to probability theory and its applications*, 2nd edition, John Wiley & Sons (New York, London, 1971), Vol. 2, section III.6, pp. 47, 48.

- [15] Uriel Frisch, *Turbulence*, Cambridge University Press (Cambridge, 1995). The legacy of A. N. Kolmogorov.
- [16] 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).
- [17] J. C. H. Fung and J. C. Vassilicos, Two-particle dispersion in turbulentlike flows, *Phys. Rev. E* (3), **57** (2, part A), 1677–1690 (1998).
- [18] I. M. Gel'fand and G. E. Shilov, *Generalized functions. Properties and operations*, Academic Press (New York, 1964), Vol. **1**, section 3.3,
- [19] Lynn W. Gelhar, *Stochastic Subsurface Hydrology*, Prentice-Hall (Englewood Cliffs, N.J., 1993).
- [20] D. Horntrop and A. Majda, An overview of Monte Carlo simulation techniques for the generation of random fields, in *Monte Carlo Simulations in Oceanography*, Proceedings of the Ninth 'Aha Huliko'a Hawaiian Winter Workshop, edited by P. Muller and D. Henderson (1997), pp. 67–79.
- [21] M. A. I. Khan, A. Pumir, and J. C. Vassilicos, Kinematic simulation of turbulent dispersion of triangles, *Phys. Rev. E* (3), **68**(2), 026313, 13 (2003).
- [22] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Vol. **23** of *Applications of Mathematics: Stochastic Modelling and Applied Probability*, Springer-Verlag (Berlin, 1992).
- [23] Dmitry Kolyukhin and Karl Sabelfeld, Stochastic flow simulation in 3D porous media, *Monte Carlo Methods Appl.* **11** (1), 15–37 (2005).
- [24] R.H. Kraichnan, Diffusion by a random velocity field, *Phys. Fluids* **13** (1), 22–31 (1970).
- [25] O. Kurbanmuradov, Weak Convergence of Approximate Models of Random Fields. *Russian Journal of Numerical Analysis and Mathematical Modelling* **10** (6), 500–517 (1995).
- [26] O. Kurbanmuradov, Weak Convergence of Randomized Spectral Models of Gaussian Random Vector Fields, *Bull. Novosibirsk Computing Center, Numerical Analysis*, issue 4, 19-25 (1993).
- [27] O. Kurbanmuradov, K. Sabelfeld, and D. Koluvin, Stochastic Lagrangian models for two-particle motion in turbulent flows. Numerical results, *Monte Carlo Methods Appl.*, **3** (3), 199–223 (1997).
- [28] N. N. Lebedev, *Special functions and their applications*, Dover (New York, 1972), chapter 1, pp. 1–15.
- [29] Marcel Lesieur, *Turbulence in fluids*, second revised edition, Number 1 in *Fluid Mechanics and its Applications*, Kluwer (Dordrecht, 1990).

- [30] 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).
- [31] Stéphane Mallat, *A wavelet tour of signal processing*, Academic Press Inc. (San Diego, CA, 1998).
- [32] Benoit B. Mandelbrot, *The fractal geometry of nature*, updated and augmented edition, W.H. Freeman (San Francisco, New York, 1983).
- [33] W. D. McComb, *The physics of fluid turbulence*, Vol. **25** of *Oxford Engineering Science Series*, Clarendon Press (New York, 1991), chapter 2.
- [34] G. A. Mikhailov, Approximate models of random processes and fields, *Russian J. Comp. Mathem. and Mathem. Physics*, **23** (3), 558–566 (1983) (in Russian).
- [35] A.S. Monin and A.M. Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence*, The M.I.T. Press (1981), Vol. **2**.
- [36] F. Poirion and C. Soize, Numerical methods and mathematical aspects for simulation of homogenous and non homogenous Gaussian vector fields. In *Probabilistic methods in applied physics*, edited by Paul Kree and Walter Wedig, *Lecture Notes in Physics*, Springer-Verlag (Berlin, 1995), Vol. **451**, pp. 17–53.
- [37] Stephen B. Pope, *Turbulent flows*, Cambridge University Press (Cambridge, 2000).
- [38] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical recipes in FORTRAN*, second edition, section 7.8, pages xxvi+963, Cambridge University Press (Cambridge, 1992), The art of scientific computing, With a separately available computer disk.
- [39] Bénédicte Puig and Fabrice Poirion, White noise and simulation of ordinary Gaussian processes, *Monte Carlo Methods Appl.*, **10** (1), 69–89 (2004).
- [40] Yoram Rubin, *Applied Stochastic Hydrogeology*, Oxford University Press (New York, 2003).
- [41] 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).
- [42] K. K. Sabelfeld and O. Kurbanmuradov, Two-particle stochastic Eulerian-Lagrangian models of turbulent dispersion, *Math. Comput. Simulation*, **47** (2-5), 429–440 (1998), IMACS Seminar on Monte Carlo Methods (Brussels, 1997).
- [43] Karl K. Sabelfeld, *Monte Carlo methods in boundary value problems*, Springer Series in Computational Physics, Springer-Verlag (Berlin, 1991), chapters 1,5, pages 31–47, 228–238.
- [44] M. Shinozuka, Simulation of multivariate and multidimensional random processes, *J. of Acoust. Soc. Am.* **49**, 357-368 (1971).

- [45] H. Sigurgeirsson and A. M. Stuart, A model for preferential concentration, *Phys. Fluids*, **14** (12), 4352–4361 (December 2002).
- [46] D. J. Thomson and B.J. Devenish, Particle pair in kinematic simulations, *Journal of Fluid Mechanics*, **526**, 277–302 (2005).
- [47] J. A. Viccelli and E. H. Canfield, Jr, Functional representation of power-law random fields and time series, *J. Comput. Phys.* **95**, 29–39 (1991).
- [48] Richard F. Voss, Random fractal forgeries, in *Fundamental algorithms for computer graphics*, edited by Rae A. Earnshaw, *NATO ASI Series F: Computer and System Sciences*, Springer-Verlag (Berlin, 1985), Vol. **17**, pp. 805–835, NATO Science Affairs Divison. .
- [49] A. M. Yaglom, *Correlation theory of stationary and related random functions. Volume I: Basic results*, Springer-Verlag (Berlin, 1987).