Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 192 (2003) 546–569
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Relative efficiency of Gaussian stochastic process sampling procedures

Chris Cameron

Department of Applied Mathematics (Computing Sciences), Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Received 14 November 2002; received in revised form 29 July 2003; accepted 29 July 2003

Abstract

Various methods for sampling stationary, Gaussian stochastic processes are investigated and compared with an emphasis on applications to processes with power law energy spectra. Several approaches are considered, including a Riemann summation using left endpoints, the use of random wave numbers to sample a the spectrum in proportion to the energy it contains, and a combination of the two. The Fourier-wavelet method of Elliott et al. is investigated and compared with other methods, all of which are evaluated in terms of their ability to sample the stochastic process over a large number of decades for a given computational cost. The Fourier-wavelet method has accuracy which increases linearly with the computational complexity, while the accuracy of the other methods grows logarithmically. For the Kolmogorov spectrum, a hybrid quadrature method is as efficient as the Fourier-wavelet method, if no more than eight decades of accuracy are required. The effectiveness of this hybrid method wanes when one samples fields whose energy spectrum decays more rapidly near the origin. The Fourier-wavelet method has roughly the same behavior independently of the exponent of the power law. The Fourier-wavelet method returns samples which are Gaussian over the range of values where the structure function is well approximated. By contrast, (multi-point) Gaussianity may be lost at the smaller length scales when one uses methods with random wave numbers.

1. Introduction

We consider the sampling of Gaussian stochastic processes with a power law energy spectrum (see [1,5,9,15]). If this spectrum is denoted by $E(k)$ then the stochastic process $u(x)$, x denoting the independent variable, has the spectral representation,

$$u(x) = \int_{-\infty}^{\infty} e^{2\pi i k x} E^{1/2}(k)\, dW(k), \tag{1}$$

a stochastic integral with respect to increments $dW(\cdot)$ of a Wiener process $W(k)$.

E-mail address: ccameron@math.lbl.gov.

0021-9991/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.jcp.2003.07.012

A method for performing this quadrature based on a wavelet expansion, the Fourier-wavelet method of Elliott et al. [6], is described and compared to other approaches, including one based on a Riemann summation using left endpoints, the use of random wave numbers to sample the spectrum in proportion to the energy it contains, and a combination of the two. These procedures are evaluated based on their ability to sample accurately a large portion of the energy spectrum $E(k)$ – in particular, on the spread between the highest and lowest wave numbers accurately modelled – and the dependence of this spread on the computational complexity involved in generating a sample. The extent to which the sampling of the field approximates a Gaussian field is also investigated.

These methods are described in detail in Sections 2–4. In Section 5, we explain how the accuracy and computational complexity of each method is assessed. Section 5.4 in particular contains qualitative predictions about the accuracy, and its dependence on the computational complexity, of each method. Finally, numerical results of testing on three examples of power law spectra are presented in Section 6. Section 7 summarizes our conclusions.

2. Fourier-wavelet method

We begin by reminding the reader of the Fourier-wavelet method; for introduction and derivation see [6].

2.1. Relation to spectral representation

Recall the spectral representation of u, (1) above. For fixed x, this is a stochastic integral of the form $I = \int f(k)\,\mathrm{d}W(k)$. By representing the integrand f in an appropriate orthonormal basis for L^2, it may be written $I = \sum c_n\gamma_n$, where the γ_n are independent Gaussian random variables, and the coefficients c_n are the inner products $c_n = \int f(k)\phi_n(k)\,\mathrm{d}k$ for functions ϕ_n whose complex conjugates are our orthonormal basis. One may formally express this representation of a stochastic integral with respect to increments $\mathrm{d}W$ of a Wiener process via the identity

$$\mathrm{d}W(k) = \sum \gamma_n\phi_n(k)\,\mathrm{d}k,$$

where (γ_n) is a sequence of independent standard normal random variables and the (ϕ_n) form a complete orthonormal basis.

In this case the coefficients $c_n = c_n(x)$ depend on x, and are the functions

$$c_n(x) = \int \mathrm{e}^{2\pi\mathrm{i}kx}E^{1/2}(k)\phi_n(k)\,\mathrm{d}k = \mathscr{F}^{-1}[E^{1/2}\phi_n](x), \tag{2}$$

where $\mathscr{F}$ denotes the Fourier transform. So the process $u(x)$ is the sum of these functions $c_n(x)$ with Gaussian random weights γ_n:

$$u(x) = \sum \gamma_n c_n(x). \tag{3}$$

A practical algorithm is constructed from the preceding representation of $u(x)$ by carefully truncating the sum so as to include only important terms.

The idea of Elliott et al. is to use Meyer wavelet functions [2–4] to create the needed orthonormal basis. The Fourier transform of the Meyer wavelet ψ has the form

$$\phi(k) = \mathscr{F}\psi(k) = A\,\mathrm{sign}(k)\mathrm{e}^{\mathrm{i}\pi k}b(|k|), \tag{4}$$

where A is a constant of normalization (it is convenient to take $A = \pm\mathrm{i}$ so that $\psi(x)$ is real; here we take $A = -\mathrm{i}$). The function $b(\cdot)$ is smooth and satisfies $b(k) = 0$ for frequencies k outside an appropriate interval.

Here is the specific form of $b(k)$, where $k \geqslant 0$ (and $v = v(\cdot)$ is a function):

$$b(k) = \begin{cases} \sin\left(\frac{\pi}{2} v(3k-1)\right), & k \in (\frac{1}{3}, \frac{2}{3}], \\ \cos\left(\frac{\pi}{2} v(3\frac{k}{2}-1)\right), & k \in (\frac{2}{3}, \frac{4}{3}), \\ 0, & \text{else.} \end{cases} \tag{5}$$

The function v is chosen to guarantee whatever smoothness is desired in the wavelet function ϕ. For things to work out it should be a nondecreasing smooth function satisfying $v(x) = 0$ when $x < 0$, $v(x) = 1$ when $x > 1$, and $v(x) + v(1-x) = 1$ for all x. In the work of Elliott et al. construction of such a function v (for $x \in [0,1]$) is achieved through the use of (the integral of) a spline function, whence v can be designed to achieve the desired smoothness, taking the form,

$$v(x) = (-1)^p \frac{4^{p-1}}{p} \left[\max(x - x_0, 0)^p + \cdots \max(x - x_p, 0)^p + 2 \sum_{j=1}^{p-1} (-1)^j \max(x - x_j, 0)^p \right], \tag{6}$$

for the interpolation points $x_j = (1/2)(\cos(((p-j)/p)\pi) + 1)$. The parameter p is the order of the spline and determines how smooth the resulting function b will be. For testing, the choice $p = 2$ was made (in accordance with the choice of Elliott et al.).

Taken together, these functions define a complex function $\phi(k) = \mathscr{F}\psi(k)$ which is the Fourier transform of the Meyer mother wavelet $\psi(x)$. The family of wavelet functions is generated from ψ by the wavelet relation (cf. [2,4])

$$\psi_{mn}(x) = 2^{-m/2} \psi(2^{-m}x - n). \tag{7}$$

Since the Meyer wavelets $\psi_{mn}(x)$ form an orthonormal basis for L^2, and since the Fourier transform preserves inner products, it follows that the corresponding functions $\phi_{mn}(k) = \mathscr{F}\psi_{mn}(k)$ themselves are a suitable orthonormal basis (in k-space) to use in the representation (3) of the stochastic integral for $u(x)$. We thus have the representation

$$\begin{aligned} u(x) &= \sum_m \sum_n \gamma_{mn} c_{mn}(x), \\ c_{mn}(x) &= \int \mathrm{e}^{2\pi\mathrm{i}kx} E^{1/2}(k) \phi_{mn}(k)\, \mathrm{d}k, \end{aligned} \tag{8}$$

where the $\phi_{mn}(k)$ are the collection of orthonormal basis functions derived from the Fourier transforms of the Meyer wavelets $\psi_{mn}(\xi)$, as defined above by (4)–(7). (No confusion need arise from the fact that the wavelets, and thus the sum, are now doubly indexed.) For this to lead to a practical algorithm, the sum must be truncated and the integrals $c_{mn}(x)$ must be evaluated, with acceptable accuracy.

2.2. Implementation

In this section let x be fixed; we wish to compute $u(x)$ approximately from the representation (8). We focus on the kernel functions $c_{mn}(x) = \mathscr{F}^{-1}[E^{1/2}\phi_{mn}](x)$. First we notice that this can be simplified; using (7), trivial manipulations show that

$$\phi_{mn}(k) = [\mathscr{F}\psi_{mn}](k) = 2^{m/2} \mathrm{e}^{-2\pi\mathrm{i}k\cdot 2^m n} [\mathscr{F}\psi](2^m k),$$

and so we have

$$c_{mn}(x) = \mathscr{F}^{-1}[E_m^{1/2}\phi](2^{-m}x - n),$$

where $E_m^{1/2}(k) = 2^{-m/2}E^{1/2}(2^{-m}k)$.

We define $f_m(\xi) = \mathscr{F}^{-1}[E_m^{1/2}\phi](\xi)$, which allows us to write for the kernels c_{mn} the expression

$$c_{mn}(x) = f_m(2^{-m}x - n).$$

Substituting this result into expression (8) for $u(x)$, we get

$$u(x) = \sum_m \sum_n \gamma_{mn} f_m(2^{-m}x - n).$$

Because of presumed built-in decay of the functions c_{mn} away from zero, equivalently of the f_m, it is advantageous to re-center this sum around the term with the smallest (in magnitude) value of the argument $2^{-m}x - n$. To this end define for each term m of the outer sum the index $\bar{\boldsymbol{n}}_m = \lfloor 2^{-m}x \rfloor$ and shift the inner index n to $n' = n - \bar{\boldsymbol{n}}_m$. This results in the expression

$$u(x) = \sum_m \sum_{n'} \gamma_{m,\bar{\boldsymbol{n}}_m + n'} f_m\Big((2^{-m}x - \bar{\boldsymbol{n}}_m) - n'\Big). \tag{9}$$

Assuming f_m to decay quickly away from zero, the bulk of its support in this expression will be concentrated on those n' values nearest to zero. For each m we suppose most of the support of f_m is concentrated between some pair of integers, $-N$ and N, so that above we may discard terms for n' outside $[-N+1, N]$ without significant loss.

In addition one must decide for which values of m to include the coefficient terms $c_{mn}(x) = f_m(2^{-m}x - n)$. The parameter m governs the octave, or spatial scale of the support of wavelets to be included; upon truncation functions $f_m(\cdot)$ whose features vary too slowly or too rapidly are discarded from the sum, so that our representation captures only the features of the process u between some minimum and maximum length scale (equivalently, within some frequency range). By rescaling the length variable x if necessary, we may therefore assume that the sum over m is truncated so as to include only terms satisfying $-M \leqslant m \leqslant 0$, for some parameter M which corresponds to the smallest-scale features of u we hope to model accurately.

The wavelet representation based on (9) is thus written as the finite sum,

$$u(x) \approx u^W(x) = \sum_{m=-M}^{0} \sum_{n=-N+1}^{N} \gamma_{m,\bar{\boldsymbol{n}}_m + n} f_m\Big((2^{-m}x - \bar{\boldsymbol{n}}_m) - n\Big), \tag{10}$$

where the cutoff choices M (how many octaves to include) and N (dictated by the support of f_m) are parameters of the method. The algorithm based on this representation of $u(x)$ is termed the Fourier-wavelet method.

Evaluation of the kernel functions $f_m(\xi) = \mathscr{F}^{-1}[E_m^{1/2}\phi](\xi)$ is accomplished by first evaluating f_m on a discrete grid using the FFT and then interpolating to obtain values of $f_m(x)$ for x not on the grid; see [6] for details. A benefit of this approach for generating many sample paths is that the grid values $f_m(x_j)$ do not depend on the particular realization and can be computed once in advance; from these the desired number of sample paths $u(x)$ can be generated.

Finally, note that the Gaussian random coefficients needed in (10) have the form $\gamma_{m,\bar{\boldsymbol{n}}_m + n}$. This depends on $\bar{\boldsymbol{n}}_m = \lfloor 2^{-m}x \rfloor$, which depends on x. For a given m, we could need in principle Gaussian coefficients γ_{mp} for p anywhere from $-N$ to $\lfloor 2^{-m}X \rfloor + N$. The total number of Gaussian random variables which could theoretically participate in the sum (10) is thus on the order of

$$\sum_{m=-M}^{0} \lfloor 2^{-m}X \rfloor \sim 2^{M+1}X.$$

Generating all of these random variables, and then storing the whole collection as a representation of one realization of a stochastic process, quickly becomes prohibitive for values of M which are large enough to capture the behavior of $u(x)$ on a reasonably small scale.

To remedy this, Elliott et al. [6] employ a reversible pseudorandom number generator (for example, see [7,10]) to sample the needed coefficients γ_{mn}. A sequence of uniform random numbers is generated, and from them Gaussian random numbers are obtained by the Box–Muller transform. When one uses a reversible pseudorandom number generator, all of the Gaussian random numbers needed for the sum (10) – and hence, for a given realization of $u(x)$ – are determined by the initial number in the sequence: any other number in the sequence is obtainable from the first number by predetermined arithmetic operations. This savings in computation comes at some cost, due to the overhead involved in arranging for efficient calculation of any given element of the pseudorandom number sequence.

3. Two simple methods

In this section we examine two basic methods for generating numerical sample paths of u. Each method is based on quadrature of the spectral representation of $u(x)$, one with fixed wave numbers and one using random wave numbers.

3.1. A fixed wave number scheme

First, assume $u \in \mathbb{R}$, so E is even. This places a restriction on $\mathrm{d}W$: Eq. (1) must now be read with the tacit understanding that $W(-k) = -\overline{W(k)}$. (In other words, $W(k)$ can be chosen freely only for $k \geqslant 0$.) Eq. (1) may be rewritten as

$$u(x) = 2^{1/2}\left[\int_0^\infty E^{1/2}(k)\cos(2\pi kx)\,\mathrm{d}W^1(k) + \int_0^\infty E^{1/2}(k)\sin(2\pi kx)\,\mathrm{d}W^2(k)\right], \tag{11}$$

where W^1 and W^2 are real, independent Wiener processes.

Since the energy spectrum E satisfies $E(k) \to 0$ as $k \to \infty$, we introduce an upper frequency cutoff K and truncate the integrals. Similarly, it will usually be necessary to introduce a low frequency cutoff c for the spectra we consider, because typically they will have a singularity at $k = 0$. (Consider the Kolmogorov spectrum $E(k) \propto k^{-5/3}$, for example.) So we replace the infinite integrals in (11) with the finite approximations

$$u(x) \approx 2^{1/2}\left[\int_c^K E^{1/2}(k)\cos(2\pi kx)\,\mathrm{d}W^1(k) + \int_c^K E^{1/2}(k)\sin(2\pi kx)\,\mathrm{d}W^2(k)\right]. \tag{12}$$

These integrals can be discretized as follows: divide the domain $[c, K]$ into N_f equally sized cells, each of size $\Delta k = ((K-c)/N_\mathrm{f})$, and approximate the integrals by Riemann sums using left endpoints:

$$u(x) \approx 2^{1/2}\left[\sum_{j=0}^{N_\mathrm{f}-1} E^{1/2}(k_j)\cos(2\pi k_j x)\Delta W_j^1 + \sum_{j=0}^{N_\mathrm{f}-1} E^{1/2}(k_j)\sin(2\pi k_j x)\Delta W_j^2\right].$$

By the definition of the Wiener process, the increments ΔW_j^i (for both $i = 1$ and $i = 2$) are independent, mean zero Gaussian random variables, with variance $k_{j+1} - k_j = \Delta k$. We can thus write each as $W_j^i = \Delta k^{1/2} Z_j^i$, where the Z_j^i are independent, identically distributed standard normal random variables. Our method now looks like:

$$u(x) \approx u^{\mathrm{F}}(x) = (2\Delta k)^{1/2} \sum_{j=0}^{N_{\mathrm{f}}-1} E^{1/2}(k_j)\left[\cos(2\pi k_j x) Z_j^1 + \sin(2\pi k_j x) Z_j^2\right], \quad k_j = c + \mathrm{j}\Delta k, \ \Delta k = \frac{K-c}{N_{\mathrm{f}}}. \tag{13}$$

We will call this the fixed wave number method because it approximates $u(x)$ by a finite sum of waves with predetermined wave numbers. (See also the Fourier method with equally spaced grid points, in [12].)

3.1.1. Discretization limitation

In order for the scheme represented by (13) to be accurate, Δk should be small. On the other hand, the smaller Δk is, the more computationally intensive the algorithm becomes. In addition, because of the periodic factor in each integrand, care should be taken to assure that in the process of discretization, the function is not undersampled.

Our goal is to sample $u(x)$ not for all x, but rather for x belonging to some finite domain, which may be taken to be $[0, X]$. Indeed, representation (13) is a finite sum of trigonometric functions, so for any Δk, there will be an upper limit X to the values of x for which we are willing to submit (13) as a fair approximation to $u(x)$.

Let us illustrate this in terms of the covariance function $R^{\mathrm{F}}(x)$ of $u^{\mathrm{F}}(x)$. Take low wave number cutoff $c = \Delta k$; it is easy to check that R^{F} is

$$R^{\mathrm{F}}(x) = \langle u^{\mathrm{F}}(x' + x) u^{\mathrm{F}}(x') \rangle = 2\Delta k \sum_{j=1}^{N_{\mathrm{f}}-1} E(\mathrm{j}\Delta k) \cos(2\pi \mathrm{j} \Delta k x).$$

So u^{F} of our method possesses a covariance function which is even and periodic with period $L = 1/\Delta k$. This periodicity does not hold for the covariance function $R(x)$, so the model will not be accurate on $[0, X]$ if $X > L/2$. Consequently, given a fixed interval $[0, X]$ on which to generate the sample, at the very least Δk should be chosen so that $\Delta k \leqslant (2X)^{-1}$.

This observation provides the starting point for choosing Δk when obtaining a sample from (13). This parameter, along with the other parameters X, c, and K, specifies the representation u^{F}.

3.2. A random wave number method

In this section we describe a method using random wave numbers, sometimes called the randomization method (see [12,14], also [13] for a variant), for sampling $u(x)$. Its choice of random wave numbers is an extension of Monte Carlo algorithms for integration. Suppose we have an integral of the form

$$J = \int_a^b g(k) r(k) \, \mathrm{d}W(k). \tag{14}$$

The approximation to this integral should be a mean-zero Gaussian random variable, with variance that well approximates the true variance $\langle |J|^2 \rangle = \int_a^b |g(k)|^2 |r(k)|^2 \, \mathrm{d}k$. The latter integral can be treated by a Monte Carlo approach using randomly chosen points in k-space. Let $f(k) = |g(k)|^2$ and $q(k) = |r(k)|^2$. If $Q = \int_a^b q(k) \, \mathrm{d}k$ is finite, then a sum of the form

$$\tilde{I} = \frac{Q}{N_{\mathrm{r}}} \sum_{i=1}^{N_{\mathrm{r}}} |g(k_i)|^2, \tag{15}$$

where the k_i are independent random variables taken according to the probability density $p(k) = |r(k)|^2 / Q$, can be expected to converge to $I = \langle |J|^2 \rangle$ as the number of Monte Carlo points N_{r} goes to infinity.

The sum

$$\tilde{J} = \left(\frac{Q}{N_r}\right)^{1/2} \sum_{i=1}^{N_r} a_i g(k_i),$$

where the k_i are as before and the a_i are independent standard normal random variables (Gaussian with mean zero and variance 1), has mean $\langle \tilde{J} \rangle = 0$ and variance

$$\langle |\tilde{J}|^2 \rangle = \frac{Q}{N_r} \sum_{i=1}^{N_r} \langle |g(k_i)|^2 \rangle = \langle \tilde{I} \rangle = I = \langle |J|^2 \rangle.$$

The variance of $\tilde{J}$ as defined above is precisely the expected value of the Monte Carlo approximation (15), and thus matches the true variance $\langle |J|^2 \rangle$. While $\tilde{J}$ is a weighted sum of Gaussian random variables, the weights are random and thus (multi-point) Gaussianity is lost. On the other hand, $\tilde{J}$ has the correct mean and the correct variance.

3.2.1. Application to spectral representation

In Section 3.1 we saw that the spectral representation for real $u(x)$ becomes (11):

$$u(x) = 2^{1/2}\left[\int_0^\infty E^{1/2}(k)\cos(2\pi k x)\,\mathrm{d}W^1(k) + \int_0^\infty E^{1/2}(k)\sin(2\pi k x)\,\mathrm{d}W^2(k)\right].$$

Write this as

$$u(x) = 2^{1/2}[J_1(x) + J_2(x)]$$

and apply the algorithm outlined in the preceding section to $J_1(x)$ and $J_2(x)$, where the spatial variable x is a parameter.

Specifically, for $J_1(x)$ we wish to estimate

$$J_1(x) = \int_0^\infty \cos(2\pi x k) E^{1/2}(k)\,\mathrm{d}W^1(k).$$

(This is analogous to (14); here $g(\cdot) = \cos(2\pi x(\cdot))$ and $r(\cdot) = E^{1/2}(\cdot)$.) We begin by letting $Q = \int_0^\infty E(k)\,\mathrm{d}k$ so as to properly normalize the probability density; however, for the energy spectra we consider $E(k)$ may not be integrable. The Kolmogorov spectrum $E(k) \propto k^{-5/3}$ has a nonintegrable singularity at $k = 0$. To get around this, introduce a low wave number cutoff c_0, so that

$$J_1(x) \approx \int_{c_0}^\infty \cos(2\pi x k) E^{1/2}(k)\,\mathrm{d}W^1(k), \tag{16}$$

and similar for $J_2(x)$. At this step in the process, by leaving out the energy contained in those wave numbers smaller than c_0, we are implicitly putting an upper limit $\sim 1/c_0$ on the values of x for which this approximation is likely to hold. It is possible to rationalize this cutoff by realizing that spectra such as the Kolmogorov spectrum are only designed to represent phenomena in an intermediate range of wave numbers.

Now sample N_r points $\{k_i\}_{i=1}^{N_r}$ randomly from the probability density $p(k) = (1/Q)E(k),\ Q = \int_{c_0}^\infty E(k)\,\mathrm{d}k$. Define the approximation $\tilde{J}_1(x)$ by

$$\tilde{J}_1(x) = \left(\frac{Q}{N_r}\right)^{1/2} \sum_{i=1}^{N_r} a_i \cos(2\pi k_i x),$$

where the a_i are independent, standard normal random variables. In a similar way define

$$\tilde{J}_2(x) = \left(\frac{Q}{N_r}\right)^{1/2} \sum_{i=1}^{N_r} b_i \sin(2\pi k_i x),$$

where the b_i are standard normal random variables, independent of the a_i. Taken together, these quadratures using random k values yield

$$u(x) \approx \left(\frac{2Q}{N_r}\right)^{1/2} \sum_{i=1}^{N_r} a_i \cos(2\pi k_i x) + b_i \sin(2\pi k_i x), \tag{17}$$

where $Q = \int_{c_0}^{\infty} E(k)\,\mathrm{d}k$ as above and the N_r discrete wave numbers k_i are randomly chosen from the wave number domain $[c_0, \infty)$ according to the probability density $p(k) = E(k)/Q$.

However, straightforward application of approximation (17) directly to the integrals is not necessarily the most desirable approach.

3.2.2. Separation of wave numbers

In the scheme (17), we have little control over the spread of wave numbers, and thus over how much small scale versus large scale detail we are including in the approximation to $u(x)$. To remedy this, we can break up the integrals into a sum of integrals over smaller domains and approximate each one separately using random wave numbers within its domain of integration. For example, let $\{I_j\}_{j=1}^{M}$ be a collection of disjoint intervals with $[c_0, \infty) = \bigcup_{j=1}^{M} I_j$. Then we can write, in (16),

$$J_1(x) \approx \int_{c_0}^{\infty} \cos(2\pi x k) E^{1/2}(k)\,\mathrm{d}W^1(k) = \sum_{j=1}^{M} \int_{I_j} \cos(2\pi x k) E^{1/2}(k)\,\mathrm{d}W^1(k) = \sum_{j=1}^{M} J_{1j}(x).$$

Now apply the preceding approach to each integral $J_{1j}(x)$. A probability density can be defined on interval I_j by $p_j(k) = E(k)/Q_j$, where $Q_j = \int_{I_j} E(k)\,\mathrm{d}k$. Sample N_j points (k_i^j), $i = 1, 2, \ldots, N_j$ from the density $p_j(x)$. Finally, form the sum

$$J_{1j}(x) \approx \left(\frac{Q_j}{N_j}\right)^{1/2} \sum_{i=1}^{N_j} a_i^j \cos(2\pi x k_i^j),$$

where the a_i^j are independent standard normal random variables.

To guarantee that high wave numbers are represented, it is convenient to take $N_j = 1$ point from each interval, and it is further convenient to set $Q_j = Q/M$ for all j, so that the integral of $E(k)$ over each interval $I_j = [K_{j-1}, K_j)$ is the same. This uniquely determines $M - 1$ points K_i ($K_M = \infty$) with

$$c_0 = K_0 \leqslant K_1 \leqslant K_2 \leqslant \cdots \leqslant K_{M-1} \leqslant K_M = \infty,$$

each interval $[K_{j-1}, K_j)$ bracketing a range of wave numbers k which contains the same amount of energy. The random wave numbers are chosen one from each interval $k_j \in [K_{j-1}, K_j)$; in particular we can be sure that the scheme contains a wave number at least as large as K_M.

Using these randomly chosen wave numbers k_j, form the sum

$$u(x) \approx u^{\mathrm{R}}(x) = \left(\frac{2Q}{N_r}\right)^{1/2} \sum_{j=1}^{N_r} a_j \cos(2\pi k_j x) + b_j \sin(2\pi k_j x), \tag{18}$$

where Q is the integral of $E(k)$ on $[c_0, \infty)$ and the a_j and b_j, $j = 1, 2, \ldots, N$, are independent standard normal random variables. The above scheme is called the randomization method, or random wave number method.

4. A hybrid scheme

In this section we describe a scheme which is a hybrid of the two preceding schemes. Since the fixed wave number method seems better for treating the low wave numbers (ensuring that they are represented well), while the random wave number method allows us to retain more of the upper end of the spectrum for a given computational cost, it makes sense to apply both schemes where their strengths lie. To this end, split the spectral representation (with low energy cutoff c) into two parts by introducing an intermediate wave number K',

$$u(x) \approx \left(\int_c^{K'} + \int_{K'}^{\infty} \right) \mathrm{e}^{2\pi \mathrm{i} k x} E^{1/2}(k)\, \mathrm{d}W(k).$$

The idea is to treat the first integral with our fixed wave number method (with low wave number cutoff c and upper wave number cutoff K', which we denote by $u^{\mathrm{F}}(c, K', N_{\mathrm{f}})$, here suppressing dependence on x), and approximate the second one by the random wave number method (this time using K' as our low wave number cutoff, and denoted by $u^{\mathrm{R}}(K', N_{\mathrm{r}})$). Thus we write

$$u \approx u^{\mathrm{H}} = u^{\mathrm{F}}(c, K', N_{\mathrm{f}}) + u^{\mathrm{R}}(K', N_{\mathrm{r}}).$$

The resulting approximation u^{H} is our hybrid method and can be written $u^{\mathrm{H}} = u^{\mathrm{H}}(c, K', N_{\mathrm{f}}, N_{\mathrm{r}})$. Its output is dependent upon c, the low wave number cutoff; K', an intermediate wave number; N_{f}, the number of grid points for the fixed wave number scheme; and N_{r}, the number of sample points for the random wave number method.

5. Method assessment

In this section, the basic strategy for assessing the accuracy of the preceding methods is outlined. For a power law spectrum it will be convenient to use for evaluation the structure function, $D(x) = \langle |u(x) - u(0)|^2 \rangle$, rather than the covariance function $R(x)$ or energy spectrum $E(k)$, because the Fourier integral for $R(x)$ diverges. Also, for a power law spectrum $E(k) = |k|^{\alpha}$, an analytical formula for $D(x)$ can be found. (For proper interpretation of Fourier integrals which only exist in a generalized sense, see [8,11].) Indeed, when $-3 < \alpha < -1$, we have

$$D(x) = \begin{cases} 2^{1-\alpha}\pi^{-1-\alpha}\Gamma(1+\alpha)\sin(\alpha\frac{\pi}{2})|x|^{-1-\alpha} & (-3 < \alpha < -1,\ \alpha \neq -2), \\ 4\pi^2|x| & (\alpha = -2). \end{cases}$$

5.1. Decades of accuracy

The structure function $D(x)$ provides a way of testing the accuracy of an algorithm for generating random fields based on such a given $E(k)$. Generate a large number M of samples $\tilde{u}_{(i)}$ $(i = 1, \ldots, M)$ using the method, and compute the approximate value of the underlying structure function $\tilde{D}(x)$ by the estimate

$$\langle |\tilde{u}(x) - \tilde{u}(0)|^2 \rangle = \tilde{D}(x) \approx \frac{1}{M} \sum_{i=1}^{M} |u^{(i)}(x) - u^{(i)}(0)|^2 = \tilde{D}_M(x).$$

Finally, compare this estimated model structure function, $\tilde{D}_M(x)$, to the true (known) structure function $D(x)$. The result provides a means of determining how accurate the method is, and over which range (of x).

We seek to extract an approximate ratio between the largest and smallest accurate length scales by comparing $\tilde{D}_M(x)$ and $D(x)$. To this end we examine the relative error between the two,

$$e_D(x) = \frac{|\tilde{D}_M(x) - D(x)|}{D(x)}$$

at the grid points of our model, and from this extract a range of such values on which $e_D(x) < \epsilon$, for some ϵ. For definiteness, say that our working grid (on which we will sample $\tilde{u}(\cdot)$) is (x_i). For testing it helps to use a logarithmic grid, $x_i = x_0 c^i$, so that a large number of different length scales can be examined with a relatively small number of grid points. On this grid, we can define a range of accuracy – a set of grid points at which the method's output is said to be accurate – by $\mathscr{A} = \{x_i \,|\, e_D(x_i) < \epsilon\}$.

Ideally the statistics of u would be well enough behaved so that the definition of the range of accuracy $\mathscr{A}$ would be insensitive to the choice of ϵ in some intermediate range of values. In practice, such a well-defined accuracy region $\mathscr{A}$ was not always observed. To get a practical measure of the extent of $\mathscr{A}$ then, we defined the maximum mesh point of the range of accuracy to be the rightmost point with a sufficiently small error: $x_{\max} = \max\{x_i \,|\, e_D(x_i) < \epsilon\}$. From this $x_{\max}$ we then defined a minimum accurate mesh point $x_{\min}$ by asserting that, of the mesh points within the resulting interval $[x_{\min}, x_{\max}]$, a certain sufficiently high fraction of them (say p) ought to also be accurate,

$$x_{\min} = \min\{x_i \,|\, e_D(x_j) < \epsilon \text{ for fraction } p \text{ of the mesh points } x_j \in [x_i, x_{\max}]\}.$$

The ratio $x_{\max}/x_{\min}$ then provides a rough guide to the number of accurate length scales in the model. From $x_{\max}$ and $x_{\min}$ we define the number of accurately modelled decades: $d = \log_{10}(r)$, $r = x_{\max}/x_{\min}$.

Throughout the testing we used $\epsilon = 0.1$ and $p = 0.9$, both defined above, for evaluating the computed d from a given method's output.

5.2. Assessment of Gaussianity

To gauge Gaussianity, we approximate the kurtosis $\tilde{\mathscr{K}}(x) = \langle \tilde{u}(x)^4 \rangle / \langle \tilde{u}(x)^2 \rangle^2$ of the model field $\tilde{u}(x)$ by taking our large number M of samples and computing the ratio

$$\tilde{\mathscr{K}}_M(x) = \tilde{m}_4(x)/(\tilde{m}_2(x))^2,$$

$$\tilde{m}_4(x) = \frac{1}{M}\sum_{i=1}^{M}\left[\tilde{u}^{(i)}(x)^4\right],$$

$$\tilde{m}_2(x) = \frac{1}{M}\sum_{i=1}^{M}\left[\tilde{u}^{(i)}(x)^2\right],$$

for all x in the model. Since the kurtosis $\tilde{\mathscr{K}}(x)$ should be 3 (for all x) for Gaussian stochastic processes, we thus use this estimate $\tilde{\mathscr{K}}_M$ as a rough measure of the Gaussianity of our method by evaluating its percent difference from 3; from this difference we extract a number of decades $d_{\mathscr{K}}$ of x values over which the model appears to be nearly Gaussian (by using the methods of Section 5.1).

5.3. Computational cost

We would also like a way of assessing the computational cost required to achieve a given accuracy (i.e. number of decades). It may be that one method is more accurate than another when a certain budget for computational cost is prescribed, but the situation becomes reversed if more (or less) operations are allowed. The main concern is how many operations are required per realization, per x value at which the realization is to be sampled.

What contributes to the operation cost of a given realization $\tilde{u}$ from a method? Typically, the model process $\tilde{u}$ is represented as a combination of the following:

1. The results of certain preprocessing calculations. (for example, in the Fourier-wavelet method, we sample the kernel functions $f_m(\cdot)$, to be evaluated once only on a discrete uniform mesh using the inverse Fast Fourier Transform).
2. A list of generated random numbers, the list being unique for each realization of the sample path.
3. Some analytical expression, which accepts an x value and, in combination with the results of the previous two items allows the value of the model sample path $\tilde{u}(x)$ to be computed at that point. (In practice this usually involves repeated evaluation of, e.g., trigonometric functions.)

The computations for the first item – done once and for all – can be neglected. The second item involves computations which are done once per sample path; if the sample path must be known at a large number N_x of spatial points, then this too becomes negligible. The primary contribution to an operation count comes from the third item, operations which must be performed in order to convert the list of random numbers - and precomputed terms – into a sample path value $\tilde{u}(x)$ for each x value.

Thus for each method, we attempt to extract a rough count of the labor required per realization, per x value. We are content to count the necessary trigonometric function calls per realization, per x (for methods for which this is applicable), and accept this as our cost. Some modification of this approach is required when treating the Fourier-wavelet method.

5.3.1. Fixed wave number method count

For (13), the main parameter is N_f, the number of grid points for the simple quadrature. Here $2N_f$ standard normal random variables are required per realization. Yet, to compute a sample path value $\tilde{u}(x)$ from this list of stored random variables (and from N_f evaluations of the energy spectrum $E(k)$, whose cost we also neglect) requires two trigonometric function evaluations (one sine, one cosine) per term in the sum. All told, the fixed wave number method thus costs $C_F = 2N_f$ such evaluations per realization per x.

5.3.2. Random wave number method count

For the random wave number method (18), we model the process $u(x)$ with a sum using a number N_r of randomly sampled wave numbers. Once the random frequencies k_i are sampled to represent a realization, the random wave number method – like the fixed wave number method – requires $2N_r$ trigonometric evaluations per x value at which $\tilde{u}(x)$ is computed. We thus have for the cost of the random wave number scheme $C_R = 2N_r$.

5.3.3. Hybrid method count

The hybrid method (4) contains a fixed wave number scheme for approximating the low end of the spectrum with some number N_f of grid points. Simultaneously, it contains a random wave number approximation to the higher end of the spectrum (from K' to ∞) using N_r random wave numbers. From the preceding two counts, one sees then that the total cost for the hybrid method is $C_H = 2(N_f + N_r)$.

5.3.4. Fourier-wavelet method count

As detailed in Section 2, computation of the kernels $f_m(\cdot)$ is achieved by first computing the values $f_m(x_j)$ on an equidistant grid once and for all, and then interpolating for needed x. This operation thus does not contribute significantly to our overall cost count.

As with the other methods, this method depends on generating Gaussian random variables to represent the stochastic process; however, the random variables are generated as needed (using a reversible pseudorandom number generator), and depend on x. All told we need to generate a total of $2N(M+1)$ Gaussian random variables, using the Box–Muller transform, for each x. Since the Box–Muller transform

obtains two Gaussian random variables from two uniform random variables, there is on average one trigonometric operation per Gaussian required. We thus take as our operation count $C_W = 2N(M+1)$ for the Fourier-wavelet method.

5.4. Qualitative behavior

Each method we consider (apart from the Fourier-wavelet method, considered separately) approximates the process $u(x)$ by a finite sum taken over a finite range of wave numbers (whether random or deterministic). Each method therefore contains a smallest wave number $k_{\min}$ and a largest wave number $k_{\max}$. It is reasonable to approximate our smallest and largest accurate length scale by $x_{\min} \sim k_{\max}^{-1}$ and $x_{\max} \sim k_{\min}^{-1}$, so that our number of accurate decades, d, is $d \sim \log_{10}(k_{\max}/k_{\min})$.

The largest and smallest wave numbers $k_{\max}$, $k_{\min}$ are obtainable from each method and from α, the exponent of the power law spectrum, so we may state the expected qualitative behavior of each method based on α and its parameters.

5.4.1. Fixed wave number

The lowest wave number represented in this method is $k_{\min} = c$; the highest is $k_{\max} = k_{N_f-1} = c + (N_f - 1)\Delta k$. We will always make the choice $c = \Delta k$, to find that

$$d \sim \log_{10}(N_f) = (\log(10))^{-1} \log(N_f)$$

for our fixed wave number method. That is, we expect the decades of accuracy to increase logarithmically with N_f, the parameter of this method, with log-linear slope $\sim 1/\log(10) \approx 0.43$. This calculation did not depend on the specific form of $E(\cdot)$.

5.4.2. Random wave number

The lowest wave number represented is no smaller than the low wave number cutoff $k_{\min} = c_0$. For $E(k) = |k|^{\alpha}$, $\alpha < -1$, the condition of equipartition of energy dictates that the highest wave number represented is at least $k_{\max} = c_0 \times N_r^{-(1/(\alpha+1))}$, where N_r is the number of random wave numbers used. We thus expect the number of decades to behave like

$$d \sim \frac{-1}{\alpha+1} \log_{10}(N_r) = \frac{-1}{\alpha+1} (\log(10))^{-1} \log(N_r).$$

So once again we expect logarithmic dependence of the accuracy of this method on its parameter N_r, although in this case with a log-linear slope of $-1/(\alpha+1)$. Note that this slope is positive for $\alpha < -1$; further, the closer $\alpha(< -1)$ is to -1, the greater this slope will be, and thus the more quickly the accuracy of the random wave number method is expected to rise with an increase in N_r. This can be understood from the observation that spectra $E(k) = |k|^{\alpha}$ with $\alpha \lesssim -1$ have relatively slowly decaying tails when compared with smaller (more negative) α; the use of random wave numbers to access the high end of the spectrum can thus be expected to pay more dividends.

For $\alpha = -5/3$, this slope is $3/(2\log(10)) \approx 0.65$. In particular, we expect this slope for the random wave number method to be about $3/2$ times that for the fixed wave number method (for $\alpha = -5/3$).

5.4.3. Hybrid method

The hybrid method combines the preceding two methods by using an intermediate wave number cutoff K'. The smallest (fixed) wave number represented is $k_{\min} = \Delta k = K'/N_f$, while the largest (random) wave number represented is $k_{\max} = K' \times N_r^{-(1/(\alpha+1))}$. The ratio gives us

$$d \sim \log_{10}(N_f) + \frac{-1}{\alpha+1}\log_{10}(N_r)$$

for the hybrid method. In other words, we expect the number of decades to be the sum of the results we would get by applying each method separately.

5.4.4. Optimal hybrid

Suppose we wish to maximize this d for our hybrid method, subject to the constraint that $N_f + N_r =$ a constant. (As we have seen, the sum $2(N_f + N_r)$ is a measure of the computational cost of this method.) When $\alpha < -1$, a nontrivial optimal choice exists, $N_r = (-1/(\alpha+1))N_f$.

In particular, for $\alpha = -5/3$ we should choose $N_r = (3/2)N_f$ (rounded to the nearest integer). In general, this observation allows us to select such an optimal balance between N_r and N_f if we know the exponent α of the energy spectrum. The closer α is to -1, the slower the tail decays, and the more random wave numbers ($N_r > N_f$) we ought to include. On the other hand, spectra with more rapidly decaying tails $\alpha < -2$ can be expected to benefit from the inclusion of more fixed wave numbers ($N_f > N_r$) at the lower end of the spectrum.

Finally, suppose that the choice $N_r = -1/(\alpha+1)N_f$ is made. The number of decades in this optimized hybrid method becomes

$$d \sim \frac{\alpha}{\alpha+1}\log_{10}(N_f) + \frac{-1}{\alpha+1}\log_{10}\left(\frac{-1}{\alpha+1}\right),$$

which for $\alpha < -1$ shows an advantage over either method separately.

When $\alpha = -5/3$, we expect a log-linear slope of $5/(2\log(10)) \approx 1.09$, or $5/2$ times greater than the slope for the fixed wave number method and $5/3$ times greater than the slope for the random wave number method.

5.4.5. Fourier-wavelet method

From the nature of the wavelets and the relationship

$$\phi_{mn}(\xi) = 2^{-m/2}\phi(2^{-m}\xi - n),$$

it is seen that the small-scale details captured in a wavelet expansion are expected to increase linearly with m, the octave number. Thus we expect a dependence of the form

$$d \propto M + 1,$$

where we recall that $M+1$ is the number of octaves included in the model (including the initial octave $m = 0$). This linear, rather than log-linear, dependence of the number of decades d on the computational complexity sets the Fourier-wavelet method apart from the other methods.

6. Numerical results

In this section are presented results obtained by running the methods above for various choices of parameters. In each case we run the method for $M = 2000$ realizations, for each chosen value of its parameters. We calculate the approximate number of decades of accuracy present in the structure function and kurtosis of each such set of samples using the heuristic formulas described in Section 5, with parameters $\epsilon = 0.1$ (agreement tolerance) and $p = 0.9$ (a fudge factor to allow for aberrant points of misfit).

For each method we model the process $u(x)$ over the interval $[0, X]$ where $X = 1$. The initial tests were performed using the energy spectrum $E(k) = k^{-5/3}$. Subsequently, the exponent was changed to test two alternate spectra, $k^{-3/2}$ and k^{-2}.

For the spectrum $k^{-5/3}$, based on our convention for the Fourier transform, we expect to see the structure function

$$D(x) = -2^{5/3}\pi^{2/3}\Gamma(-2/3)x^{2/3} \approx 27.3655 \times x^{2/3}.$$

Finally, each modelled process $\tilde{u}(\cdot)$ was sampled on a logarithmically spaced grid

$$x_i = x_0 \times r^i, \quad i = 0, 1, 2, \ldots, N_x - 1,$$

where N_x is the total number of spatial points and $r = (X/x_0)^{(1/(N_x-1))}$ is the common ratio between points. The smallest point x_0 is chosen so as to make sure to represent the smallest possible length scale in the model. In our tests we had $N_x = 100$ and $x_0 = 10^{-12}$, so that up to 12 decades could be resolved. Thus $r = 10^{12/99}$.

6.1. Results for the fixed wave number method

For the fixed wave number method (13), the number of grid points chosen was related to K (the upper wave number cutoff) by a heuristic $N_f \simeq \beta KX$, based on the need to adequately sample the spectrum to avoid aliasing (we took $\beta = 5$). The lower wave number cutoff was taken to be $c = \Delta k$. The method was run for 41 values of K ranging between $K = 100$ and $K = 10,000$, increasing logarithmically.

Fig. 1 shows the number of decades of accuracy in the structure function $R(x)$ plotted against K (in logarithmic scale). The result appears to follow a roughly log-linear dependence, with approximate slope $m = 0.53$. Recall from Section 5.4 that we expected a log-linear dependence, with slope approximately $(\log(10))^{-1} \approx 0.43$, based only on rough estimates of the smallest and largest length scales included.

From the graph, for $K \sim 10{,}000$ we expect no more than four decades of accuracy in the results.

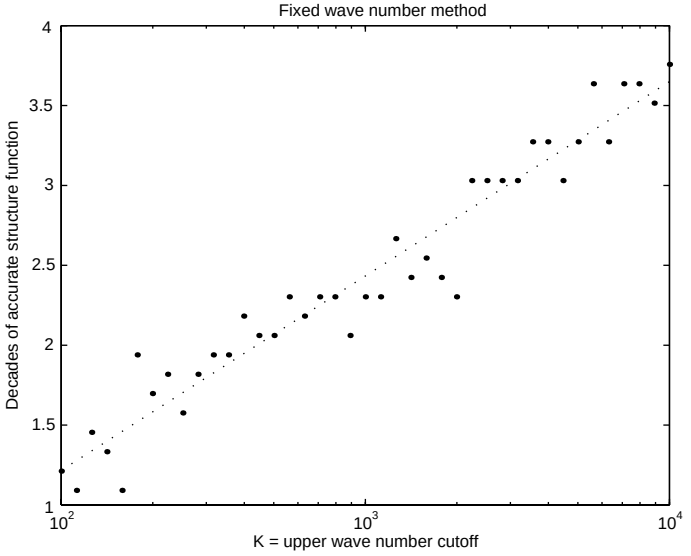

Fig. 1. Fixed wave number method results: Dependence of decades of accurate structure function on the parameter K, the upper wave number cutoff. The slope of the log-linear least-squares line (dashed) is $m \approx 0.53$.

6.2. Results for the random wave number method

The random wave number method (18) was applied for various values of N_r, the number of random wave numbers, ranging between 10 and 2000. Between $N_r = 10$ and $N_r = 400$, a test was performed for every multiple of 10; beyond this N_r was increased more dramatically due to the length of time required by the calculations, and also because for these high values of N_r the number of decades was not observed to increase as consistently, and was more sporadic.

Fig. 2 displays the decades of accuracy present in the both the structure function and kurtosis, simultaneously. The results are (very) approximately log-linear, with least-squares slopes $m = 0.67$ and $m = 0.74$, respectively. Compare with the approximate slope $m \approx 0.5$ of the results of the fixed wave number method, to see that the accuracy of the random wave number method increases faster with the number of grid points (and hence, computations) performed.

This is also consistent with the discussion of Section 5.4, in which it was argued that we could expect a slope of $3/(2\log(10)) \approx 0.65$ for this method.

On the other hand, it is seen from Fig. 2 that the number of accurate decades in the kurtosis is consistently lower than the number of accurate decades in the structure function. This coincides with our earlier observation that the random wave number method will fail to exhibit (multi-point) Gaussianity at the high wave numbers; indeed, it was seen that for small values x (hence, high wave numbers k) the computed kurtosis was significantly greater than 3.

6.3. Results for the hybrid method

For our hybrid method (4) the two main parameters are N_f (the number of fixed wave numbers at the lower end of the spectrum) and N_r (the number of random wave numbers at the higher end). The hybrid scheme also depends on an intermediate wave number K', which plays the simultaneous role of upper cutoff for the fixed wave number method and lower cutoff for the random wave number method. By our rule of

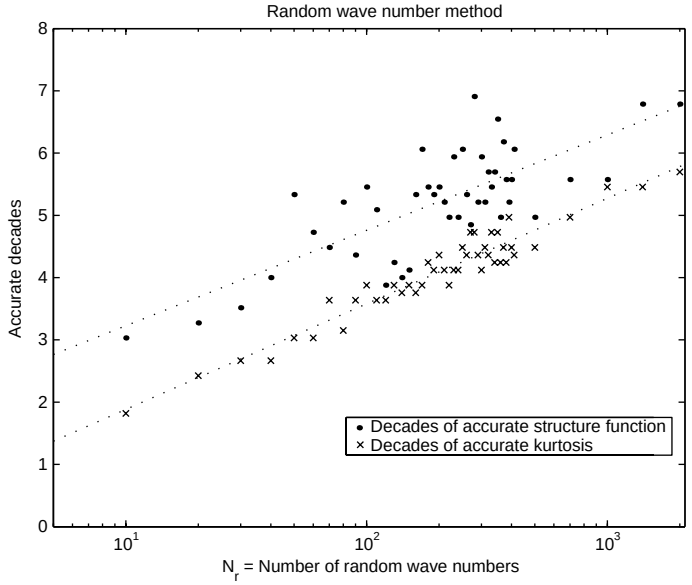

Fig. 2. Random wave number method results: Decades of structure function and kurtosis plotted simultaneously. Dashed lines are least-squares fits with approximate slopes $m = 0.67$ and $m = 0.74$, respectively.

thumb for the fixed wave number method, we have $N_f \approx 5K'$. The parameters K' and N_r, therefore, are equally sufficient to specify the method.

Nine logarithmically equidistant values of K' were selected; they corresponded to the range of values of K used to test the fixed wave number method previously and ranged between $K' = 10$ and $K' \approx 4642$. Similarly, ten values of N_r, the random wave number parameter, were chosen, ranging between $N_r = 10$ and 1000 (also logarithmically equidistant). The hybrid method was then run for all 90 pairs (K', N_r) coming from these parameter values.

Fig. 3 represents the computed accurate decades of both the structure function and the kurtosis using the hybrid method, represented as a surface plot above the two-dimensional grid of parameter values. When a corresponding surface was fit to the resulting data set for the structure function, the approximate slopes $m_F = 0.55$ and $m_R = 0.71$ were obtained. This appears to coincide reasonably well with our estimates of the log-linear dependence of each method (used individually) on its parameter, in the previous two sections. A similar fit was performed for the kurtosis data, revealing the coefficients $m_F = 0.49$ and $m_R = 0.80$, respectively.

As can be seen from the consistent gap between the two surfaces in Fig. 3, the hybrid method inherits the property of the random wave number method that it loses its multi-point Gaussianity (as measured by the kurtosis) at high wave numbers. The kurtosis accurate decades are 1–2 less than the structure function accurate decades, more or less consistently.

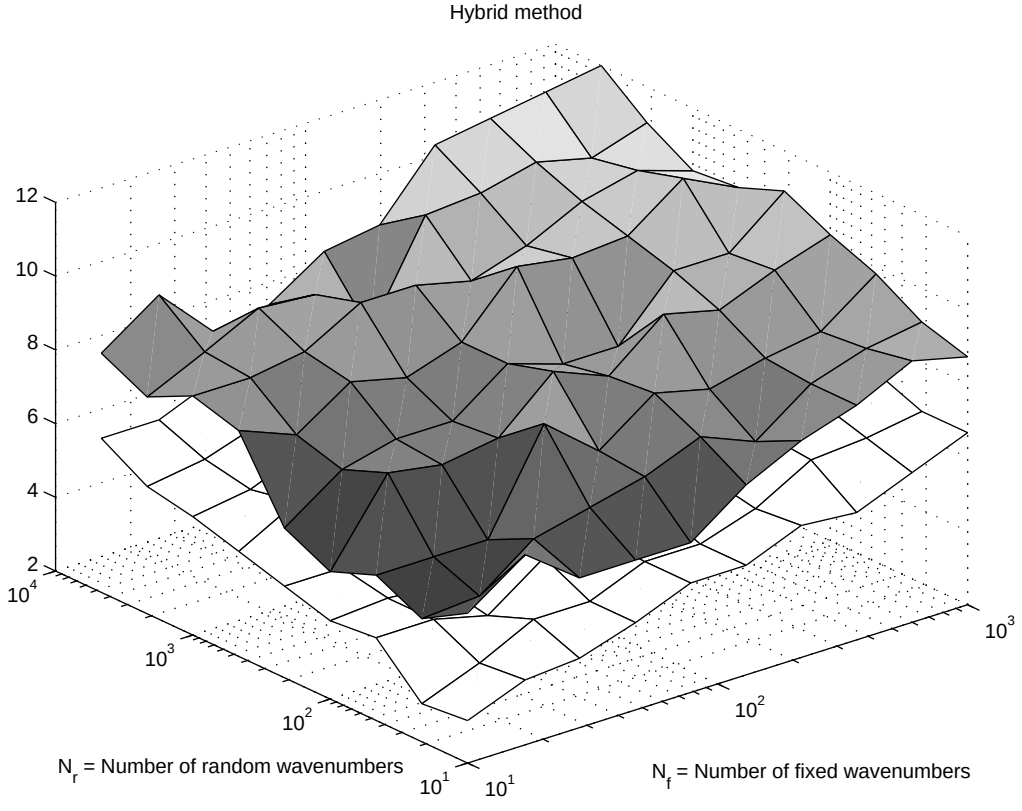

Fig. 3. Surface plot of decades of accuracy of hybrid method, versus each of its parameters K' and N_r (in log scale). Solid surface represents the structure function decades; mesh surface below it represents the decades of accurate kurtosis. Log-linear fit in direction of K' and N_r reveals slopes of 0.55 and 0.71, respectively, for the structure function. For kurtosis the corresponding slopes are 0.49 and 0.80.

On the other hand, it is seen that for the highest parameters tested, this method was capable of representing between 11 and 12 decades of accuracy, in the structure function. It will be seen that the total cost for doing so is comparable with the cost of running either method separately and obtaining far fewer decades of accuracy.

6.3.1. Optimizing the accuracy

Recall from Section 5.3 that our approximate cost (per realization, per x value) for the hybrid method depends only on the sum $2(N_f + N_r)$. The cost is therefore constant along the lines $N_f + N_r = (\text{constant})$. However, the number of decades is not constant and appears to have a maximum for some nontrivial combination of these parameter values. This suggests a strategy for optimizing the hybrid algorithm by selecting, for a given total cost, the particular values of N_f and N_r which maximize the number of decades seen in Fig. 3. In Section 5.4 we gave a heuristic argument for the existence and choice of such an optimal balance between the two methods.

And from the planar fit which was performed, resulting in coefficients m_F and m_R in the K' (or N_f, since N_f depends linearly on K') and N_r directions, respectively, a direction of largest increase can be obtained by following the gradient $(m_F, m_R)^T$. Maximizing the fitted surface with respect to the constraint $N_f + N_r = (\text{constant})$ reveals that the optimal choices should satisfy

$$N_r = \frac{m_R}{m_F} N_f.$$

In other words, the amount of each method one should include in the hybrid construction (as measured by their comparable parameters N_f and N_r, respectively) ought to be in proportion to the respective slopes m_R and m_F which describe their accuracy versus cost dependence.

From Fig. 3, we see slopes of approximately $m_R \approx 3/4$ and $m_F \approx 1/2$, respectively. The ratio between the two is $3/2$, which suggests we take $N_r = \frac{3}{2} N_f$. Compare this with the discussion in Section 5.4, where this same balance of

$$N_r = \frac{-1}{\alpha + 1} N_f \left(= \frac{3}{2} N_f \right)$$

was obtained analytically based on heuristic considerations.

Consequently, a subsequent set of runs was performed for 19 logarithmically equidistant choices of the parameter N_f, ranging from 5 to 3476, and then choosing the random wave number parameter N_r according to the optimal relation (6.3.1). The results are presented in Fig. 4.

Notice the same result we saw in Fig. 3, that up to 12 decades of accuracy were achieved. (Results for the kurtosis followed the same basic pattern as in the more extensive test, lagging 1–2 decades behind the structure function.) This suggests that we succeeded in following the line of steepest ascent in Fig. 3. The approximate slope is $m = 1.18$, which represents a marked improvement over either of the other two methods employed separately. (Note that since $N_r \propto N_f$, the computational cost C is linearly related to N_f, so the slope is unchanged if we plot d against $\log(C)$ instead of $\log(N_f)$.) Compare with the expected result from Section 5.4 of $5/(2\log(10)) \approx 1.09$.

6.4. Results for the Fourier-wavelet method

In our tests for the Fourier-wavelet method (10) with $E(k) = k^{-5/3}$, following [6] we chose $N = 10$ as our cutoff for the sum over n, which ranged from $-N + 1$ to N, because the kernel function $f_0(\cdot)$ appeared to have most of its support in the interval $[-10, 10]$. The main parameter of the method was thus M, the number of octaves of detail to include in the approximation (technically, this number is $M + 1$ because

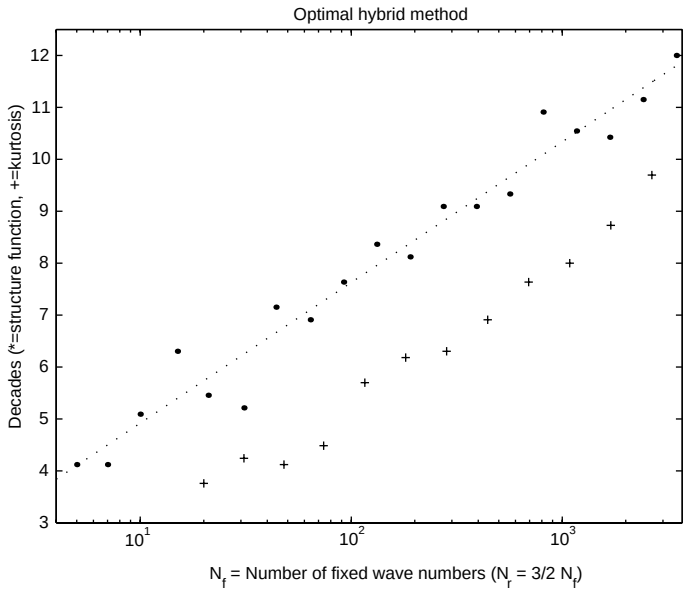

Fig. 4. Results for the hybrid method optimized to maximize accuracy versus complexity. The x axis is N_f (in log scale), the number of mesh points for the fixed wave number method. The number of random wave numbers is determined by the guideline $N_r = (3/2)N_f$. The dotted line is the best log-linear fit of the structure function decades and has slope 1.18.

$m = 0$ represents the largest octave included). Tests were performed for $M = 10, 11, 12, \ldots, 40$. The results for the structure function are shown in Fig. 5.

The approximate slope of the fitted line is $m = 0.29$. For the largest value $M = 40$, the method was capable of modelling between 10 and 11 accurate decades.

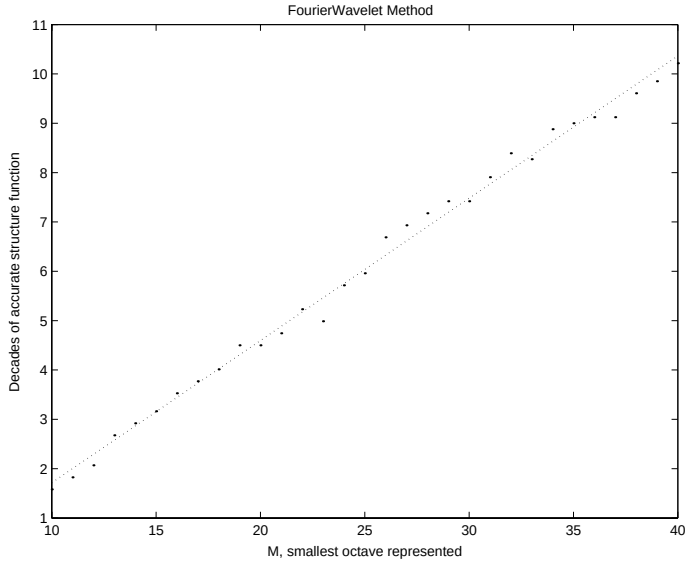

Fig. 5. Results for Fourier-wavelet method of Elliott et al. Note the linear, rather than log-linear, dependence of accurate decades on M, which is proportional to the computational complexity. The approximate slope of the dependence, represented by the dashed line, is $m = 0.29$.

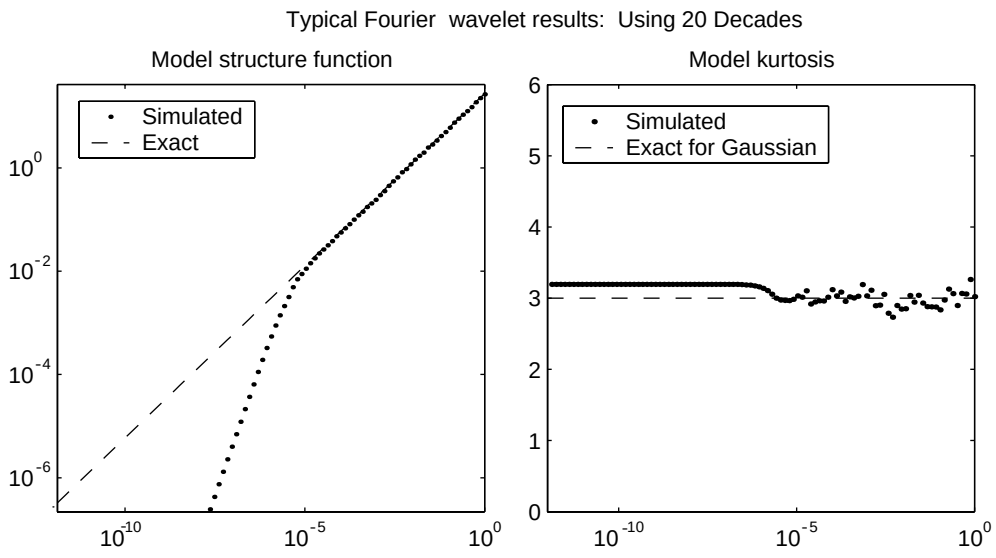

Fig. 6. Illustration of Fourier-wavelet results using 20 decades and 2000 realizations.

It was observed that the results were roughly Gaussian ($\mathscr{K} \approx 3$) over the entire spatial range of the model. Furthermore, for x below a certain threshold a plot of $\mathscr{K}$ was relatively flat and still close to 3, to within the error of our setup (see Fig. 6). This highly Gaussian behavior is likely due to the multiscale nature of the wavelet expansion: at sufficiently small scales, there is roughly only one octave of functions $f_m(\cdot)$ (the lowest, $m = -M$) which introduces significant variation into the expansion for $\tilde{u}(x)$. Thus for all such sufficiently small x, the representation is almost a sum of independent Gaussians, and so the Gaussianity (though not the structure function, of course) is almost exact.

6.5. Cost comparison

To compare the accuracy versus computational cost C of the preceding methods, we plot decades of accuracy versus C for each method (using the cost calculations for each method derived previously) on a single graph. The results are displayed in Fig. 7.

The Fourier-wavelet method begins to overtake the random wave number scheme when somewhere around 5 decades of accuracy are required in the structure function. This is consistent with the statements of the authors [6,12]. The optimal hybrid method, by contrast, evidently performs better than the random wave number method at the same cost level, with a higher (log-linear) slope. It is inevitably overtaken by the linear payoff displayed by the Fourier-wavelet method, although the crossover point is higher; based on this set of tests, one can see that for situations requiring roughly 9 or more decades of accuracy, there is a clear advantage to using the Fourier-wavelet method to sample a Gaussian random field with spectrum $E(k) = k^{-5/3}$ over a simpler method of more modest construction such as the optimal hybrid scheme described herein.

6.6. Other spectra

Finally let us examine how the preceding picture changes when the spectrum $E(k)$ is varied. We consider here two alternative spectra of the form $E(k) = |k|^{\alpha}$, where the choices $\alpha = -3/2$ and $\alpha = -2$ were made. Recall from the discussion in Section 5.4 how the behavior of each method can be expected to change with the exponent α. The fixed wave number method appears insensitive to this choice. For the random wave

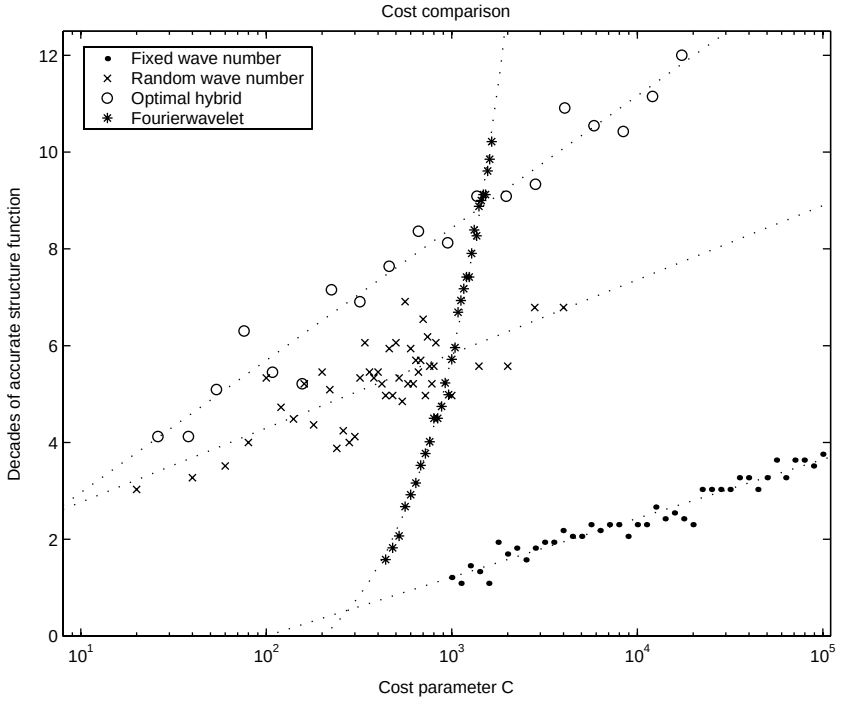

Fig. 7. Comparison of accuracy versus cost for the methods. Horizontal axis is logarithmically scaled. The curve for the Fourier-wavelet method crosses that of the random wave number scheme at roughly 5.8 decades, and that of the optimal hybrid scheme at 8.8 decades.

number method we expect the accuracy to depend on computational cost with a log-linear slope which is $|1+\alpha|^{-1}$ times that for the fixed wave number method. Consequently, the closer $\alpha(<-1)$ is to -1, the better we expect the random wave number method to perform compared with using fixed wave numbers.

For the optimized hybrid method, which combines aspects of both of these methods in an optimal way, the predicted slope was $\alpha/(\alpha+1)$, which improves upon that of the random wave number method by a factor of $|\alpha|$. As a result, when $\alpha \ll -1$ we expect the hybrid method to perform significantly better than the random wave number method. At the same time, when $\alpha \lesssim -1$, although the accuracy of the hybrid method will have a slope similar to that of the random wave number method, we still expect a shift upward in the graph due to a constant term of $|1+\alpha|^{-1}\log|1+\alpha|$ which arose in our heuristic considerations.

Finally, as before we expect the accuracy of the Fourier-wavelet method to depend linearly on its cost (which is proportional to the number of octaves included in the wavelet expansion).

In what follows we present the results of tests to create plots which are equivalent to Fig. 7 for each of the two alternative spectra.

6.6.1. Alternative spectrum I

Consider the spectrum

$$E(k) = k^{-3/2},$$

which has a more slowly decaying tail, and thus more energy contained in high wave numbers, than the Kolmogorov spectrum $k^{-5/3}$.

The results of testing for this spectrum are displayed in Fig. 8. As expected, the random wave number method appears to be doing marginally better for the larger (closer to -1) exponent α, in capturing the

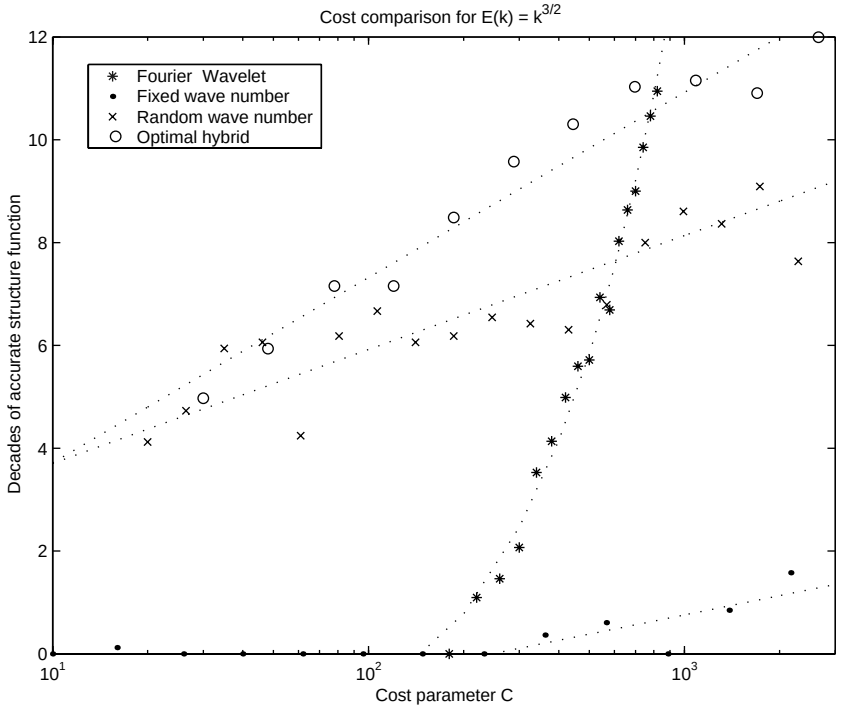

Fig. 8. Cost comparison for $E(k) = k^{-3/2}$. Log-linear slopes for fixed, random, and hybrid methods are 0.54, 0.96, and 1.56, respectively. Fourier-wavelet method begins to outperform the hybrid method when between 10 and 11 decades of accuracy are required.

more slowly decaying tail. The hybrid method also displays a significant increase in its slope with the exponent closer to -1.

For the three methods (fixed, random, and hybrid) we obtained the log-linear slopes of 0.54, 0.96, and 1.56, respectively. Compare with the results of Section 5.4; for $\alpha = -3/2$ we expect these slopes to be $\log(10)^{-1} \approx 0.43$, $2\log(10)^{-1} \approx 0.87$, and $3\log(10)^{-1} \approx 1.30$. The results are comparable, particularly in terms of the ratios between each of the three successively better methods.

As before, due to its linear behavior the Fourier-wavelet method becomes significantly more efficient than any of the others, including the hybrid method, when a sufficient number of decades of accuracy are required. For this spectrum we see such a crossover when between 10 and 11 decades (compared with 8–9 for the original spectrum) are required in the structure function. While changing α has affected the log-linear slopes associated with the random wave number and hybrid methods, the linear slope of the Fourier-wavelet method (about 0.30, as before) was unaffected.

Overall, the increase in the exponent α (thus, in the thickness of the tail of the spectrum) has had the overall effect of making both the hybrid and random wave number methods more relatively efficient than before, as compared with the Fourier-wavelet method.

6.6.2. Alternative spectrum II

Now consider the spectrum

$$E(k) = k^{-2},$$

which has a tail that decays more rapidly, and more energy near the origin $k \approx 0$, than the Kolmogorov spectrum $k^{-5/3}$.

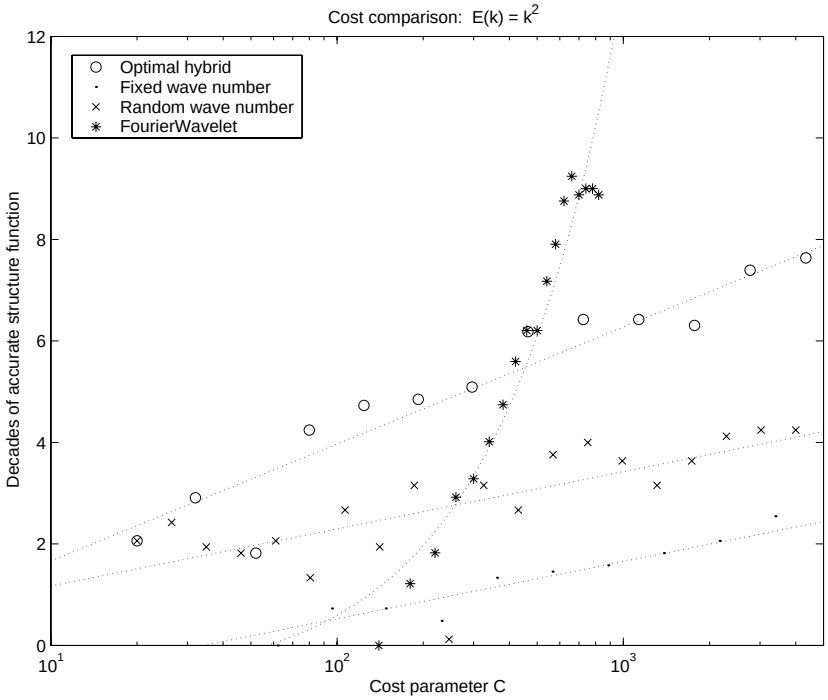

Fig. 9. Cost comparison for $E(k) = k^{-2}$. Log-linear slopes for fixed, random, and hybrid methods are 0.49, 0.49, and 1.00, respectively. Fourier-wavelet method begins to outperform the hybrid method when between 5 and 6 decades of accuracy are required.

From the discussion in Section 5.4, we expect this decrease in α to reduce the effectiveness of the random wave number and hybrid methods (although the latter should remain a significant improvement over the former). In particular we expect to see log-linear slopes of $\log(10)^{-1}$, $\log(10)^{-1}$, and $2\log(10)^{-1}$ for the fixed, random, and hybrid methods, respectively.

The results are displayed in Fig. 9. As expected, for this spectrum the effectiveness of the random wave number and hybrid methods has been reduced. The slopes of 0.49, 0.49, and 1.00 for the fixed, random, and hybrid methods are comparable to what was expected from Section 5.4.

By contrast, the Fourier-wavelet method shows no corresponding loss of effectiveness in dealing with this spectrum. Consequently the crossover point has been reduced, to between 5 and 6 decades. This can be attributed to the wavelet construction of this method; when M octaves are included in the expansion, the Fourier-wavelet method includes features from the corresponding length scale of $2^{-M}X$, regardless of the energy spectrum being approximated. For this spectrum it is evident that the other methods, by their construction, are required to focus most of their attention on wave numbers near $k \approx 0$ where most of the energy is contained. Since the Fourier-wavelet method lacks this necessity, one sees that it is an especially appealing method for sampling energy spectra with a large amount of energy contained in low wave numbers and rapidly decaying tails.

7. Conclusions

A method for sampling stationary Gaussian stochastic processes has been tested and compared with certain other methods of simple construction and/or wide use. The Fourier-wavelet method was found to

differ markedly from the other methods in that its effective accuracy, measured in terms of decades, increases linearly (rather than logarithmically) with computational complexity.

This linearly increasing efficiency of the Fourier-wavelet method placed it in stark contrast to the other methods, whose constructions all entail a built-in logarithmically increasing efficiency. On the other hand, as its authors point out, the advantage of the Fourier-wavelet method is not immediate, but rather shows up – dramatically so – for situations necessitating a certain number of decades of accuracy in the solution.

For the well-known Kolmogorov energy spectrum $E(k) = k^{-5/3}$, the results for a hybrid scheme displayed up to 8 or more decades of accuracy, at a comparable cost to that of the Fourier-wavelet method. Although the accuracy of the hybrid method, like that of the ones on which it was built, displayed a log-linear dependence on computational cost, we saw that it had a greater marginal payoff than either the random or fixed wave number methods, particularly for power law spectra with rapidly decaying tails $\alpha \ll -1$. At the same time, for slowly decaying tails $\alpha \lesssim -1$, the hybrid method is still expected to perform consistently better than the random wave number method, due to the fact that it allows higher wave numbers to be accessed for the same cost. One caveat is that any hybrid method using random wave numbers will inherit the inability of such randomization methods to consistently display multi-point Gaussianity at the highest wave numbers.

Testing for power law spectra with other exponents produced mixed results. Because the random wave numbers accessed by the hybrid method depend on the exponent of the energy spectrum, we saw that it can be expected to perform better for spectra with $\alpha \lesssim -1$ than for spectra with more rapidly decaying tails, such as $\alpha \lesssim -2$. This is to be contrasted with the Fourier-wavelet method, whose behavior remained roughly the same, regardless of α. As a result, the crossover point – the point at which the Fourier-wavelet method begins to outperform the hybrid method – depended quite sensitively on the exponent α. For $\alpha = -3/2$ this crossover occurred when about 10 or 11 decades of accuracy are required in the structure function; for $\alpha = -2$, the same crossover point occurred when only 5 or 6 decades are needed.

Overall, it was clear that for situations calling for a sufficiently large number of decades of accuracy, the Fourier-wavelet method shows a distinct advantage over the other methods. We have also seen that it is possible to construct a simpler method which would give comparable accuracy for a comparable cost for situations requiring no more than a certain amount of accuracy, particularly for power spectra with relatively slowly decaying tails. Because of the somewhat trickier implementation of the Fourier-wavelet method, due in no small part to the necessity to generate Gaussian random variables on an as needed basis, a rough knowledge of crossover points may provide useful when considering these methods.

Acknowledgements

I would like to thank Prof. A.J. Chorin for his very helpful comments. This work was supported in part by the Office of Science, office of Advanced Scientific Computing Research, Mathematical, Information, and Computational Sciences Division, Applied Mathematical Sciences Subprogram, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 and in part by the National Science Foundation under grant number DMS89-19074.

References

[1] L. Arnold, Stochastic Differential Equations, Wiley, New York, 1974.
[2] C. Blatter, Wavelets: A Primer, A.K. Peters, Natick, MA, 1998.
[3] C. Chui, An Introduction to Wavelets, Academic Press, San Diego, 1992.
[4] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.
[5] J.L. Doob, Stochastic Processes, Wiley and Sons, New York, 1953.

[6] F. Elliott Jr., D. Horntrop, A. Majda, A Fourier-Wavelet Monte Carlo method for fractal random fields, J. Comput. Phys. 132 (1997) 384–408.
[7] F. Elliott, A. Majda, A wavelet Monte Carlo method for turbulent diffusion with many spatial scales, J. Comput. Phys. 113 (1994) 82.
[8] I.M. Gel'fand, G.E. Shilov, Generalized Functions: vol. 1, Properties and Operations, Academic Press, New York, 1964.
[9] I.I. Gikhman, A.V. Skorokhod, The Theory of Stochastic Processes, Springer, 1979.
[10] A.E. Koniges, C.E. Leith, Parallel processing of random number generation for Monte Carlo turbulence simulation, J. Comput. Phys. 81 (1989) 230–235.
[11] M.J. Lighthill, Introduction to Fourier Analysis and Generalised Functions, Cambridge University Press, New York, 1958.
[12] A. Majda, P. Kramer, Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena, Phys. Rep. 314 (1999) 237–574.
[13] A. McCoy, A numerical study of turbulent diffusion (PhD thesis), Department of Mathematics, University of California at Berkeley, 1975.
[14] K. Sabelfeld, Monte Carlo Methods, Springer-Verlag, New York, 1991.
[15] A.M. Yaglom, An Introduction to the Theory of Stationary Random Functions, Prentice-Hall, Englewood Cliffs, NJ, 1962.