ELSEVIER

PII: S0309-1708(96)00064-4

*Advances in Water Resources* 21 (1998) 385–399

Printed in Great Britain
0309-1708/98/$19.00

# An efficient multivariate random field generator using the fast Fourier transform

**Feng Ruan & Dennis McLaughlin**

*Ralph M. Parsons Laboratory, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

(Received 31 May 1996; accepted 16 November 1996)

It is often convenient to use synthetically generated random fields to study the hydrologic effects of spatial heterogeneity. Although there are many ways to produce such fields, spectral techniques are particularly attractive because they are fast and conceptually straightforward. This paper describes a spectral algorithm for generating sets of random fields which are correlated with one another. The algorithm is based on a discrete version of the Fourier–Stieltjes representation for multidimensional random fields. The Fourier increment used in this representation depends on a random phase angle process and a complex-valued spectral factor matrix which can be readily derived from a specified set of cross-spectral densities (or cross-covariances). The inverse Fourier transform of the Fourier increment is a complex random field with real and imaginary parts which each have the desired coveriance structure. Our complex-valued spectral formulation provides an especially convenient way to generate a set of random fields which all depend on a single underlying (independent) field, provided that the fields in question can be related by space-invariant linear transformations. We illustrate this by generating multi-dimensional mass conservative groundwater velocity fields which can be used to simulate solute transport through heterogeneous anisotropic porous media. © 1998 Elsevier Science Ltd. All rights reserved.

*Key words:* multivariate random field, random field generation, fast Fourier transform, spectral analysis.

## 1 INTRODUCTION

Synthetically generated random fields are often used to study the effects of heterogeneity in situations where field investigations may be inconclusive or prohibitively expensive. An important example is the simulation of contaminant transport through a porous medium with spatially variable log hydraulic conductivity. If the log conductivity varies in a complex way it may be reasonable to treat it as a random field with specified statistical properties. Solute concentrations derived from such a random field can provide insight about the relationship between geological variability and solute dispersion.[3,35] Other examples of interest to hydrologists and soil scientists include the generation of correlated spatially variable unsaturated soil properties (e.g. the coefficients used to describe moisture retention and relative permeability curves) or correlated spatially variable land surface variables (e.g. vegetative cover, soil moisture and temperature).

In some applications, it may be useful to generate several different random fields that are correlated with one another by virtue of their dependence on some common physical property. The contaminant transport simulation mentioned above may, for example, rely on a synthetic velocity field whose components are related in a well-defined way to hydraulic conductivity. If this field is to be mass conservative the various components of velocity must be correlated so as to ensure that the divergence of the velocity vector is zero. This requirement is easier to satisfy if the hydraulic conductivity dependence is explicitly incorporated into the velocity generation algorithm.

In this paper we present a convenient method for generating correlated random fields which are related by a set of specified spectral densities. Our method is based on a discrete complex-valued spectral representation which is easy to implement with readily available fast Fourier transform (FFT) algorithms. There are many other techniques for generating random fields, some designed only for univariate applications and some for more general multivariate applications. These

techniques can be broadly divided into covariance-based methods and spectrally-based methods. We provide a brief review of the available alternatives in the following paragraphs.

Many of the covariance-based random field generators in current use obtain spatially correlated random fields from linear combinations of a large number of uncorrelated random variables (which are easy to generate). The various alternatives differ primarily with respect to the methods used to derive the required linear transformations. Some examples include the techniques proposed by Smith and Freeze,[30,31] Smith and Schwartz,[32] and Davis,[9] Random field generators based on the linear transformation concept are typically limited to problems where no more than a few hundred random field values are required. This makes such generators inappropriate for large multi-dimensional flow or transport simulations (which can easily require millions of correlated random field values). A more efficient alternative is to use a sequential conditioning approach which derives each new random field value from a small number of neighboring values. Examples include the generators described by Journel and Huijbregts,[20] Delhomme,[10] Hernandez,[19] Rubin,[29] and Cushy *et al.*[7] The conditioning weights used in these sequential algorithms are computed from the covariances and cross-covariances of the fields to be generated. In some applications, such as our multivariate velocity example, it may be quite difficult to derive the required covariances.

The most popular spectral approaches for generating random fields are probably the turning bands method[11,12,22,23,33,34] and the direct fast Fourier transform method.[5,16,27] The turning bands algorithm generates a number of one-dimensional line processes by inverse Fourier transforming independent random Fourier increments derived from an appropriate one-dimensional (radial) spectral density function. A multi-dimensional random field is then obtained by projecting and superimposing values from the various lines into a discrete computational grid. In order to apply the turning bands method, the radial spectral density must be derived from the spectral density function of the multidimensional random field. The success of this method depends on the feasibility of adequately reproducing multidimensional spectral properties from spectral line processes.

The direct fast Fourier transform (FFT) method produces stationary random fields by inverse Fourier transforming independent random Fourier increments generated on a multi-dimensional wave number grid.[5,16,17,27] The procedure is straightforward in the univariate case but somewhat more difficult when several mutually-correlated random fields are required. This difficulty is related to the fact that the cross-spectral densities between correlated random fields are generally complex-valued, even when the fields are real. FFT methods that can only work with real-valued spectra, such as those described by Robin *et al.*[27] and Gutjahr *et al.*,[17] must partition complex cross-spectra into real and imaginary parts. This complicates implementation of the multivariate FFT algorithm.

Here, we propose a new FFT method that is based on a more general complex-valued spectral representation which readily accommodates complex-valued cross-spectra. The complex-valued representation greatly facilitates the generation of multiple correlated random fields, especially when the number of fields is large. If the individual fields are related to a single underlying field by a space-invariant linear transformation, they may be derived by operating on a single scalar sequence of spectral increments with a set of complex-valued spectral domain transfer functions. As a side benefit, the complex-valued spectral representation enables us to generate two statistically identical but uncorrelated random replicates with the same effort as real-valued methods use to generate a single replicate.

In the following sections, we present the theoretical basis for our random field generator and briefly discuss discrete implementation with a fast Fourier transform algorithm. The algorithm we propose is particularly convenient for generating correlated fields which must obey physical constraints. We illustrate this with a series of examples which yield correlated log hydraulic conductivity and velocity fields. These fields reproduce ensemble statistics derived from linearized groundwater flow and Darcy equations. Moreover, the individual velocity replicates conserve mass, as predicted by a theoretical analysis. We conclude with a brief review of the capabilities and limitations of our random field generator.

## 2 GENERAL THEORY

We are interested in developing a numerical method for generating $M$ correlated zero-mean stationary random fields. These fields are related in the space domain by a set of cross-covariance functions and in the spectral (wave-number) domain by a corresponding set of cross-spectral density functions. The covariance and spectral density functions may be specified directly or derived indirectly from physical laws. Each of the $M$ random fields may be expressed as a Fourier–Stieltjes integral over a random Fourier increment $\mathrm{d}Z_{Y_i}(\mathbf{k})$ (Refs 21, 26)

$$Y_i(\mathbf{x}) = \int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot\mathbf{x}}\,\mathrm{d}Z_{Y_i}(\mathbf{k}) \qquad i = 1,\ldots,M \tag{1}$$

The random increment must satisfy the following orthogonality conditions (Ref. 26, p. 245)

$$\begin{aligned} E[\mathrm{d}Z_{Y_i}(\mathbf{k})] &= 0 \\ E[\mathrm{d}Z_{Y_i}(\mathbf{k})\,\mathrm{d}Z^*_{Y_j}(\mathbf{k}')] &= 0 \qquad \mathbf{k} \neq \mathbf{k}' \quad i,j = 1,\ldots,M \\ E[\mathrm{d}Z_{Y_i}(\mathbf{k})\,\mathrm{d}Z^*_{Y_j}(\mathbf{k})] &= S_{ij}(\mathbf{k})\,\mathrm{d}\mathbf{k} \end{aligned} \tag{2}$$

where $E[\cdot]$ indicates mathematical expectation, $*$ indicates the complex conjugate operator, $\imath = \sqrt{-1}$ and $S_{ij}(\mathbf{k})$ is the cross-spectral density between the two random fields $Y_i(\mathbf{x})$ and $Y_j(\mathbf{x})$. The vector differential $\mathrm{d}\mathbf{k}$ represents the $p$ dimensional differential wave number volume element $\mathrm{d}k_1, \ldots, \mathrm{d}k_p$, where $p$ is the dimension of the spatial vector $\mathbf{x}$.

Equation (2) specifies that the two Fourier increments $\mathrm{d}Z_{Y_i}(\mathbf{k})$ and $\mathrm{d}Z_{Y_j}(\mathbf{k}')$ are statistically correlated only when $\mathbf{k} = \mathbf{k}'$. The cross-spectral densities among all $M$ fields can be assembled in an $M$ by $M$ spectral density matrix which is generally complex-valued, even when the $Y_i(\mathbf{x})$ are real. When the spectral density matrix is Hermitian (i.e. when $S_{ij} = S_{ji}^*$) and positive definite, its elements can be factored as follows[18]

$$S_{ij}(\mathbf{k}) = H_{il}(\mathbf{k})H_{jl}^*(\mathbf{k}) \qquad i, j = 1, \ldots, M \tag{3}$$

where repeated indices imply summation from 1 to $M$. This spectral factorization operation can be tedious in the general case since it must be carried out iteratively.

If we can generate a complex-valued Fourier increment process $\mathrm{d}Z_{Y_i}(\mathbf{k})$ having the statistics specified in (2) then we can use (1) to generate the random field $Y_i(\mathbf{x})$. In practice this requires that $\mathrm{d}Z_{Y_i}(\mathbf{k})$ be discretized so that the integration in (1) can be carried out numerically. The discrete approximation to the random Fourier increment can be constructed in a number of different ways. The version we use is given by the following expression

$$\mathrm{d}Z_{Y_i}(\mathbf{k}) \approx \Delta Z_{Y_i}(\mathbf{k}) = H_{il}(\mathbf{k})\, \mathrm{e}^{\imath\theta_l(\mathbf{k})}|\Delta\mathbf{k}|^{1/2} \qquad i = 1, \ldots, M \tag{4}$$

where $H_{il}(\mathbf{k})$ is the complex-valued deterministic matrix obtained from the spectral factorization of (3) and the phase angle $\theta_l(\mathbf{k})$ is a random process in the discretized wave number domain. Phase angle values at different wave numbers or with different subscripts are required to be statistically independent. Also, each of the $M$ phase angle processes is uniformly distributed over the range $[0, 2\pi]$ at any given wave number. The wave number discretization is performed on a $p$-dimensional rectangular computational grid. The scalar $|\Delta\mathbf{k}| = \prod_{j=1}^{p} \Delta k_j$ is the volume of the grid element associated with a particular discrete wave number $\mathbf{k}$.

When (4) is substituted into (1) we obtain a complex-valued spectral representation for $Y_i(\mathbf{x})$ which is suitable for discrete computation

$$\begin{aligned} Y_i(\mathbf{x}) &= \int_{-\infty}^{\infty} \mathrm{e}^{\imath\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}Z_{Y_i}(\mathbf{k}) \\ &\approx \sum_n \mathrm{e}^{\imath\mathbf{k}_n\cdot\mathbf{x}} H_{il}(\mathbf{k}_n)\, \mathrm{e}^{\imath\theta_l(\mathbf{k}_n)}|\Delta\mathbf{k}_n|^{1/2} \\ & \qquad i = 1, \ldots, M \end{aligned} \tag{5}$$

The summation of this expression is understood to be over all of the discrete wave numbers included in the $p$-dimensional computational grid. A more detailed discussion of the wave number integration is provided later in this paper. It should be noted that the synthetic random fields generated by the discrete representation of (5) are normally distributed, by virtue of the central limit theorem. Gutjahr *et al.*[17] provide a more detailed discussion of this point.

The discrete representation of (5) is valid only if $\Delta Z_{Y_i}(\mathbf{k})$ and $\Delta Z_{Y_j}(\mathbf{k})$ can be shown to satisfy the orthogonality conditions of (2) in the limit as the discretization becomes finer (i.e. as the wave number volume elements shrink in size). The first (zero-mean) condition can be confirmed by noting that the expectation of $\Delta Z_{Y_i}(\mathbf{k})$ is an integral over $p_{\theta_l}(\theta_l; \mathbf{k})$, the univariate probability density of the $\theta_l(\mathbf{k})$ process. This integral may be written as

$$\begin{aligned} E[\Delta Z_{Y_i}(\mathbf{k})] &= H_{il}(\mathbf{k})|\Delta\mathbf{k}|^{1]2} \int_{-\infty}^{\infty} \mathrm{e}^{\imath\zeta} p_{\theta_l}(\zeta; \mathbf{k})\, \mathrm{d}\zeta \\ &= H_{il}(\mathbf{k})|\Delta\mathbf{k}|^{1/2} \int_0^{2\pi} \mathrm{e}^{\imath\zeta} \frac{1}{2\pi}\, \mathrm{d}\zeta \\ &= 0 \end{aligned} \tag{6}$$

This discrete expansion approaches $E[\mathrm{d}Z_{Y_i}(\mathbf{k})]$ in the limit as $|\Delta\mathbf{k}|$ approaches zero.

The second and third conditions of (2) can be confirmed by substituting (4) for $\Delta Z_{Y_i}(\mathbf{k})$ and $\Delta Z_{Y_j}^*(\mathbf{k}')$ and taking the expectation of the product. Since $\theta_l(\mathbf{k})$ and $\theta_m(\mathbf{k}')$ are independent for $l \neq m$ or $\mathbf{k} \neq \mathbf{k}'$, the result is

for $\mathbf{k} \neq \mathbf{k}'$

$$\begin{aligned} E[\Delta Z_{Y_i}(\mathbf{k})\Delta Z_{Y_j}^*(\mathbf{k}')] &= H_{il}(\mathbf{k})H_{jm}^*(\mathbf{k}')|\Delta\mathbf{k}|^{1/2} \\ &\quad \times |\Delta\mathbf{k}'|^{1/2} E[\mathrm{e}^{\imath\theta_l(\mathbf{k}) - \imath\theta_m(\mathbf{k}')}] \\ &= H_{il}(\mathbf{k})H_{jm}^*(\mathbf{k}')|\Delta\mathbf{k}|^{1/2} \\ &\quad \times |\Delta\mathbf{k}'|^{1/2} E[\mathrm{e}^{\imath\theta_l(\mathbf{k})}] E[\mathrm{e}^{-\imath\theta_m(\mathbf{k}')}] \\ &= 0 \end{aligned} \tag{7}$$

for $\mathbf{k} = \mathbf{k}'$

$$\begin{aligned} E[\Delta Z_{Y_i}(\mathbf{k})\Delta Z_{Y_j}^*(\mathbf{k}')] &= H_{il}(\mathbf{k})H_{jm}^*(\mathbf{k})|\Delta\mathbf{k}|^{1/2} \\ &\quad \times |\Delta\mathbf{k}|^{1/2} E[\mathrm{e}^{\imath\theta_l(\mathbf{k}) - \imath\theta_m(\mathbf{k}')}] \\ &= H_{il}(\mathbf{k})H_{jm}^*(\mathbf{k})|\Delta\mathbf{k}|\delta_{lm} \\ &= H_{il}(\mathbf{k})H_{jl}^*(\mathbf{k})|\Delta\mathbf{k}| \\ &= S_{ij}(\mathbf{k})|\Delta\mathbf{k}| \end{aligned} \tag{8}$$

where $\delta_{lm}$ is equal to 1 when $l = m$ and equal to 0 when $l \neq m$. The deterministic discrete expression $S_{ij}(\mathbf{k})|\Delta\mathbf{k}|$ approaches $S_{ij}(\mathbf{k})\,\mathrm{d}\mathbf{k}$ as $|\Delta\mathbf{k}|$ approaches zero.

The properties of (2) may be used to show that the

cross-covariance $R_{ij}(\mathbf{x}-\mathbf{x}') = E[Y_i(x)Y_j^*(\mathbf{x}')]$ is the inverse Fourier transform of the cross-spectral density $S_{ij}(\mathbf{k})$:

$$R_{ij}(\mathbf{x}-\mathbf{x}') = E[Y_i(\mathbf{x})Y_j^*(\mathbf{x}')]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot\mathbf{x}-i\mathbf{k}'\cdot\mathbf{x}'} E[dZ_{Y_i}(\mathbf{k})\,dZ_{Y_j}^*(\mathbf{k}')]$$
$$= \int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} S_{ij}(\mathbf{k})\,dk \qquad (9)$$

Similarly, the cross-spectral density is the Fourier transform of the cross-covariance. Note that $R_{ij}(\mathbf{x}-\mathbf{x}')$ may be complex if $Y_i(\mathbf{x})$ and/or $Y_j(\mathbf{x}')$ are complex. Also, $S_{ij}(\mathbf{k})$ may be complex for $i \neq j$ even when $R_{ij}(\mathbf{x}-\mathbf{x}')$ is real.

In hydrogeologic applications the synthetic random fields of most interest are usually real fields which have real cross-covariances $R_{ij}(\mathbf{x}-\mathbf{x}')$ for $i,j = 1,\ldots,M$. The particular spectral representation presented in (5) produces complex fields even when the specified cross-covariances are real. However, real fields with the correct cross-covariance can be readily obtained from these complex $Y_i$. To see this, note that $Y_i(\mathbf{x})$ can always be decomposed into two real random fields as follows

$$Y_i(\mathbf{x}) = Y_{Ri}(\mathbf{x}) + iY_{Ii}(\mathbf{x}) \qquad l = 1,\ldots,M \qquad (10)$$

where the R and I subscripts indicate the real and imaginary parts of $Y_i(\mathbf{x})$, respectively. We show in Appendix A that $Y_{Ri}(\mathbf{x})$ nd $Y_{Ij}(\mathbf{x})$ satisfy the following properties when $R_{ij}(\mathbf{x}-\mathbf{x}')$ is real

$$E[Y_{Ri}(\mathbf{x})Y_{Ij}(\mathbf{x}')] = E[Y_{Ii}(\mathbf{x})Y_{Rj}(\mathbf{x}')] = 0$$
$$i,j = 1,\ldots,M$$
$$E[Y_{Ri}(\mathbf{x})Y_{Rj}(\mathbf{x}')] = E[Y_{Ii}(\mathbf{x})Y_{Ij}(\mathbf{x}')] = \tfrac{1}{2}R_{ij}(\mathbf{x}-\mathbf{x}')$$
$$i,j = 1,\ldots,M \qquad (11)$$

This result indicates that the real and imaginary parts of $\sqrt{2}Y_i(\mathbf{x})$ are uncorrelated replicates, each with cross-covariance $R_{ij}(\mathbf{x}-\mathbf{x}')$. So each complex field produced by our algorithm yields two real fields with the desired cross-covariance.

The above analysis shows that we can generate correlated random fields in four steps.

*Step 1.* Specify a set of cross-covariance or cross-spectral densities which characterize the statistical structure of the desired random fields.

*Step 2.* Perform a matrix factorization of the cross-spectral matrix, as indicated by eqn (3).

*Step 3.* Generate a set of $M$ random Fourier increments from eqn (4) using a set of $M$ independent random phase angle processes uniformly distributed over $[0, 2\pi]$. The phase angle values at different wave numbers should be independent.

*Step 4.* Take the inverse Fourier transforms of the random Fourier increments to obtain a set of $M$ complex or $2M$ real random fields.

In practice, it is most convenient to use a fast Fourier transform algorithm to evaluate the inverse transforms.

## 3 SIMPLIFICATION FOR RANDOM FIELDS RELATED BY LINEAR TRANSFORMATIONS

The general procedure outlined above can be simplified considerably if $M-1$ of the random fields (the dependent variables) can be related to the remaining field (the independent variable) by space-invariant linear transformations $L_i[Y_1(\mathbf{x})]$ having the following general form

$$Y_i(\mathbf{x}) = L_i(Y_1(\mathbf{x})], \qquad i = 2,\ldots,M \qquad (12)$$

where $Y_1(\mathbf{x})$ is the independent variable. An example considered in more detail later in this paper is simultaneous generation of correlated log hydraulic conductivity, longitudinal velocity and transverse velocity fields for investigations of multi-dimensional groundwater flow and transport. In this case, the log conductivity is the independent variable and the velocity components are the dependent variables. The linear transformations relating the dependent variables to the independent variable may be derived from linearized infinite-domain approximations to Darcy's law and the groundwater flow equation.

Since the linear transformations in (12) are space-invariant, this equation may be Fourier transformed to give the following spectral relationship between stationary Fourier increments of the dependent and independent variables

$$dZ_{Y_i}(\mathbf{k}) = T_i(\mathbf{k})\,dZ_{Y_1}(\mathbf{k}) \qquad i = 2,\ldots,M \qquad (13)$$

where $T_i(\mathbf{k})$ is a transfer function which can be identified directly from the Fourier transformed version of (12). In the example considered below the $T_i$ are ratios of simple polynomials in $\mathbf{k}$.

Note that when we substitute the representation of (13) into the final expression in (2) we obtain the following simplified spectral factorization for $S_{ij}(\mathbf{k})$

$$S_{ij}(\mathbf{k})\,dk = E[dZ_{Y_i}(\mathbf{k})\,dZ_{Y_j}^*(\mathbf{k})]$$
$$= T_i(\mathbf{k})T_j^*(\mathbf{k})S_{11}(\mathbf{k})\,d\mathbf{k}$$
$$= T_i(\mathbf{k})T_j^*(\mathbf{k})H_{11}(\mathbf{k})H_{11}^*(\mathbf{k})\,d\mathbf{k}$$
$$i,j = 1,\ldots,M \qquad (14)$$

where $H_{11}(\mathbf{k})$ is the scalar function obtained from a spectral factorization of the scalar independent variable spectral density $S_{11}(\mathbf{k})$ (see eqn (3)). Since $S_{11}(\mathbf{k})$ is real, $H_{11}(\mathbf{k})$ is just its square root. Equation (14) shows that the cross-spectral density matrix may be factored into the product of a complex vector $T_i(\mathbf{k})$ and its transpose when $M-1$ of the random fields are linear transformations of the remaining field.

If we use eqn (4) to generate a discrete approximation to $dZ_{Y_i}(\mathbf{k})$, we can approximate (13) by

$$dZ_{Y_i}(\mathbf{k}) \approx \Delta Z_{Y_i}(\mathbf{k}) = T_i(\mathbf{k}) H_{11}(\mathbf{k})\, e^{\imath\theta_1(\mathbf{k})} |\Delta\mathbf{k}|^{1/2} \quad i = 1, \ldots, M \tag{15}$$

where it is understood that $T_1(\mathbf{k}) = 1$. Note that we now require only a single phase angle process $\theta_1(\mathbf{k})$.

When the discrete Fourier increment expression of (15) is substituted into (1) we obtain a simplified spectral representation for $Y_i(\mathbf{x})$ which is suitable for discrete computation (compare with eqn (5))

$$\begin{aligned} Y_i(\mathbf{x}) &= \int_{-\infty}^{\infty} e^{\imath\mathbf{k}\cdot\mathbf{x}}\, dZ_{Y_i}(\mathbf{k}) \\ &\approx \sum_n e^{\imath\mathbf{k}_n\cdot\mathbf{x}} T_i(\mathbf{k}_n) H_{11}(\mathbf{k}_n)\, e^{\imath\theta_1(\mathbf{k}_n)} |\Delta\mathbf{k}_n|^{1/2} \\ &\quad i = 1, \ldots, M \end{aligned} \tag{16}$$

For this special case our random field generator can be simplified as follows.

*Step 1.* Specify the covariance or spectral density of the independent variable $Y_1(\mathbf{x})$ as well as the space-invariant linear transformations which relate this variable to the $M - 1$ dependent variables.

*Step 2.* Derive $M - 1$ spectral transfer functions from the specified linear transformations.

*Step 3.* Set $H_{11}(\mathbf{k})$ equal to the square root of the independent variable spectral density $S_{11}(\mathbf{k})$.

*Step 4.* Generate a set of $M$ random Fourier increments from (15) using a single random phase angle process uniformly distributed over $[0, 2\pi]$. The phase angle values at different wave numbers should be independent.

*Step 5.* Take the inverse Fourier transforms of the random Fourier increments as before.

This is the version of the algorithm used in our examples. Note that the simplified spectral representation of eqn (15) is mathematically equivalent to the more general one presented in eqn (5). The advantage of using eqn (15) is that we need only to evaluate one spectral factor matrix ($H_{11}$) rather than many (all the $H_{il}$). This reduces the amount of preparatory work needed to apply the FFT algorithm to multivariate problems.

## 4 DISCRETE IMPLEMENTATION

The most efficient way to compute the inverse transforms required in our random field generation procedure is to evaluate the summations of eqns (5) and (16) with a fast Fourier transform algorithm. This algorithm discretizes $\mathbf{x}$ and $\mathbf{k}$ over uniform $p$ dimensional grids. In order to simplify notation, we assume here that $p = 2$ so that the computational grid points in the space and wave number domains are given by

$$\begin{aligned} x_{m1} &= m_1 \Delta x_1, & m_1 &= 1, \ldots, L_1 \\ x_{m2} &= m_2 \Delta x_2, & m_2 &= 1, \ldots, L_2 \\ k_{n_1} &= n_1 \Delta k_1, & n_1 &= 1, \ldots, L_1 \\ k_{n2} &= n_2 \Delta k_2, & n_2 &= 1, \ldots, L_2 \end{aligned} \tag{17}$$

where $m_j$ and $n_j$ are, respectively, the discrete wave number and location indices for coordinate $j$ and $j = 1, 2$; $L_j$ is the number of grid points in coordinate $j$ for both the space and wave number grids; $\Delta k_j$ and $\Delta x_j$ are the (constant) space and wave number grid spacing in coordinate $j$. The 'bit-reversal' version of the fast Fourier transform algorithm which we use requires that $L_1$ and $L_2$ be integer powers of 2. When this particular discretization is applied to the general expression given in eqn (5) the result is

$$\begin{aligned} Y_i(x_{m1}, x_{m2}) &= Y_i(m_1\Delta x_1, m_2 \Delta x_2) \\ &= \sum_{n_1=1}^{L_1} \sum_{n_2=0}^{L_2} \exp[\imath(n_1\Delta k_1 m_1 \Delta x_1 \\ &\quad + n_2 \Delta k_2 m_2 \Delta x_2)] \\ &\quad \times \Delta Z_{Y_i}[\zeta_1(n_1\Delta k_1), \zeta_2(n_2\Delta k_2)] \\ &= \sum_{n_1=0}^{L_1} \sum_{n_2=0}^{L_2} \exp[\imath(n_1\Delta k_1 m_1 \Delta x_1 \\ &\quad + n_2 \Delta k_2 m_2 \Delta x_2)] \cdot \\ &\quad \times \sum_{l=1}^{M} [\exp\{\imath\theta_l[\zeta_1(n_1\Delta k_1), \zeta_2(n_2\Delta k_2)]\}] \\ &\quad \times H_{il}[\zeta_1(n_1\Delta k_1), \zeta_2(n_2\Delta k_2)] \\ &\quad \times [\Delta k_1 \Delta k_2]^{1/2}], \qquad i = 1, \ldots, M \end{aligned} \tag{18}$$

where the arguments of $\Delta Z_{Y_i}$, $\theta_l$ and $H_{il}$ are given by

$$\zeta_j(n_j\Delta k_j) = \begin{cases} (n_j - 1)\Delta k_j & \text{for } 1 \le n_j < \dfrac{L_j}{2} + 2 \\ (n_j - 1 - L_j)\Delta k_j & \text{for } \dfrac{L_j}{2} + 2 \le n_j \le L_j \end{cases} \tag{19}$$

Note that this formulation presumes that the cross-spectral densities and the associated spectral factors (the $H_{ij}$s) are defined over the discrete wave number range from $-L_j/2$ to $+L_j/2$. The $\zeta_j(n_j)$ variables provide a conversion between the arguments of these spectral quantities and the wave number range 1 to $L_j$ used in the summation of eqn (18). Note that the conversion superimposes the end points at $-L_j/2$ and $+L_j/2$. This conversion is part of the 'bit reversal' algorithm. In cases where the $L_j$ are not integer powers of 2 we can either use a different fast Fourier transform algorithm or zero padding, as described in Press *et al.*[25] Note that the fields generated by the discrete fast Fourier transform

algorithm are periodic with a spatial period equal to $L_j$ in direction $j$.

The fast Fourier transform approach requires that the maximum (or Nyquist) wave number in direction $j$ must be $k_j^{\max} = L_j \Delta k_j = 2\pi/\Delta x_j$, which implies that

$$\Delta k_j = \frac{2\pi}{L_j \Delta x_j} \tag{20}$$

It is important to choose $k_j^{\max}$ large enough to ensure that all of the $H_{ij}$s are adequately sampled. Given this requirement, there are a number of trade-offs to be considered when selecting the space and wave number grid spacing. If we wish to obtain fine spatial resolution we can decrease $\Delta x_j$. If $L_j$ is increased by a comparable amount the size $L_j \Delta x_j$ of the spatial domain remains the same and eqn (20) is satisfied but the computational demand of the algorithm increases. If $L_j$ is held fixed $\Delta k_j$ must be increased. In this case we may not be able properly to resolve variations in the shape of the input cross-spectral density functions. In practice, $\Delta x_j$ and $\Delta k_j$ need to be selected to provide a reasonable balance between spatial resolution, spectral resolution, and computational effort.

## 5 EXAMPLES OF SYNTHETICALLY GENERATED CROSS-CORRELATED RANDOM FIELDS

As mentioned in Section 1, numerical studies of flow and transport in heterogeneous porous media frequently rely on synthetically generated log hydraulic conductivity and velocity fields. Physically realistic velocity components are cross-correlated since they all ultimately depend on the log conductivity. One obvious way to obtain a realistic velocity field is to generate a synthetic log conductivity field and then solve the resulting groundwater flow equation (for hydraulic head) and Darcy equation (for velocity). The primary disadvantage of this direct approach is the computational effort required to obtain an accurate solution when the log conductivity field is moderately heterogeneous and anisotropic. Iterative flow solvers such as the popular preconditioned conjugate gradient method tend to converge very slowly for problems with large numbers of nodes if the log conductivity variance is of order 1.0 (Refs 2, 24).

A less computationally demanding option is to derive the cross-spectral densities (or transfer functions) between the log conductivity and the $p$ velocity components, using a linearized version of the flow and Darcy equations.[14] In this case, the algorithm of eqn (16) may be used to generate a set of synthetic random fields which are correlated in the proper way. The primary disadvantage of this approach is its dependence on the linearization approximation, which tends to break down as the log conductivity variance increases much beyond 1.0 (Ref. 1). Since the direct and spectrally-based approaches both degrade in performance for large log conductivity variances, the choice between them should be based on a careful evaluation of accuracy and computational requirements. Here we illustrate how the spectral alternative can be applied to problems of moderate log hydraulic conductivity variance.

The following approximate relationship between the Fourier increments of log conductivity and steady-state velocity may be derived from a linear infinite domain spatially invariant approximation of Darcy's law and the steady-state groundwater flow equation[13,14]

$$\mathrm{d}Z_{v_i} = \frac{K_G}{n}\left(J_i - \frac{J_j k_i k_j}{k^2}\right)\mathrm{d}Z_f \qquad i,j = 1,\ldots,p \tag{21}$$

where $f$ and $v_i$ represent log conductivity and component $i$ of the Darcy velocity, respectively, $K_G$ is the geometric mean of the log conductivity, $J_i$ is the mean head gradient (assumed constant), $k_i$ is the wave number in direction $i$, $k^2 = \mathbf{k}\cdot\mathbf{k}$, and $p$ is the number of spatial dimensions. If the mean gradient is aligned with the longitudinal ($i = 1$) direction, $J_1 = J$, and the transfer functions introduced in eqn (13) are

$$\begin{aligned} T_1(\mathbf{k}) &= T_f(\mathbf{k}) = 1 \\ T_2(\mathbf{k}) &= T_{v_1}(\mathbf{k}) = \frac{K_G J}{nk^2}\sum_{j=2}^{p} k_j^2 \\ T_i(\mathbf{k}) &= T_{v_{i-1}}(\mathbf{k}) = -\frac{K_G J}{nk^2} k_1 k_{i-1} \qquad i = 3,\ldots,p+1 \end{aligned} \tag{22}$$

These transfer functions may be substituted, together with $H_{11}(\mathbf{k})$, the square root of the log conductivity spectral density function, directly into eqn (16).

In the following paragraphs we examine a set of two-dimensional cross-correlated log conductivity and velocity fields generated from eqn (16). The synthetic generation process begins with the specification of a set of cross-covariances or cross-spectral densities. Closed-form expressions for two-dimensional cross covariances between log conductivity and velocity are available when the log hydraulic conductivity is described by either an exponential or a Whittle-A covariance.[3,28] Here we use the Whittle-A alternative, which is characterized by the following log conductivity spectral density function[15]

$$S_{ff}(\mathbf{k}) = H_{11}(\mathbf{k})H_{11}^*(\mathbf{k}) = \frac{2\sigma_f^2 a^2 \mathbf{k}^2}{\pi(a^2 + \mathbf{k}^2)^3} \tag{23}$$

where $a = \pi/4\lambda$, $\lambda$ is the spatial correlation scale, and $\sigma_f^2$ is the variance of the log hydraulic conductivity field. Figure 1 shows surface plots for the covariances of the two velocity components $v_1$ and $v_2$ and for the cross-covariance between $v_1$ and $v_2$. All three covariances are anisotropic. The $v_1$ covariance function has the longest correlation length in the $x_1$ (longitudinal) direction and

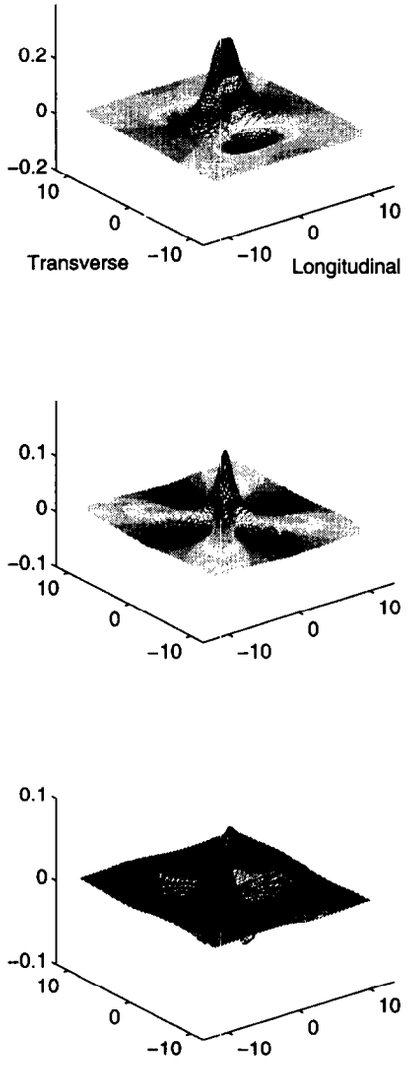


**Fig. 1.** Spatial covariance functions for the longitudinal velocity component $v_1$ (top), the transverse velocity component $v_2$ (middle), and the spatial cross-covariance function between $v_1$ and $v_2$ (bottom).

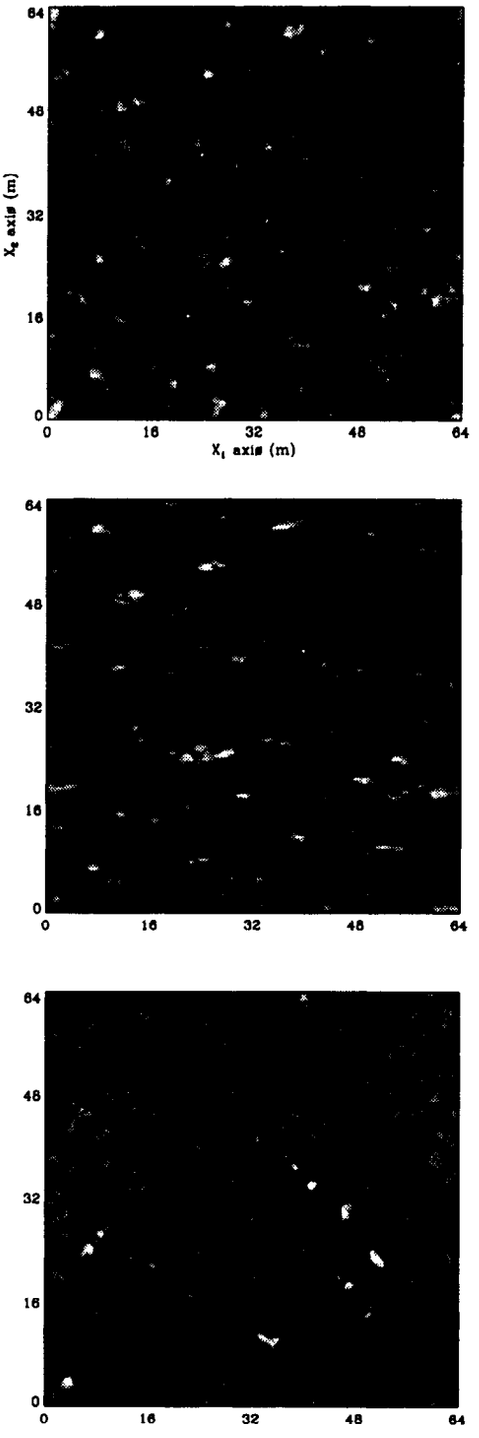


**Fig. 2.** Spatially cross correlated isotropic hydraulic conductivity field (top), $x_1$-component of velocity field (middle), and $x_2$-component of velocity field (bottom), all generated with the Whittle spectral density function.

the shortest correlation length in the $x_2$ (transverse) direction. The $v_2$ covariance function has the longest correlation length along the diagonals $x_1 = \pm x_2$ with shorter correlation lengths in the longitudinal and transverse directions. The cross-covariance between $v_1$ and $v_2$ exhibits maximum positive correlation along the diagonal $x_1 = x_2$ and maximum negative correlation along the diagonal $x_1 = -x_2$. These covariances are displayed here for informational purposes. Only the log conductivity spectral density and the log conductivity–velocity transfer functions are actually used in the random field generator.

The contour plots in Fig. 2 show the three components (log conductivity $f$, longitudinal velocity $v_1$, and transverse velocity $v_2$) of a typical multivariate replicate obtained from the Whittle-A log conductivity spectrum. The three fields are generated on a uniform spatial grid spacing of $\Delta x_1 = \Delta x_2 = 0{\cdot}5$ m with a ratio of correlation length to grid spacing of $\lambda/\Delta x_1 = 2$ and a grid size of $128 \times 128$. The log conductivity realization appears to be isotropic, as expected. Longitudinal persistence is significantly greater than transverse persistence in the $v_1$ realization. In the $v_2$ realization persistence is apparent in the diagonal $x_1 = \pm x_2$ directions. These features are

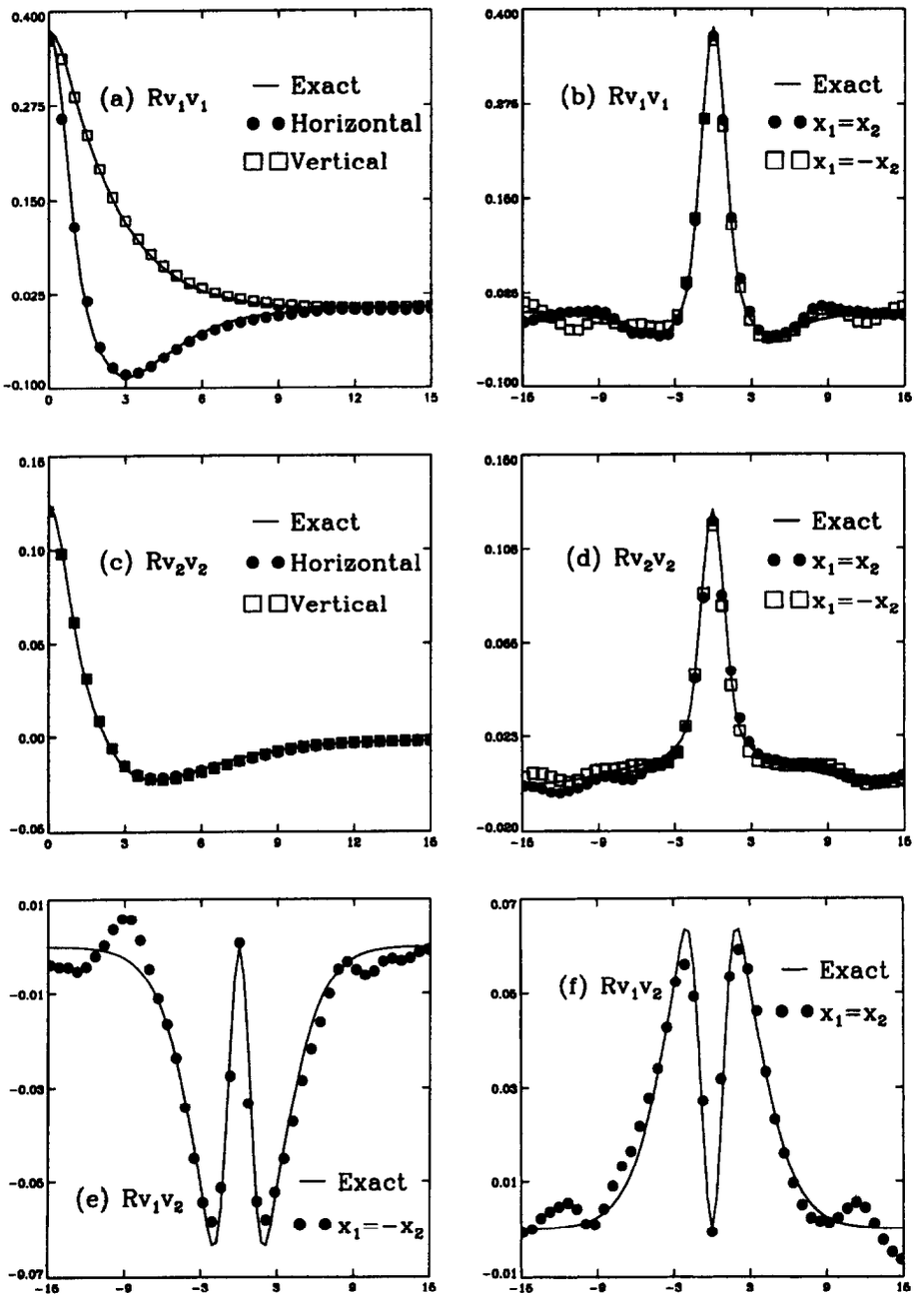


**Fig. 3.** Comparison of ensemble averages 100 realizations of the normalized empirical covariance functions $n^2R/K_G^2J^2$ to the analytical covariance functions, where $\lambda_1/\Delta x_1 = \lambda_2/\Delta x_2 = 2$, $L_1/\lambda_1 = L_2/\lambda_2 = 64$, and a size of $128 \times 128$. The abscissa is the dimensionless separation distance and the ordinate is (a) $R_{v_1}$ in the $x_1$ and $x_2$ directions, (b) $R_{v_1}$ in the direction along $x_1 = x_2$, (c) $R_{v_2}$ in the $x_1$ direction, (d) $R_{v_2}$ in the direction $x_1 = x_2$, (e) and (f) $R_{v_1v_2}$ in the directions $x_1 = \pm x_2$.

consistent with the velocity covariance functions plotted in Fig. 1.

In order to check the statistical validity of the synthetic fields generated by our algorithm we need to compare ensemble statistics such as the sample cross-covariances to their theoretical counterparts. Figure 3 displays the sample covariances obtained with a Monte Carlo analysis based on an ensemble of 100 log conductivity and velocity realizations. All realizations were generated with the parameter set used for Fig. 2. Figure 3(a) plots one-dimensional $v_1$ sample covariances along the $x_1$ and $x_2$ directions with circles and squares, respectively. These should be compared with the corresponding theoretical covariances, shown with solid lines. Figure 3(b) plots the one-dimensional $v_1$ sample covariance along $x_1 = \pm x_2$ with circles and squares, and the corresponding theoretical covariance with a solid line. Figure 3(c) and 3(d) compare the ensemble and analytical covariance functions for $v_2$ along the $x_1$ and $x_2$ directions and along the diagonal $x_1 = \pm x_2$, respectively. Finally, Figs 3(e) and 3(f) compare the ensemble and analytical cross-covariance functions

between $v_1$ and $v_2$ along the diagonals $x_1 = \pm x_2$. All of these comparisons indicate that the sample statistics converge to their theoretical counterparts for an ensemble of 100 replicates. Statistical sampling theory suggests that the difference between sample and theoretical covariance values should slowly decrease as the number of replicates increases, particularly at large lags. This is consistent with the behavior we obtained when we gradually increased the number of replicates in the problem of Fig. 3 up to 100. Note that the increase of $\lambda/\Delta x_1$ ratio should increase the grid resolution in derived random fields, however this also increases the computational costs required to obtain results similar to those shown in this figure.

A synthetically generated groundwater velocity field should conserve mass if it is physically meaningful, even if it is a random replicate. In particular, the divergence of the velocity vector, which is itself a random field, should be zero everywhere. This implies that the mean and variance of the divergence must both be equal to zero. Graham and McLaughlin[15] show that this requirement is met for a velocity field with the covariances plotted in Fig. 2. Although velocity fields generated from discrete approximations will never have divergences identically equal to zero we should expect computed point divergences to be small compared with $\bar{v}_1/\Delta x_1$, the ratio of the mean velocity to the local grid spacing. Ruan[28] investigated this point in a Monte Carlo analysis of the velocity ensemble. This analysis shows that the mean and variance of the velocity divergence converge to zero at the discrete grid spacings $\Delta x_1$ and $\Delta x_2$ approach zero.

Spectrally-based random field generators can readily be modified to produce stationary fields with different statistical structures. As an illustration we show in Fig. 4 a set of cross correlated hydraulic conductivity and velocity fields obtained from the following two-dimensional Gaussian log conductivity spectrum[15]

$$S_{ff}(\mathbf{k}) = H_{11}(\mathbf{k})H_{11}^{*}(\mathbf{k}) = \frac{1}{2\pi}\sigma_f^2\lambda_1\lambda_2 \exp[-(k_1^2\lambda_1^2 + k_2^2\lambda_2^2)/2] \quad (24)$$

The left half of Fig. 4 shows a set of log conductivity and velocity fields for the isotropic case, with $\lambda_1 = \lambda_2$. The right half shows a set of log conductivity and velocity fields for an anisotropic case with the $\lambda_1 = 4\lambda_2$. In both cases the grid spacing is $\Delta x_1 = \Delta x_2 = 0{\cdot}5$ m, the ratio of $x_2$ correlation length to $x_2$ grid spacing is $\lambda_2/\Delta x_2 = 2$, and the grid size is $128 \times 128$. A comparison between Fig. 2 and the left half of Fig. 4 shows that the isotropic Gaussian log conductivity field is smoother and has more regular features than an isotropic Whittle-A field with a comparable correlation length. The elongated features in the $x_1$ direction of the $f$ and $v_1$ fields plotted in the right half of Fig. 4 clearly reveal the effects of anisotropy in the log conductivity spectrum. Anisotropy also changes the angle of the dominant diagonal axis in the $v_2$ field.

Cross-correlated three-dimensional fields can be readily generated with our algorithm, although the computational effort required (which grows as $\log_2(N)$) can be significant if the spatial grid becomes very large. Figure 5 illustrates a typical three-dimensional application. In this case we have generated four cross-correlated fields: the log hydraulic conductivity and three components of velocity. The log conductivity field has a Gaussian spectral density function[36] with an anisotropy ratio of $\lambda_1 = \lambda_2 = 4\lambda_3$, a grid spacing of $\Delta x_1 = \Delta x_2 = \Delta x_3 = 0{\cdot}5$ m, a normalized correlation length of $\lambda_1/\Delta x_1 = 8$, and a grid size of $128 \times 128 \times 64$. Figure 5 indicates that the three-dimensional velocity fields share some qualitative features with the two-dimensional velocity fields plotted in Fig. 4. In particular, there is longitudinal persistence in $v_1$ as well as diagonal persistence along the $x_1 = \pm x_2$ plane in $v_2$ and along the $x_1 = \pm x_3$ plane in $v_3$. Similar features have also been reported by Cushy *et al.*[7]

The cross-spectral densities used in the above examples are derived from linear, spatially-invariant approximations to the groundwater flow equation and Darcy's law. While these approximations enable us to obtain stationary spectral representations for the individual components of velocity they also introduce some restrictions. In particular, the synthetically-generated stationary velocities will be valid only when log conductivity variability is relatively small (i.e. $\sigma_f < 1.0$) and the mean hydraulic gradient is nearly constant over several log conductivity correlation scales. Bellin *et al.*[3] have compared synthetic velocities obtained from a random field generator based on linearization and stationarity approximations with random velocities derived from an exact numerical solution to the flow equation. They note that the two methods produce the same spatial probability distribution only when $\sigma_f^2 \leq 0{\cdot}2$ for log hydraulic conductivity fields. When $\sigma_f$ is high, the approximate solution yields normally distributed $v_1$ while the exact solution yields a log normal probability density (this issue is also discussed by Zhang and Neuman).[38] In such cases, the negative values of the approximate normal $v_1$ are systematically smaller than that of the exact lognormal $v_1$. It is possible that a better velocity approximation could be obtained (at some computational expense) if the spectral generator were based on higher-order poly-spectral density functions.[26] These could be derived from higher-order perturbation approximations to the flow equation and Darcy's law.

Although the spectrally-based synthetic velocities described in this example are admittedly approximations, they provide mass conservative flow fields which are very convenient for investigations of solute transport through heterogeneous media. These fields are much easier to generate at high resolution than exact solutions to the stochastic flow equation, especially in

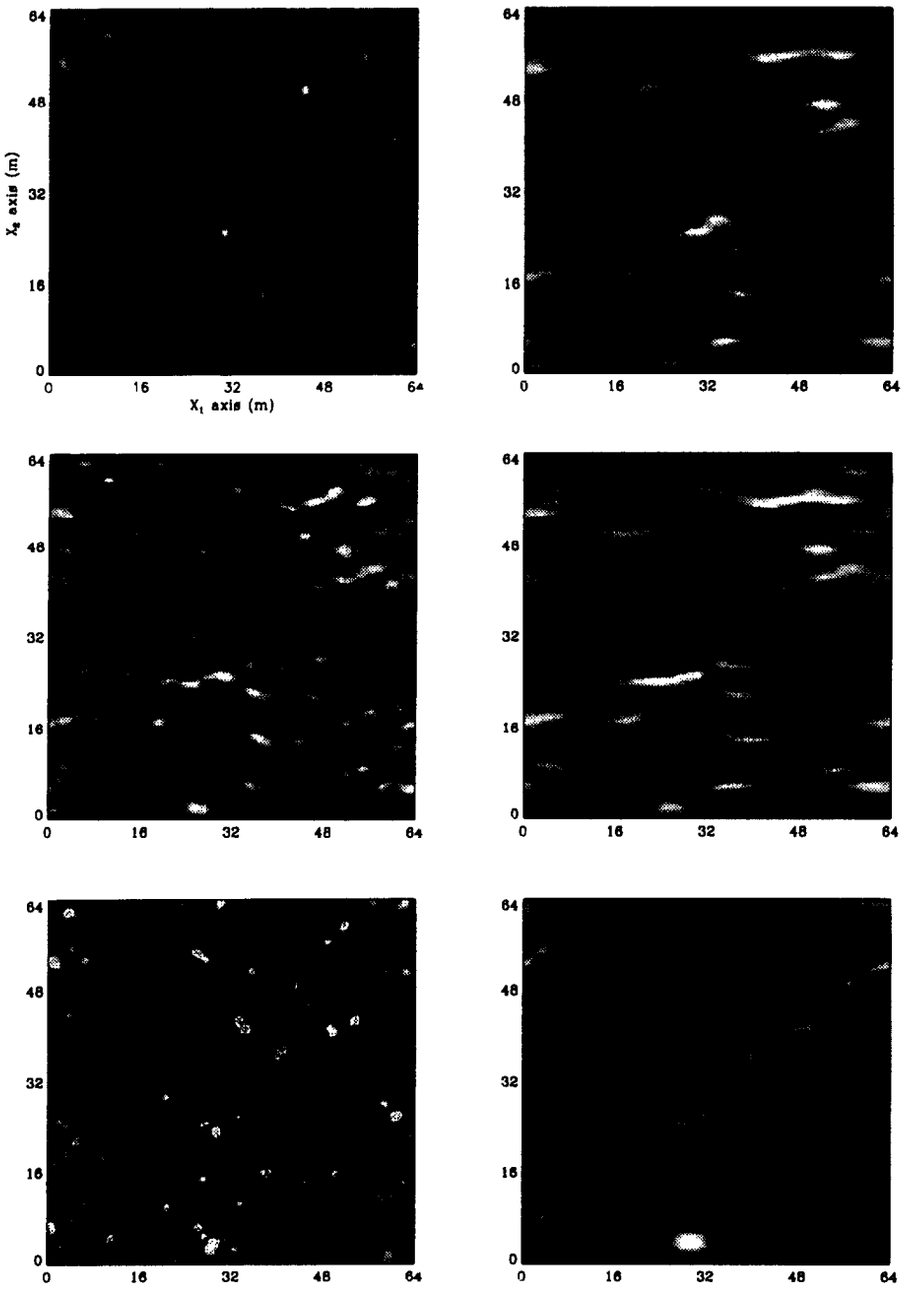


**Fig. 4.** (a) (Left column) Spatially cross-correlated isotropic hydraulic conductivity (top), $x_1$-component of velocity (middle), and $x_2$-component of velocity (bottom). (b) (Right column) Spatially cross correlated anisotropic hydraulic conductivity (top), $x_1$-component of velocity (middle), and $x_2$-component of velocity (bottom), all generated with a Gaussian spectral density function.

three-dimensions. The choice between the approximate spectral approach and the exact direct solution approach for obtaining synthetic velocities is ultimately application-dependent.

## 6 CONCLUSIONS

In this paper, we propose a new spectral approach for generating cross-correlated random fields. This approach is sufficiently general to be able to produce large numbers of complex stationary random fields which are related by any consistent set of rational cross-spectral densities. Our random field generator is based on a discrete complex-valued approximation to the classic Fourier–Stieltjes representation of a random field. The Fourier increment used in this representation depends on a random phase angle process and a matrix of spectral factors which can be readily derived from a specified set of cross-spectral densities (or cross-covariances). The random field values at any given location are obtained from an inverse fast Fourier transform of the discrete Fourier increment. When the specified cross-covariances (the $R_{ij}$s) are real, the real and imaginary parts of the

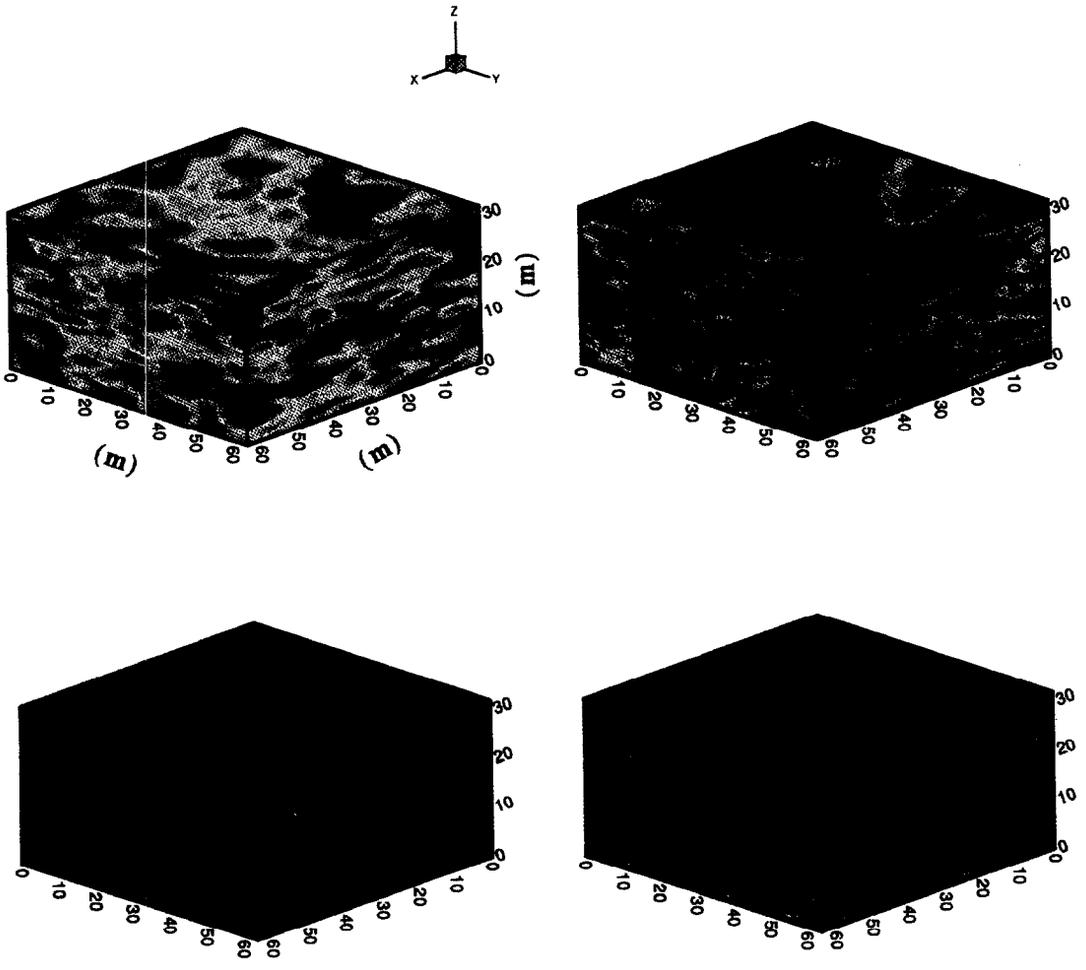


**Fig. 5.** Three dimensional spatially cross-correlated anisotropic hydraulic conductivity (top left), $x_1$-component of velocity (top right), $x_2$ component of velocity (bottom left), $x_3$ component of velocity, all generated with a Gaussian spectral density function.

complex field generated by our algorithm are independent (real) replicates with cross-covariances equal to $\sqrt{2}R_{ij}$ (for $i, j = 1, \ldots, M$).

Our algorithm simplifies considerably if $M - 1$ of the $M$ fields to be generated can be related to the remaining field by a space-invariant linear transformation. In this case, the spectral factors can be constructed directly from the linear transformation and the spectral density of the independent field. Also, only one (rather than $M$) random phase angle process is required. A useful application of the simplified algorithm is the generation of a vector of two or three groundwater velocity components which depend on a scalar log hydraulic conductivity field. The spectral factors for this problem are derived from linearized approximations to the groundwater flow equation and Darcy's law. The example we present in Section 5 confirms that our algorithm generates mass conservative velocity fields which reproduce specified ensemble covariances.

The primary advantages of a complex-valued spectrally-based approach to multivariate random field generation are its simplicity and convenience. The basic concepts can be explained in a few equations (e.g. eqn (1) through (5) of this paper) and the algorithm can be readily programmed. The procedure can be formulated entirely in terms of cross-spectral densities, which generally have much simpler forms than the corresponding cross-covariances. It is also easy to accommodate a wide variety of different spectral densities since the required spectral factorization is either trivial (for the case where several dependent random fields are related to a common independent field) or can be performed numerically. Since the approach is based on a complex-valued representation, no special manipulations are required to convert the complex spectra in order to be consistent with the real spectral representation.

The groundwater velocity example presented here is just one of many hydrologically-relevant applications of our algorithm. Other examples include the generation of correlated spatially heterogeneous unsaturated soil properties (such as the parameters used to describe relative permeability and moisture-retention curves) and the generation of correlated spatially heterogeneous land surface variables such as vegetative cover, soil

moisture and temperature. In each case, a description of the relevant physical processes can be used to derive the cross-spectral densities required by the generator. This capability makes it possible to carry out synthetic experiments that properly account for the physical processes which influence heterogeneity.

## References


1. Ababou, R., McLaughlin, D. B. & Gelhar, L. W., Three-dimensional flow in random porous media, Technical Report 318, R. M. Parsons Lab., Dept. of Civil Engineering, MIT, Cambridge, 1988.
2. Beckie, R., Wood, E. F. & Aldama, A. A., Mixed finite element simulation of saturated groundwater flow using a multigrid accelerated domain decomposition technique. *Water Resour. Res.*, **29** (1993) 3145–3157.
3. Bellin, A., Salandin, P. & Rinaldo, A., Simulation of dispersion in heterogeneous porous formations: statistical first-order theories, convergence of computations. *Water Resour. Res.*, **28** (1992) 2211–2227.
4. Bendat, J. S. & Piersol, A. G., *Random Data*, 2nd Edition. John Wiley & Sons, New York, 1986.
5. Borgman, L., Taheri, M. & Hagan, R., Three-dimensional frequency domain simulations of geological variables. In *Geostatistics for Natural Resources Characterization, Part I*, eds G. Verly *et al.* Kluwer Academic, M.A., 1984, pp. 517–541.
6. Brigham, E. O., *The Fast Fourier Transform and Its Applications*, Signal Processing Series, Prentice-Hall, Englewood Cliffs, N.J., 1988.
7. Cushy, M., Bellin, A. & Rubin, Y., Generation of three dimensional flow fields for statistically anisotropic heterogeneous porous media. *Stochastic Hydrology and Hydraulics*, **9** (1995) 89–104.
8. Dagan, G., Stochastic modeling of groundwater by unconditional and conditional probabilities, 1, Conditional simulation and the direct problem. *Water Resour. Res.*, **18** (1982) 813–833.
9. Davis, M W., Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Math. Geol.*, **19** (1987) 91–107.
10. Delhomme, J. P., Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. *Water Resour. Res.*, **15** (1979) 269–280.
11. Dietrich, C. R. & Newsam, G. N., A fast and exact method for multi-dimensional Gaussian stochastic simulations. *Water Resour. Res.*, **29** (1993) 2861–2869.
12. Dietrich, C. R., A simple and efficient space domain implementation of the turning bands method. *Water Resour. Res.*, **31** (1995) 147–156.
13. Gelhar, L. W. & Axeness, G. L., Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resour. Res.*, **19** (1983) 161–180.
14. Gelhar, L. W., Stochastic subsurface hydrology from theory to applications. *Water Resour. Res.*, **25** (1986) 135S–145S.
15. Graham, W. & McLaughlin, D., Stochastic analysis of nonstationary subsurface solute transport, 1, Unconditional moments. *Water Resour. Res.*, **25** (1989) 215–232.
16. Gutjahr, A. L., Fast Fourier transform for random field generation. Project report for Los Alamos grant, contract 4-R58-2690R, N. M. Inst. of Min. and Technol., Socorro, 1989.
17. Gutjahr, A. L., Bullard, B., Hatch, S. & Hughson, L., Joint conditional simulations and the spectral approach for flow modeling. *Stochastic Hydrology and Hydraulics*, **8** (1994) 70–108.
18. Harris, T. J. & Davis, J., An iterative method for matrix spectral factorization. *SIAM J. Numer. Anal.*, **13** (1992) 531–540.
19. Hernandez, G. J., A stochastic approach to the simulation of block conductivity fields conditioned upon data measured at a smaller scale. PhD dissertation, Stanford University, 1991.
20. Journel, A. G. & Huijbregts, C. J., *Mining Geostatistics*, Academic Press, London, 1978.
21. Lumley, J. L. & Panofsky, H. A., *The Structure of Atmospheric Turbulence*, John Wiley, New York, 1965, 239pp.
22. Mantoglou, A. & Wilson, J. L., The turning bands method for the simulation of random fields using line generation by a spectral method. *Water Resour. Res.*, **18** (1982) 1379–1394.
23. Mantoglou, A., Digital simulation of multivariate two- and three-dimensional stochastic processes with a spectral turning bands method. *Math. Geol.*, **19** (1987) 129–149.
24. Meyer, P. D., Valocchi, A. J., Ashby, S. F. & Saylor, P. E., A numerical investigation of the Conjugate Gradient method as applied to three-dimensional groundwater flow problems in randomly heterogeneous porous media. *Water Resour. Res.*, **25** (1989) 1440–1446.
25. Press, W., Flannery, B., Teukolsky, S. & Vetterling, W., *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, New York, 1986.
26. Priestley, M. B., *Spectral Analysis and Time Series*. Academic, San Diego, CA, 1981.
27. Robin, M. J., Gutjahr, A. L., Sudicky, E. A. & Wilson, J. L., Cross-correlated random field generation with the direct Fourier transform method. *Water Resour. Res.*, **29** (1993) 2385–2397.
28. Ruan, F., An efficient multivariate multidimensional random field generator using the Fast Fourier Transform. Master Thesis, MIT, May, 1994.
29. Rubin, Y., Stochastic modeling of macrodispersion in heterogenous media. *Water Resour. Res.*, **26** (1990) 133–141.
30. Smith, L. & Freeze, R. A., Stochastic analysis of steady state groundwater flow in a bounded domain, 1, One-dimensional simulations. *Water Resour. Res.*, **15** (1979) 521–528.
31. Smith, L. & Freeze, R. A., Stochastic analysis of steady state groundwater flow in a bounded domain, 2, Two-dimensional simulations. *Water Resour. Res.*, **15** (1979) 1543–1559.
32. Smith, L. & Schwartz, F. W., Mass transport, 2, Analysis of uncertainty predictions. *Water Resour. Res.*, **17** (1981) 351–369.
33. Tompson, A. F. B., Ababou, R. & Gelhar, L. W., Applications and use of three-dimensional turning bands generator: single realization problems. Technical Report 313, R. M. Parsons Lab., Dept. of Civil Engineering, MIT, Cambridge, 1987.
34. Tompson, A. F. B., Ababou, R. & Gelhar, L. W., Implementation of the three-dimensional turning bands random field generator. *Water Resour. Res.*, **25** (1989) 2227–2243.
35. Tompson, A. F. B. & Gelhar, L. W., Numerical simulation of solute transport in three-dimensional, randomly heterogenous porous media. *Water Resour. Res.*, **26** (1990) 2541–2562.
36. Vomvoris, E. G. & Gelhar, L. W., Stochastic analysis of the concentration variability in a three-dimensional

heterogeneous aquifer. *Water Resour. Res.*, **26** (1990) 2591–2602.
37. Zhang, D. X. & Neuman, S. P., Comment on 'A note on head and velocity covariances in three-dimensional flow through heterogeneous anisotropic porous media' by Y. Rubin and G. Dagan. *Water Resour. Res.*, **28** (1992) 3343–3344.
38. Zhang, D. X. & Neuman, S. P., Eulerian–Lagrangian analysis of transport conditioned on hydraulic data, 1, Analytical-numerical approach. *Water Resour. Res.*, **31** (1995) 39–51.

## APPENDIX

### Cross-covariances of the real and imaginary parts of synthetically generated complex random fields

In eqn (5) we propose the following discrete representation for the complex random field $Y_i(\mathbf{x})$

$$Y_i(\mathbf{x}) = \int_{-\infty}^{\infty} e^{\imath \mathbf{k}\cdot\mathbf{x}}\, dZ_{Y_i}(\mathbf{k}) \approx \sum_n e^{\imath \mathbf{k}_n\cdot\mathbf{x}} H_{il}(\mathbf{k}_n)\, e^{\imath\theta_l(\mathbf{k}_n)} |\Delta\mathbf{k}_n|^{1/2} \quad i = 1,\ldots,M \tag{A1}$$

where the summation is understood to be over all of the discrete wave numbers included in the $p$-dimensional computational grid. In this Appendix we investigate the statistical properties of the real and imaginary parts of the fields $Y_i(\mathbf{x})$ and $Y_j(\mathbf{x}')$ when their cross-covariance $R_{ij}(\mathbf{x}-\mathbf{x}')$ is real.

The cross-spectral spectral density matrix $S_{ij}(\mathbf{k})$ used to derive the $H_{ij}(\mathbf{k})$ weights in (A1) is generally complex when $i \neq j$. If we indicate the real and imaginary parts of $H_{ij}(\mathbf{k})$ with R and I subscripts the cross-spectral density can be written as

$$\begin{aligned} S_{ij}(\mathbf{k}) &= H_{il}H^*_{jl} \\ &= [H_{\mathrm{R}il} + \imath H_{\mathrm{I}il}][H_{\mathrm{R}jl} - \imath H_{\mathrm{I}jl}] \\ &= c_{ij}(\mathbf{k}) - \imath g_{ij}(\mathbf{k}) \end{aligned} \tag{A2}$$

where $c_{ij}(\mathbf{k})$ (the cospectrum) and $g_{ij}(\mathbf{k})$ (the quadrature spectrum) are given by

$$c_{ij}(\mathbf{k}) = H_{\mathrm{R}il}(\mathbf{k})H_{\mathrm{R}jl}(\mathbf{k}) + H_{\mathrm{I}il}(\mathbf{k})H_{\mathrm{I}jl}(\mathbf{k})$$

$$g_{ij}(\mathbf{k}) = H_{\mathrm{R}il}(\mathbf{k})H_{\mathrm{I}jl}(\mathbf{k}) - H_{\mathrm{I}il}(\mathbf{k})H_{\mathrm{R}jl}(\mathbf{k})$$

Equation (9) indicates that the cross-covariance $R_{ij}(\mathbf{x}-\mathbf{x}')$ is the inverse Fourier transform of $S_{ij}(\mathbf{k})$. It follows that the real part of this cross-covariance is

$$\begin{aligned} \mathrm{Re}\,[R_{ij}(\mathbf{x}-\mathbf{x}')] &= \mathrm{Re}\left[\int_{-\infty}^{\infty} e^{\imath\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} S_{ij}(\mathbf{k})\, d\mathbf{k}\right] \\ &= \int_{-\infty}^{\infty} \{\cos[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}) \\ &\quad + \sin[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]\, g_{ij}(\mathbf{k})\}\, d\mathbf{k} \end{aligned} \tag{A3}$$

while the imaginary part is

$$\begin{aligned} \mathrm{Im}\,[R_{ij}(\mathbf{x}-\mathbf{x}')] &= \mathrm{Im}\left[\int_{-\infty}^{\infty} e^{\imath\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} S_{ij}(\mathbf{k})\, d\mathbf{k}\right] \\ &= \int_{-\infty}^{\infty} \{\sin[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}) \\ &\quad - \cos[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]\, g_{ij}(\mathbf{k})\}\, d\mathbf{k} \end{aligned} \tag{A4}$$

When $R_{ij}(\mathbf{x}-\mathbf{x}')$ is real we obtain

$$\int_{-\infty}^{\infty} \{\cos[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}) + \sin[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]\, g_{ij}(\mathbf{k})\}\, d\mathbf{k} = R_{ij}(\mathbf{x}-\mathbf{x}') \tag{A5}$$

and

$$\int_{-\infty}^{\infty} \{\sin[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}) - \cos[\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')]\, g_{ij}(\mathbf{k})\}\, d\mathbf{k} = 0 \tag{A6}$$

With these relationships established we can investigate the properties of the real and imaginary parts of $Y_i(\mathbf{x})$.

The complex random field $Y_i(\mathbf{x})$ can be divided into real and imaginary parts as follows

$$Y_i(\mathbf{x}) = Y_{\mathrm{R}i}(\mathbf{x}) + \imath Y_{\mathrm{I}i}(\mathbf{x}) \tag{A7}$$

If we expand (A1) we can write the discrete approximations to $Y_{\mathrm{R}i}(\mathbf{x})$ and $Y_{\mathrm{I}i}(\mathbf{x})$ as

$$Y_{\mathrm{R}i}(\mathbf{x}) = \sum_m \{\cos[\mathbf{k}_m\cdot\mathbf{x} + \theta_l(\mathbf{k}_m)]H_{\mathrm{R}il}(\mathbf{k}_m) - \sin[\mathbf{k}_m\cdot\mathbf{x} + \theta_l(\mathbf{k}_m)]H_{\mathrm{I}il}(\mathbf{k}_m)\}|\Delta\mathbf{k}_m|^{1/2} \tag{A8}$$

$$Y_{\mathrm{I}i}(\mathbf{x}) = \sum_m \{\sin[\mathbf{k}_m\cdot\mathbf{x} + \theta_l(\mathbf{k}_m)]H_{\mathrm{R}il}(\mathbf{k}_m) + \cos[\mathbf{k}_m\cdot\mathbf{x} + \theta_l(\mathbf{k}_m)]H_{\mathrm{I}il}(\mathbf{k}_m)\}|\Delta\mathbf{k}_m|^{1/2} \tag{A9}$$

The spatial cross-covariance between $Y_{\mathrm{R}i}(\mathbf{x})$ and $Y_{\mathrm{R}j}(\mathbf{x}')$ may be obtained directly from (A8)

$$E[Y_{\mathrm{R}i}(\mathbf{x})Y_{\mathrm{R}j}(\mathbf{x}')] \tag{A10}$$

$$\begin{aligned} &= E\Bigg[\sum_n\sum_q \{\cos[\mathbf{k}_n\cdot\mathbf{x} + \theta_l(\mathbf{k}_n)]H_{\mathrm{R}il}(\mathbf{k}_n) \\ &\quad - \sin[\mathbf{k}_n\cdot\mathbf{x} + \theta_l(\mathbf{k}_n)]H_{\mathrm{I}il}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n|^{1/2} \cdot \\ &\quad \times \{\cos[\mathbf{k}'_q\cdot\mathbf{x}' + \theta_m(\mathbf{k}'_q)]H_{\mathrm{R}jm}(\mathbf{k}'_q) \\ &\quad - \sin[\mathbf{k}'_q\cdot\mathbf{x}' + \theta_m(\mathbf{k}'_q)]H_{\mathrm{I}jm}(\mathbf{k}'_q)\}|\Delta\mathbf{k}'_q|^{1/2}\Bigg] \end{aligned} \tag{A11}$$

Since the phase angles are uncorrelated for $\mathbf{k} \neq \mathbf{k}'$ and $l \neq m$ the double summation over $n$ and $q$ reduces to a single summation over $n$. After expanding the product in (A11), applying standard trigonometric identities, and noting that the phase angle is uniformly distributed over $[0, 2\pi]$ the cross-covariance can be written as follows if

$R_{ij}$ is real

$$
\begin{aligned}
&E[Y_{\mathrm{R}i}(\mathbf{x})Y_{\mathrm{R}j}(\mathbf{x}')] \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n) \\
&\quad + H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n)] \\
&\quad + \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n) \\
&\quad - H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n)]\}|\Delta\mathbf{k}_n| \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}_n) \\
&\quad + \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')]g_{ij}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n| \\
&\approx \tfrac{1}{2}R_{ij}(\mathbf{x}-\mathbf{x}')
\end{aligned}
\tag{A12}
$$

The final approximate equality follows from (A5). As $\Delta\mathbf{k}$ approaches zero the cross-covariance between the real fields $Y_{\mathrm{R}i}(\mathbf{x})$ and $Y_{\mathrm{R}j}(\mathbf{x}')$ approaches one-half of $R_{ij}(\mathbf{x}-\mathbf{x}')$, the specified real covariance between the complex fields $Y_i(\mathbf{x})$ and $Y_j(\mathbf{x}')$.

A similar derivation may be used to obtain the cross-covariance between $Y_{\mathrm{I}i}(\mathbf{x})$ and $Y_{\mathrm{I}j}(\mathbf{x}')$ when $R_{ij}$ is real

$$
\begin{aligned}
&E[Y_{\mathrm{I}i}(\mathbf{x})Y_{\mathrm{I}j}(\mathbf{x}')] \\
&= E\Bigg[\sum_n\sum_q \{\sin[\mathbf{k}_n\cdot\mathbf{x}+\theta_l(\mathbf{k}_n)]H_{\mathrm{R}il}(\mathbf{k}_n) \\
&\quad + \cos[\mathbf{k}_n\cdot\mathbf{x}+\theta_l(\mathbf{k}_n)]H_{\mathrm{I}il}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n|^{1/2}\cdot \\
&\quad \times \{\sin[\mathbf{k}'_q\cdot\mathbf{x}'+\theta_m(\mathbf{k}'_q)]H_{\mathrm{R}jm}(\mathbf{k}'_q) \\
&\quad + \cos[\mathbf{k}'_q\cdot\mathbf{x}'+\theta_m(\mathbf{k}'_q)]H_{\mathrm{I}jm}(\mathbf{k}'_q)\}|\Delta\mathbf{k}'_q|^{1/2}\Bigg]
\end{aligned}
\tag{A14}
$$

After applying the statistical properties of the phase angles and appropriate trigonometric identities, and noting that the phase angle is uniformly distributed over $[0, 2\pi]$ this cross-covariance can be written as

$$
\begin{aligned}
&E[Y_{\mathrm{I}i}(\mathbf{x})Y_{\mathrm{I}j}(\mathbf{x}')] \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n) \\
&\quad + H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n)] \\
&\quad + \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n) \\
&\quad - H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n)]\}|\Delta\mathbf{k}_n| \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}_n) \\
&\quad + \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')]g_{ij}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n| \\
&\approx \tfrac{1}{2}R_{ij}(\mathbf{x}-\mathbf{x}')
\end{aligned}
\tag{A15}
$$

This demonstrates that the cross-covariance between the real fields $Y_{\mathrm{I}i}(\mathbf{x})$ and $Y_{\mathrm{I}j}(\mathbf{x}')$ also approaches one-half of $R_{ij}(\mathbf{x}-\mathbf{x}')$.

Next, we note that the real part of $Y_i(\mathbf{x})$ is uncorrelated with the imaginary part of $Y_j(\mathbf{x}')$ (and vice versa) when $R_{ij}$ is real. To see this, first consider the cross-covariance between $Y_{\mathrm{R}i}(\mathbf{x})$ and $Y_{\mathrm{I}j}(\mathbf{x}')$

$$
\begin{aligned}
&E[Y_{\mathrm{R}i}(\mathbf{x})Y_{\mathrm{I}j}(\mathbf{x}')] \\
&= E\Bigg[\sum_n\sum_q \{\cos[\mathbf{k}_n\cdot\mathbf{x}+\theta_l(\mathbf{k}_n)]H_{\mathrm{R}il}(\mathbf{k}_n) \\
&\quad - \sin[\mathbf{k}_n\cdot\mathbf{x}+\theta_l(\mathbf{k}_n)]H_{\mathrm{I}il}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n|^{1/2}\cdot \\
&\quad \times \{\sin[\mathbf{k}'_q\cdot\mathbf{x}'+\theta_m(\mathbf{k}'_q)]H_{\mathrm{R}jm}(\mathbf{k}'_q) \\
&\quad + \cos[\mathbf{k}'_q\cdot\mathbf{x}'+\theta_m(\mathbf{k}'_q)]H_{\mathrm{I}jm}(\mathbf{k}'_q)\}|\Delta\mathbf{k}'_q|^{1/2}\Bigg]
\end{aligned}
\tag{A17}
$$

This becomes

$$
\begin{aligned}
&E[Y_{\mathrm{R}i}(\mathbf{x})Y_{\mathrm{I}j}(\mathbf{x}')] \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n) \\
&\quad + H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n)] \\
&\quad + \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n) \\
&\quad + H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n)]\}|\Delta\mathbf{k}_n| \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')]g_{ij}(\mathbf{k}_n) \\
&\quad - \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')]c_{ij}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n| \\
&\approx 0
\end{aligned}
\tag{A18}
$$

The final approximate equality follows from (A6).

Next, consider the cross-covariance between $Y_{\mathrm{I}i}(\mathbf{x})$ and $Y_{\mathrm{R}j}(\mathbf{x}')$ for real $R_{ij}$

$$
\begin{aligned}
&E[Y_{\mathrm{I}i}(\mathbf{x})Y_{\mathrm{R}j}(\mathbf{x}')] \\
&= E\Bigg[\sum_n\sum_q \{\sin[\mathbf{k}_n\cdot\mathbf{x}+\theta_l(\mathbf{k}_n)]H_{\mathrm{R}il}(\mathbf{k}_n) \\
&\quad + \cos[\mathbf{k}_n\cdot\mathbf{x}+\theta_l(\mathbf{k}_n)]H_{\mathrm{I}il}(\mathbf{k}_n)\}|\Delta\mathbf{k}_n|^{1/2}\cdot \\
&\quad \times \{\cos[\mathbf{k}'_q\cdot\mathbf{x}'+\theta_m(\mathbf{k}'_q)]H_{\mathrm{R}jm}(\mathbf{k}'_q) \\
&\quad - \sin[\mathbf{k}'_q\cdot\mathbf{x}'+\theta_m(\mathbf{k}'_q)]H_{\mathrm{I}jm}(\mathbf{k}'_q)\}|\Delta\mathbf{k}'_q|^{1/2}\Bigg]
\end{aligned}
\tag{A20}
$$

This becomes

$$
\begin{aligned}
&E[Y_{\mathrm{I}i}(\mathbf{x})Y_{\mathrm{R}j}(\mathbf{x}')] \\
&= \frac{1}{2}\sum_n \{\cos[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n) \\
&\quad - H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n)] \\
&\quad + \sin[\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{x}')][H_{\mathrm{R}il}(\mathbf{k}_n)H_{\mathrm{R}jl}(\mathbf{k}_n) \\
&\quad + H_{\mathrm{I}il}(\mathbf{k}_n)H_{\mathrm{I}jl}(\mathbf{k}_n)]\}|\Delta\mathbf{k}_n|
\end{aligned}
$$

$$= \frac{1}{2}\sum_n \{-\cos[\mathbf{k}_n \cdot (\mathbf{x} - \mathbf{x}')]\, g_{ij}(\mathbf{k}_n) + \sin[\mathbf{k}_n \cdot (\mathbf{x} - \mathbf{x}')] c_{ij}(\mathbf{k}_n)\} |\Delta \mathbf{k}_n|$$

$$\approx 0 \qquad \text{(A21)}$$

This is the last of the four possible cross-covariances between the real and imaginary parts of $Y_i(\mathbf{x})$ and $Y_j(\mathbf{x}')$.