

Representations of Inverse Covariances by Differential Operators

Qin XU*

*National Oceanic and Atmospheric Administration/National Severe Storms Laboratory,
Norman, Oklahoma 73069*

(Received 12 April 2004; revised 15 November 2004)

ABSTRACT

In the cost function of three- or four-dimensional variational data assimilation, each term is weighted by the inverse of its associated error covariance matrix and the background error covariance matrix is usually much larger than the other covariance matrices. Although the background error covariances are traditionally normalized and parameterized by simple smooth homogeneous correlation functions, the covariance matrices constructed from these correlation functions are often too large to be inverted or even manipulated. It is thus desirable to find direct representations of the inverses of background error correlations. This problem is studied in this paper. In particular, it is shown that the background term can be written into $\int d\mathbf{x} |\mathbf{D}v(\mathbf{x})|^2$, that is, a squared L_2 norm of a vector differential operator \mathbf{D} , called the D-operator, applied to the field of analysis increment $v(\mathbf{x})$. For autoregressive correlation functions, the D-operators are of finite orders. For Gaussian correlation functions, the D-operators are of infinite order. For practical applications, the Gaussian D-operators must be truncated to finite orders. The truncation errors are found to be small even when the Gaussian D-operators are truncated to low orders. With a truncated D-operator, the background term can be easily constructed with neither inversion nor direct calculation of the covariance matrix. D-operators are also derived for non-Gaussian correlations and transformed into non-isotropic forms.

Key words: differential operator, inverse background covariance, data assimilation

1. Introduction

Various variational formulations have been developed and used in meteorological data analyses and assimilation since the pioneering studies of Sasaki (1958, 1970). As a powerful mathematical tool, the variational framework has great flexibility to allow various types of constraints to be incorporated into the cost function and weighted differently. For practical applications, various differential operators have sometimes been used as smoothness constraints to suppress noise and improve the smoothness of the analyses. Theoretically, however, it is not obvious how different types of constraints should be selected and weighted unless the variational formulations are derived formally based on Bayesian probabilistic principles such as those in three- and four-dimensional variational data assimilation (Jazwinski, 1970; Wahba and Wendelberger, 1980, Lorenc 1986; Cohn 1997). The paper by Wahba and Wendelberger (1980) was, perhaps, the first in meteorology that not only extended Sasaki's variational formalism rigorously in several aspects but also put it in

the context of Bayesian estimation of conditional expectation. Based on Bayesian probabilistic principles, each term in the costfunction should be weighted by the inverse of its associated error covariance matrix. How to estimate these error covariances and represent their inverses then becomes the central issue. This paper concerns representations of the inverse covariances, especially the inverse background error covariances in three-dimensional variational data assimilation.

Qualitatively, it has then been well recognized that the smoothing effect of the background term can be mimicked by a smoothing penalty term in which a Laplacian or a polynomial of a Laplacian is applied to the analysis increment field as a weak constraint (Wahba and Wendelberger, 1980; Purser, 1986; Xu et al., 2001a). In the seminal paper by Wahba and Wendelberger (1980), the differential operator in the spline formulation was related to an (implicit) prior background covariance, while the averaged background error variance (inversely related to the background

*E-mail: Qin.Xu@noaa.gov

weight) and the required smoothness (related to the order of the differential operator) were estimated by the Generalized Cross Validation method from the data being analyzed. Motivated by the above and other subsequent studies, it should be interesting to explore quantitatively whether and how a differential operator can be constructed concisely and precisely in consistency with a given background error correlation among all admissible forms. This problem is addressed in this paper by using a functional approach.

Due to lack of sufficient knowledge about the true covariance structures, simplifications often have to be made in order to obtain stable error statistics as well as to facilitate computations. In this respect, background error covariances are often assumed to be describable by a small number of parameters (such as the variance and the characteristic spatial scale or truncated spectral coefficients describing the gross features of the shape of the covariance function). Then, the parameters can be estimated by fitting the parameterized covariance function to observation innovations (Gandin, 1963; Hollingsworth and Lönnberg, 1986; Lönnberg and Hollingsworth, 1986; Thiebaut et al., 1986; Bartello and Mitchell, 1992; Dévényi and Schlatter, 1994; Xu et al., 2001c; Xu and Wei, 2001, 2002) or to forecast differences (Parrish and Derber, 1992; Derber and Bouttier, 1999). Although the covariance functions are parameterized with relatively simple analytical forms, the covariance matrices constructed from these functions are often too large and too complex to be inverted into the weight matrix in the background term. Thus, it is necessary to find an equivalent or approximate representation of the background covariance matrix or its inverse in the variational analysis to reduce the computational demand.

Traditionally, as reviewed in the next section, the above task was accomplished by introducing an intermediate state vector through a linear transformation of the analysis increment vector so that the cost function can be reformulated and minimized with respect to the intermediate state vector without inverting the background error covariance matrix. This paper considers a direct approach in which the inverse covariance matrix is represented by the inverse of its associated correlation function (as an operator) in the limit of infinitely high resolution for continuous fields of analysis. In this case, as an example in the one-dimensional space of x , the inverse relationship between a homogeneous univariate correlation function $C(x)$ and its inverse $Q(x)$ is defined by

$$\int dx' C(x-x')Q(x') = \delta(x), \quad (1.1)$$

where $\delta(\cdot)$ is the delta function (Courant and Hilbert, 1962) and the integral $\int dx'$ is over the entire space.

The associated background term can then be expressed by

$$J_b = \int dx \int dx' v(x)Q(x-x')v(x'), \quad (1.2)$$

where $v(x)$ is the increment analysis field normalized by the background error standard deviation. By applying the generalized Fourier transformation (Lighthill, 1958) to Eq. (1.1), the inverse relationship is expressed concisely in the space of wavenumber k by

$$S(k)G(k) = (2\pi)^{-1}, \quad (1.3)$$

where $S(k)$ and $G(k)$ are the Fourier transformations of $C(x)$ and $Q(x)$, respectively. As power spectra, $S(k)$ and $G(k)$ are real and even functions of k , so $G(k)$ can be derived from $S(k)$ [or directly from $C(x)$ as shown in the appendix] as a Taylor expansion in terms of k^2 , say, $G(k) = \sum_0 g_n k^{2n}$, where the summation \sum_0 is over $n (= 0, 1, 2, \dots)$. The inverse Fourier transformation of this expansion gives $Q(x) = (2\pi)^{1/2} \sum_0 (-1)^n g_n (d/dx)^{2n} \delta(x)$. Substituting this into Eq. (1.2) gives

$$J_b = \int dx |Dv(x)|^2, \quad (1.4)$$

where integration by parts is used, and

$$D = (2\pi)^{1/4} [g_0^{1/2}, g_1^{1/2} d/dx, \dots, g_n^{1/2} (d/dx)^n, \dots]^T$$

is a vector differential operator, called the D-operator.

The functional approach outlined above in Eqs. (1.1)–(1.4) indicates that the background term can be written into a squared L_2 norm of a D-operator applied to the field of analysis increment. Similar D-operator formulations can be derived for homogeneous and isotropic univariate correlation functions in two- and three-dimensional space, and the details will be presented in sections 4–5 of this paper. Depending on the form of $S(k)$ or $C(x)$, the D-operator can be of either finite or infinite order. For practical applications, a D-operator of infinite order must be truncated to a finite order. With a truncated D-operator, the background term can be constructed directly just like a smoothing penalty term except that the smoothing is controlled precisely by the background correlation (rather than arbitrarily or empirically). In this case, neither inversion nor direct calculation of the background error covariance matrix is required. The background term and its gradient can be directly and easily computed.

The D-operator is very different from the generalized diffusive operator of Weaver and Courtier (2001) in correlation modeling. Their approach is to model the background correlation (rather than its inverse) by

the inverse of a generalized diffusive operator (rather than by the operator itself). As the inversion is implemented by a recursive algorithm, their approach can be viewed as a recursive filter (see section 2.2). This paper takes an opposite but straightforward approach, that is, to model the inverse of the background correlation directly by a differential operator. This functional approach is similar to that used in Problems 7.7 and 7.9 of Tarantola (1987) for exponentially decaying correlation functions in one-dimensional and two-dimensional spaces. These exponential functions, however, are not differentiable at zero correlation distance, so they are not suitable for smooth background error covariances. For example, the exponential correlation function has the form of $C(x) = \exp(-|x|/L)$ in one-dimensional space, which is simply the first-order autoregressive correlation. By using the approach outlined above in Eqs. (1.1)–(1.4), D-operators can be derived for any high-order autoregressive correlation functions. The analytical forms of these high-order functions can be found in Jenkins and Watts (1968), and here we only need to show their D-operators. The n -th order autoregressive correlation is defined by

$$C(x - x') = \langle \xi(x)\xi(x') \rangle / \sigma^2,$$

where $\langle (\cdot) \rangle$ denotes the probabilistic mean, σ^2 is the variance of the random field $\xi(x)$ generated by $\sum_{(0,n)} a_{n'} (d/dx)^{n'} \xi(x) = w(x)$ from the white-noise field $w(x)$ for which $\langle w(x)w(x') \rangle = \delta(x - x')$, $a_{n'}$ is the coefficient for the n' -th order term [as in Eq.(A.1) of Thiebaut et al., 1986], and the summation $\sum_{(0,n)}$ is over $n' = (0, 1, 2, \dots, n)$. This definition leads to

$$\begin{aligned} \delta(x) = & \left[\sum_{(0,n)} a_{n'} (d/dx)^{n'} \right] \\ & \times \left[\sum_{(0,n)} a_{n'} (-d/dx)^{n'} \right] C(x) \sigma^2 \end{aligned}$$

or, equivalently,

$$\begin{aligned} (2\pi)^{-1/2} = & \left[\sum_{(0,n)} a_{n'} (ik)^{n'} \right] \\ & \times \left[\sum_{(0,n)} a_{n'} (-ik)^{n'} \right] S(k) \sigma^2, \end{aligned}$$

so $S(k)$ is known explicitly. Substituting this $S(k)$ into Eq. (1.3) gives

$$G(k) = \sum_{(0,n)} g_{n'} k^{2n'},$$

where $g_{n'} = [a_{n'}^2 + 2 \sum'' (-1)^{n''} a_{n''} a_{2n'-n''}] \sigma^2 (2\pi)^{-1/2}$, and the summation \sum'' is over $n'' (= 1, 2, \dots, n' - 1)$. Thus, for the n -th order autoregressive correlation, the D-operator in Eq. (1.4) is only of a finite order up to n . When $n = 1$, the D-operator recovers the result in Problem 7.7 of Tarantola (1987).

This paper considers the widely used Gaussian correlations first and then deals with other more general and complex situations. The Gaussian correlations

are infinitely differentiable and their inverses are D-operators of unlimited high order, so proper truncations must be considered and justified by the smallness of the truncation errors. This problem is addressed in this paper. Furthermore, when a background correlation is estimated from innovation data, its spectrum (Fourier transformation) is often in a discrete form rather than a closed analytical form (Xu et al., 2001c; Xu and Wei, 2001, 2002). Thus, it will be interesting and practically useful to derive a properly truncated D-operator for a correlation whose Fourier transformation is not analytically available. This problem is also addressed in this paper with the detailed derivation of general D-operator formulations presented in the appendix.

D-operators will be derived first and essentially for homogeneous and isotropic univariate correlations. Via proper variable and coordinate transformations, the derived D-operators can be extended and used for multivariate, non-homogeneous and non-isotropic covariances (although the cross-correlation between the transformed variables may not be truly negligible and the transformed coordinate may allow limited degrees of non-homogeneity and non-isotropy). To demonstrate this, we need to briefly review the traditional multivariate approaches in the next section, and then introduce the representer method for continuous fields of multivariate analyses in section 3 to show that the analyses are at least as smooth as the multivariate correlations and thus possess the differentiability required by the associated D-operator. D-operator formulations are derived subsequently for one- and two-dimensional Gaussian correlation functions in section 4, and for three-dimensional ellipsoidal Gaussian correlation functions in section 5. General D-operator formulations for non-Gaussian and non-isotropic correlation functions are presented in the appendix. The results are discussed with conclusions in section 6.

2. Review of traditional approaches

During each cycle of three-dimensional variational data assimilation, observations are linearly combined into the background field provided by the prediction of a numerical model. Errors in the background fields are often assumed to be Gaussian random with zero mean, while errors in the updated observations (with biases removed) are also Gaussian random with zero mean and independent of the background errors (see Chapter 4 of Daley, 1991). Under these assumptions and based on Bayesian probabilistic principles (see Lorenc, 1986; Cohn, 1997), the maximum likelihood estimate (optimal analysis) of the true state can be obtained by

minimizing the following increment form of the cost function:

$$J = J_b + J_o = \Delta \mathbf{a}^T \mathbf{B}^{-1} \Delta \mathbf{a} + (\mathbf{H} \Delta \mathbf{a} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{H} \Delta \mathbf{a} - \mathbf{d}). \quad (2.1)$$

Here, $\Delta \mathbf{a} = \mathbf{a} - \mathbf{b}$ is the analysis increment which is taken to be the control vector in the discrete model space, $\mathbf{d} = \mathbf{o} - \mathbf{H}(\mathbf{b})$ is the observation innovation vector in the observation space, \mathbf{b} is the background state vector, \mathbf{a} is the analysis vector, \mathbf{o} is the observation vector, \mathbf{B} and \mathbf{R} are the background and observation error covariance matrices, respectively, \mathbf{H} is a tangent linear approximation of the (nonlinear) observation operator H that maps the model space to the observation space, and $(\cdot)^T$ denotes the transpose of (\cdot) . The optimal analysis (maximum likelihood estimate) is given by $\mathbf{a} = \mathbf{b} + \Delta \mathbf{a}$ at the minimum of J . The value of $\Delta \mathbf{a}$ at the minimum of J is called the minimizer of J but, for simplicity, it will still be denoted by $\Delta \mathbf{a}$ as long as the meaning is clearly understood from the derivations developed in this paper.

To facilitate the presentation, the model space is decomposed into subspaces and each subspace corresponds to one variable only. Denote by $\xi_k(\mathbf{x})$ the random field of background error for the k -th variable (in the k -th subspace), where $\mathbf{x} = (x, y, z)$ denotes a point in the three-dimensional coordinate system used by the model. Then,

$$\langle \xi_k(\mathbf{x}) \xi_{k'}(\mathbf{x}') \rangle \equiv B_{kk'}(\mathbf{x}, \mathbf{x}') \quad (2.2a)$$

defines the background error covariance function associated with the k -th and k' -th subspaces, where $\langle (\cdot) \rangle$ denotes the expectation (probabilistic mean) of (\cdot) . When the field is discretized on a grid, then $\xi_k(\mathbf{x})$ reduces to $\boldsymbol{\xi}_k$ —a random vector representing the grid field of background error in the k -th subspace. Then, Eq. (2.2a) reduces to

$$\langle \boldsymbol{\xi}_k \boldsymbol{\xi}_{k'}^T \rangle \equiv \mathbf{B}_{kk'}. \quad (2.2b)$$

Here, $\mathbf{B}_{kk'}$ is the (k, k') -th submatrix of \mathbf{B} associated with the k -th and k' -th subspaces. The components of $\boldsymbol{\xi}_k$ are given by $\xi_k(\mathbf{x}_i)$, where \mathbf{x}_i denotes the i -th grid point. The elements of $\mathbf{B}_{kk'}$ are given by

$$B_{kk'}(\mathbf{x}_i, \mathbf{x}_{i'}) \equiv \langle \xi_k(\mathbf{x}_i) \xi_{k'}(\mathbf{x}_{i'}) \rangle.$$

In principle, the minimizer of J can be obtained by solving the linear algebraic equations derived from

$$\partial J / \partial \Delta \mathbf{a} = 0.$$

In practice, however, the background error covariance matrix is often too large to be inverted. To avoid large matrix inversion, different approaches have been developed (Lorenc, 1986; Cohn, 1997; Courtier et al., 1998; Daley and Barker, 2001). One common strategy is to introduce an intermediate state vector, \mathbf{c} , through a

linear transformation $\Delta \mathbf{a} = \mathbf{T} \mathbf{c}$, so the cost function can be reformulated in terms of \mathbf{c} without inverting \mathbf{B} . Thus, previous approaches can be classified based on their selected types of \mathbf{T} , as reviewed briefly in the following subsections.

2.1 Observation space approach

By choosing $\mathbf{T} = \mathbf{B} \mathbf{H}^T$, the minimizer of J can be obtained in the observation space by solving the linear algebraic equation derived from $\partial J / \partial \mathbf{c} = 0$, that is,

$$(\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R}) \mathbf{c} = \mathbf{d}. \quad (2.3)$$

The minimizer of J in the original model space is then given by

$$\Delta \mathbf{a} = \mathbf{B} \mathbf{H}^T \mathbf{c}. \quad (2.4)$$

This is the statistical interpolation scheme often called optimal interpolation (OI) (see Daley, 1991). Often the observation space is still too large to allow Eq. (2.3) be solved directly. The traditional OI approach is to partition the observation space into partially overlapped subspaces so that $\mathbf{H} \mathbf{B} \mathbf{H}^T$ and $\mathbf{B} \mathbf{H}^T$ become small enough in each subspace. The final solution is not truly global as it is pieced by local solutions (minimizers in subspaces). In this case, the global solution requires an iterative approach with a descending algorithm to find the minimum of the cost function. This leads to the three-dimensional variational (3D Var) scheme with different approaches, including the Physical-space Statistical Analysis Scheme (PSAS). The PSAS is an observation space approach in which the cost function is reformulated from Eq. (2.3) with $\mathbf{T} = \mathbf{B} \mathbf{H}^T$ (Cohn et al., 1998; Daley and Barker, 2001). Other 3D Var approaches are reviewed in the following subsections.

2.2 Model space approach

By choosing $\mathbf{T} = \mathbf{B}$, the cost function in Eq. (2.1) can be reformulated as a function of $\mathbf{c} = \mathbf{B}^{-1} \Delta \mathbf{a}$. As \mathbf{B}^{-1} is absorbed into \mathbf{c} , the inversion of \mathbf{B} is avoided and the reformulated cost function is preconditioned (Lorenc, 1988; Derber and Rosati, 1989). Although \mathbf{B}^{-1} disappears in the reformulated cost function (not shown), the size of the problem is still often very large in the model space and an iterative method must be resorted to for the minimization. The most costly part in each iteration is the multiplication of \mathbf{c} by \mathbf{B} , and even a single multiplication can become prohibitively expensive if carried out with the full matrix \mathbf{B} . One remedy, as adopted by Gaspari and Cohn (1999), is to constrain the covariance functions to zero beyond certain distances (compactly supported) and thus to make \mathbf{B} sparse. The computational cost can also be reduced by choosing $\mathbf{T} = \mathbf{B}^{1/2}$ while $\mathbf{B}^{1/2}$ is modeled by a recursive filter (Purser and McQuigg, 1982;

Lorenc, 1992; Hayden and Purser, 1995; Huang, 2000; Purser et al., 2003a, b) or by inverting a generalized diffusive operator (Weaver and Courtier, 2001). However, when the covariance function in Eq. (2.2) is non-isotropic (Xu and Gong, 2003), $\mathbf{B}^{1/2}$ cannot be easily computed or modeled. In this case, it is convenient and efficient to choose $\mathbf{T} = \mathbf{B}$ and model \mathbf{B} by a recursive filter (Purser et al., 2003a, b).

2.3 Spectral-space approach

Assume that the analysis increment is horizontally periodic (on either a flat or spherical domain). Denote by \mathbf{F} the matrix operator of the discrete Fourier transformation in the horizontal (or Legendre-Fourier transformation on the spherical surface). Denote by $\Delta\mathbf{a}_k$ the component of $\Delta\mathbf{a}$ in the k -th subspace. By choosing $\mathbf{T} = \mathbf{F}^{-1}$ in each subspace, we have $\Delta\mathbf{a}_k = \mathbf{F}^{-1}\mathbf{c}_k$, where $\mathbf{c}_k = \mathbf{F}\Delta\mathbf{a}_k$ is the Fourier transformation of $\Delta\mathbf{a}_k$. Substituting this into the cost function in Eq. (2.1) gives

$$J = \sum \sum' \mathbf{c}_k^T (\mathbf{S}^{-1})_{kk'} \mathbf{c}_{k'} + \left(\sum \mathbf{H}_k \mathbf{F}^{-1} \mathbf{c}_k - \mathbf{d} \right)^T \mathbf{R}^{-1} \left(\sum \mathbf{H}_k \mathbf{F}^{-1} \mathbf{c}_k - \mathbf{d} \right), \quad (2.5)$$

where \sum and \sum' are summations indexed by k and k' , respectively, \mathbf{H}_k is the k -th block of columns of \mathbf{H} that maps $\Delta\mathbf{a}_k$ from the k -th subspace to the observation space, $(\mathbf{S}^{-1})_{kk'}$ denotes the (k, k') -th submatrix of \mathbf{S}^{-1} , and the (k, k') -th submatrix of \mathbf{S} is given by

$$\mathbf{S}_{kk'} = \mathbf{F} \mathbf{B}_{kk'} \mathbf{F}^T = \langle \mathbf{s}_k \mathbf{s}_{k'}^T \rangle. \quad (2.6)$$

Here, $\mathbf{B}_{kk'}$ is as in Eq. (2.2b), $\mathbf{s}_k = \mathbf{F}\boldsymbol{\xi}_k$ and $\mathbf{s}_{k'} = \mathbf{F}\boldsymbol{\xi}_{k'}$ are the random vectors of the background errors in the k -th and k' -th spectral subspaces, respectively.

When the vector field of horizontal velocity is represented by the scalar fields of the streamfunction and velocity potential, all the analysis variables are scalars. The associated normalized scalar background errors can be assumed to be jointly homogeneous and isotropic in the horizontal, so spectral components of different horizontal wavenumbers are not correlated [see section 2.9 of Panchev (1971) for a one-dimensional example, and Boer (1983) for the spherical spectral representation]. Thus, each $\mathbf{S}_{kk'} = \langle \mathbf{s}_k \mathbf{s}_{k'}^T \rangle$ is block-diagonal with each diagonal block associated with a horizontal wavenumber. As each diagonal block is a small matrix in the space of vertical levels or modes, it can be easily converted into a diagonal form by its eigenvectors and so does $\mathbf{S}_{kk'}$. This simplifies the inversion of \mathbf{S} in Eq. (2.5). For large-scale analyses (Parrish and Derber, 1992; Courtier et al., 1998), the background errors are often assumed to be not correlated between different scalar

variables (except for the geostrophically coupled error fields of the streamfunction and geopotential height), so $\mathbf{S}_{kk'} = \langle \mathbf{s}_k \mathbf{s}_{k'}^T \rangle = 0$ for $k \neq k'$ except that k and k' are associated with a pair of streamfunction and geopotential height fields. This further simplifies the inversion of \mathbf{S} .

2.4 Reformulated spectral-space approach

Recently, Derber and Bouttier (1999) proposed a comprehensive yet very compact reformulation of \mathbf{S} in the spectral-space approach which allows for the separation of the univariate and multivariate components of \mathbf{S} into distinct operators. In this formulation, the intermediate state vector $\mathbf{c} = \mathbf{F}\Delta\mathbf{a}$ in the spectral space is partitioned into balanced and unbalanced parts except for one variable (that is, the vorticity or streamfunction) which is balanced-only (or taken in totality). This balanced-only variable and all the unbalanced variables are assumed to be univariate in the sense that their background errors are not cross-correlated, so their error covariance matrix, denoted by \mathbf{U} [or \mathbf{B}_u as in Eq. (7) of Derber and Bouttier (1999)], is block-diagonal with each diagonal block associated with one of the univariate variables. All the remaining balanced variables are related to the univariate variables by a linear operator, denoted by \mathbf{K}' . Thus, $\mathbf{c} = \mathbf{K}\mathbf{u}$ where $\mathbf{K} = \mathbf{I} + \mathbf{K}'$ as in Eq. (10) of Derber and Bouttier (1999) and \mathbf{u} is the vector that represents all the univariate variables in the spectral space. With this \mathbf{K} transformation, \mathbf{S} takes the form of $\mathbf{S} = \mathbf{K}\mathbf{U}\mathbf{K}^T$ and Eq. (2.5) can be written into

$$J = \mathbf{u}^T \mathbf{U}^{-1} \mathbf{u} + (\mathbf{H}\mathbf{F}^{-1}\mathbf{K}\mathbf{u} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{F}^{-1}\mathbf{K}\mathbf{u} - \mathbf{d}). \quad (2.7)$$

For the same reason as explained in section 2.3, each diagonal block of \mathbf{U} is a block-diagonal matrix and can be easily converted into a diagonal form. Since \mathbf{U} is univariate-block-diagonal but \mathbf{S} is not, \mathbf{U}^{-1} can be computed more easily than \mathbf{S}^{-1} in Eq. (2.5).

3. Representer method for continuous fields of analyses

The difficulties in computing the background term (and its gradient) can be greatly reduced by using the spectral-space approach as shown in Eq. (2.5) and (2.7). The spectral-space approach is particularly suitable for spectral models over horizontally periodic domains or on spherical surfaces. When the domain is not periodic and is discretized by a grid in the horizontal, other approaches (see sections 2.1 and 2.2) are often used to avoid the inversion of \mathbf{B} . In this

paper, the inverse of \mathbf{B} is directly represented by a Taylor series of differential operators in the functional space for each univariate variable. Applying this functional approach to Eq. (2.1) requires \mathbf{B} to be reformulated into a block-diagonal form with each diagonal block associated with a univariate variable. This is done by a \mathbf{K} transformation as in Eq. (2.7) but in the grid space. In this case, $\mathbf{u} = \mathbf{K}^{-1}\mathbf{c} = \mathbf{K}^{-1}\mathbf{F}\Delta\mathbf{a}$ and $\mathbf{U} = \mathbf{K}^{-1}\mathbf{S}\mathbf{K}^{-\top} = \mathbf{K}^{-1}\mathbf{F}\mathbf{B}\mathbf{F}^{\top}\mathbf{K}^{-\top}$ are replaced by $\mathbf{u} = \mathbf{K}^{-1}\Delta\mathbf{a}$ and $\mathbf{U} = \mathbf{K}^{-1}\mathbf{B}\mathbf{K}^{-\top}$, respectively. Denote by Λ^2 the diagonal part of \mathbf{U} . Then, $\mathbf{C} = \Lambda^{-1}\mathbf{U}\Lambda^{-1}$ is the error correlation matrix for \mathbf{u} or, equivalently, the error covariance matrix for the normalized vector $\mathbf{v} = \Lambda^{-1}\mathbf{u}$. As in Eq. (2.7), \mathbf{U} is block-diagonal and so is \mathbf{C} . Their k -th diagonal blocks are linked by $\mathbf{C}_k = \Lambda_k^{-1}\mathbf{U}_k\Lambda_k^{-1}$, where Λ_k is the k -th diagonal component of Λ in the k -th subspace associated with the k -th univariate variable. Denote by \mathbf{u}_k the k -th vector component of \mathbf{u} in the k -th subspace and thus $\mathbf{v}_k = \Lambda_k^{-1}\mathbf{u}_k$ is the k -th vector component of $\mathbf{v} = \Lambda^{-1}\mathbf{u}$. With the above \mathbf{K} transformation, the cost function in Eq. (2.1) can be rewritten into the following univariate form:

$$J = \sum \mathbf{v}_k^{\top} \mathbf{C}_k^{-1} \mathbf{v}_k + \left(\sum \mathbf{H}_k \sum' \mathbf{K}_{kk'} \Lambda_{k'} \mathbf{v}_{k'} - \mathbf{d} \right)^{\top} \times \mathbf{R}^{-1} \left(\sum \mathbf{H}_k \sum' \mathbf{K}_{kk'} \Lambda_{k'} \mathbf{v}_{k'} - \mathbf{d} \right), \quad (3.1)$$

where the summations \sum and \sum' are over k and k' , respectively, \mathbf{H}_k is the k -th block of columns of \mathbf{H} that maps $\Delta\mathbf{a}_k$ from the k -th subspace of $\Delta\mathbf{a}$ to the observation space, and $\mathbf{K}_{kk'}\Lambda_{k'}$ is the (k, k') -th submatrix of $\mathbf{K}\Lambda$ that maps $\mathbf{v}_{k'}$ (from the k' -th subspace of \mathbf{v}) to $\Delta\mathbf{a}_k$ (in the k -th subspace of $\Delta\mathbf{a}$). It is easy to see that $\Delta\mathbf{a}_k = \sum' \mathbf{K}_{kk'}\mathbf{u}_{k'} = \sum' \mathbf{K}_{kk'}\Lambda_{k'}\mathbf{v}_{k'}$.

After the above preparations, the aforementioned functional approach can be introduced in two steps. First, we show that in the limit of infinitely high resolution (for continuous fields of analysis) the background term can be expressed as a sum of squared norms of the normalized univariate variables in reproducing kernel Hilbert spaces, and the reproducing kernels are $C_k(\mathbf{x}, \mathbf{x}')$ ($k = 1, 2, \dots$) [see Eqs. (3.2b) and (3.7)], so the optimal analysis is a smooth function constructed by $C_k(\mathbf{x}, \mathbf{x}')$ [see Eqs. (3.10) and (3.16)]. This first step is presented in this section. Then, based on the smoothness of the solution and by assuming that $C_k(\mathbf{x}, \mathbf{x}')$ is homogeneous, we show in the next three sections that the inverse of $C_k(\mathbf{x}, \mathbf{x}')$ can be represented by a Taylor series of differential operators applied to the delta function in the functional space for each normalized univariate variable. Consequently, the background term can be written into a squared Sobolev norm for each normalized univariate variable [see Eqs. (4.17) and (5.10)].

In the continuous limit, we have

$$\mathbf{v}_k = \Lambda_k^{-1}\mathbf{u}_k \rightarrow v_k(\mathbf{x}) = u_k(\mathbf{x})/\sigma_k(\mathbf{x}), \quad (3.2a)$$

$$\mathbf{C}_k \rightarrow C_k(\mathbf{x}, \mathbf{x}') \equiv \langle \zeta_k(\mathbf{x})\zeta_k(\mathbf{x}') \rangle, \quad (3.2b)$$

$$\mathbf{K}_{kk'} \rightarrow K_{kk'}(\mathbf{x}, \mathbf{x}'), \quad (3.2c)$$

$$\xi_k = \sum' \mathbf{K}_{kk'} \Lambda_{k'} \zeta_{k'} \rightarrow \sum' \int d\mathbf{x}' K_{kk'}(\mathbf{x}, \mathbf{x}') \times \sigma_{k'}(\mathbf{x}') \zeta_{k'}(\mathbf{x}') = \xi_k(\mathbf{x}), \quad (3.2d)$$

$$\Delta\mathbf{a}_k = \sum' \mathbf{K}_{kk'} \Lambda_{k'} \mathbf{v}_{k'} \rightarrow \sum' \int d\mathbf{x}' K_{kk'}(\mathbf{x}, \mathbf{x}') \times \sigma_{k'}(\mathbf{x}') v_{k'}(\mathbf{x}') = \Delta a_k(\mathbf{x}). \quad (3.2e)$$

Here, $v_k(\mathbf{x})$ and $\Delta a_k(\mathbf{x})$ are the continuous fields of analysis increment for the normalized univariate variable and original variable, respectively, in the k -th subspace; $\zeta_k(\mathbf{x})$ and $\xi_k(\mathbf{x})$ denote random fields of background errors for $v_k(\mathbf{x})$ and $\Delta a_k(\mathbf{x})$, respectively; and ξ_k and ζ_k are random vectors composed of $\xi_k(\mathbf{x}_i)$ and $\zeta_k(\mathbf{x}_i)$, respectively (where \mathbf{x}_i denotes the i -th grid point and $i = 1, 2, \dots$). The integral $\int d\mathbf{x}'$ is over the entire space. Substituting Eq. (3.2d) into Eq. (2.2a) and using Eq. (3.2b) yield

$$B_{kk'}(\mathbf{x}, \mathbf{x}') = \left\langle \sum'' \int d\mathbf{x}'' K_{kk''}(\mathbf{x}, \mathbf{x}'') \sigma_{k''}(\mathbf{x}'') \zeta_{k''}(\mathbf{x}'') \times \int d\mathbf{x}''' K_{k''k'''}(\mathbf{x}', \mathbf{x}''') \sigma_{k'''}(\mathbf{x}''') \zeta_{k'''}(\mathbf{x}''') \right\rangle = \sum'' \int d\mathbf{x}'' \int d\mathbf{x}''' K_{kk''}(\mathbf{x}, \mathbf{x}'') \sigma_{k''}(\mathbf{x}'') \times C_{k''k'''}(\mathbf{x}'', \mathbf{x}''') \sigma_{k'''}(\mathbf{x}''') K_{k''k'''}(\mathbf{x}', \mathbf{x}'''). \quad (3.3)$$

After the above preparations, the k -th subspace for $v_k(\mathbf{x})$ can be completed as a Hilbert space or, specifically, as a Sobolev space [since $C_k(\mathbf{x}, \mathbf{x}')$ is smooth], denoted by \mathbb{H}_k , with the following inner product:

$$\{\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x})\}_k \equiv \int d\mathbf{x} \int d\mathbf{x}' \varphi_1(\mathbf{x}) Q_k(\mathbf{x}, \mathbf{x}') \varphi_2(\mathbf{x}') \quad (3.4)$$

for any functions φ_1 and φ_2 in \mathbb{H}_k , where $Q_k(\mathbf{x}, \mathbf{x}')$ is the inverse of $C_k(\mathbf{x}, \mathbf{x}')$ defined by

$$\int d\mathbf{x}' C_k(\mathbf{x}, \mathbf{x}') Q_k(\mathbf{x}', \mathbf{x}'') = \delta(\mathbf{x} - \mathbf{x}''), \quad (3.5)$$

and $\delta(\cdot)$ denotes the delta function. Using Eqs. (3.4)–(3.5), one can verify that

$$\{C_k(\mathbf{x}, \mathbf{x}'), v_k(\mathbf{x}')\}_k = v_k(\mathbf{x}), \quad (3.6a)$$

$$\{C_k(\mathbf{x}, \mathbf{x}'), C_k(\mathbf{x}', \mathbf{x}'')\}_k = C_k(\mathbf{x}, \mathbf{x}''), \quad (3.6b)$$

so $C_k(\mathbf{x}, \mathbf{x}')$ is the reproducing kernel and \mathbb{H}_k is a reproducing kernel Hilbert space. Using Eq. (3.2b) and (3.4)–(3.5), one can verify that

$$J_b = \sum \mathbf{v}_k^{\top} \mathbf{C}_k^{-1} \mathbf{v}_k \rightarrow \sum \{v_k(\mathbf{x}), v_k(\mathbf{x})\}_k \quad (3.7)$$

in the continuous limit.

The observation term in the cost function is given by the second term in Eq. (3.1), where \mathbf{H}_k is the k -th submatrix of \mathbf{H} that maps $\Delta \mathbf{a}_k$ from the k -th subspace of $\Delta \mathbf{a}$ to the observation space. Note that

$$\sum \mathbf{H}_k \Delta \mathbf{a}_k \rightarrow \sum \int d\mathbf{x} \mathbf{h}_k(\mathbf{x}) \Delta a_k(\mathbf{x}) \quad (3.8)$$

in the continuous limit, where $\mathbf{h}_k(\mathbf{x})$ is the vector kernel function of the observation operator that maps $\Delta a_k(\mathbf{x})$ to the observation space. Substituting Eq. (3.2e) into Eq. (3.8) gives

$$\begin{aligned} & \sum \mathbf{H}_k \sum' \mathbf{K}_{kk'} \mathbf{A}_{k'} \mathbf{v}_{k'} \rightarrow \sum \sum' \int d\mathbf{x} \mathbf{h}_k(\mathbf{x}) \\ & \times \int d\mathbf{x}' \mathbf{K}_{kk'}(\mathbf{x}, \mathbf{x}') \sigma_{k'}(\mathbf{x}') v_{k'}(\mathbf{x}') \\ & = \sum \{ \mathbf{L}_k(\mathbf{x}), v_k(\mathbf{x}) \}_k, \end{aligned} \quad (3.9)$$

where Eq. (3.6a) is used in the last step with the subscripts k and k' exchanged and

$$\begin{aligned} \mathbf{L}_k(\mathbf{x}) & \equiv \sum' \int d\mathbf{x}'' \mathbf{h}_{k'}(\mathbf{x}'') \\ & \times \int d\mathbf{x}' \mathbf{K}_{k'k}(\mathbf{x}'', \mathbf{x}') \sigma_k(\mathbf{x}') C_k(\mathbf{x}', \mathbf{x}) \end{aligned} \quad (3.10)$$

is the observation operator that maps the k -th subspace \mathbf{H}_k to the observation space. Substituting Eqs. (3.9) and (3.7) into Eq. (3.1) gives

$$\begin{aligned} J & = \sum \{ v_k(\mathbf{x}), v_k(\mathbf{x}) \}_k + \left(\sum \{ \mathbf{L}_k(\mathbf{x}), v_k(\mathbf{x}) \}_k - \mathbf{d} \right)^T \\ & \times \mathbf{R}^{-1} \left(\sum \{ \mathbf{L}_k(\mathbf{x}), v_k(\mathbf{x}) \}_k - \mathbf{d} \right) \end{aligned} \quad (3.11)$$

in the continuous limit.

As a function in \mathbf{H}_k , each $v_k(\mathbf{x})$ in Eq. (3.11) can be expressed in general by

$$v_k(\mathbf{x}) = \sum \mathbf{c}^T \mathbf{L}_k(\mathbf{x}) + p_k(\mathbf{x}), \quad (3.12)$$

where \mathbf{c} is an intermediate vector to be determined in the observation space and $p_k(\mathbf{x})$ is an element in \mathbf{H}_k perpendicular to every component function of $\mathbf{L}_k(\mathbf{x})$ [see Eq. (3.10)]. Substituting Eq. (3.12) into Eq. (3.11) gives

$$\begin{aligned} J & = \mathbf{c}^T \mathbf{LCL}^T \mathbf{c} + \sum \{ p_k, p_k \} \\ & + (\mathbf{LCL}^T \mathbf{c} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{LCL}^T \mathbf{c} - \mathbf{d}), \end{aligned} \quad (3.13a)$$

where

$$\mathbf{LCL}^T = \sum \{ \mathbf{L}_k(\mathbf{x}), \mathbf{L}_k^T(\mathbf{x}) \}_k$$

is a symmetric matrix in the observation space. Using Eqs. (3.3)–(3.5) and (3.10), one can verify that

$$\mathbf{LCL}^T = \sum \sum' \int d\mathbf{x} \int d\mathbf{x}' \mathbf{h}_k(\mathbf{x}) B_{kk'}(\mathbf{x}, \mathbf{x}') \mathbf{h}_{k'}^T(\mathbf{x}'). \quad (3.13b)$$

For a point observation of $\Delta a_{k'}(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_{j'}$, $\mathbf{h}_k(\mathbf{x})$ reduces to $\delta_{kk'} \delta(\mathbf{x}_{j'} - \mathbf{x})$ which gives

$$\int d\mathbf{x} \sum \mathbf{h}_k(\mathbf{x}) \Delta a_k(\mathbf{x}) = \Delta a_{k'}(\mathbf{x}_{j'})$$

in Eq. (3.8). Corresponding to a pair of point observations, namely, $\Delta a_k(\mathbf{x}_j)$ and $\Delta a_{k'}(\mathbf{x}_{j'})$, the element of \mathbf{LCL}^T in Eq. (3.13b) is just $B_{kk'}(\mathbf{x}_j, \mathbf{x}_{j'})$. Thus, for point observations, \mathbf{LCL}^T is identical to \mathbf{HBH}^T in Eq. (2.3).

Clearly, $\{ p_k, p_k \}$ is non-negative and thus must be zero if $v_k(\mathbf{x})$ minimizes J in Eq. (3.11). This implies that the unobservable field $p_k(\mathbf{x})$ must be discarded, so the cost function in Eq. (3.13a) reduces to the following finite-dimensional form:

$$\begin{aligned} J & = \mathbf{c}^T \mathbf{LCL}^T \mathbf{c} \\ & + (\mathbf{LCL}^T \mathbf{c} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{LCL}^T \mathbf{c} - \mathbf{d}). \end{aligned} \quad (3.14)$$

The minimizer of J is solved from the following linear algebraic equation derived from $\partial J / \partial \mathbf{c} = 0$:

$$(\mathbf{LCL}^T + \mathbf{R}) \mathbf{c} = \mathbf{d}. \quad (3.15)$$

As explained with Eq. (3.13b), \mathbf{LCL}^T is identical to \mathbf{HBH}^T for point observations, so the system of Eq. (3.15) is identical to Eq. (2.3). The solution for the increment field, however, is a continuous function given by Eq. (3.12) with $p(\mathbf{x}) = 0$, that is,

$$v_k(\mathbf{x}) = \sum \mathbf{c}^T \mathbf{L}_k(\mathbf{x}). \quad (3.16)$$

This approach belongs to the representer method [see Chapters 1–2 of Wahba (1990); Chapter 5 of Bennett (1992)].

4. D-operators for Gaussian correlations in one- and two-dimensional spaces

Assume that $C_k(\mathbf{x}, \mathbf{x}')$ is homogeneous in the horizontal, that is,

$$C_k(\mathbf{x}, \mathbf{x}') = C_k(\mathbf{x}_h - \mathbf{x}'_h, z, z')$$

where $\mathbf{x}_h = (x, y)$ denotes the horizontal component of $\mathbf{x} = (x, y, z)$. In a properly transforming vertical coordinate, $C_k(\mathbf{x}, \mathbf{x}')$ can be further assumed approximately to have a homogeneous form in the vertical (Franke, 1999; Franke and Barker, 2000). In this case, we have $C_k(\mathbf{x}, \mathbf{x}') = C_k(\mathbf{x}_h - \mathbf{x}'_h, z - z') = C_k(\mathbf{x} - \mathbf{x}')$, where and hereafter z is the transformed vertical coordinate over $(-\infty, \infty)$. For each normalized univariate variable, the background term in Eq. (3.7) reduces to

$$J_b = \int d\mathbf{x} \int d\mathbf{x}' v(\mathbf{x}) Q(\mathbf{x} - \mathbf{x}') v(\mathbf{x}'), \quad (4.1)$$

$$\int d\mathbf{x}' C(\mathbf{x} - \mathbf{x}') Q(\mathbf{x}') = \delta(\mathbf{x}), \quad (4.2)$$

where the subscript k is dropped because we only need to consider a single normalized univariate variable in this section and the results apply to any of the normalized univariate variables. To facilitate the derivations in this section, we assume that $C(\mathbf{x})$ is separable between the horizontal and vertical, that is, $C(\mathbf{x}) = C_h(\mathbf{x}_h)C_z(z)$. In this case, $Q(\mathbf{x})$ is also separable, that is, $Q(\mathbf{x}) = Q_h(\mathbf{x}_h)Q_z(z)$, and thus the integrals in Eqs. (4.1) and (4.2) can be separated as follows:

$$J_b = \int d\mathbf{x}_h \int d\mathbf{x}'_h Q_h(\mathbf{x}_h - \mathbf{x}'_h) J_z(\mathbf{x}_h, \mathbf{x}'_h), \quad (4.3a)$$

$$J_z(\mathbf{x}_h, \mathbf{x}'_h) = \int dz \int dz' v(\mathbf{x}_h, z) Q_z(z - z') v(\mathbf{x}'_h, z'), \quad (4.3b)$$

$$\int d\mathbf{x}'_h C_h(\mathbf{x}_h - \mathbf{x}'_h) Q_h(\mathbf{x}'_h) = \delta(\mathbf{x}_h), \quad (4.4a)$$

$$\int dz' C_z(z - z') Q_z(z') = \delta(z). \quad (4.4b)$$

The D-operator formulations will be derived for the background term first in the vertical as a one-dimensional example and then in the horizontal as a two-dimensional example in the following two subsections. Three-dimensional D-operators will be examined in section 5. As will be seen, with the D-operator, the double integral in Eq. (4.3a) [or Eq. (4.3b)] will reduce to a single integral with an infinite sum in the integrand, and the infinite sum can be truncated to a finite one.

4.1 One-dimensional D-operator

As a typical example, the vertical correlation is assumed to have a Gaussian form, say, given by $C_z(z) = \exp(-z^2/2L_z^2)$ where L_z is the vertical decorrelation length. The associated spectrum is the Fourier transformation of $C_z(z)$ given by

$$\begin{aligned} S_z(k_z) &= (2\pi)^{-1/2} \int dz C_z(z) \exp(-ik_z z) \\ &= L_z \exp(-k_z^2 L_z^2/2). \end{aligned}$$

By substituting this into the Fourier transformation of Eq. (4.4b), that is, $S_z(k_z)G_z(k_z) = (2\pi)^{-1}$ as in Eq. (1.3), we obtain

$$\begin{aligned} G_z(k_z) &= (2\pi L_z)^{-1} \exp(k_z^2 L_z^2/2) \\ &= (2\pi L_z)^{-1} \sum_0 (n!2^n)^{-1} (k_z^2 L_z^2)^n, \end{aligned} \quad (4.5)$$

where

$$G_z(k_z) = (2\pi)^{-1/2} \int dz Q_z(z) \exp(-ik_z z)$$

is the Fourier transformation of $Q_z(z)$, and the summation \sum_0 is over n from 0 to ∞ . Limited by the ordinary test functions introduced in Lighthill (1958,

page 5), the generalized Fourier transformation and its inverse cannot be directly applied to exponentially growing functions (to the infinity in the spaces of z and k_z , respectively). However, one can apply the generalized inverse Fourier transformation term-by-term to any high-order terms in the summation of Eq. (4.5) and then let the sum of the transformed terms go to infinity. This leads to

$$Q_z(z) = (2\pi L_z^2)^{-1/2} \sum_0 (n!2^n)^{-1} (-L_z^2 \partial_z^2)^n \delta(z). \quad (4.6)$$

Substituting this into Eq. (4.3b) and using the property of $\partial_z^{2n} \delta(z - z') = (-1)^n \partial_z^n \partial_{z'}^n \delta(z - z')$ and integration by parts, that is,

$$\begin{aligned} &\int dz \int dz' v(\mathbf{x}_h, z) [\partial_z^n \partial_{z'}^n \delta(z - z')] v(\mathbf{x}'_h, z') \\ &= \int dz \int dz' [\partial_z^n v(\mathbf{x}_h, z)] \delta(z - z') [\partial_{z'}^n v(\mathbf{x}'_h, z')] \\ &= \int dz [\partial_z^n v(\mathbf{x}_h, z)] [\partial_z^n v(\mathbf{x}'_h, z)], \end{aligned}$$

we obtain the following D-operator formulation:

$$J_z(\mathbf{x}_h, \mathbf{x}'_h) = \int dz [\mathbf{D}_z v(\mathbf{x}_h, z)]^T [\mathbf{D}_z v(\mathbf{x}'_h, z)], \quad (4.7a)$$

where $v(\mathbf{x}_h, z) = v_k(\mathbf{x})$ is the continuous field of analysis increment for the normalized univariate variable [see Eq. (3.2a)], and

$$\begin{aligned} \mathbf{D}_z &\equiv (2\pi L_z^2)^{-1/4} \{1, 2^{-1/2} L_z \partial_z, (2!)^{-1/2} (L_z \partial_z)^2, \\ &\dots, (n!2^n)^{-1/2} (L_z \partial_z)^n, \dots\}^T \end{aligned} \quad (4.7b)$$

is the D-operator (of infinite order). Note from Eq. (3.12) that $v(\mathbf{x}_h, z)$ belongs to H_k (Sobolev space of infinite order), so $v(\mathbf{x}_h, z)$ and all its derivatives approach 0 as $|z| \rightarrow \infty$. Thus, all the boundary terms vanish when integration by parts is used in the above derivation.

The background term is formulated in Eq. (4.7a) into a squared L_2 norm of $\mathbf{D}_z v$ or, equivalently, a squared Sobolev norm of v (Adams, 1975). The order of the D-operator is unlimitedly high and so is the order of the associated Sobolev space. This Sobolev space of infinite order is nontrivial (not empty) according to Theorem 2.1 of Dubinskij (1986). The D-operator of infinite order in Eq. (4.6) is thus the inverse of the Gaussian correlation operator $C_z(z)$ in this space. For a practical application, \mathbf{D}_z must be truncated to a finite order. Denote by $\mathbf{D}_z^{(n)}$ the n -th order truncation of \mathbf{D}_z that contains only the first $n + 1$ terms in Eq. (4.7b). The associated n -th order truncation of $G_z = G_z(k_z)$ is denoted by $G_z^{(n)}$. Since all the terms of G_z in Eq. (4.5) are positive, we have

$G_z > G_z^{(n)} > 0$ and $G_z > G_z - G_z^{(n)} > 0$ for $n > 0$. The truncation error can be estimated by

$$|G_z^{(n)} - G_z| = G_z - G_z^{(n)} < [(n+1)!]^{-1} \times (k_z^2 L_z^2 / 2)^{n+1} G_z, \quad (4.8a)$$

where the inequality $(n+1)![(n+1+n')!]^{-1} < (n')^{-1}$ is used. For any fixed k_z , $|G_z^{(n)} - G_z| \rightarrow 0$ as $n \rightarrow \infty$. In association with $G_z^{(n)}$, the truncated correlation spectrum is given by $S_z^{(n)} = (2\pi G_z^{(n)})^{-1}$ according to Eq. (1.3). The truncation error is

$$\begin{aligned} S_z^{(n)} - S_z &= [2\pi G_z^{(n)}]^{-1} - S_z \\ &= L_z \left\{ \left[\sum_{(0,n)} (n')^{-1} (k_z^2 L_z^2 / 2)^{n'} \right]^{-1} \right. \\ &\quad \left. - \exp(-k_z^2 L_z^2 / 2) \right\}, \end{aligned} \quad (4.8b)$$

where $S_z^{(n)} = (2\pi G_z^{(n)})^{-1} > (2\pi G_z)^{-1} = S_z > 0$ is used and the summation $\sum_{(0,n)}$ is over n' from 0 to n . Note that the truncation error estimated for $G_z^{(n)}$ in Eq. (4.8a) increases as a power function of k_z^2 , but the truncation error for $S_z^{(n)}$ in Eq. (4.8b) decreases rapidly as k_z increases.

The truncated correlation is given by the inverse Fourier transform of $S_z^{(n)}$:

$$\begin{aligned} C_z^{(n)}(z) &= (2\pi)^{-1/2} \int dk_z S_z^{(n)} \exp(ik_z z) \\ &= (2\pi)^{-1/2} L_z \int dk_z \left[\sum_{(0,n)} (n')^{-1} \right. \\ &\quad \left. \times (k_z^2 L_z^2 / 2)^{n'} \right]^{-1} \cos(k_z z), \end{aligned} \quad (4.9)$$

where the integral $\int dk_z$ is over the entire wavenumber space of k_z . The truncation error is bounded by

$$\begin{aligned} |C_z^{(n)}(z) - C_z(z)| &= \left| (2\pi)^{-1/2} \int dk_z [S_z^{(n)} - S_z] \cos(k_z z) \right| \\ &\leq (2\pi)^{-1/2} \int dk_z [S_z^{(n)} - S_z] = C_z^{(n)}(0) - C_z(0) \\ &= (2\pi)^{-1/2} \int d\eta \left[\sum_{(0,n)} (n')^{-1} (\eta^2 / 2)^{n'} \right]^{-1} - 1 \\ &\equiv \varepsilon_1(n), \end{aligned} \quad (4.10)$$

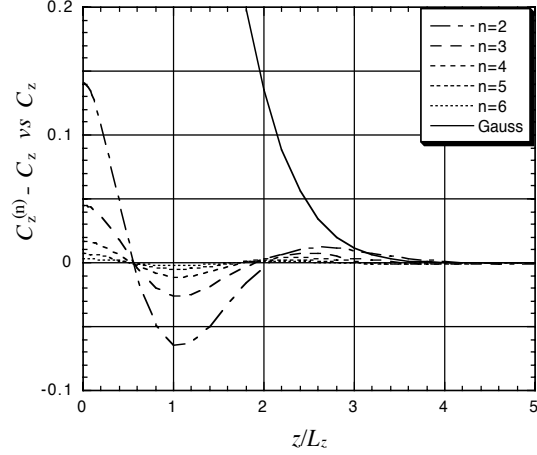


Fig. 1. $C_z^{(n)} - C_z$ plotted as functions of z/L_z for $n = 2, 3, 4, 5$ and 6 . The bottom portion of the one-dimensional Gaussian correlation function C_z is also shown (solid curve). The top of this solid curve is at $C_z(0) = 1$ beyond the plotted range.

Table 1. Truncation errors estimated by Eqs. (4.10) and (4.14) for one-dimensional Gaussian correlation function and related single-observation analysis, respectively.

n	2	3	4	5	6
$\varepsilon_1(n)$	0.141	0.045	0.017	0.007	0.003
$e_1(n)$	0.066	0.022	0.008	0.0035	0.0015

where $\eta = k_z L_z$, Eq. (4.8b) and $C_z(0) = 1$ are used. This result shows that the truncation error is an even function of z and reaches the upper bound $\varepsilon_1(n)$ at $z = 0$. Setting $k_z = 0$ in the Fourier integral of $C_z^{(n)}(z) - C_z(z)$ gives

$$\begin{aligned} &\int dz [C_z^{(n)}(z) - C_z(z)] \\ &= (2\pi)^{1/2} [S_z^{(n)}(0) - S_z(0)] = 0, \end{aligned} \quad (4.11)$$

where $S_z^{(n)}(0) = S_z(0) = L_z$ is used. This result indicates that the domain average of the truncation error is zero. The truncation errors are plotted as functions of $|z|$ for $n = 2, \dots, 6$ in Fig. 1. As shown, these functions decrease to zero around $|z| = 0.6L_z$, reach negative peaks around $|z| = 1.1L_z$ and then diminish rapidly as $|z|$ increases beyond $3L_z$. The upper bound $\varepsilon_1(n)$ decreases rapidly as n increases (see Table 1).

Using the above results, we can further estimate the truncation-caused error in the analysis of a single observation. According to Eqs. (3.10) and (3.12), $v_k(\mathbf{x})$ is a convolution of $K_{k'k}(\mathbf{x}_j, \mathbf{x}')$ with $C_k(\mathbf{x}', \mathbf{x})$, so $v_k(\mathbf{x})$ is at least as smooth as $C_k(\mathbf{x}', \mathbf{x})$ even when $K_{k'k}(\mathbf{x}_j, \mathbf{x}')$ reduces to $\delta(\mathbf{x}_j - \mathbf{x}')\delta_{kk'}$ which is the least differentiable case considered below. In this case, for a single observation of the k -th univariate variable $u_k(\mathbf{x})$

at $\mathbf{x} = \mathbf{x}_j$, the observation space is one-dimensional, $\mathbf{L}_k(\mathbf{x})$ in Eq. (3.10) reduces to $\sigma_k(\mathbf{x}_j)C_k(\mathbf{x}_j, \mathbf{x})$, and \mathbf{LCL}^T in Eq. (3.13) reduces to $\{\mathbf{L}_k(\mathbf{x}), \mathbf{L}_k^T(\mathbf{x})\}_k = [\sigma_k(\mathbf{x}_j)]^2$. Without loss of generality, we may choose $\mathbf{x}_j = 0$ and drop the subscript k . Substituting $v_k(\mathbf{x}) = v(\mathbf{x}_h, z)$ together with $\sigma_k(\mathbf{x}_j) = \sigma_k(0) = \sigma$ and $C_k(\mathbf{x}_j, \mathbf{x}) = C(\mathbf{x}) = C_h(\mathbf{x}_h)C_z(z)$ into Eqs. (3.15)–(3.16) yields

$$[\sigma^2 C_z(0) + \sigma_{\text{ob}}^2]c = d_o, \quad (4.12a)$$

$$v(\mathbf{x}_h, z) = c\sigma^2 C_h(\mathbf{x}_h)C_z(z), \quad (4.12b)$$

where c is a scalar reduced from vector \mathbf{c} in Eq. (3.15), and σ_{ob}^2 and d_o are the observation error variance and observation innovation at $\mathbf{x}_j = 0$, respectively. Note that $C_z(0) = 1$ but the symbol is retained as a tracer in Eq. (4.12a). When \mathbf{D}_z is truncated to $\mathbf{D}_z^{(n)}$ in the background formulation Eq. (4.7a), then Eq. (4.12) is replaced by

$$[\sigma^2 C_z^{(n)}(0) + \sigma_{\text{ob}}^2]c^{(n)} = d_o, \quad (4.13a)$$

$$v^{(n)}(\mathbf{x}_h, z) = c^{(n)}\sigma^2 C_h(\mathbf{x}_h)C_z^{(n)}(z), \quad (4.13b)$$

where $C_z^{(n)}(0) = 1 + \varepsilon_1(n) > 1$ and $v^{(n)}(\mathbf{x}_h, z)$ is the truncation-affected analysis. The associated truncation error is estimated by

$$\begin{aligned} & |v^{(n)}(\mathbf{x}_h, z) - v(\mathbf{x}_h, z)| |v(\mathbf{x}_h, 0)|^{-1} = \\ & |c^{(n)}C_z^{(n)}(z) - cC_z(z)| \sigma^2 C_h(\mathbf{x}_h) |v(\mathbf{x}_h, 0)|^{-1} \\ & \leq \varepsilon_1(n) \sigma_{\text{ob}}^2 \{\sigma^2 [1 + \varepsilon_1(n)] + \sigma_{\text{ob}}^2\}^{-1} \equiv e_1(n) \end{aligned} \quad (4.14)$$

where Eq. (4.10) and

$$|v(\mathbf{x}_h, 0)| = c\sigma^2 |C_h(\mathbf{x}_h)C_z(0)| = c\sigma^2 |C_h(\mathbf{x}_h)|$$

are used. For $\sigma = \sigma_{\text{ob}}$, we have

$$e_1(n) = \varepsilon_1(n) [2 + \varepsilon_1(n)]^{-1}.$$

As shown in Table 1, $e_1(n)$ decreases rapidly as n increases, and $\mathbf{D}_z^{(n)}$ can be a good approximation of \mathbf{D}_z for $n \geq 3$ or even $n = 2$.

4.2 Two-dimensional D-operator

The generalized Fourier transformation of Eq. (4.4a) yields

$$S_h(\mathbf{k}_h)G_h(\mathbf{k}_h) = (2\pi)^{-2}, \quad (4.15)$$

where

$$S_h(\mathbf{k}_h) = (2\pi)^{-1} \int d\mathbf{x}_h C_h(\mathbf{x}_h) \exp(-i\mathbf{k}_h \cdot \mathbf{x}_h),$$

$$G_h(\mathbf{k}_h) = (2\pi)^{-1} \int d\mathbf{x}_h Q_h(\mathbf{x}_h) \exp(-i\mathbf{k}_h \cdot \mathbf{x}_h),$$

$\mathbf{k}_h = (k_x, k_y)$ is the horizontal vector wavenumber, and \cdot denotes the dot-product. In addition to the assumed homogeneity in Eq. (4.4a), the correlation considered here is isotropic and has a Gaussian form, say,

given by

$$C_h(\mathbf{x}_h) = \exp(-|\mathbf{x}_h|^2/2L_h^2)$$

where L_h is the horizontal decorrelation length. The associated spectrum is $S_h(\mathbf{k}_h) = L_h^2 \exp(-|\mathbf{k}_h|^2 L_h^2/2)$. Substituting this into Eq. (4.15) gives

$$G_h(\mathbf{k}_h) = (2\pi L_h)^{-2} \sum_0 (n!2^n)^{-1} (L_h |\mathbf{k}_h|)^{2n}. \quad (4.16a)$$

Applying the generalized inverse Fourier transformation term-by-term to Eq. (4.16a) gives

$$Q_h(\mathbf{x}_h) = (2\pi L_h^2)^{-1} \sum_0 (n!2^n)^{-1} (-L_h^2 \Delta_h)^n \delta(\mathbf{x}_h), \quad (4.16b)$$

where $\Delta_h = \partial_x^2 + \partial_y^2$. Substituting this with Eq. (4.7a) into Eq. (4.3a) and integrating by parts yield the following D-operator formulation:

$$J_b = \int d\mathbf{x}_h \int dz [\mathbf{D}_h \mathbf{D}_z^T v(\mathbf{x}_h, z)] \cdot [\mathbf{D}_h \mathbf{D}_z^T v(\mathbf{x}_h, z)]. \quad (4.17a)$$

Here, $\cdot\cdot$ denotes the Schur product defined by $\mathbf{A} \cdot\cdot \mathbf{B} \equiv \sum_i \sum_{i'} A_{ii'} B_{ii'}$ where the double summation is over i and i' , and $A_{ii'}$ and $B_{ii'}$ denote the elements of \mathbf{A} and \mathbf{B} , respectively. As a matrix differential operator, $\mathbf{D}_h \mathbf{D}_z^T$ is constructed by the outer-product of \mathbf{D}_h and \mathbf{D}_z , where \mathbf{D}_z is as in Eq. (4.7b) and \mathbf{D}_h is defined by

$$\begin{aligned} \mathbf{D}_h \equiv & (2\pi)^{-3/4} (L_h^2 L_z)^{-1/2} \{1, 2^{-1/2} L_h \mathbf{d}_{h1}, (2!)^{-1/2} \\ & \times L_h^2 \mathbf{d}_{h2}, \dots, (n!2^n)^{-1/2} L_h^n \mathbf{d}_{hn}, \dots\}^T, \end{aligned} \quad (4.17b)$$

$$\left\{ \begin{aligned} \mathbf{d}_{h1} &= \{\partial_x, \partial_y\}, \\ \mathbf{d}_{h2} &= \{\partial_x^2, 2^{1/2} \partial_x \partial_y, \partial_y^2\}, \\ \dots & \\ \mathbf{d}_{hn} &= \{\partial_x^n, n^{1/2} \partial_x^{n-1} \partial_y, \dots, \\ & (\alpha_{n,p})^{1/2} \partial_x^{n-p} \partial_y^p, \dots, \partial_y^n\}, \\ \dots & \end{aligned} \right. \quad (4.17c)$$

and $\alpha_{n,p} = n![(n-p)!p!]^{-1}$ for $0 \leq p \leq n$.

Denote by $\mathbf{D}_h^{(n)}$ the n -th order truncation of \mathbf{D}_h that contains only the first $n+1$ terms up to $(1/n!)^{1/2} (L_h 2^{-1/2})^n \mathbf{d}_{hn}$ in Eq. (4.17b), and by $G_h^{(n)}$ the n -th order truncation of G_h in Eq. (4.16a). The truncated correlation spectrum is given by $S_h^{(n)} = [(2\pi)^2 G_h^{(n)}]^{-1}$ according to Eq. (4.15). The truncation error can be estimated for $G_h^{(n)}$ as in Eq. (4.8a) except that $k_z L_z$ is replaced by $|\mathbf{k}_h| L_h$. The truncation error for $S_h^{(n)}$ is

$$\begin{aligned} S_h^{(n)} - S_h &= [(2\pi)^2 G_h^{(n)}]^{-1} - S_h \\ &= L_h^2 \left\{ \left[\sum_{(0,n)} (n!)^{-1} (|\mathbf{k}_h|^2 L_h^2/2)^{n'} \right]^{-1} \right. \\ & \quad \left. - \exp(-|\mathbf{k}_h|^2 L_h^2/2) \right\}, \end{aligned} \quad (4.18)$$

where $S_h^{(n)} > S_h > 0$ is used and $\sum_{(0,n)}$ is as in Eq. (4.8b).

The truncated correlation is the inverse Fourier transform of $S_h^{(n)}$:

$$\begin{aligned} C_h^{(n)}(\mathbf{x}_h) &= (2\pi)^{-1} \int d\mathbf{k}_h S_h^{(n)} \cos(\mathbf{k}_h \cdot \mathbf{x}_h) \\ &= \int_0^\infty k_h dk_h S_h^{(n)}(k_h) J_0(k_h r_h), \end{aligned} \quad (4.19)$$

where $k_h = |\mathbf{k}_h|$, $r_h = |\mathbf{x}_h|$, $S_h^{(n)}(k_h)$ denotes the same spectrum as $S_h^{(n)}(\mathbf{k}_h)$ but expressed as a function of k_h , and $J_0(\cdot)$ is the zero-th order Bessel function of (\cdot) . The truncation error is bounded by

$$\begin{aligned} &|C_h^{(n)}(\mathbf{x}_h) - C_h(\mathbf{x}_h)| \\ &= |(2\pi)^{-1} \int d\mathbf{k}_h [S_h^{(n)} - S_h] \cos(\mathbf{k}_h \cdot \mathbf{x}_h)| \\ &\leq \int_0^\infty k_h dk_h [S_h^{(n)}(k_h) - S_h(k_h)] = C_h^{(n)}(0) - C_h(0) \\ &= \int_0^\infty d\eta \left[\sum_{(0,n)} (n!)^{-1} \eta^{n'} \right]^{-1} - 1 \equiv \varepsilon_h(n), \end{aligned} \quad (4.20)$$

where $\eta = (k_h L_h)^2/2$, Eq. (4.18) and $C_h(0) = 1$ are used. Setting $\mathbf{k}_h = 0$ in the Fourier integral of $C_h^{(n)}(\mathbf{x}_h) - C_h(\mathbf{x}_h)$ gives

$$\begin{aligned} &\int d\mathbf{x}_h [C_h^{(n)}(\mathbf{x}_h) - C_h(\mathbf{x}_h)] \\ &= 2\pi [S_h^{(n)}(0) - S_h(0)] = 0, \end{aligned} \quad (4.21)$$

so the domain average of the truncation error is zero. As shown in Fig. 2, the errors decrease to zero around $r_h = 0.8L_h$, reach negative peak values around $r_h = 1.3L_h$ and then diminish rapidly as r_h increases beyond $3L_h$. The upper bound $\varepsilon_h(n)$ decreases rapidly as n increases (see Table 2).

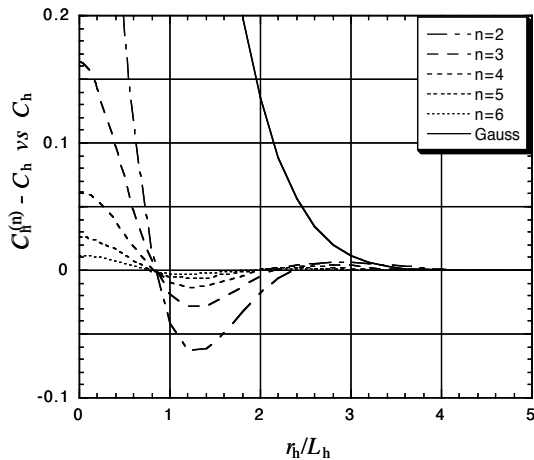


Fig. 2. As in Fig. 1 but for $C_h^{(n)} - C_h$ plotted as functions of r_h/L_h for the two-dimensional case.

Table 2. Truncation errors estimated by Eqs. (4.20) and (4.22) for two-dimensional Gaussian correlation function and related single-observation analysis, respectively.

n	2	3	4	5	6
$\varepsilon_h(n)$	0.57	0.16	0.062	0.026	0.012
$e_h(n)$	0.22	0.074	0.030	0.013	0.006

As explained earlier, for a single observation at the point of $\mathbf{x} = \mathbf{x}_j = 0$, the analysis is given by Eq. (4.12b). When \mathbf{D}_h is truncated to $\mathbf{D}_h^{(n)}$ in Eq. (4.17a), the analysis is given by

$$v^{(n)}(\mathbf{x}_h, z) = c^{(n)} \sigma^2 C_h^{(n)}(\mathbf{x}_h) C_z(z),$$

where $c^{(n)}$ is obtained as in Eq. (4.13a) except that $C_z^{(n)}(0)$ is replaced by $C_h^{(n)}(0)$. The truncation-caused analysis error relative to $|v(0, z)|$ is bounded by

$$\begin{aligned} &|v^{(n)}(\mathbf{x}_h, z) - v(\mathbf{x}_h, z)| |v(0, z)|^{-1} \\ &= [c^{(n)} C_h^{(n)}(\mathbf{x}_h) - c C_h(\mathbf{x}_h)] \sigma^2 C_z(z) |v(0, z)|^{-1} \\ &\leq \varepsilon_h(n) \sigma_{\text{ob}}^2 \{ \sigma^2 [1 + \varepsilon_h(n)] + \sigma_{\text{ob}}^2 \}^{-1} \equiv e_h(n), \end{aligned} \quad (4.22)$$

where Eq. (4.20) and $|v(0, z)| = c \sigma^2 |C_h(0) C_z(z)| = c \sigma^2 |C_z(z)|$ are used. For $\sigma = \sigma_{\text{ob}}$, we have $e_h(n) = \varepsilon_h(n) [2 + \varepsilon_h(n)]^{-1}$. As shown in Table 2, $e_h(n)$ decreases rapidly as n increases, and $\mathbf{D}_h^{(n)}$ can be a good approximation of \mathbf{D}_h for $n \geq 4$ or even $n = 3$.

5. D-operators in three-dimensional space

5.1 D-operators for correlation functions separable between horizontal and vertical

In the previous section, the correlation function is assumed to be separable between the horizontal and vertical. The D-operator is also separable and is given by $\mathbf{D}_h \mathbf{D}_z^T$ as shown in Eq. (4.17a). When $\mathbf{D}_h \mathbf{D}_z^T$ is truncated to $\mathbf{D}_h^{(n)} \mathbf{D}_z^{(n')T}$, the truncation error can be estimated by

$$\begin{aligned} &|C_h^{(n)}(\mathbf{x}_h) C_z^{(n')}(z) - C_h(\mathbf{x}_h) C_z(z)| \\ &\leq |C_z^{(n')}(z) - C_z(z)| |C_h^{(n)}(\mathbf{x}_h)| \\ &\quad + |C_h^{(n)}(\mathbf{x}_h) - C_h(\mathbf{x}_h)| |C_z(z)| \\ &\leq \varepsilon_1(n') |C_h^{(n)}(\mathbf{x}_h)| + \varepsilon_h(n) |C_z(z)| \\ &\leq \varepsilon_1(n') + \varepsilon_h(n), \end{aligned} \quad (5.1)$$

where Eqs. (4.10) and (4.20) are used.

As in Eqs. (4.14) and (4.22), the truncation error in the analysis of a single observation is bounded by

$$\begin{aligned} &|v^{(n,n')}(\mathbf{x}_h, z) - v(\mathbf{x}_h, z)| \leq [\varepsilon_h(n) + \varepsilon_1(n')] \sigma_{\text{ob}}^2 \\ &\quad \times \{ \sigma^2 [1 + \varepsilon_h(n) + \varepsilon_1(n')] + \sigma_{\text{ob}}^2 \}^{-1} |v(0, 0)|. \end{aligned} \quad (5.2)$$

The results in Tables 1–2 indicate that $\mathbf{D}_h^{(n)} \mathbf{D}_z^{(n')\text{T}}$ can be a good approximation of $\mathbf{D}_h \mathbf{D}_z^{\text{T}}$ for $n \geq 4$ and $n' \geq 3$.

5.2 *D-operators for correlation functions separable between vertical modes*

When the background errors are not correlated between different vertical modes, we have

$$C(\mathbf{x}) = \sum C_m(\mathbf{x}_h) Z_m(z), \quad (5.3)$$

$$Q(\mathbf{x}) = \sum Q_m(\mathbf{x}_h) Z_m(z), \quad (5.4)$$

where the summation is over m , $C_m(\mathbf{x}_h)$ and $Q_m(\mathbf{x}_h)$ are the projections of $C(\mathbf{x})$ and $Q(\mathbf{x})$, respectively, onto the m -th vertical mode, and the vertical modes are assumed to be orthonormal, that is, $\int dz Z_m(z) Z_{m'}(z) = \delta_{mm'}$. Similarly, the analysis field can be expressed by

$$v(\mathbf{x}) = \sum v_m(\mathbf{x}_h) Z_m(z), \quad (5.5)$$

where $v_m(\mathbf{x}_h)$ is the projection of $v(\mathbf{x})$ onto the m -th vertical mode. Substituting Eqs. (5.3)–(5.4) into Eqs. (4.1)–(4.2) gives

$$J_b = \sum J_m, \quad (5.6a)$$

$$J_m = \int d\mathbf{x}_h \int d\mathbf{x}'_h v_m(\mathbf{x}_h) Q_m(\mathbf{x}_h - \mathbf{x}'_h) v_m(\mathbf{x}'_h), \quad (5.6b)$$

$$\int d\mathbf{x}'_h C_m(\mathbf{x}_h - \mathbf{x}'_h) Q_m(\mathbf{x}'_h) = \delta(\mathbf{x}_h). \quad (5.6c)$$

The two-dimensional D-operator in section 4.2 can be applied to each J_m in Eq. (5.6a), since J_m involves only horizontal correlation as shown in Eqs. (5.5b)–(5.5c).

5.3 *D-operators for ellipsoidal correlations*

The generalized Fourier transformation of Eq. (4.2) is

$$S(\mathbf{k})G(\mathbf{k}) = (2\pi)^{-3}, \quad (5.7)$$

where

$$S(\mathbf{k}) = (2\pi)^{-3/2} \int d\mathbf{x} C(\mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x}),$$

$$G(\mathbf{k}) = (2\pi)^{-3/2} \int d\mathbf{x} Q(\mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x}),$$

and $\mathbf{k} = (k_x, k_y, k_z)$ is the three-dimensional vector wavenumber. In Eq. (4.1), the correlation is assumed to be homogenous but not necessarily isotropic or separable. The correlation considered here is homogenous and has a Gaussian ellipsoidal form, say, given by

$$C(\mathbf{x}) = \exp(-|\boldsymbol{\lambda} \mathbf{E} \mathbf{x}|^2 / 2L^2) = \exp(-|\mathbf{x}'|^2 / 2L^2).$$

Here, \mathbf{E} is the matrix that rotates the \mathbf{x} -coordinates to coincide with the ellipsoidal axes, $\boldsymbol{\lambda} = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$

is the diagonal matrix that transforms the Gaussian ellipsoid into a sphere, and L is the decorrelation length for the Gaussian function in the transformed coordinate system of $\boldsymbol{\lambda} \mathbf{E} \mathbf{x} = \mathbf{x}' = (x', y', z')^{\text{T}}$. Without loss of generality, $\det(\boldsymbol{\lambda}) = \lambda_1 \lambda_2 \lambda_3 = 1$ is chosen, so $\det(\boldsymbol{\lambda} \mathbf{E}) = 1$ and the transformation conserves volumes. The correlation is isotropic in \mathbf{x}' . The associated spectrum is given by

$$\begin{aligned} S(\mathbf{k}) &= (2\pi)^{-3/2} \int d\mathbf{x} C(\mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x}) \\ &= (2\pi)^{-3/2} \int d\mathbf{x}' \exp(-|\mathbf{x}'|^2 / 2L^2) \exp(-i\mathbf{k}' \cdot \mathbf{x}') \\ &= L^3 \exp(-|\mathbf{k}'|^2 L^2 / 2) \\ &= L^3 \exp(-|\boldsymbol{\lambda}^{-1} \mathbf{E} \mathbf{k}|^2 L^2 / 2), \end{aligned} \quad (5.8)$$

where $\det(\partial \mathbf{x} / \partial \mathbf{x}') = \det(\boldsymbol{\lambda} \mathbf{E}) = 1$ and $\mathbf{k}' \cdot \mathbf{x}' = (\boldsymbol{\lambda}^{-1} \mathbf{E} \mathbf{k})^{\text{T}} (\boldsymbol{\lambda} \mathbf{E} \mathbf{x}) = \mathbf{k} \cdot \mathbf{x}$ are used, and $\mathbf{k}' = \boldsymbol{\lambda}^{-1} \mathbf{E} \mathbf{k}$ is the transformed wavenumber in association with \mathbf{x}' . Substituting Eq. (5.8) into Eq. (5.7) gives

$$G(\mathbf{k}) = (2\pi L)^{-3} \sum_0 (n! 2^n)^{-1} (|\boldsymbol{\lambda}^{-1} \mathbf{E} \mathbf{k}| L)^{2n}. \quad (5.9a)$$

Applying the generalized inverse Fourier transformation term-by-term to Eq. (5.9a) gives

$$Q(\mathbf{x}) = (2\pi L^2)^{-3/2} \sum_0 (n! 2^n)^{-1} (-L^2 \Delta')^n \delta(\mathbf{x}), \quad (5.9b)$$

where $\Delta' = \nabla' \cdot \nabla'$, $\nabla' = \{\partial_{x'}, \partial_{y'}, \partial_{z'}\}^{\text{T}} = \boldsymbol{\lambda}^{-1} \mathbf{E} \nabla$ and $\nabla = (\partial_x, \partial_y, \partial_z)^{\text{T}}$. Substituting Eq. (5.9b) into Eq. (4.1) and integrating by parts lead to the following D-operator formulation:

$$J_b = \int d\mathbf{x} |\mathbf{D} v(\mathbf{x})|^2, \quad (5.10a)$$

where

$$\begin{aligned} \mathbf{D} \equiv & (2\pi L^2)^{-3/4} \{1, 2^{-1/2} L \mathbf{d}_1, (2! 2)^{-1/2} L^2 \mathbf{d}_2, \\ & \dots, (n! 2^n)^{-1/2} L^n \mathbf{d}_n, \dots\}^{\text{T}}, \end{aligned} \quad (5.10b)$$

$$\left\{ \begin{aligned} \mathbf{d}_1 &= \{\partial_{x'}, \partial_{y'}, \partial_{z'}\}, \\ \mathbf{d}_2 &= \{\partial_{x'}^2, \partial_{y'}^2, \partial_{z'}^2, 2^{1/2} \partial_{x'} \partial_{y'}, 2^{1/2} \partial_{x'} \partial_{z'}, 2^{1/2} \partial_{y'} \partial_{z'}\}, \\ &\dots \\ \mathbf{d}_n &= \{\partial_{x'}^n, \partial_{y'}^n, \partial_{z'}^n, n^{1/2} \partial_{x'}^{n-1} \partial_{y'}, \dots, \\ &(\beta_{n,p,q})^{1/2} \partial_{x'}^{n-p-q} \partial_{y'}^p \partial_{z'}^q \dots\}, \\ &\dots \end{aligned} \right. \quad (5.10c)$$

$\beta_{n,p,q} = n! [(n-p-q)! p! q!]^{-1}$, p and q are integers between 0 and n , and their sum is bounded by $p+q \leq n$. The derivatives in (5.10c) are given in the transformed coordinates. They can be transformed back to the

original coordinates by using $\nabla' = \boldsymbol{\lambda}^{-1} \mathbf{E} \nabla$ or, equivalently,

$$\{\partial_{x'}, \partial_{y'}, \partial_{z'}\} = (\lambda_1 \mathbf{e}_1 \cdot \nabla, \lambda_2 \mathbf{e}_2 \cdot \nabla, \lambda_3 \mathbf{e}_3 \cdot \nabla)$$

where $\mathbf{e}_1, \mathbf{e}_2$ and \mathbf{e}_3 are the three row vectors of \mathbf{E} , that is, $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \mathbf{E}^T$. This gives

$$\partial_{x'}^{n-p-q} \partial_{y'}^p \partial_{z'}^q = (\lambda_1 \mathbf{e}_1 \cdot \nabla)^{n-p-q} (\lambda_2 \mathbf{e}_2 \cdot \nabla)^p (\lambda_3 \mathbf{e}_3 \cdot \nabla)^q$$

in Eq. (5.10c).

Denote by $\mathbf{D}^{(n)}$ the n -th order truncation of \mathbf{D} in Eq. (5.10b), and by $G^{(n)}$ the n -th order truncation of G in Eq. (5.9a). The truncated correlation spectrum is given by $S^{(n)} = [(2\pi)^3 G^{(n)}]^{-1}$ according to Eq. (5.7). The truncation error is

$$\begin{aligned} S^{(n)} - S &= [(2\pi)^3 G^{(n)}]^{-1} - S \\ &= L^3 \left\{ \left[\sum_{(0,n)} (n!)^{-1} (|\mathbf{k}'|^2 L^2 / 2)^{n'} \right]^{-1} \right. \\ &\quad \left. - \exp(-|\mathbf{k}'|^2 L^2 / 2) \right\}, \end{aligned} \quad (5.11)$$

where $\sum_{(0,n)}$ is as in Eq. (4.8b).

The truncated correlation is given by the inverse Fourier transform of $S^{(n)}$:

$$\begin{aligned} C^{(n)}(\mathbf{x}) &= (2\pi)^{-3/2} \int d\mathbf{k} S^{(n)} \cos(\mathbf{k} \cdot \mathbf{x}) \\ &= (2\pi)^{-3/2} \int d\mathbf{k}' S^{(n)}(|\mathbf{k}'|) \cos(\mathbf{k}' \cdot \mathbf{x}') \\ &= (2/\pi)^{1/2} \int_0^\infty k'^2 dk' S^{(n)}(k') [\sin(k' r') / (k' r')], \end{aligned} \quad (5.12)$$

where $k' = |\mathbf{k}'|$, $r' = |\mathbf{x}'|$, $S^{(n)}(k') = S^{(n)}(|\mathbf{k}'|)$ denotes the same spectrum as $S^{(n)}(\mathbf{k})$ but expressed as a function of $k' = |\mathbf{k}'|$. The truncation error is bounded by

$$\begin{aligned} &|C^{(n)}(\mathbf{x}) - C(\mathbf{x})| \\ &= |(2\pi)^{-3/2} \int d\mathbf{k}' [S^{(n)}(|\mathbf{k}'|) - S(|\mathbf{k}'|)] \cos(\mathbf{k}' \cdot \mathbf{x}')| \\ &\leq (2/\pi)^{1/2} \int_0^\infty k'^2 dk' [S^{(n)}(k') - S(k')] \\ &= C^{(n)}(0) - C(0) \\ &= (2/\pi)^{1/2} \int_0^\infty \eta^2 d\eta \left[\sum_{(0,n)} (n!)^{-1} (\eta^2 / 2)^{n'} \right]^{-1} - 1 \\ &\equiv \varepsilon(n), \end{aligned} \quad (5.13)$$

where $\eta = Lk'$, Eq. (5.11) and $C(0) = 1$ are used. Similarly to Eq. (4.11) and (4.21), one can verify that

$$\begin{aligned} &\int d\mathbf{x} [C^{(n)}(\mathbf{x}) - C(\mathbf{x})] \\ &= (2\pi)^{3/2} [S^{(n)}(0) - S(0)] = 0. \end{aligned} \quad (5.14)$$

The truncation errors are plotted in Fig. 3. As shown, the errors decrease to zero around $r' = 1.1L$, reach negative peaks around $r' = 1.4L$, and then diminish

rapidly as r' increases beyond $3L$. The upper bound $\varepsilon(n)$ decreases rapidly as n increases (see Table 3).

When \mathbf{D} is replaced by $\mathbf{D}^{(n)}$ in the background formulation (5.10a), the analysis of a single observation is given by $v^{(n)}(\mathbf{x}_h, z) = c^{(n)} \sigma^2 C^{(n)}(\mathbf{x})$, where $c^{(n)}$ is obtained as in Eq. (4.13a) except that $C_z^{(n)}(0)$ is replaced by $C^{(n)}(0) = 1 + \varepsilon(n)$. The truncation error in the analysis is bounded by

$$\begin{aligned} &|v^{(n)}(\mathbf{x}) - v(\mathbf{x})| / |v(0)| \\ &\leq \varepsilon(n) \sigma_{\text{ob}}^2 \{ \sigma^2 [1 + \varepsilon(n)] + \sigma_{\text{ob}}^2 \}^{-1} \\ &\equiv e(n). \end{aligned} \quad (5.15)$$

For $\sigma = \sigma_{\text{ob}}$, $e(n) = \varepsilon(n)[2 + \varepsilon(n)]^{-1}$. As shown in Table 3, $e(n)$ decreases rapidly as n increases, and $\mathbf{D}^{(n)}$ can be a good approximation of \mathbf{D} for $n \geq 5$ or even $n = 4$. General D-operator formulations for non-Gaussian correlations are derived in the appendix.

6. Discussion and conclusion

It is shown in this paper that the continuous-form background term in the cost function of three- or four-dimensional variational data assimilation can be written into a squared Sobolev norm (Adams, 1975) or, equivalently, a squared L_2 norm of a vector differential operator (D-operator) applied to the continuous

Table 3. Truncation errors estimated by Eqs. (5.13) and (5.15) for three-dimensional Gaussian correlation function and related single-observation analysis, respectively.

n	2	3	4	5	6
$\varepsilon(n)$	1.8	0.43	0.160	0.066	0.030
$e(n)$	0.47	0.177	0.074	0.032	0.015

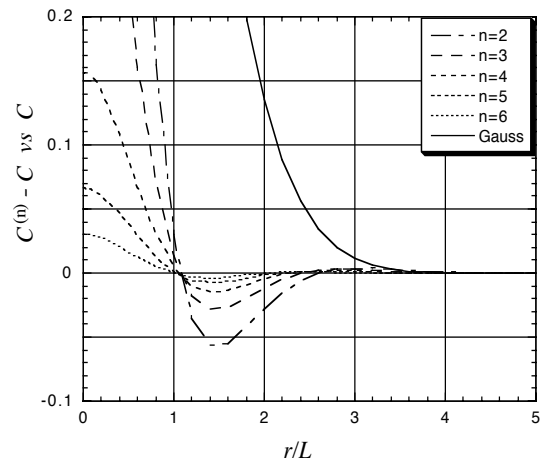


Fig. 3. As in Fig. 1 but for $C^{(n)} - C$ plotted as functions of r/L for the three-dimensional case.

field of analysis increment [see Eq. (5.10) or Eq. (A.6)]. In particular, D-operators are derived for Gaussian correlation functions in one-, two- and three-dimensional spaces. These Gaussian D-operators have unlimitedly high order, so their associated norms belong to Sobolev spaces of infinite order (Dubinskij, 1986), which means that the analysis increment fields are infinitely smooth.

For practical applications, the Gaussian D-operators should be truncated to finite orders. Errors in the truncated correlations are estimated analytically and found to be small (< 0.07 , see line 1 in Tables 1–3) even when the Gaussian D-operators are truncated to the third, fourth and fifth orders in one-, two- and three-dimensional spaces, respectively. Truncation-caused errors are also estimated for single-observation analyses obtained with truncated D-operators. These errors are found to be smaller than those for the truncated correlations (compare line 2 with line 1 in Tables 1–3). Since the domain-averaged errors are zero for all the truncated correlations [see Eqs. (4.11), (4.21) and (5.14)], positive and negative errors caused by a truncated correlation applied to a large number of observations may largely cancel out. The actual errors caused by using truncated D-operators in the analyses should be lower than estimated in this paper (and this is confirmed by numerical results reported in follow-up papers). Furthermore, when truncated D-operators are used for the background term in four-dimensional variational data assimilation, truncation-caused errors in the solution might be further reduced because the smoothness of the solution is controlled not only by the background error correlation but also by the prediction model equations, although the truncation-caused errors could also be amplified due to various instabilities in the prediction model.

D-operator formulations are also derived for non-Gaussian and non-isotropic correlations in the appendix. The associated truncation errors (if any) can be estimated in the same way as demonstrated for the truncated Gaussian D-operators in sections 4–5. It is necessary to point out, however, that many non-Gaussian D-operators have finite orders only and exactly, so there is no need for truncation. This is especially true for the autoregressive correlations for finite orders, including the widely used second-order autoregressive (SOAR) function (Thiebaut, 1976; Gaspari and Cohn, 1999). As shown in the introduction, the D-operator is exactly of n -th order for the n -th order autoregressive correlation in the one-dimensional space. This result remains true when it is extended to two- and three-dimensional spaces (not shown). As mentioned in section 2b for the model space approach, various compactly supported correlation functions (con-

strained to zero beyond certain distances) have been proposed (Gaspari and Cohn, 1999) to reduce the computational cost in the multiplication of the state vector by the covariance matrix (and to improve the analyses in the Tropics). With the D-operator formulation, the state vector is “multiplied” by the D-operator instead of the covariance matrix, so the computational cost is affected mainly by the order and the (discrete) form of the D-operator regardless whether the correlation function is compactly supported. In a spectral model, the D-operator is simply a power series in wavenumber, so the background term and its gradient can be computed very efficiently similarly to the spectral-space approach reviewed in section 2.3. In a grid model, the background term and its gradient can be computed efficiently if the D-operator is truncated to a relatively low order and discretized by low-order finite-difference schemes.

It is assumed in this paper that the correlations are homogeneous over the entire space of analysis. This assumption can be relaxed and applied to local areas. In this case, the parameters that describe the shape and orientation (if not isotropic) of a specified correlation are not necessarily constant but can slowly vary in the analysis space. The correlation is thus only locally homogeneous. D-operator formulations can be derived similarly by applying the generalized Fourier transformation approximately within the concerned ranges of length scales or wavenumbers. The formulation will have the same form as derived in this paper except that the D-operator coefficients [such as g_n in Eq. (A.6c)] will be slowly varying functions in the analysis space. Furthermore, by using a proper curvilinear coordinate transformation, the D-operator can be formulated first for a homogeneous correlation function in the transformed space, and then transformed back to the original space to represent the inverse of a desired and admissible non-homogeneous correlation function. This technique is worthy of further investigations.

As mentioned in the introduction, differential operators have sometimes been used to construct penalty terms to filter noise and improve the smoothness of the analysis. For example, weak vorticity and divergence constraints were used by Xu et al. (1994, 1995) to suppress spurious strong divergence and vorticity caused by shortwave noise in the data, and the associated differential operators can be reviewed as a horizontal Laplacian applied to the streamfunction and velocity potential fields. In general, differential operators used in smoothing penalty terms can be viewed as D-operators and their implied “correlations” can be examined by using the general inverse relationship derived in Eqs. (A.4)–(A.10). The D-operators can also

be viewed as generalized spline filters (Wahba, 1990). Owing to the duality between the spline model solution and Bayes estimate, the spline model solution can be obtained by the representer method in section 3 except that the null space of the differential operator may no longer be empty and should be excluded from the solution space [see Chapter 1 of Wahba (1990)]. The D-operator formulations derived in this paper have demonstrated how to construct a generalized spline filter whose response function is precisely consistent with the background error correlation.

Quadratic bi-spline basis functions have been used to express the fields of analysis on coarse finite-element meshes to filter shortwave noise and reduce the analysis space dimension (Xu et al., 2001b). The filter property, however, was controlled only implicitly by the finite-element meshes. Using the formulations derived in this paper, D-operators can be conveniently constructed with spline basis functions to achieve the desired filter property (in consistency with the estimated background error covariances). This approach was used recently in Doppler radar data assimilation with encouraging results (Xu et al., 2001a), although the D-operator was simply a Laplacian constructed with quadratic bi-spline basis functions on coarse finite-element meshes in the horizontal. The utilities and potential merits of D-operator formulations for practical applications to discrete models will be examined in follow-up papers.

Acknowledgments. The author is thankful to Dr. Jiandong Gong for his help in producing the figures and to Drs. David Parrish, James Purser and Stephen Cohn and anonymous reviewers for their comments and suggestions that improved the exposition of the results. The research work was supported by the FAA contract IA# DTFA03-01-X-9007 to NSSL, and by the ONR Grants N000140310822 and N000140410312 to the University of Oklahoma.

APPENDIX

D-operators for non-Gaussian correlations

For simplicity, \mathbf{x} is used to represent a point in any of the three spaces and \mathbf{k} is a point in the associated wavenumber space. Denote by $C(|\lambda\mathbf{E}\mathbf{x}|) = C(|\mathbf{x}'|)$ a non-Gaussian correlation which is isotropic in the transformed space of $\mathbf{x}' = \lambda\mathbf{E}\mathbf{x}$, where the transformed matrix $\lambda\mathbf{E}$ is similar to that in section 5.3 but is used here for both two- and three-dimensional spaces and reduces to 1 for one-dimensional space. This correlation function is elliptical or ellipsoidal in the original two- or three-dimensional space of \mathbf{x} . The associated

correlation spectrum is given by

$$\begin{aligned} & (2\pi)^{-n_d/2} \int d\mathbf{x} C(|\lambda\mathbf{E}\mathbf{x}|) \exp(-i\mathbf{k}\cdot\mathbf{x}) \\ &= (2\pi)^{-n_d/2} \int d\mathbf{x}' C(|\mathbf{x}'|) \exp(-i\mathbf{k}'\cdot\mathbf{x}') \\ &= S(|\mathbf{k}'|), \end{aligned} \quad (\text{A.1})$$

where $\mathbf{k}' = \lambda^{-1}\mathbf{E}\mathbf{k}$ and $n_d (= 1, 2, 3)$ is the number of dimensions. The function forms of C and S are related to each other by

$$S(k') = \int_0^\infty r'^{n_d-1} dr' C(r') P(k'r'), \quad (\text{A.2a})$$

$$C(r') = \int_0^\infty k'^{n_d-1} dk' S(k') P(k'r'), \quad (\text{A.2b})$$

where $k' = |\mathbf{k}'|$, $r' = |\mathbf{x}'|$, $P(k'r')$ is given by $(2/\pi)^{1/2} \cos(k'r')$ for $n_d = 1$, by $J_0(k'r')$ for $n_d = 2$, and by $(2/\pi)^{1/2} (k'r')^{-1} \sin(k'r')$ for $n_d = 3$. As a correlation spectrum, $S(k')$ is an even and non-negative function of k' . Assume further that $S(k') > 0$ for $k' < \infty$. According to Eqs. (1.3), (4.15) and (5.7), we have

$$S(k')G(k') = (2\pi)^{-n_d}. \quad (\text{A.3})$$

This implies that $G(k')$ is an even, positive and non-singular function for $k' < \infty$ and thus can be expressed by the following Taylor expansion

$$G(k') = \sum_0 g_n k'^{2n} = \sum_0 g_n |\mathbf{k}'|^{2n}, \quad (\text{A.4})$$

where $g_n \equiv (n!)^{-1} \partial^n G / \partial \eta^n |_{\eta=0}$, $\partial^n G / \partial \eta^n |_{\eta=0}$ is the n -th order derivative of $G = G(k') = G(\eta^{1/2})$ with respect to $\eta = k'^2 = |\mathbf{k}'|^2$ evaluated at $\eta = 0$, and \sum_0 is as in Eq. (4.6). Applying the generalized inverse Fourier transformation term-by-term to (A.4) gives

$$Q(\mathbf{x}) = (2\pi)^{n_d/2} \sum_0 (-1)^n g_n |\mathbf{d}_n|^2 \delta(\mathbf{x}), \quad (\text{A.5})$$

where \mathbf{d}_n ($n = 1, 2, \dots$) are as in Eq. (4.7b) but with z replaced by x for $n_d = 1$, as in Eq. (4.17c) but with \mathbf{x}_h replaced by \mathbf{x}' for $n_d = 2$, and as in (5.10c) for $n_d = 3$. Substituting (A.5) into (4.1) and integrating by parts gives

$$J_b = (2\pi)^{n_d/2} \int d\mathbf{x} \sum_0 g_n |\mathbf{d}_n v(\mathbf{x})|^2. \quad (\text{A.6a})$$

This result can be written into the following compact D-operator form:

$$J_b = \int d\mathbf{x} |\mathbf{D}v(\mathbf{x})|^2, \quad (\text{A.6b})$$

where

$$\mathbf{D} \equiv (2\pi)^{n_d/4} \{1, g_1^{1/2} \mathbf{d}_1, g_2^{1/2} \mathbf{d}_2, \dots, g_n^{1/2} \mathbf{d}_n, \dots\}^T. \quad (\text{A.6c})$$

Note that the n -th component in Eq. (A.6c) becomes imaginary when g_n is negative, but the coefficients in Eq. (A.6a) are always real.

Denote by $\mathbf{D}^{(n)}$ the n -th order truncation of \mathbf{D} in Eq. (A.6c), and by $G^{(n)}$ the n -th order truncation of G in Eq. (A.4). The associated correlation spectrum is given by $S^{(n)} = [(2\pi)^{n_d} G^{(n)}]^{-1}$ according to Eq. (5.3). The truncation error is

$$\begin{aligned} S^{(n)} - S &= [(2\pi)^{n_d} G^{(n)}]^{-1} - S \\ &= (2\pi)^{-n_d} \left[\sum_{(0,n)} g_{n'} |\mathbf{k}'|^{2n'} \right]^{-1} - S(|\mathbf{k}'|). \end{aligned} \quad (\text{A.7})$$

The truncated correlation, denoted by $C^{(n)}(\mathbf{x})$, is the inverse Fourier transform of $S^{(n)}$. The associated truncation error is bounded by

$$\begin{aligned} &|C^{(n)}(|\mathbf{x}'|) - C(|\mathbf{x}'|)| \\ &= |(2\pi)^{-n_d/2} \int d\mathbf{k}' [S^{(n)}(|\mathbf{k}'|) - S(|\mathbf{k}'|)] \cos(\mathbf{k}' \cdot \mathbf{x}')| \\ &= \left| \int_0^\infty k'^{n_d-1} dk' [S^{(n)}(k') - S(k')] P(k'r') \right| \\ &\leq \int_0^\infty k'^{n_d-1} dk' [S^{(n)}(k') - S(k')] = C^{(n)}(0) - C(0) \\ &= \int_0^\infty k'^{n_d-1} dk' P(0) \left[\sum_{(0,n)} g_{n'} k'^{2n'} \right]^{-1} - 1 \\ &\equiv \varepsilon(n), \end{aligned} \quad (\text{A.8})$$

where $P(k'r')$ is as in Eq. (A.2). If the correlation has a Gaussian form as previously assumed, then $g_n = (2\pi)^{-n_d} (n!)^{-1} (L^2/2)^n$ and Eq. (A.8) reduces to Eq. (4.10) for $n_d = 1$, to Eq. (4.20) for $n_d = 2$, and to (5.13) for $n_d = 3$.

The truncation error in Eq. (A.8) depends on the specific form of the correlation. To estimate the error, it is necessary to calculate $g_n (n = 0, 1, \dots)$. If an analytical form is available for $S(k')$, then g_n can be derived from the n -th order derivative of Eq. (A.3) with respect to $\eta = k'^2 = |\mathbf{k}'|^2$ evaluated at $\eta = 0$. The results are

$$\begin{cases} g_0 = (2\pi)^{-n_d} / S_0, \\ g_1 = (2\pi)^{-n_d} [-S_1 / S_0^2], \\ g_2 = (2!)^{-1} (2\pi)^{-n_d} [-S_2 / S_0^2 + 2S_1^2 / S_0^3], \\ g_3 = (3!)^{-1} (2\pi)^{-n_d} [-S_3 / S_0^2 \\ \quad + 6S_2 S_1 / S_0^3 - 6S_1^3 / S_0^4], \\ \dots \end{cases} \quad (\text{A.9a})$$

where

$$S_n = \partial^n S / \partial \eta^n |_{\eta=0}. \quad (\text{A.9b})$$

If an analytical form is not available for $S(k')$, then S_n can be calculated in terms of the moments of $C(r')$ and then g_n can be obtained from Eq. (A.9a). The detailed formulations for this two-step approach are omitted here. There is also a simple one-step approach to calculate g_n as described below.

Substituting the Taylor expansion of $P(k'r')$ into Eq. (A.2a) and using the fact that $C(r')$ is an even function, we obtain

$$S(k') = \sum_0 (-1)^n s_n k'^{2n}, \quad (\text{A.10})$$

where the summation \sum_0 is over n from 0 to ∞ , and

$$\begin{cases} s_n = [(2n-1)!!(2n)!!]^{-1} \int_0^\infty dr' C(r') r'^{2n} \\ \quad \text{for } n_d = 1, \\ s_n = [(2n)!!(2n)!!]^{-1} \int_0^\infty dr' C(r') r'^{2n} \\ \quad \text{for } n_d = 2, \\ s_n = [(2n)!!(2n+1)!!]^{-1} \int_0^\infty dr' C(r') r'^{2n+2} \\ \quad \text{for } n_d = 3. \end{cases} \quad (\text{A.11})$$

Substituting Eq. (A.4) and Eq. (A.10) into Eq. (A.3) and collecting terms of the same powers of k' , we obtain

$$\begin{cases} s_0 g_0 = (2\pi)^{-n_d}, \\ s_0 g_1 - s_1 g_0 = 0, \\ s_0 g_2 - s_1 g_1 + s_2 g_0 = 0, \\ \dots \\ s_0 g_n - s_1 g_{n-1} + s_2 g_{n-2} + \dots + (-1)^n s_n g_0 = 0, \\ \dots \end{cases} \quad (\text{A.12})$$

so g_n can be calculated from s_n recursively (for $n = 0, 1, 2, \dots$). When the correlations are represented by a truncated spectral expansion (Hollingsworth and Lönnberg, 1986; Lönnberg and Hollingsworth, 1986; Xu et al., 2001c; Xu and Wei, 2001, 2002), $S(k')$ is discrete but $C(r')$ is given analytically by the spectral expansion. In this case, s_n can be obtained either from $C(r')$ by using Eq. (A.11) or by directly fitting a truncated form of Eq. (A.10) to the estimated discrete values of $S(k')$. Thus, truncated D-operators can always be obtained from Eq. (A.12) even when $S(k')$ and $C(r')$ are given in discrete forms.

Note that the differential operator $\sum_0 (-1)^n g_n |\mathbf{d}_n|^2$ defined by $Q(\mathbf{x})$ in Eq. (A.5) is elliptical because its characteristic function is positive-definite (Courant and Hilbert, 1962). By substituting Eq. (A.5) into Eq. (4.2), one can see that the Green's response function of this differential operator is just the correlation. This further explains the inverse relationship defined in Eq. (3.5) or Eq. (4.2).

REFERENCES

- Adams, R. A., 1975: *Sobolev Spaces*. Academic Press, 268pp.
 Bartello, P., and H. L. Mitchell, 1992: A continuous three-dimensional model of short-range forecast error covariances. *Tellus*, **44A**, 217–235.
 Bennett, A. F., 1992: *Inverse Method in Physical Oceanography*. Cambridge University Press, 346pp.

- Boer, G. J., 1983: Homogeneous and Isotropic Turbulence on the Sphere. *J. Atmos. Sci.*, **40**, 154–163.
- Cohn, S., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan*, **75**, 257–288.
- Cohn, S., A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, 1998: Assessing the effects of data selection with the DAO Physical-Space Statistical Analysis System. *Mon. Wea. Rev.*, **126**, 2913–2926.
- Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics*. Vol. II, Interscience Publ., 830pp.
- Courtier, P., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1807.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457pp.
- Daley, R., and E. Barker, 2001: NAVDAS: Formulation and Diagnostics. *Mon. Wea. Rev.*, **129**, 869–883.
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- Derber, J., and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221.
- Dévényi, D., and T. W. Schlatter, 1994: Statistical properties of three-hour prediction “errors” derived from the mesoscale analysis and prediction system. *Mon. Wea. Rev.*, **122**, 1263–1280.
- Dubinskij, J. A., 1986: *Sobolev Spaces of Infinite Order and Differential Equations*. Kluwer Academic Publishers, 161pp.
- Franke, R., 1999: Three-dimensional covariance functions for NOGAPS data. *Mon. Wea. Rev.*, **127**, 2293–2308.
- Franke, R., and E. Barker, 1999: Vertical correlation functions for temperature and relative humidity errors. *Mon. Wea. Rev.*, **128**, 3962–3981.
- Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translation, 242pp. (in Russian)
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15.
- Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136.
- Huang, X.-Y., 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.*, **128**, 2588–2600.
- Jazwinski, A. H., 1970: *Stochastic processes and filtering theory*. Academic Press, 376pp.
- Jenkins, G. M., and D. G. Watts, 1968: *Spectral Analysis and Its Applications*. Holden-Day, 525pp.
- Lighthill, M. J., 1958: *Introduction to Fourier Analysis and Generalized Functions*. Cambridge University Press, 79pp.
- Lönnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, **38A**, 137–161.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194.
- Lorenc, A. C., 1988: Optimal nonlinear objective analysis. *Quart. J. Roy. Meteor. Soc.*, **114**, 205–240.
- Lorenc, A. C., 1992: Interactive analysis using covariance functions and filters. *Quart. J. Roy. Meteor. Soc.*, **118**, 569–591.
- Panchev, S., 1971: *Random Functions and Turbulence*. Pergamon Press, 444pp.
- Parrish, D., and J. Derber, 1992: The National Meteorological Center’s spectral statistical analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Purser, R. J., 1986: Bayesian optimal analysis for meteorological data. *Variational Method in Geosciences*, Y. S. Sasaki, Ed., Elsevier, 167–172.
- Purser, R. J., and R. McQuigg, 1982: A successive correction analysis scheme using recursive numerical filters. Met O 11 Tech. Note No. 154, British Meteorological Office, UK, 17pp.
- Purser, R. J., W.-S Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535.
- Purser, R. J., W.-S Wu, D. F. Parrish, and N. M. Roberts, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially Inhomogeneous and Anisotropic General Covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.
- Sasaki, Y., 1958: An objective analysis based on the variational method. *J. Meteor. Soc. Japan*, **36**, 77–88.
- Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875–883.
- Tarantola, A., 1987: *Inverse Problem Theory: Methods for Data Fitting and Parameter Estimation*. Elsevier, 613pp.
- Thiebaux, H. J., 1976: Anisotropic correlation functions for objective analysis. *Mon. Wea. Rev.*, **104**, 994–1002.
- Thiebaux, H. J., H. L. Mitchell, and D. W. Shantz, 1986: Horizontal structure of hemispheric forecast error correlations for geopotential and temperature. *Mon. Wea. Rev.*, **114**, 1048–1066.
- Wahba, G., 1990: *Spline Models for Observation Data*. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, Society for Industrial and applied Mathematics, Philadelphia, 169pp.
- Wahba, G., and J. Wendelberger, 1980: Some new mathematical methods for variational objective analysis using splines and cross validation. *Mon. Wea. Rev.*, **108**, 1122–1143.
- Weaver, A., and P. Courtier, 2001: Correlation modeling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.*, **127**, 1815–1846.
- Xu, Q., and J. Gong, 2003: Background error covariance functions for Doppler radial-wind analysis. *Quart. J. Roy. Meteor. Soc.*, **129**, 1703–1720.

- Xu, Q., C. Qiu, and J. Yu, 1994: Adjoint-method retrievals of low-altitude wind fields from single-Doppler reflectivity measured during Phoenix II. *J. Atmos. Oceanic Technol.*, **11**, 275–288.
- Xu, Q., C. Qiu, H. D. Gu, and J. Yu, 1995: Simple adjoint retrievals of microburst winds from single-Doppler radar data. *Mon. Wea. Rev.*, **123**, 1822–1833.
- Xu, Q., H. Gu, and W. Gu, 2001a: A variational method for Doppler radar data assimilation. Preprints. *Fifth Symposium on Integrated Observing Systems*, 14–19 January 2001, Albuquerque, New Mexico, Amer. Meteor. Soc., 118–121.
- Xu, Q., H. Gu, and S. Yang, 2001b: Simple adjoint method for three-dimensional wind retrievals from single-Doppler radar. *Quart. J. Roy. Meteor. Soc.*, **127**, 1053–1067.
- Xu, Q., L. Wei, A. Van Tuyl, and E. H. Barker, 2001c: Estimation of three-dimensional error covariances. Part I: Analysis of height innovation vectors. *Mon. Wea. Rev.*, **129**, 2126–2135.
- Xu, Q., and L. Wei, 2001: Estimation of three-dimensional error covariances. Part II: Analysis of wind innovation vectors. *Mon. Wea. Rev.*, **129**, 2939–2954.
- Xu, Q., and L. Wei, 2002: Estimation of three-dimensional error covariances. Part III: Height-wind error correlation and related geostrophy. *Mon. Wea. Rev.*, **130**, 1052–1062.