Measuring information content from observations for data assimilations: utilities of spectral formulations demonstrated with radar observations

By QIN XU[1]* and LI WEI[2], [1]*NOAA/National Severe Storms Laboratory, Norman, Oklahoma, USA*;
[2]*Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma, USA*

ABSTRACT

Utilities of the spectral formulations for measuring information content from observations are explored and demonstrated with real radar data. It is shown that the spectral formulations can be used (i) to precisely compute the information contents from one-dimensional radar data uniformly distributed along the radar beam, (ii) to approximately estimate the information contents from two-dimensional radar observations non-uniformly distributed on the conical surface of radar scan and thus (iii) to estimate the information losses caused by super-observations generated by local averaging with a series of successively coarsened resolutions to find an optimally coarsened resolution for radar data compression with zero or near-zero minimal loss of information. The results obtained from the spectral formulations are verified against the results computed accurately but costly from the singular-value formulations. As the background and observation error power spectra can be derived analytically for the above utilities, the spectral formulations are computationally much more efficient and affordable than the singular-value formulations, even and especially when the background space and observation space are both extremely large and too large to be computed by the singular-value formulations.

1. Introduction

Remotely sensed observations, especially those scanned from radars, are often much denser than the model background and analysis grids used in data assimilations for numerical weather predictions. Excessive observation resolutions not only impose unnecessary computational burdens on a data analysis system but can also cause the analysis system to be ill conditioned. To reduce or eliminate observation resolution redundancy for data assimilation, it is necessary to compress densely distributed observations into fewer super-observations with minimum possible loss of information. Such observational data compression is called super-Obbing. Some general super-Obbing techniques were proposed by Purser et al. (2000) to minimize the information loss measured by the Shannon entropy difference with the observations batched locally around each super-observation. With these techniques, super-observations can be generated efficiently by weighted local averaging, but the information loss is minimized locally rather than globally. By converting the general formulations of relative entropy and Shannon entropy difference into

singular-value forms for measuring information content from observations, Xu (2007, referred to as X07 hereafter) examined the above issue without batching the observations locally. With the singular-value formulations, the information redundancy from densely distributed observations can be explicitly quantified in connection with the rank deficiency of the scaled observation operator, $\mathbf{M} \equiv \mathbf{R}^{-1/2}\mathbf{HB}^{1/2}$, and super-observations can be constructed with zero or minimized information loss globally. However, since the super-observations and their covariance are constructed in the truncated singular-vector space of $\mathbf{M}$, their connections to the original observations are non-local and non-explicit in the physical space, and the singular-value decomposition (SVD) of $\mathbf{M}$ will become too expensive or impractical to compute if the background space and observation space are both very large, as often seen in radar data assimilation (Gao et al., 2004; Xu et al., 2010).

To overcome the above shortcomings, Xu (2011, referred to as X11 hereafter) transformed the singular-value formulations into spectral forms in the wavenumber space for situations in which the observations are locally uniform and the observation and background error covariances or correlations (i.e. the covariances scaled by their respective variances) are locally homogeneous in each direction in the analysis domain. The spectral formulations exhibited the following advantages over their counterpart singular-value formulations: (i) The information contents

*Corresponding author: National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, Oklahoma 73072-7326, USA.
e-mail: Qin.Xu@noaa.gov

from densely distributed observations can be calculated very efficiently and analysed easily for each wavenumber. (ii) Super-observations can be constructed by spectral truncations in the wavenumber space with zero or minimum loss of information and explicitly related to the original observations in the physical space. These advantages, however, have not been tested and demonstrated yet with real observations. As a follow-up of X11, this paper will design and perform these tests with radar observations and explore the utilities of the spectral formulations for radar data compression.

As shown and commented in section 4.1 of X11, constructing super-observations by a spectral truncation requires the observations to be precisely uniformly distributed, and this is a serious limitation for super-Obbing by spectral truncation. On the other hand, super-Obbing by local averaging or weighted local averaging can be used easily and adaptively for any densely distributed observations without the above limitation (see e.g. Purser et al., 2000; Liu et al., 2005; Alpert and Kumar 2007; Lu et al. 2011, section 3.3). According to the theoretical analysis based on the spectral formulations in section 4.3 and remarks in section 4.4 of X11, super-Obbing by local averaging will cause no loss of information (measured globally by the dispersion part of relative entropy or the Shannon entropy difference) for uniformly distributed observations with unbiased and spatially uncorrelated errors unless the super-observation resolution becomes coarser than the effective background resolution (Frehlich, 2008) or the background error covariance resolution (determined by one half of the wavelength associated with the cut-off wavenumber of the background error power spectrum). This is an attractive global property for super-Obbing by local averaging, but it is not clear whether and to what extent this global property can be retained for non-uniformly distributed observations. It is also not clear whether and how the information loss (if any) caused by local averaging for non-uniformly distributed observations can be estimated by the spectral formulations. These problems will be examined in this paper. In particular, we will explore how to apply the spectral formulations to densely but non-uniformly distributed radar observations as well as to their compressed super-observations by local averaging, so the information loss (if any) caused by local averaging can be estimated efficiently as a function of the successively coarsened super-observation resolution to find an optimally coarsened resolution for super-Obbing by local averaging that causes zero or near-zero minimal loss of information.

The paper is organized as follows. The next section will examine the utility and accuracy of the spectral formulations, in comparison with the singular-value formulations, for computing the information contents from one-dimensional radar observations uniformly distributed along the radar beam and the information losses caused by spectral truncation, uniform thinning and local averaging. Section 3 will explore and demonstrate the utility of the spectral formulations, in comparison with the singular-value formulations, for estimating the information contents from two-dimensional radar observations non-uniformly distributed on the conical surface of radar scan and the information losses caused by local averaging with a series of successively coarsened resolutions. Conclusions follow in Section 4.

2. Applications to one-dimensional radar data along the radar beam

2.1. Descriptions of the one-dimensional radar data and background field

The selected radar data are the same as those described in section 5.2 of X07, and they are the radial-component velocities (along the radar beam) scanned by the NSSL phased array radar from 2100 to 2200 UTC on 2 June 2004 with a resolution of 240 m in the radial direction and a resolution of 1.6° in the azimuthal direction. The data were processed through quality control and then thinned to the coarsened resolution of $\Delta x_o = 3$ km along the radar beam (see fig. 1 of X07). After the thinning, the observation errors become spatially uncorrelated (Xu et al., 2007) to facilitate the tests in Section 2.5 for super-Obbing by local averaging to cause zero or minimal information loss. The observation error variance is $\sigma_0^2 = 6.4$ m^2 s^{-2}, as estimated in Xu et al. (2005). The single beam of radial-velocity observations that was used for the illustrative examples in section 5 of X07 is used here again. This beam was scanned at 2108 UTC along 0.75° elevation angle and 97.8° azimuthal angle (positive clockwise from the north), and it contains 40 (thinned) observations (spaced every $\Delta x_o = 3$ km). The observation error covariance matrix is thus given by $\mathbf{R} = \sigma_0^2 \mathbf{I}_M$ with $M = 40$.

The background velocity field is produced by the Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS, Hodur, 1997) on a 6 km grid and then projected onto the aforementioned radar beam on a one-dimensional grid of $\Delta x = 6$ km resolution with $N = 20$, where N denotes the number of grid points. This one-dimensional grid is similar to that described in section 5.2 of X07, except that the resolution is 6 km instead of 5 km and thus the number of grid points is 20 instead of 24. In this case, the background radial-velocity error covariance still can be modelled approximately by the Gaussian function, that is, $\sigma^2 \exp[-x^2/(2L^2)]$ according to (2.7) and (3.2) of Xu and Gong (2003), where x is the distance between two correlated points along the radar beam, σ^2 and L denote the background radial-velocity error variance and de-correlation length scale, respectively. As in section 5.4 of X07, $\sigma^2 = 70$ m^2 s^{-2} and $L = 15$ km are selected here for the illustrative purpose. However, to use the spectral formation, the above one-dimensional domain of length $D = M\Delta x_o = N\Delta x$ (=120 km) needs to be extended periodically as described in section 2.2 of X11. In particular, the background error covariance function can be extended periodically into

$$B(x) = \sigma^2 \exp\left\{-(|x| - \text{Int}\,[|x|/D + 1/2]\,D)^2/(2L^2)\right\},$$

(1a)

where Int[ ] denotes the integer part of [ ]. This periodic co-variance function recovers exactly $B(x) = \sigma^2 \exp[-x^2/(2L^2)]$ for $|x| \leq D/2$ in the original domain, but it is continuous only to the zeroth order and thus non-smooth at $|x| = (1/2 + i)D$ for $i = 0$, $1, 2, \ldots$. For $D \gg L$, the background error covariance function can be also extended periodically into

$$B(x) = (\sigma^2/A_0) \sum_{i=-\infty}^{\infty} \exp\left[-(x - iD)^2/(2L^2)\right], \quad (1b)$$

where $A_0 = \sum_{i=-\infty}^{\infty} \exp[-(iD)^2/(2L^2)]$ to ensure $B(0) = \sigma^2$. For the above selected $D = 120$ km and $L = 15$ km, $A_0 = 1 + 2.5 \times 10^{-14} \approx 1$ and thus (1b) recovers $B(x) = \sigma^2 \exp[-x^2/(2L^2)]$ almost exactly for $|x| \leq D/2$ in the original domain. Since the periodic covariance function in (1b) is smooth to infinitely high order and its associated power spectrum has a simple analytical form [see (10)], (1b) will be used in this section as the periodically extended covariance function. The background error covariance matrix **B** can be constructed from $B(x)$ in (1b) with the nth element in the n'th column of **B** given by $B(n\Delta x - n'\Delta x)$ for n and n' in the range of $[N_-, N_+]$, where $N_- \equiv \text{Int}[(1 - N)/2] = -19$ and $N_+ \equiv \text{Int}[N/2] = 20$ for $N = 40$.

2.2. Applications of spectral formulations versus singular-value formulations

According to (8) of Xu et al. (2009) and (22)–(24) of X11, the spectral formations for the signal part of the relative entropy (denoted by Sg), the dispersion part of the relative entropy (denoted by Ds), the Shannon entropy difference (denoted by SD), and the degrees of freedom for signal (denoted by DFS), can be written into the following forms:

$$Sg = \sum_{i=\mu_-}^{\mu_+} Sg_i, \quad (2)$$

$$Ds = \sum_{i=\mu_-}^{\mu_+} Ds_i, \quad (3)$$

$$SD = \sum_{i=\mu_-}^{\mu_+} SD_i, \quad (4)$$

$$DFS = \sum_{i=\mu_-}^{\mu_+} DFS_i. \quad (5)$$

Here, $\mu_- \equiv \text{Int}[(1 - \mu)/2]$, $\mu_+ \equiv \text{Int}[\mu/2]$ and $\mu \equiv \min(N, M)$;

$$Sg_i = [(|c_i|^2/C_i)|\gamma_i|^2/(1 + |\gamma_i|^2)^2]/2$$
$$= \left[|c_i|^2/(2\beta\Pi_i)\right]|\gamma_i|^4/(1 + |\gamma_i|^2)^2, \quad (6)$$

$$Ds_i = SD_i - DFS_i/2, \quad (7)$$

$$SD_i = \ln(1 + |\gamma_i|^2)/2 \quad (8)$$

and

$$DFS_i = |\gamma_i|^2\left(1 + |\gamma_i|^2\right), \quad (9)$$

are the ith components of Sg, Ds, SD and DFS, respectively, in the spectral space; and c_i is the ith component of the normalized discrete Fourier transform (DFT) of the innovation d_m [see (10)–(11) of X11]; C_i is the observation error power spectrum [see (14) of X11]; Π_i is the collapsed background error power spectrum (into the range of the observation error spectrum) if $N > M$ but becomes identical to the background error power spectrum, denoted by S_i, if $N \leq M$ [see (13) and (25) of X11]; $|\gamma_i|^2 = \beta\Pi_i/C_i$ and $\beta = \Delta x/\Delta x_o$.

For the data described in Section 2.1, the number of observations ($M = 40$) is larger than the number of the background grid points ($N = 20$) and the observation errors are spatially uncorrelated, so $\Pi_i = S_i$, $C_i = \sigma_0^2$ and $|\gamma_i|^2 = \beta S_i/\sigma_0^2$. By using the fast Fourier transform (FFT), S_i and c_i can be computed efficiently. For $B(x)$ given in (1b), S_i can also be obtained analytically by using the following generalized Fourier Transformation:

$$S(k) = \int_{-\infty}^{\infty} dx B(x) \exp(-jkx)$$
$$= \left(\sigma^2/A_0\right) L(2\pi)^{1/2} \exp(-k^2 L^2/2) \sum_{-\infty}^{\infty} \exp(-jikD),$$
$$= (\sigma^2/A_0) L(2\pi)^{1/2} \exp(-k^2 L^2/2) \sum_{i=-\infty}^{\infty} \delta(k - k_i), \quad (10)$$

where $j \equiv (-1)^{1/2}$ is the imaginary unit, k is the wavenumber, $\delta(\bullet)$ denotes the delta function of $(\bullet)$, $\int_{-\infty}^{\infty} dx \exp[-(x - iD)^2/(2L^2)] = L(2\pi)^{1/2} \exp(-k^2 L^2/2) \exp(-jikD)$ is used in the second step, and the identity $\sum_{i=-\infty}^{\infty} \exp(-jikD) = (2\pi/D) \sum_{i=-\infty}^{\infty} \delta(k - 2\pi i/D)$ is used with $k_i = i\Delta k$ and $\Delta k \equiv 2\pi/D$ in the third step (see page 68 of Lighthill, 1958). Note that the Fourier integral that defines $S(k)$ in (10) can be discretized into $\sum_{n=N_-}^{N_+} B(x_n) \exp(-jk_i x_n)\Delta x = S_i \Delta x$ over the periodic domain D, where $S_i = S(k_i)$ is precisely the same discrete power spectrum as that defined in (12)–(13) of X11. On the other hand, the analytical form of $S(k)$ derived in (10) can be discretized directly into $\int_{(i-1/2)\Delta k}^{(i+1/2)\Delta k} dk S(k)$. According to (10), the above two discretized forms should be equal to each other, and this leads to the following analytical expression of S_i

$$S_i = \int_{(i-1/2)\Delta k}^{(i+1/2)\Delta k} dk S(k)/\Delta x$$
$$= \left(\sigma^2/A_0\right)(L/\Delta x)(2\pi)^{1/2} \exp\left(-k_i^2 L^2/2\right). \quad (11)$$

As shown above, S_i can be obtained either numerically by using the DFT in (12) of X11 [which is equivalent to the discrete cosine transform in (13) of in X11] or analytically by using (11). For the data described in Section 2.1, these two ways give almost identical S_i and thus almost identical $|\gamma_i| = (\beta S_i)^{1/2}/\sigma_o$. For $|i| \leq$

6, $|\gamma_i|$ is obtained identically at least to the 5th effective digit with the exact symmetry of $|\gamma_{-i}| = |\gamma_i|$, and the obtained values are $|\gamma_0| = 11.848$, $|\gamma_1| = 10.155$, $|\gamma_2| = 6.3936$, $|\gamma_3| = 2.9572$, $|\gamma_4| = 1.0048$, $|\gamma_5| = 0.25078$ and $|\gamma_6| = 0.045982$. For $|i| \geq 7$, $|\gamma_i|$ is also obtained with the exact symmetry of $|\gamma_{-i}| = |\gamma_i|$, but the values obtained by the two ways become slightly different, and the numerically (or analytical) obtained values are $|\gamma_7| = 6.1935 \times 10^{-3}$ (or 6.1934×10^{-3}), $|\gamma_8| = 6.1191 \times 10^{-4}$ (or 6.1281×10^{-4}), $|\gamma_9| = 5.1068 \times 10^{-5}$ (or 4.4542×10^{-5}), and $|\gamma_{10}| = 3.3613 \times 10^{-6}$ (or 2.3783×10^{-6}). Here, since $M = 40$ and $N = 20$, the wavenumber index i ranges from $\mu_- = -9$ to $\mu_+ = 10$ in (2)–(5), and the index range length is $\mu = \mu_+ - \mu_- + 1 = 20$. Due to the symmetry of $|\gamma_{-i}| = |\gamma_i|$, there are only 11 independent values of $|\gamma_i|$ as listed above for $0 \leq i \leq 10$.

In the singular-value formulations [see (3.6)–(3.10) of X07], the singular-value index i ranges from 1 to μ, so the index range length is also $\mu = 20$. Since $|\gamma_i| = (\beta S_i)^{1/2}/\sigma_o$ is a decreasing function of $|i|$ as shown in (11), $|\gamma_i|$ for $i > 0$ (or $i \leq 0$) in the spectral formulations is equivalent to λ_{2i} (or λ_{2i+1})—the $(2i)$th [or $(1 + 2i)$th] singular value in the singular-value formulations. This equivalence was explained in the conclusion of X11 and can be verified here numerically by comparing the above obtained $|\gamma_i|$ for $i > 0$ (or $i \leq 0$) with λ_{2i} (or λ_{2i+1}) computed from the SVD of the scaled observation operator $\mathbf{M} \equiv \mathbf{R}^{-1/2}\mathbf{HB}^{1/2}$ [see (3.6) of X07]. The computed $\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_{19}$ are found to be identical (at least to the 5th effective digit) to the above analytically derived $|\gamma_0|, |\gamma_1|, |\gamma_{-1}|, \ldots |\gamma_{-9}|$, respectively. Only the last and smallest singular value λ_{20} ($= 3.3635 \times 10^{-6}$) becomes different from the above obtained $|\gamma_{10}|$ mainly due to computer round-off errors. This well-verified equivalence is also shown in Fig. 1 by the two coincided curves, that is, (i) the grey dashed curve labelled 'SVD-Svi' for the SVD-computed λ_i/λ_1 (called the relative amplitude of λ_i), and (ii) the thin solid curve labelled 'Spc-Svi' for the spectral-derived $|\gamma_i/\gamma_0|$ (called the relative amplitude of $|\gamma_i|$) plotted in the wavenumber index sequence of $i = 0, 1, -1, \ldots, 7, -7$ (not shown) corresponding to the singular-value index sequence from $i = 1$ to 15 (as shown explicitly). According to the above comparisons, the spectral-derived S_i in (11) is sufficiently accurate and will be used for the applications considered in this section.

With S_i given by (11) and c_i given by the DFT of the innovation d_m, the signal part of the relative entropy Sg and its partitioned component, denoted by Sg_i for the ith wavenumber k_i, can be computed from (2) and (6), respectively. The computed Sg_i/Sg is plotted in Fig. 1 by the dark dotted curve labelled 'Spc-Sgi' in the wavenumber index sequence of $i = 0, 1, -1, \ldots 7, -7$ (not shown) corresponding to the singular-value index sequence from $i = 1$ to 15 (as shown explicitly). The relative amplitude Sg_i/Sg is also computed by using the singular-value formulation {see (5.4) of X07 but note that Sg_i was already normalized by Sg and there was a typo in the definition of Sg, that is, $[\lambda_i^2(1 + \lambda_i^2)^{-2}]^{-1}$ should be $\lambda_i^2(1 + \lambda_i^2)^{-2}$ inside the summation for the
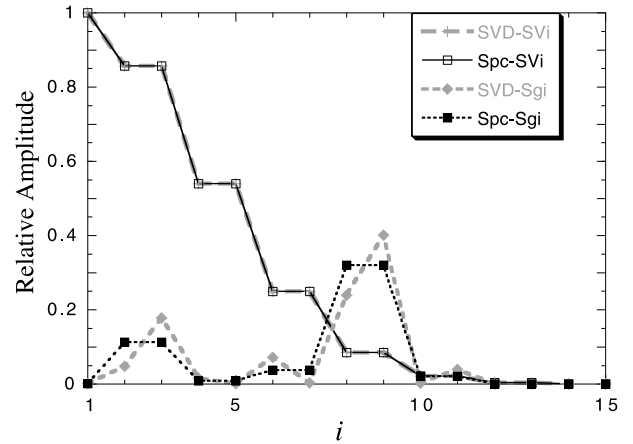

Fig. 1. Relative amplitudes of SVD-computed λ_i and Sg_i plotted as functions of the singular-value index i by the grey dashed curve (labelled SVD-SVi) and grey dotted curve (labelled SVD-Sgi), respectively. Relative amplitudes of Spectral-derived $|\gamma_i|$ and Sg_i plotted by the thin solid curve (labelled Spc-SVi) and dark dotted curve (labelled Spc-Sgi), respectively, in the wavenumber index sequence of $i = 0, 1, -1, \ldots, 7, -7$ (not shown) corresponding to the singular-value index sequence from $i = 1$ to 15 (as shown explicitly). When the singular-value index i increases beyond 15 toward its upper bound of $\mu \equiv \min(N, M) = 20$ (outside the plotted range), all the relative amplitudes become extremely and increasingly close to zero (not shown). The relative amplitudes of λ_i, $|\gamma_i|$ and Sg_i are defined by λ_i/λ_1, $|\gamma_i/\gamma_0|$ and Sg_i/Sg, respectively.

second equation in (5.4) of X07}, and the result is plotted by the grey dotted curve labelled 'SVD-Sgi' as a function of the singular-value index i in Fig. 1. As shown by the Spc-Sgi curve, the spectral-computed Sg_i satisfies the symmetry of $Sg_i = Sg_{-i}$ where i is the wavenumber index, but the SVD-Sgi curve does not. The difference is caused by the non-uniqueness of the SVD of $\mathbf{M}$. As shown in (7)–(9) and (20)–(21) of X11, the SVD presented by $\mathbf{M} = \mathbf{U}\Lambda\mathbf{V}^T$ in X07 can be generalized to a complex SVD (although $\mathbf{M}$ is always a real matrix), so the squared absolute value of the complex singular value can be given by $|\gamma_i|^2 = \beta S_i/\sigma_o^2$ for the ith wavenumber k_i and the complex counterpart of $\mathbf{U}$ can be given by the unitary matrix F_M^H, where $\mathbf{F}_M$ is the normalized DFT matrix in the observation space [see (11) of X11], and $()^H \equiv ()^{*T}$, $()^*$ and $()^T$ denote the Hermit transport, complex conjugate and real transpose of $()$, respectively. With this generalization, the transformed innovation vector $\mathbf{d}' = \mathbf{U}^T\mathbf{R}^{-1/2}\mathbf{d}$ in the singular-vector space [see (3.6) of X07] is replaced by its spectral representation $\mathbf{C}^{-1/2}\mathbf{c} = \mathbf{F}_M\mathbf{R}^{-1/2}\mathbf{d}$ [see (22) of X11]. Clearly, $\mathbf{F}_M$ is different from $\mathbf{U}^T$, and this explains the difference between the SVD-Sgi and Spc-Sgi curves in Fig. 1. As shown by the Spc-Sgi curve, the spectral-computed Sg_i/Sg and Sg_{-i}/Sg have the same value (due to the symmetry of $Sg_i = Sg_{-i}$) for the ith paired (positive and negative) wavenumbers and this value is the same as the averaged value of the two counterpart points
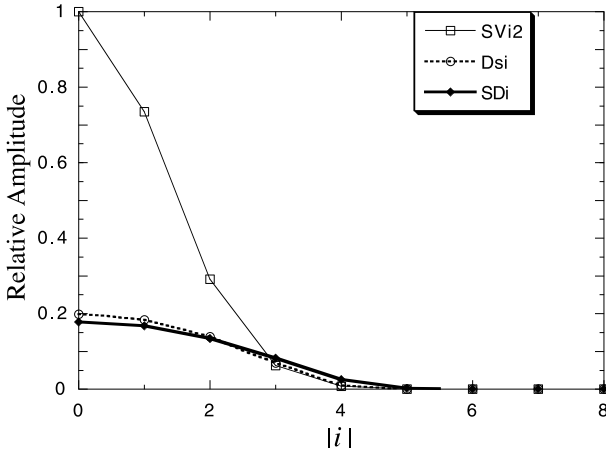
Fig. 2. Relative amplitudes of spectral-derived $|\gamma_i|^2$, Ds_i and SD_i plotted by the thin solid curve (labelled SVi2), dotted curve (labelled Dsi) and thick solid curve (labelled SDi), respectively, as functions of the absolute value of the wavenumber index over the range of $0 \leq |i| < 8$. The relative amplitudes of $|\gamma_i|^2$, Ds_i and SD_i are defined by $|\gamma_i/\gamma_0|^2$, Ds_i/Ds and SD_i/SD, respectively. The SVi2 curve is duplicated from the square of the Spc-SVi curve in Fig. 1 but plotted here as a function of the absolute value of the wavenumber index (instead of the singular-value index).

on the SVD-Sgi curve, so the two curves yield exactly the same summed value of Sg (=5.994949) in (2).

The dispersion part of the relative entropy Ds and its partitioned component Ds_i for k_i are formulated in (3) and (7), respectively. The Shannon entropy difference SD and its partitioned component SD_i for the ith wavenumber k_i are formulated in (4) and (8), respectively. The spectral-derived Ds_i/Ds (called the relative amplitude of Ds_i) and SD_i/SD (called the relative amplitude of SD_i) are plotted in Fig. 2 by the dark dotted and dark solid curves, respectively, as functions of the absolute value of the wavenumber index over the range of $0 \leq |i| < 8$. The Spc-SVi curve in Fig. 1 is squared and recast into the thin solid curve labelled 'SVi2' for the spectral-derived $|\gamma_i/\gamma_0|^2$ as a function of $|i|$, where i is the wavenumber index (instead of the singular-value index in Fig. 1). Note from (7)–(8) that Ds_i and SD_i are determined solely by $|\gamma_i|$, and $|\gamma_i|$ for $i > 0$ (or $i \leq 0$) is equivalent to λ_{2i} (or λ_{2i+1}) as shown by the coincided Spc-SVi and SVD-SVi curves in Fig. 1, so the spectral-derived Ds_i and SD_i for $i > 0$ (or $i \leq 0$) are equivalent to the SVD-computed Ds_{2i} (or Ds_{2i+1}) and SD_{2i} (or SD_{2i+1}), respectively. In view of the above explained and verified equivalences, we only need to present the results from the spectral formulations in the remaining part of this section.

Note that $|\gamma_i/\gamma_0|^2 \leq 1.5 \times 10^{-5}$, $Sg_i/Sg \leq 1.29 \times 10^{-4}$, $Ds_i/Ds \leq 1.1 \times 10^{-7}$ and $SD_i/SD \leq 7.6 \times 10^{-5}$ when $|i| \geq 6$ in Fig. 2 (or, equivalently, $i \geq 12$ in Fig. 1). Therefore, the upper limit of the wavenumber index i can be truncated from $\mu_+ = 10$ to 5 for the summations (over $i > 0$) in the spectral

formulations of Sg, Ds and SD [see (2)–(4)] with virtually no loss of information. This implies that the 40 observations can be compressed into 10 uniformly spaced super-observations with the observation resolution coarsened from $\Delta x_o = 3$ km to $\Delta x_s = 4\Delta x_o = 12$ km with no loss of information. This implication will be verified in Section 2.4.

2.3. Information loss caused by uniform thinning

According to the analysis in section 4.2 of X11, uniformly thinning the observations will always cause a loss of information and the information loss increases as the thinned observation resolution becomes increasingly coarse. The information losses for the signal and dispersion parts of the relative entropy and the Shannon entropy difference can be measured relative to $Sg(\Delta x_o)$, $Ds(\Delta x_o)$ and $SD(\Delta x_o)$, respectively, by

$$SIL = 1 - Sg(\Delta x_s)/Sg(\Delta x_0), \tag{12}$$

$$DIL = 1 - Ds(\Delta x_s)/Ds(\Delta x_0) \tag{13}$$

and

$$SDIL = 1 = SD(\Delta x_s)/SD(\Delta x_0). \tag{14}$$

Here, Δx_s ($> \Delta x_o$) denotes the thinned observation resolution; $Sg(\Delta x_o)$, $Ds(\Delta x_o)$ and $SD(\Delta x_o)$ denote the information contents from the original observations; $Sg(\Delta x_s)$, $Ds(\Delta x_s)$ and $SD(\Delta x_s)$ denote the information contents from the thinned observations. The relative information losses defined by SIL, DIL and SDIL in (12)–(14) may also be called information losses for short.

For the data described in Section 2.1, $Sg(\Delta x_o)$, $Ds(\Delta x_o)$ and $SD(\Delta x_o)$ are computed from the spectral formulations in (2)–(4) with $\mu_- = -9$ and $\mu_+ = 10$ for $\mu \equiv \min(N, M) = 20$. By reducing the number of observations from $M = 40$ to $M_s = 20$, 10, 8 and 4 and thus reducing β ($\equiv M/N = \Delta x/\Delta x_o$) = 2 to β_s ($\equiv M_s/N = \Delta x/\Delta x_s$) = 1, 1/2, 2/5 and 1/5 in (2)–(4), $Sg(\Delta x_s)$, $Ds(\Delta x_s)$ and $SD(\Delta x_s)$ are computed for four types of thinned observations with the observation resolution Δx_o coarsened to $\Delta x_s = 2\Delta x_o$, $4\Delta x_o$, $5\Delta x_o$ and $10\Delta x_o$, respectively. The resulting SIL, DIL and SDIL calculated from (12)–(14) for the thinned observations (with $\Delta x_s/\Delta x_o = 2, 4, 5$ and 10) are plotted in Fig. 3 by the thin solid curve labelled 'SIL-thin', light dotted curve labelled 'DIL-thin' and thin dashed curve labelled 'SDIL-thin', respectively. As shown, the information losses are significant and increase quasi-linearly with $\Delta x_s/\Delta x_o$. This exemplifies the analysis in section 4.2 of X11.

2.4. Super-Obbing by spectral truncation

According to the analysis in section 4.1 of X11, the original observations can be compressed into fewer super-observations by a proper spectral truncation with zero or minimum loss of information. This compression, called super-Obbing by spectral truncation, requires the observations to be uniformly distributed. The
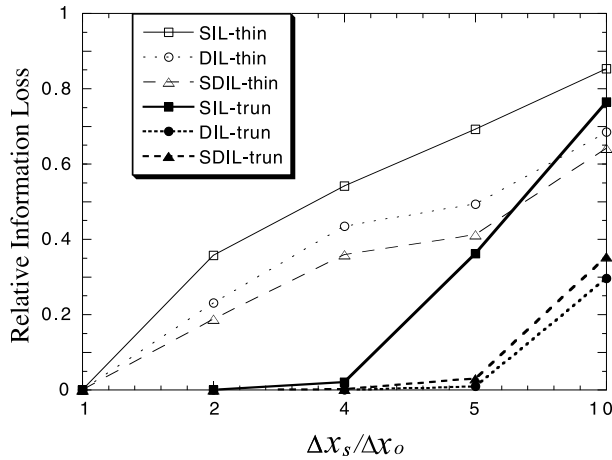
Fig. 3. Relative information losses SIL, DIL and SDIL caused by thinning plotted as functions of $\Delta x_s/\Delta x_o$ by the thin solid curve (labelled SIL-thin), light dotted curve (labelled DIL-thin") and thin dashed curve (labelled SDIL-thin), respectively; and relative information losses SIL, DIL and SDIL caused by spectral truncation plotted as functions of $\Delta x_s/\Delta x_o$ by the thick solid curve (labelled SIL-trun), dark dotted curve (labelled DIL-trun) and thick dashed curve (labelled SDIL-trun), respectively.

information losses caused by this super-Obbing can be measured by the same SIL, DIL and SDIL as those defined in (12)–(14) except that Δx_s ($>\Delta x_o$) represents now the super-observation resolution, and thus $Sg(\Delta x_s)$, $Ds(\Delta x_s)$ and $SD(\Delta x_s)$ denote the information contents from the super-observations. In this case, $Sg(\Delta x_o)$, $Ds(\Delta x_o)$ and $SD(\Delta x_o)$ remain the same as those in Section 2.3, but $Sg(\Delta x_s)$, $Ds(\Delta x_s)$ and $SD(\Delta x_s)$ are obtained for four types of super-observations with Δx_o coarsened to $\Delta x_s = 2\Delta x_o$, $4\Delta x_o$, $5\Delta x_o$ and $10\Delta x_o$ by truncating the observation spectral space dimension from $M = 40$ to $M_s = 20, 10, 8$ and 4, respectively, (without changing β) as explained in section 4.1 of X11.

The information losses SIL, DIL and SDIL caused by the above super-observations are computed and plotted as functions of $\Delta x_s/\Delta x_o$ (=2, 4, 5 and 10) in Fig. 3 by the thick solid curve labelled 'SIL-trun', dark dotted curve labelled 'DIL-trun' and thick dashed curve labelled 'SDIL-trun', respectively. As shown, the information losses are virtually zero for $\Delta x_s/\Delta x_o = 2$ (because $\Delta x_s = \Delta x$ in this case) and remain negligibly small (=0.02, 0.0001 and 0.002 for SIL, DIL and SDIL, respectively) as $\Delta x_s/\Delta x_o$ increases to 4. This verifies the implication stated at the end of Section 2.2. As $\Delta x_s/\Delta x_o$ further increases to 5, the dispersion-part information loss DIL (or SDIL) remains small and below 0.01 (or 0.03), but the signal-part information loss SIL becomes significant and increases to 0.36. Note that the signal-part information content depends on the sampled individual realization of the innovation and so does the signal-part information loss. Note also from in Fig. 1 that the Spc-Sgi curve reaches its peak value at the 4th paired wavenumbers k_4 and

k_{-4} (corresponding to the singular-value indices at $i = 8$ and 9, respectively, in Fig. 1). When the wavenumber index increases from $i = 0$ to 4, $|\gamma_i|$ decreases rapidly from 11.84 to 1.01 but $|\gamma_i|^4/(1 + |\gamma_i|^2)^2$ decreases slowly from 0.96 to 0.26. Since $Sg_i = [|c_i|^2/(2\beta S_i)]|\gamma_i|^4/(1 + |\gamma_i|^2)^2$ for $N \leq M$ as shown in (6), the peak of the Spc-Sgi curve for Sg_i/Sg can be largely explained by the peak of $|c_i|^2/(\beta S_i)$ at $|i| = 4$, while the latter peak is tied up with the smallness of S_i and local (secondary) peak of $|c_i|$ at $|i| = 4$ in the innovation spectrum (not shown). When $\Delta x_s/\Delta x_o$ increases to 5, the wavenumber k_4 is marginally retained but the wavenumber k_{-4} is truncated. This truncation is significant and it explains why SIL increases rapidly (from 0.02 to 0.36) as $\Delta x_s/\Delta x_o$ increases from 4 to 5 in Fig. 3.

Based on the equivalence between the spectral and generalized singular-value formulations explained and verified in Section 2.2, the above super-Obbing by spectral truncation can be viewed as a special and yet generalized form of super-Obbing by singular-value truncation (see section 4 of X07) for uniformly distributed observations. An advantage of this super-Obbing is that super-observations generated in the truncated spectral space can be easily transformed back and explicitly related to the original observations in the physical space. In particular, as shown in (41)–(42) of X11, super-observations can be expressed in the physical space by the original observations convoluted with a super-Obbing operator [see (40) of X11], and their error covariance can be given explicitly by the original observation error covariance convoluted with the super-Obbing operator. In this case, the original observations are required to be precisely uniformly distributed (in each direction) so that the spectral truncation can be accurately implemented for super-Obbing to avoid or minimize information loss. The required uniform distribution, however, can rarely be satisfied in real data assimilation, and this is a serious limitation for super-Obbing by spectral truncation, as commented at the end of section 4.1 of X11 and mentioned in the introduction of this paper.

2.5. Super-Obbing by local averaging

According to the analyses in sections 4.3 and 4.4 of X11, uniformly distributed observations with unbiased and spatially uncorrelated errors can be compressed into super-observations by local averaging with no loss of information until the super-observation resolution becomes coarser than the effective background resolution (Frehlich, 2008) or the background error covariance resolution (determined by π/k_c where k_c is the cutoff wavenumber of the background error power spectrum). In this case, the original observations are simply averaged over each Δx_s, where Δx_s is the super-observation resolution. To verify and illustrate the above property, four types of super-observations are generated by local averaging with Δx_s set to $\Delta x_s = 2\Delta x_o, 4\Delta x_o, 5\Delta x_o$ and $10\Delta x_o$, respectively. The information losses caused by super-Obbing can be measured by the same SIL, DIL and SDIL as defined in (12)–(14) except that $Sg(\Delta x_s)$,
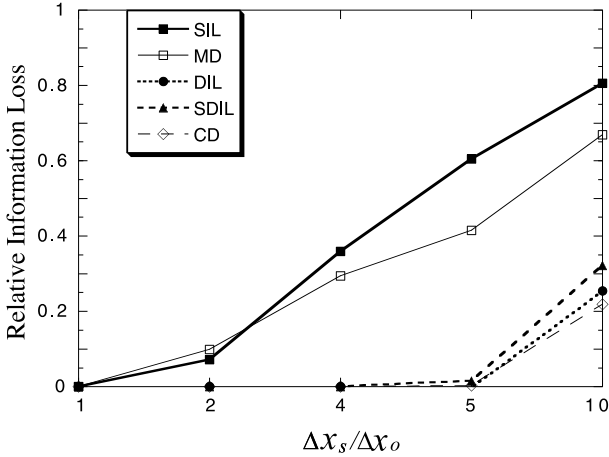

1020 Q. XU AND L. WEI

Fig. 4. Relative information losses SIL, DIL and SDIL caused by local averaging plotted as functions of $\Delta x_s/\Delta x_o$ by the dark solid, dark dashed and dark dotted curves, respectively; and relative degradations of the analysis mean and covariance, defined by MD/MD_0 and CD/CD_0, respectively, plotted as functions of $\Delta x_s/\Delta x_o$ by the thin solid curve (labelled MD) and thin dashed curve (labelled CD), respectively.

$Ds(\Delta x_s)$ and $SD(\Delta x_s)$ are now the information contents from super-observations generated by local averaging. The super-observation error variance is estimated by $\sigma_s^2 = \sigma_o^2 \Delta x_o/\Delta x_s$, where the sampling error in the original observations is assumed to remain constant and not yet start to increase with Δx_s [see (43) and remark (ii) in section 4.4 of X11].

The computed SIL, DIL and SDIL are plotted in Fig. 4 by the thick solid curve, dark dotted curve and thick dashed curve, respectively. As shown, DIL and SDIL are zero for $\Delta x_s/\Delta x_o = 2$ and remain negligibly small as $\Delta x_s/\Delta x_o$ increases to 5, but SIL is non-zero (0.07) for $\Delta x_s/\Delta x_o = 2$ and increases rapidly as $\Delta x_s/\Delta x_o$ increases beyond 2. For the original observations, $Ds(\Delta x_o)$ [or $SD(\Delta x_o)$] is the summation of Ds_i (or SD_i) over the wavenumber index range of $-9 = \mu_- \leq i \leq \mu_+ = 10$, and Ds_i (or SD_i) is determined solely by $|\gamma_i|^2 = \beta S_i/\sigma_o^2$. For $\Delta x_s = 2\Delta x_o$ $(=\Delta x)$, the number of super-observations is $M_s = 20$ $(=N)$, so $Ds(\Delta x_s)$ [or $SD(\Delta x_s)$] is still the summation of Ds_i (or SD_i) over $-9 \leq i \leq 10$, and Ds_i (or SD_i) is still determined solely by $|\gamma_i|^2 = \beta_s S_i/\sigma_s^2 = \beta S_i/\sigma_0^2$, where $\beta_s \equiv \Delta x/\Delta x_s$ and $\sigma_s^2 = \sigma_0^2 \Delta x_o/\Delta x_s$ are used. Thus, as revealed by the analysis in section 4.3 of X11, $Ds(\Delta x_s)$ and $SD(\Delta x_s)$] remain the same as $Ds(\Delta x_o)$ and $SD(\Delta x_o)$, respectively, for $\Delta x_s = 2\Delta x_o = \Delta x$. This explains why DIL and SDIL are zero for $\Delta x_s/\Delta x_o = 2$ in Fig. 4.

When $\Delta x_s/\Delta x_o$ increases to 5, the number of super-observations reduces to $M_s = 8$ $(<N)$, so $Ds(\Delta x_s)$ [or $SD(\Delta x_s)$] is the summation of Ds_i (or SD_i) over $-3 = \mu_{s-} \leq i \leq \mu_{s+} = 4$, where $\mu_{s-} \equiv \text{Int}[(1 - \mu_s)/2]$, $\mu_{s+} \equiv \text{Int}[\mu_s/2]$ and $\mu_s \equiv \min(N, M_s)$. Note that Ds_i (or SD_i) is determined solely by $|\gamma_i|^2 = \beta_s \Pi_i/\sigma_s^2 = \beta \Pi_i/\sigma_0^2$, where Π_i is the collapsed background error power spectrum. According to (25) of X11, we have $\Pi_{-3} =$

$S_{-3} + S_5$, $\Pi_{-2} = S_{-2} + S_6$, $\Pi_{-1} = S_{-9} + S_{-1} + S_7$, $\Pi_0 = S_{-8} + S_0 + S_8$, $\Pi_1 = S_{-7} + S_1 + S_9$, $\Pi_2 = S_{-6} + S_2 + S_{10}$, $\Pi_3 = S_{-5} + S_3$, and $\Pi_4 = S_{-4} + S_4$ for $N = 20$ and $M_s = 8$ (with $\Delta x_s = 5\Delta x_o$). Note also that $|\gamma_i/\gamma_0|^2 = S_i/S_o$ for $M \geq N$, and S_i/S_o decreases rapidly to 7.2×10^{-3}, 4.5×10^{-4}, 1.5×10^{-5} and virtually zero as $|i|$ increases from 0 to 4, 5, 6 and beyond (as shown by the SVi2 curve in Fig. 2). Therefore, $S_i/S_o <$ or $\ll 4.5 \times 10^{-4}$ for $|i| > 4$, and we only need to consider the case of $|i| = 4$. Note that $\Pi_4 = S_{-4} + S_4 = 2S_4$ due to the symmetry of $S_{-i} = S_i$, so S_{-4} is collapsed into Π_4 and this doubles Ds_4 and SD_4. The original contribution of Ds_{-4} (or SD_{-4}) to $Ds(\Delta x_o)$ [or $SD(\Delta x_o)$] in the summation over $-9 \leq i \leq 10$ for the original observations is thus largely retained by the doubled Ds_4 (or SD_4) in the summation over $-3 \leq i \leq 4$ for super-observations generated with $\Delta x_s = 5\Delta x_o$. This explains why the dispersion-part information loss, measured by DIL (or SDIL), remains negligibly small as $\Delta x_s/\Delta x_o$ increases to 5 in Fig. 4.

The signal-part information loss, measured by SIL, depends mainly on the difference between the original innovation spectrum c_i and associated super-innovation spectrum, denoted by c_{si}. The local averaging acts like a sinc-function filter in the spectral space [see the discussion of (40) in X11], so $|c_{si}|$ is not exactly the same as $|c_i|$ even when $\Delta x_s = 2\Delta x_o = \Delta x$ [and the summation of Sg_i in (2) for the super-observations is still over the same range of $-9 \leq i \leq 10$ as for the original observations]. This explains why SIL is non-zero (0.07) for $\Delta x_s/\Delta x_o = 2$ as shown in Fig. 4. When $\Delta x_s/\Delta x_o$ increases to 4, 5 and 10, the number of super-observations reduces to $M_s = 10$, 8 and 4, respectively. In this case, the difference between $|c_{si}|$ and $|c_i|$ becomes increasingly large and the summation of Sg_i in (2) for the super-observations is confined to the successively narrowed ranges of $-4 \leq i \leq 5$, $-3 \leq i \leq 4$ and $-1 \leq i \leq 2$. This explains why SIL increases rapidly as $\Delta x_s/\Delta x_o$ increases to 4 and beyond as shown in Fig. 4.

As explained in section 5.1 of X07, the signal-part information loss measures the information loss caused by super-Obbing that degrades the analysis mean, and the analysis mean degradation (MD) can be quantified by

$$MD = |\mathbf{a} - \mathbf{a}_s|, \qquad (15)$$

where $\mathbf{a}$ and $\mathbf{a}_s$ are the state vectors of the analysis means obtained from the original observations and super-observations, respectively. On the other hand, the dispersion-part information loss measures the information loss caused by super-Obbing that degrades the analysis covariance and, the analysis covariance degradation (CD) can be quantified by

$$CD = ||\mathbf{A} - \mathbf{A}_s||_F, \qquad (16)$$

where $\mathbf{A}$ and $\mathbf{A}_s$ are the analysis covariance matrices obtained from the original observations and super-observations, respectively, and $||\mathbf{A} - \mathbf{A}_s||_F$ denotes the Frobenius norm of $\mathbf{A}_s - \mathbf{A}$ defined by the square root of the sum of the squared absolute values of all the elements in $\mathbf{A}_s - \mathbf{A}$ [see (2.2–4) of Golub and Van

Tellus 63A (2011), 5

Loan 1983]. If the observations are not used at all, then the analysis mean and covariance matrix reduce to the background mean and covariance matrix, denoted by **b** and **B**, respectively. In this case, the degradations of the analysis mean and covariance reach their maxima $MD_0 \equiv |\mathbf{a} - \mathbf{b}|$ and $CD_0 \equiv \|\mathbf{A} - \mathbf{B}\|_F$, respectively. The relative degradations of the analysis mean and covariance can be then measured by MD/MD_0 and CD/CD_0, respectively, and they are expected to increase with $\Delta x_s/\Delta x_o$ similarly to the relative information losses SIL and DIL (or SDIL), respectively. As shown in Fig. 4, these expected similarities are indeed verified by the closeness of the thin solid curve for MD/MD_0 to the thick solid curve for SIL and the closeness of the thin dashed curve for CD/CD_0 to the thick dotted curve for DIL (or thick dashed curve for SDIL).

3. Applications to two-dimensional radar data on the conical surface of radar scan

3.1. Descriptions of the two-dimensional radar data and background field

The two-dimensional radar data used in this section are the radial velocities on the conical surface at 0.75º elevation angle scanned by the same phased array radar around 2108 UTC as described in Section 2.1. As these original radar data are dense and have high resolutions (240 m in the radial direction and 1.6º in the azimuthal direction) as mentioned in Section 2.1, the number of observations is quite large (=10134) over the $(108 \times 60 \text{ km}^2)$ area covered by the analysis domain on the conical surface at 0.75º elevation angle. The original radar data are thinned to the coarsened resolutions of $\Delta r_o = 1.2$ km in the radial direction and $\Delta \theta_o = 3.2º$ in the azimuthal direction. After the thinning, the observation errors become essentially uncorrelated (Xu et al., 2007) to facilitate the tests in Section 3.3 for super-Obbing by local averaging to cause near-zero minimal information loss. The observation error variance is $\sigma_0^2 = 6.4 \text{ m}^2$ s^{-2}, as stated in Section 2.1. The analysis domain is a 108×60 km² rectangular area on the conical surface at 0.75º elevation angle, and it ranges from 30 to 138 km in x-direction and from –30 to 30 km in y-direction relative to the radar. The analysis grid consists of 18×10 grid boxes with a uniform resolution of $\Delta x = \Delta y = 6$ km. The thinned radial-velocity data within the analyses domain are plotted in Fig. 5a. These thinned data will be used as the input observations for the analysis in this section, and the number of the observations is $M_o = 1012$.

The background velocity field is produced by the COAMPS on the same 6 km grid as described in Section 2.1, and then projected onto the above two-dimensional analysis grid on the conical surface at 0.75º elevation angle. The projected background radial-velocity field is shown in Fig. 5b. According to (5.1) of Xu and Gong (2003), the background radial-velocity error covariance between two points $\mathbf{x}_2$ and $\mathbf{x}_1$ on this surface

can be modelled by

$$B(\mathbf{r}) = \cos(\Delta\beta)\sigma^2 \exp\left[-r^2/(2L^2)\right], \quad (17a)$$

where $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$, $r = |\mathbf{r}|$, $\Delta\beta$ is the angle between the two distance vectors represented by $\mathbf{x}_2$ and $\mathbf{x}_1$ in the radar-centred Cartesian coordinate system (see fig. 1 of Xu and Gong, 2003). Here, as in (1), σ^2 and L denote the background radial-velocity error variance and decorrelation length scale, respectively. Again, as in Section 2.1, $\sigma^2 = 70 \text{ m}^2 \text{ s}^{-2}$ and $L = 15$ km are selected for the illustrative purpose in this section.

3.2. Applications of spectral formulations and singular-value formulations

As shown in Fig. 5a, the input observations are not uniformly distributed and contain data-void areas. In this case, the signal part of the relative entropy cannot be easily or directly estimated by the two-dimensional spectral formulation in (29) of X11, because the signal part depends on the sampled observations and the observations must be properly mapped on a uniform grid in each direction in order to use the two-dimensional spectral formulation (which is an unexplored approach beyond this paper). However, the dispersion part of the relative entropy, Ds, and the Shannon entropy difference, SD, do not depend on the individual realization of the sampled set of observations although they depend on the observation distribution and error covariance. Thus, Ds and SD can be estimated approximately by the two-dimensional spectral formulations in (30)–(31) of X11, as long as the input observations in Fig. 5a can be treated approximately as uniformly distributed with an averaged resolution in each direction and the background error covariance function in (17a) can be simplified into the following homogeneous and isotropic form:

$$B(r) = \sigma^2 \exp\left[-r^2/(2L^2)\right]. \quad (17b)$$

The analysis domain in Fig. 5b can be extended periodically as described in section 2.2 of X11. The simplified background error covariance function in (17b) can be extended periodically in each direction, similar to that in (1b), into the following form:

$$B(x, y) = (\sigma^2/A_0) \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty}$$
$$\times \exp\{-[(x - iD_x)^2 + (y - jD_y)^2]/(2L^2)\}, \quad (17c)$$

where $A_0 = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \exp[-(i^2 D_x^2 + j^2 D_y^2)/(2L^2)]$ to ensure $B(0, 0) = \sigma^2$. For the analysis grid described in Section 3.1, we have $\Delta x = \Delta y = 6$ km, $N_x = 18$, $N_y = 10$, $D_x = 108$ km, $D_y = 60$ km and $A_0 = 1 + 6.7 \times 10^{-4}$ (with $L = 15$ km). The background error power spectrum can be derived from (17c) by using the generalized Fourier transformation in each direction similar to that in (10)–(11). The derived background
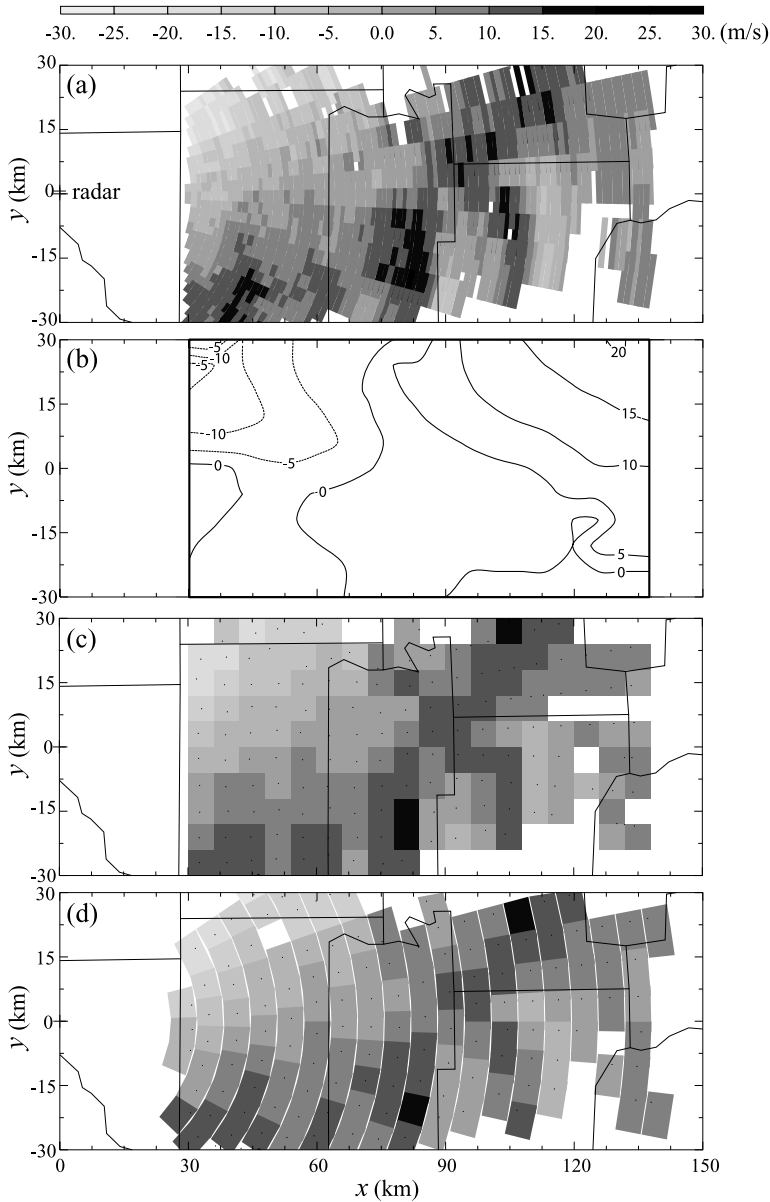
Fig. 5. (a) Input radial-velocity observations plotted by grey image pixels (with $\Delta r_o = 1.2$ km and $\Delta \theta_o = 3.2^\circ$) on the conical surface at 0.75° elevation angle scanned by the radar, (b) background radial-velocity field plotted by contours of every 5 m s^{-1}, (c) S1 super-observations plotted by grey image pixels (with $\Delta x_s = \Delta y_s = 6$ km), and (d) S2 super-observations plotted by grey image pixels [with $\Delta r_s = 6$ km and $\Delta \theta_s = (\Delta r_s / r)(180^\circ/\pi)$]. The greyscale of every 5 m s^{-1} is shown on the top. The (x, y)-coordinate system is centred at the radar site. The $+$ sign marks the radar site in panel (a) and the thin lines show the county boundaries in panels (a), (c) and (d). The dark rectangular boundary lines highlight the analysis domain in panel (b). The tiny black dot within each super-observation grey image pixel marks the location of the super-observation.

error power spectrum has the following discrete form:

$$S_{ij} = \left(\sigma^2 / A_0\right) 2\pi L^2 \left(\Delta x \Delta y\right)^{-1} \exp\left[-\left(k_i^2 + k_j^2\right) L^2 / 2\right]$$

$$\text{for } N_{x-} \leq i \leq N_{x+} \text{ and } N_{y-} \leq j \leq N_{y+}, \qquad (18)$$

where $k_i = i\Delta k_x$, $k_j = j\Delta k_y$, $\Delta k_x \equiv 2\pi/D_x$, $\Delta k_y \equiv 2\pi/D_y$, $D_x = N_x\Delta x$, $D_y = N_y\Delta y$, $N_{x-} \equiv \text{Int}[(1 - N_x)/2]$, $N_{x+} \equiv \text{Int}[N_x/2]$, $N_{y-} \equiv \text{Int}[(1 - N_y)/2]$ and $N_{y+} \equiv \text{Int}[N_y/2]$.

For the input observations in Fig. 5a, the averaged resolutions can be represented by $\Delta x_o = 1.2$ km in the x direction and $\Delta y_o = 5$ km in the y direction. With these averaged resolutions, the observation space can be approximated by a $M_x \times M_y = 90 \times 12$ uniform grid over the 108×60 km^2 rectangular area of the

above analysis domain. The observation error covariance is σ_0^2 multiplied by an identity matrix in the approximated observation space of $R^M = R^{Mx} \otimes R^{My}$ where $M = M_x \times M_y = 90 \times 12$. The observation error power spectrum is thus a constant in the approximated observation-spectral space and is given by

$$C_{ij} = C(i\Delta k_x, j\Delta k_y) = \sigma_0^2$$

$$\text{for } M_{x-} \leq i \leq M_{x+} \text{ and } M_{y-} \leq j \leq M_{y+}, \qquad (19)$$

where $M_{x-} \equiv \text{Int}[(1 - M_x)/2]$, $M_{x+} \equiv \text{Int}[M_x/2]$, $M_{y-} \equiv \text{Int}[(1 - M_y)/2]$ and $M_{y+} \equiv \text{Int}[M_y/2]$.

By substituting the derived background error power spectrum in (18) and approximated observation error power spectrum in

Table 1. Spectral-estimated and SVD-computed information contents from the input observations in Fig. 5a.

	Sg	Ds	SD
Spectral-estimated	N/A	57.0	75.2
SVD-computed with B(**r**) in (17a):	66.7	61.6	83.6
SVD-computed with B(r) in (17b):	70.2	58.4	78.7
SVD-computed with B(x, y) in (17c):	83.9	53.5	71.0

(19) into the spectral formulations in (30)–(32) of X11, we can estimate Ds and SD approximately for the dispersion part of the information content extracted from the input observations in Fig. 5a by the analysis with the background field in Fig. 5b. As listed in the first row of Table 1, the estimated values are $Ds = 57.0$ and $SD = 75.2$, while the observation space dimension is approximated by $M = M_x \times M_y = 90 \times 12 = 1080$. These spectral-estimated values are close to their counterpart values computed by the singular-value formulations, called SVD-computed values, as shown below.

With the singular-value formulations [see (3.6)–(3.10) of X07], the background error covariance matrix **B** can be constructed directly by discretizing $B(\mathbf{r})$ in (17a), $B(r)$ in (17b) or $B(x, y)$ in (17c) on the analysis grid, while the observation error covariance matrix **R** is simply given by $\sigma_0^2 \mathbf{I}_{Mo}$, where $\mathbf{I}_{Mo}$ is an identity matrix in the observation space R^{Mo} and M_o (=1012) is the number of the input observations in Fig. 5a. The innovation vector, denoted by $\mathbf{d} \in R^{Mo}$, in the observation space can be obtained by subtracting the background values in Fig. 5b from the observations in Fig. 5a. By substituting this **d** with the above constructed **B** and **R** into (3.6) of X07, we can compute Sg for the signal part of the information content from the input observations in Fig. 5a. As listed in the first column of Table 1, the computed value is $Sg = 66.7$ for **B** constructed from the original $B(\mathbf{r})$ in (17a), and the value is inflated slightly (by 5.2%) to 70.2 when **B** is constructed from the simplified $B(r)$ in (17b), and inflated further (by 19.5%) to 83.9 when **B** is constructed from the periodically extended $B(x, y)$ in (17c). By substituting the above **B** and **R** into (3.9)–(3.10) of X07, we can compute Ds and SD for the input observations in Fig. 5a. As listed in Table 1, the computed values are $Ds = 61.6$ and $SD = 83.6$ when **B** is constructed from the original $B(\mathbf{r})$ in (17a), and the values are deflated slightly (by 5.2% and 5.9%) to $Ds = 58.4$ and $SD = 78.7$, respectively, when **B** is constructed from $B(r)$ in (17b), and deflated further (by 2.4% and 9.8%) to $Ds = 53.5$ and $SD = 71.0$, respectively, when **B** is constructed from $B(x, y)$ in (17c).

Note that $\cos(\Delta\beta)$ (≤ 1) is a non-isotropic damping factor for $B(\mathbf{r})$ in (17a). When this factor is neglected by the simplified $B(r)$ in (17b), the spatial correlation is enhanced in $B(r)$ relative to that in $B(\mathbf{r})$. When $B(r)$ is extended periodically into $B(x, y)$ in (17c), the spatial correlation is further enhanced. Thus, when **B**

is constructed from $B(r)$ in place of $B(\mathbf{r})$ and then from $B(x, y)$ in place of $B(r)$, the off-diagonal elements of **B** are inflated consecutively but the diagonal elements still have the constant value of σ^2. In this case, the sum of λ_i^2 is not changed [because $\sum_{i=1}^{N} \lambda_i^2 = $ trace$(\mathbf{MM}^{\mathrm{T}}) = $ trace$(\mathbf{M}^{\mathrm{T}}\mathbf{M}) = $ trace$(\mathbf{B})/\sigma_0^2 = N\sigma^2/\sigma_0^2$ for $N = N_x \times N_y = 18 \times 10 < M_o = 1012$ and $\mathbf{R} = \sigma_0^2 \mathbf{I}_{Mo}$ where $\mathbf{H}^{\mathrm{H}}\mathbf{H} = \mathbf{I}_N$ is used], but λ_i^2 (that is, the eigenvalue of $\mathbf{M}^{\mathrm{T}}\mathbf{M} = \mathbf{B}/\sigma_0^2$) is a consecutively steepened decreasing function of i when **B** is constructed from $B(r)$ in place of $B(\mathbf{r})$ and then from $B(x, y)$ in place of $B(r)$. The incremental change of λ_i^2 caused by each consecutive steepening is thus a zero-sum function of i, and this function increases positively (or negatively) as i decreases toward 1 (or increases toward N). This zero-sum incremental change of λ_i^2 explains why the SVD-computed Ds and SD are deflated consecutively when **B** is constructed from $B(r)$ in place of $B(\mathbf{r})$ and then from $B(x, y)$ in place of $B(r)$, as shown in the last three rows of Table 1.

The ith component of Sg is computed by $Sg_i = d_i'^2/\lambda_i^2(1+\lambda_i^2)^{-2}/2$ where d_i' is the ith component of the transformed innovation vector $\mathbf{d}' = \mathbf{U}^{\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{d}$ in the singular-vector space [see (3.6) of X07]. For **d** computed from the observations in Fig. 5a and the background field in Fig. 5b, the transformed innovation vector $\mathbf{d}'$ has its large-amplitude (measured by $|d'_i|$) components mostly within the index range of $1 \leq i < N/2$, and this feature becomes increasingly distinctive when **B** is constructed from $B(r)$ in place of $B(\mathbf{r})$ and then from $B(x, y)$ in place of $B(r)$. This feature in combination with the aforementioned zero-sum incremental change of λ_i^2 explains why Sg is inflated consecutively when **B** is constructed from $B(r)$ in place of $B(\mathbf{r})$ and then from $B(x, y)$ in place of $B(r)$, as shown in the first column of Table 1.

The spectral-estimated Ds (or SD) in the first row of Table 1 can be verified against the SVD-computed Ds (or SD) in the last row of Table 1, because they both use the periodically extended background error covariance $B(x, y)$ in (17c). When the observation error power spectrum is derived in (19), the observations are assumed to cover the analysis domain uniformly with no data-void area. The observation space dimension is approximated by $M = M_x \times M_y = 90 \times 12 = 1080$ in the derivation of (19) which is slightly larger than that ($M_o = 1012$) for the input observations in Fig. 5a, while the observation error variance remains the same ($\sigma_0^2 = 6.4$ m^2 s^{-2}). This slightly increased number of observations (with the same σ_0^2) explains why the spectral-estimated Ds (or SD) in the first row is slightly larger than the SVD-computed Ds (or SD) in the last row of Table 1. Consequently, the spectral-estimated Ds (or SD) is very close to and only slightly smaller, by merely 4.4% (or 2.4%), than the SVD-computed Ds (or SD) with **B** constructed from $B(r)$ in the third row of Table 1. Therefore, $B(r)$ will be used as an intermediate approximation of the original $B(\mathbf{r})$, which is between $B(x, y)$ and $B(\mathbf{r})$, to compute the information losses caused by super-observations and verify the spectral-estimated information losses in the next subsection.

Table 2. Super-observation resolution Δl_s, its ratio $\Delta l_s/\Delta l_o$ to the input observation resolution ($\Delta l_o = 1.2$ km) and its associated super-observation space dimension M_s listed (in the first five rows) for each set (in each column) of S1 super-observations, S2 super-observations and spectral-approximated super-observations. The last two rows list the spectral-approximated super-observation resolutions (Δx_s, Δy_s).

Δl_s (km)	3	6	12	15	24	30
$\Delta l_s/\Delta l_o$	2.5	5	10	12.5	20	25
S1 M_s	410	153	44	30	15	8
S2 M_s	337	136	50	34	19	14
Spectral M_s	432	180	45	32	15	8
Spectral Δx_s (km)	3	6	12	13.5	21.6	27
Spectral Δy_s (km)	5	6	12	15	24	30

3.3. Super-Obbing by local averaging

In this subsection, two types of super-observations, called S1 and S2, are generated from the input observations in Fig. 5a by local averaging. Each S1 super-observation covers a square area of $\Delta x_s \times \Delta y_s$ with $\Delta x_s = \Delta y_s$ set to a prescribed resolution, denoted by Δl_s. Each S2 super-observation covers a sector area with the radial-range interval Δr_s set to a prescribed resolution Δl_s in the radar-centred polar coordinate system (on the conical surface of radar scan), while the azimuth span of the sector area is determined by $\Delta\theta_s = (\Delta r_s/r)(180^{\circ}/\pi)$ where r is the radial range distance from the radar to the far side of the sector area. Each super-observation of either type is computed by averaging the input observations locally in its covered area. The location of each super-observation is assigned to the centroid of the input data points in its covered local area. The error variance of each super-observation is estimated by $\sigma_s^2 = \sigma_0^2/n_s$, where n_s is the number of input observations in the super-observation covered local area and the input observation error variance σ_0^2 is assumed to be constant and not affected by the increase of Δx_s [see (43) and remark (ii) in section 4.4 of X11].

Six sets of S1 super-observations are generated with $\Delta x_s = \Delta y_s = \Delta l_s$ prescribed to 3, 6, 12, 15, 24 and 30 km, respectively, as listed in the first row of Table 2. The associated ratios with respect to the input observation resolution $\Delta x_o = \Delta l_o$ (=1.2 km) in the x-direction are $\Delta l_s/\Delta l_o = 2.5, 5, 10, 12.5, 20$ and 25, respectively, as listed in the second row of Table 2. These six sets contain $M_s = 410, 153, 44, 30, 15$ and 8 super-observations, respectively, as listed in the third row of Table 2, and an example is shown in Fig. 5c for the set obtained with $\Delta x_s = \Delta y_s = \Delta l_s = 6$ km. Six sets of S2 super-observations are generated with $\Delta r_s = \Delta l_s$ prescribed to 3, 6, 12, 15, 24 and 30 km, respectively, as listed again in the first row of Table 2. The associated ratios with respect to Δl_o are the same as those listed in the second row of Table 2. These six sets contain $M_s = 337, 136, 50, 34, 19$ and 14 super-observations, respectively, as listed in the fourth row of Table 2, and an example is shown in Fig. 5d for the set obtained with $\Delta r_s = \Delta l_s = 6$ km. As shown in Fig. 5c (or 5d),

super-observation areas do not overlap and their combined area covers all the input data points in Fig. 5a.

For each set of S1 or S2 super-observations listed in Table 2, Sg, Ds and SD can be computed by the singular-value formulations in the same way as described in the previous subsection for the input observations. The background error covariance matrix $\mathbf{B}$ is still constructed directly from either $B(\mathbf{r})$ in (17a) or $B(r)$ in (17b), while the super-observation error covariance matrix $\mathbf{R}$ is simply given by $\sigma_s^2 \mathbf{I}_{Ms}$ (since the super-observation errors are still spatially uncorrelated) where $\mathbf{I}_{Ms}$ is an identity matrix in the super-observation space R^{Ms} and M_s is the number of super-observations in the concerned set. The super-innovation vector, $\mathbf{d} \in R^{Ms}$, is obtained by subtracting the background values in Fig. 5b from the super-observations in the concerned set. By substituting the above $\mathbf{d}$, $\mathbf{B}$ and $\mathbf{R}$ into (3.6) and (3.9)–(3.10) of X07, Sg, Ds and SD are obtained for each set of S1 or S2 super-observations with $\mathbf{B}$ constructed from either $B(\mathbf{r})$ or $B(r)$.

The relative information losses caused by each set of super-observations with respect to the input observations in Fig. 5a can be measured and computed by SIL, DIL and SDIL in the same way as those in (12)–(14) except that $Sg(\Delta x_o)$, $Ds(\Delta x_o)$ and $SD(\Delta x_o)$ are replaced by $Sg(\Delta l_o)$, $Ds(\Delta l_o)$ and $SD(\Delta l_o)$, respectively, to denote the information contents from the input observations in Fig. 5a, and $Sg(\Delta x_s)$, $Ds(\Delta x_s)$ and $SD(\Delta x_s)$ are replaced by $Sg(\Delta l_s)$, $Ds(\Delta l_s)$ and $SD(\Delta l_s)$, respectively, to denote the information contents from each set of super-observations listed in Table 2. Here, Δl_o represents (Δr_o, $\Delta\theta_o$) with $\Delta r_o = \Delta l_o$ (=1.2 km) and $\Delta\theta_o = 3.2^{\circ}$ for the input observations in Fig. 5a, and Δl_s represents (Δx_s, Δy_s) with $\Delta x_s = \Delta y_s = \Delta l_s$ for S1 super-observations or represents (Δr_s, $\Delta\theta_s$) with $\Delta r_s = \Delta l_s$ and $\Delta\theta_s = (\Delta l_s/r)(180^{\circ}/\pi)$ for S2 super-observations.

The SVD-computed SIL and DIL for the six sets of S1 (or S2) super-observations with $\mathbf{B}$ constructed from $B(\mathbf{r})$ in (17a) are plotted in Fig. 6 as functions of $\Delta l_s/\Delta l_o$ by the dark long-dashed (or dark short-dashed) curve labelled SIL-S1a (or SIL-S2a) and dark dot-dashed (or dark dotted) curves labelled DIL-S1a (or DIL-S2a), respectively. The counterpart SIL and DIL computed with $\mathbf{B}$ constructed from $B(r)$ in (17b) are plotted in Fig. 6 by the grey long-dashed (or grey short-dashed) curve labelled SIL-S1b (or SIL-S2b) and grey dot-dashed (or grey dotted) curve labelled DIL-S1b (or DIL-S2b), respectively. The SVD-computed SDIL curves (not shown) are very similar to and only slightly higher than their counterpart DIL curves in Fig. 6. Note that $\Delta l_s/\Delta l_o = 1$ corresponds to the case of no super-Obbing and thus no information loss, so all the curves start from zero at $\Delta l_s/\Delta l_o = 1$ in Fig. 6. All the grey curves are very close to their counterpart dark curves, so the simplified $B(r)$ in (17b) can be indeed used as an approximation of the original $B(\mathbf{r})$ in (17a) to facilitate the use of the spectral formulation to estimate the information losses as shown below.

When the spectral formulations in (30)–(32) of X11 are used to estimate Ds (or SD) for each set of super-observations, the super-observation space is approximated by a $M_{xs} \times M_{ys}$
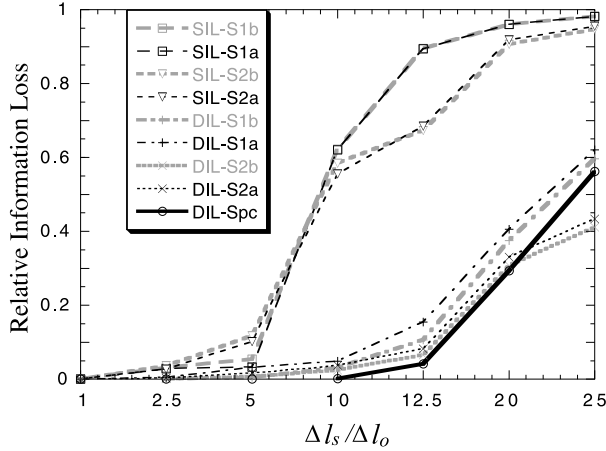
Fig. 6. SVD-computed SIL and DIL for S1 (or S2) super-observations with **B** constructed from $B(\mathbf{r})$ in (17a) plotted as functions of $\Delta l_s/\Delta l_o$ by the dark long-dashed (or dark short-dashed) curve labelled SIL-S1a (or SIL-S2a) and dark dot-dashed (or dark dotted) curves labelled DIL-S1a (or DIL-S2a), respectively; SVD-computed SIL and DIL for S1 (or S2) super-observations with **B** constructed from $B(r)$ in (17b) plotted as functions of $\Delta l_s/\Delta l_o$ by the grey long-dashed (or grey short-dashed) curve labelled SIL-S1b (or SIL-S2b) and grey dot-dashed (or grey dotted) curve labelled DIL-S1b (or DIL-S2b), respectively; and spectral-estimated DIL plotted as a function of $\Delta l_s/\Delta l_o$ by the dark solid curve labelled DIL-Spc. See the first two rows and last two row in Table 2 for the super-observation resolutions Δl_s and $(\Delta x_s, \Delta y_s)$ associated with $\Delta l_s/\Delta l_o$.

uniform grid over the $D_x \times D_y = 108 \times 60$ km^2 rectangular area of the analysis domain, where $M_{xs} = \text{Int}[D_x/\Delta x_s + 1/2]$ and $M_{ys} = \text{Int}[D_y/\Delta y_s + 1/2]$. For the six sets of S1 or S2 super-observations, the super-observation space dimensions are approximated by $M_s = M_{xs} \times M_{ys} = 36 \times 12 = 432$, $18 \times 10 = 180$, $9 \times 5 = 45$, $8 \times 4 = 32$, $5 \times 3 = 15$ and $4 \times 2 = 8$, respectively, as listed in the fifth row of Table 2. The approximated super-observation resolutions are $(\Delta x_s, \Delta y_s) = (3, 5)$, $(6, 6)$, $(12, 12)$, $(13.5, 15)$, $(21.6, 24)$ and $(27, 30)$ km, respectively, as listed in the last two rows of Table 2. In this case, the background error power spectrum is still given by (18), but the observation error power spectrum in (19) is replaced by the following super-observation error power spectrum

$$C_{ij} = C(i\Delta k_x, j\Delta k_y) = \sigma_s^2 = \sigma_s^2/n_s$$

$$\text{for } M_{xs-} \leq i \leq M_{xs+} \text{ and } M_{ys-} \leq j \leq M_{ys+}, \quad (20)$$

where $n_s = \Delta x_s \Delta y_s/(\Delta x_o \Delta y_o)$ is the averaged number of input observations in the local area covered by a super-observation, $M_{xs-} \equiv \text{Int}[(1 - M_{xs})/2]$, $M_{xs+} \equiv \text{Int}[M_{xs}/2]$, $M_{ys-} \equiv \text{Int}[(1 - M_{ys})/2]$ and $M_{ys+} \equiv \text{Int}[M_{ys}/2]$. By substituting (18) and (20) into (30)–(32) of X11, we can obtain the spectral-estimated D_s (or SD) for each set of super-observations and then estimate the associated information loss, DIL (or SDIL), relative to the spectral-estimated D_s (or SD) for the input observations (listed in the first row of Table 1). The spectral-estimated DIL is plotted

as functions of $\Delta l_s/\Delta l_o$ by the dark solid curve labelled DIL-Spc in Fig. 6, where Δl_o represents $(\Delta x_o, \Delta y_o)$ with $\Delta x_o = \Delta l_o = 1.2$ km and $\Delta y_o = 5$ km for the approximated observation resolutions used in the derivation of (19), and Δl_s represents $(\Delta x_s, \Delta y_s)$, with $\Delta x_s = \Delta l_s$ for $\Delta l_s \leq 12$ km and $\Delta y_s = \Delta l_s$ for $\Delta l_s > 12$ km, as shown by the approximated super-observation resolutions listed in the last two rows of Table 2. As shown in Fig. 6, the spectral-estimated DIL-Spc curve follows closely the SVD-computed DIL-S1b curve with **B** constructed from $B(r)$.

All the curves in Fig. 6 increase with $\Delta l_s/\Delta l_o$ monotonically, and the SVD-computed SIL curves increase more rapidly than the SVD-computed DIL curves especially when $\Delta l_s/\Delta l_o$ increases from 5 to 10 and beyond. Note that $\Delta x_o = 1.2$ km and $\Delta l_s/\Delta l_o = 5$ corresponds to $\Delta x_s = \Delta y_s = \Delta l_s = 6$ km (or $\Delta r_s = \Delta l_s = 6$ km) for the S1 (or S2) super-observations, so the super-observation resolution becomes the same as the background resolution ($\Delta x = \Delta y = 6$ km) when $\Delta l_s/\Delta l_o = 5$ in Fig. 6. For the one-dimensional observations in Section 2, $\Delta x_o = 3$ km and the super-observation resolution becomes the same as the background resolution ($\Delta x = 6$ km) when $\Delta x_s/\Delta x_o = 2$ in Fig. 4. Thus, the rapid increase of SIL after $\Delta l_s/\Delta l_o > 5$ in Fig. 6 is similar to the rapid increase of SIL (shown by the dark solid curve) after $\Delta x_s/\Delta x_o > 2$ in Fig. 4. This rapid increase of SIL can be explained by the reduced super-observation space dimension (below the background space dimension) plus the reduced amplitude of the transformed super-innovation vector caused by the local averaging (that acts like a sinc-function filter in the spectral space), which is similar to that explained for the one-dimensional super-observations in Section 2.5.

As shown in Fig. 6, the SVD-computed DIL curves increase slowly from zero to various levels below 0.05 when $\Delta l_s/\Delta l_o$ increases from 1 to 10. Note that $\Delta l_s/\Delta l_o = 10$ corresponds to $\Delta x_s = \Delta y_s = 12$ km (or $\Delta r_s = 12$ km) for the S1 (or S2) super-observations and this super-observation resolution ($=12$ km) is still finer than the background error de-correlation length scale $L = 15$ km. This explains why the SVD-computed DIL curves all remain very low (< 0.05) for $\Delta l_s/\Delta l_o \leq 10$ in Fig. 6. Note also that the S1 and S2 super-observations are not exactly uniformly distributed (because the location of each super-observation is assigned to the centroid of the input data points in its covered local area as described in Section 3.1 and shown in Figs. 5c and d), and they both contain data-void areas due to the incomplete coverage of the input data as shown in Fig. 5. This explains why the SVD-computed DIL curves increase slowly but notably above zero as $\Delta l_s/\Delta l_o$ increases from 1 to 10 in Fig. 6. On the other hand, the spectral-estimated DIL curve (dark solid labelled DIL-Spc) remains virtually zero for $\Delta l_s/\Delta l_o \leq 10$ in Fig. 6. Similarly, the dark dotted DIL curve computed for the uniform one-dimensional super-observations is also virtually zero for $\Delta x_s/\Delta x_o \leq 5$ as shown in Fig. 4, and this is simply because $\Delta x_s/\Delta x_o = 5$ in Fig. 4 and $\Delta l_s/\Delta l_o = 10$ in Fig. 6 both correspond to the same super-observation resolution of $\Delta x_s = 12$ km and thus the same ratio of $\Delta x_s/L$ ($=0.8$).

As listed in Table 2, the number of S1 super-observations is larger (or smaller) than that of S2 super-observations for $\Delta l_s/\Delta l_o \leq 5$ (or $\Delta l_s/\Delta l_o \geq 10$). This explains why the SIL-S1 curves are lower (or higher) than their respective SIL-S2 curves for $\Delta l_s/\Delta l_o \leq 5$ (or $\Delta l_s/\Delta l_o \geq 10$) as shown in Fig. 6. Note also that the number of S1 super-observations is substantially smaller than that of S2 super-observations for $\Delta l_s/\Delta l_o \geq 20$) as listed in Table 2. This explains why the DIL-S1 curves become significantly higher than their respective DIL-S2 curves when $\Delta l_s/\Delta l_o \geq 20$ as shown in Fig. 6. Also as listed in Table 2, the number of super-observations approximated by $M_s = M_{xs} \times M_{ys}$ for the super-observation error power spectrum in (20) is much closer to the number of S1 super-observations than the number of S2 super-observations and even becomes identical to the number of S1 super-observations when $\Delta l_s/\Delta l_o \geq 20$. This explains why the spectral-estimated DIL curve follows the SVD-computed DIL-S1b curve more closely than the DIL-S2b curve especially when $\Delta l_s/\Delta l_o \geq 20$, as shown in Fig. 6. A similar close comparison is seen between the spectral-estimated SDIL curve and the SVD-computed SDIL-S1b curve (not shown) for the S1 observations with $\mathbf{B}$ constructed from $B(r)$ in (17b). These close comparisons indicate that the spectral formulations can be used to estimate the information loss caused by local averaging for the purpose of finding an optimally coarsened resolution for radar radial-velocity data compression with zero or near-zero minimal information loss.

As mentioned at the end of the previous subsection, the simplified $B(r)$ in (17b) can be used as an intermediate approximation of the original background radial-velocity error covariance $B(\mathbf{r})$ in (17a) to compute the information losses caused by radial-velocity super-observations. This is indeed the case as shown in this subsection. In addition to this proxy use, $B(r)$ can be also used directly to model the pure-scalar error covariance without approximation for the background reflectivity field if σ^2 in (17b) is viewed as the background reflectivity error variance. In this case, the radial-velocity observations in Fig. 5a and background field in Fig. 5b can be viewed as reflectivity observations and background field, respectively. Then, the SVD-computed information losses with $\mathbf{B}$ constructed from $B(r)$ become accurate results (without approximation) for reflectivity super-observations, and their close comparisons with the spectral-estimated information losses (shown by the DIL-S1b and DIL-Spc curves in Fig. 6) ought to indicate that the spectral formulations can have the same (or even better) utility for radar reflectivity observations as demonstrated for the radar radial-velocity observations in this section.

4. Conclusions

In this paper, the spectral formulations derived in X11 for measuring information contents from uniformly distributed observations are applied to radar observations to test and demonstrate the advantages of the spectral formulations in comparison with the singular-value formulations derived in X07 and to explore the utilities of the spectral formulations for radar data compression. The main results are summarized below with conclusions:

(i) For uniformly distributed observations, the spectral formulations are shown to have precise correspondences to the singular-value formulations, as speculated in the conclusion of X11. As the SVD of the scaled observation operator $\mathbf{M}$ is generalized to complex domain in the spectral formulations, the absolute singular values can be computed very efficiently or even derived analytically in the spectral space [see Fig. 1 and (11)]. In this case, the spectral formulations are not only very accurate but also much more efficient than the singular-value formulations in computing the information contents. This advantage and related utilities of the spectral formulations for computing the information losses (if any) caused by uniform thinning, spectral truncation and local averaging are demonstrated (in Sections 2.2–2.5) with one-dimensional radar radial-velocity observations uniformly distributed along the radar beam.

(ii) For densely but non-uniformly distributed observations, the spectral formulations are shown to be able to estimate the dispersion part of the information content (measured globally by the dispersion part of relative entropy Ds or the Shannon entropy difference SD) with the observation error power spectrum constructed approximately based on the averaged observation resolution in each direction [see (19)]. The spectral formulations can also estimate the dispersion part of the information content from super-observations generated by local averaging with the super-observation error power spectrum constructed based on the super-observation resolution in each direction [see (20)]. These utilities of the spectral formulations are demonstrated with two-dimensional radar radial-velocity observations densely but non-uniformly distributed on the conical surface of radar scan (see Fig. 5a). The spectral-estimated information contents are closely verified by the SVD-computed counterpart values (see Table 1).

(iii) The background error power spectrum can be derived analytically [see (18)] in addition to the observation and super-observation error power spectra, so the information content (measured globally by Ds or SD) from the original observations and the information loss caused by their compressed super-observations through local averaging can be estimated very efficiently by the spectral formulations. The high-computational efficiency is an important advantage for the spectral formulations over the singular-value formulations, especially when the background and observation spaces are both extremely large and too large to be computed by the singular-value formulations, as often seen in radar data assimilation (Gao et al., 2004; Xu et al. 2010). Thus, regardless of the sizes of the observation and background spaces, the spectral formulations can always efficiently estimate the information loss caused by local averaging as a function of successively coarsened super-observation resolution, so an optimally coarsened super-observation resolution can be found for

radar data compression to minimize the information loss to an insignificant level. This utility is demonstrated with the aforementioned two-dimensional radar radial-velocity observations in section 3.3 (see Fig. 6).

In addition to the results summarized above, the spectral formulations can also estimate the dispersion part of information content from thinned non-uniform observations and the associated information loss (not shown), while the thinned observation error power spectrum can be constructed based on the thinned observation resolution in each direction similarly to that for the original observations [see (19)]. Moreover, as explained at the end of Section 3.3, the original non-isotropic background radial-velocity error covariance [in (17a)] is simplified into an isotropic form [in (17b)] to facilitate the use of the spectral formulations and such an isotropic form can be used directly to model the pure-scalar error covariance without approximation for the background reflectivity field (or any other pure-scalar background field). Thus, although the spectral formulations are tested only with the radial-velocity observations, they can be equally well or even better tested with radar reflectivity observations for similar utilities. The spectral formulations are also expected to have similar utilities for other densely and remotely sensed observations (such as those from GOES satellites).

As explained in Sections 2.1 and 3.1, the original radar observations are thinned to facilitate the tests in Sections 2.5 and 3.3 for super-Obbing by local averaging to cause zero or minimal information loss. In addition, the thinning also reduced the computational cost of the singular-value formulations drastically (especially if the number of observations exceeds the order of 10^3) for the two-dimensional case tested in Section 3. Ideally for real-data applications (without using the expensive singular-value formulations for verifications), the spectral formulations should be applied to the original radar observations directly without thinning to avoid thinning-caused information loss (see section 4.2 of X11 and Section 2.3 of this paper). In this case, we need to consider correlated radar observation errors between neighbouring gates or neighbouring beams (Xu et al., 2007) and use their associated non-constant observation error power spectrum [more complex than that in (19)] in the spectral formulations to estimate the information contents from the original observations and evaluate the information loss caused by local averaging. This problem will be examined in a follow-up study.

5. Acknowledgments

References

Alpert, J. and Kumar, K. 2007. Radial wind super-obs from the WSR-88D radars in the NCEP operational assimilation system. *Mon. Wea. Rev.* **135**, 1090–1109.

Frehlich, R. 2008. Atmospheric turbulence component of the innovation error covariance. *Quart. J. Roy. Meteor. Soc.* **134**, 931–940.

Gao, J., Xue, M., Brewster, K. and Droegemeier, K. K. 2004. A three-dimensional variational data assimilation method with recursive filter for single-Doppler radar. *J. Atmos. Oceanic. Technol.* **21**, 457–469.

Golub, G. H. and Van Loan, C. F. 1983. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 476 pp.

Hodur, R. M. 1997. The Naval Research Laboratory's coupled ocean/atmosphere mesoscale prediction system (COAMPS). *Mon. Wea. Rev.* **125**, 1414–1430.

Lighthill, M. J. 1958. *Introduction to Fourier Analysis and Generalized Functions*. Cambridge University Press, New York, 79 pp.

Liu, S., Xue, M., Gao, J. and Parrish, D. 2005. Analysis and impact of super-obbed Doppler radial velocity in the NCEP grid-point statistical interpolation (GSI) analysis system. In: *Proceedings of the 17th Conference on Numerical Weather Prediction*. American Meteorological Society, Washington DC, 13A.4.

Lu, H., Xu, Q., Yao, M. and Gao, S. 2011. Time-Expanded sampling for ensemble-based filters: assimilation experiments with real radar observations. *Adv. Atmos. Sci.* **28**, doi:10.1007/s00376-010-0021-4.

Purser, R. J., Parrish, D. F. and Masutani, M. 2000. Meteorological observational data compression; An alternative to conventional "super-Obbing". *Office Note 430*, National Centers for Environmental Prediction, Camp Springs, MD, 12 pp. Available at: http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html#2000 (last accessed 20 Apr 2011).

Xu, Q. 2007. Measuring information content from observations for data assimilation: relative entropy versus Shannon entropy difference. *Tellus* **59A**, 198–209.

Xu, Q. 2011. Measuring information content from observations for data assimilation: spectral formulations and their implications to observational data compression. *Tellus* **63A**, 793–804.

Xu, Q. and Gong, J. 2003. Background error covariance functions for Doppler radial-wind analysis. *Quart. J. Roy. Meteor. Soc.* **129**, 1703–1720.

Xu, Q., Nai, K. and Wei, L. 2007. An innovation method for estimating radar radial-velocity observation error and background wind error covariances. *Quart. J. Roy. Meteor. Soc.* **133**, 407–415.

Xu, Q., Nai, K., Wei, L., Zhang, P., Wang, L., and co-authors. 2005. Progress in Doppler radar data assimilation. In: *Proceedings of the 32nd Conference on Radar Meteorology*, American Meteorological Society., Albuquerque, New Mexico, JP1J7.

Xu, Q., Wei, L., Gu, W., Gong, J. and Zhao, Q. 2010. A 3.5-dimensional variational method for Doppler radar data assimilation and its application to phased-array radar observations. *Adv. Meteorol.* **2010**, 14, doi:10.1155/2010/797265.

Xu, Q., Wei, L. and Healy, S. 2009. Measuring information content from observations for data assimilations: connection between different measures and application to radar scan design. *Tellus* **61A**, 144–153.