applied optics

Analysis of weighted subspace fitting and subspace-based eigenvector techniques for frequency estimation for the coherent Doppler lidar

Yanwei Wu, Pan Guo,* Siying Chen, He Chen, Yinchao Zhang, and Xunbao Rui

School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China
**Corresponding author: guopan@bit.edu.cn*

Received 24 July 2017; revised 15 October 2017; accepted 22 October 2017; posted 23 October 2017 (Doc. ID 303171); published 15 November 2017

Since the periodogram maximum (PM) algorithm fails to provide consistent estimates, more robust techniques are developed, especially in a low signal-to-noise ratio (SNR) regime. The methods are formulated in a subspace fitting-based framework, such as the eigenvector (EV) method and the proposed weighted subspace fitting (WSF) method by introducing an optimal weighting matrix, which exploits the low-rank properties of the covariance matrix of the coherent Doppler lidar echo data. Simulation results reveal that the number of the reliable estimates by the WSF method is more than the other two methods, and the standard deviation is the smallest. Furthermore, the predicted best-fit Gaussian model for the probability density function of the estimates has a narrower spectral width than that of PM and EV methods. Experimental results also validate the simulation results, which show that the WSF approach outperforms the PM and EV algorithms in the furthest detectable range. The proposed method improves the detection range approximately up to 14.2% and 26.6% when compared to the EV method and the PM method, respectively. In conclusion, the proposed method can reduce the statistical uncertainties and enhance the accuracy in wind estimation specifically for a low SNR regime.

OCIS codes: (280.3640) Lidar; (280.1100) Aerosol detection; (280.1350) Backscattering; (070.4340) Nonlinear optical signal processing; (030.1640) Coherence.

https://doi.org/10.1364/AO.56.009268

1. INTRODUCTION

The analysis of lidar echo signal appears to be of importance in the measurement of wind velocity by means of a coherent Doppler lidar (CDL). The received signal is a noise-corrupted superposition of a signal backscattered by atmospheric aerosol particles [1]. The challenge of estimating the signal parameters, such as the mean frequency by locating the maximum power, has received much attention, especially in a low signal-to-noise ratio (SNR) environment. The radial velocity is derived from Doppler frequency estimated from the power spectral density (PSD).

For the CDL echo signals, numerous approaches have been developed to extract the corresponding Doppler frequencies. Conventionally, the periodogram maximum (PM) method and the pulse pair method are the most widely used [2,3], which may fail to give an accurate frequency estimate in the low SNR regime. Concerning this point, it is urgent to develop algorithms for a better performance. The linear prediction approach [4,5], which is used for real-time processing implementation, suffers from a limited resolution at finite SNR. Such improvements can be obtained by introducing the geometric approaches for analyzing the CDL signal, which marks the beginning of a wide interest in subspace or eigenstructure methods.

A previous study [6] illustrates that the subspace-based spectral estimation method, which is applied to the autocorrelation function, appears to be a good alternative to Fourier-like techniques when echoes are buried in a noisy environment. Meanwhile, model-based spectral estimation methods [6–8] have shown greater spectral resolution and improvement of SNR. The most famous algorithms exploring signal subspace decomposition, such as multiple signal classification (MUSIC) [9] and estimation of signal parameters by rotational invariance techniques (ESPRIT) [10], are categorized as a parametric line

power spectra estimation method [11]. The key idea of these methods consists of a decomposition of the observation space in two subspaces: the signal subspace containing the echoes, and the noise subspace. This decomposition is realized by computing the covariance matrix of the signal, which is then decomposed into its eigenvectors (EVs). The EV derived from the MUSIC method is used to obtain the PSD from the synthesized time-series signal for wind parameter estimation [12]. Although the EV method exhibits the advantage of spectral resolution, the spectrum produced by this method is sharply peaked at the signal frequency, and the retrieval of the actual spectrum width cannot be well observed in many cases, compared with standard Fourier estimates [13].

Subspace techniques are of great interest since, in general, no satisfactory solution exists for the maximum likelihood (ML) estimation procedure [14], which is known to be asymptotically efficient and achieves the Cramer–Rao bound (CRB) on the estimation error variance. Like the deterministic ML [14], as a member of the same general class of subspace fitting-based algorithms, subspace fitting formulation of the ML method is referred to as the weighted subspace fitting (WSF) algorithm, which is proposed to exploit the low-rank properties of the covariance matrix of the CDL data. The covariance matrix of the estimation error for the general subspace fitting method is then minimized in respect to the optimal weighting matrix. The rank properties are essential when determining optimal weighting matrices. It is shown that the WSF method gives lower estimation error variance than the ML method [15]. To our knowledge, this method is used only in single- or multi-tone recognition of the direction-of-arrival estimation in spatial filtering and array signal-processing environment, and the idea has not been adapted and studied for the CDL signal analysis, or tested on actual data. It identifies the frequency components of interest embedded in noisy observation time series by the pseudospectrum. The pseudospectrum does not represent the actual PSD, but simply reveals the locations (frequencies) of signals of interest.

In this paper, the covariance matrix model of the CDL data is presented by introducing an optimal weighting matrix, which is the basis of the proposed WSF algorithm. Next, simulation results show the potential of the WSF method by the model of probability density function (PDF) when that is compared with the PM and EV algorithms in various SNRs. Moreover, the interpretation of the experimental results of the WSF method is compared with those by the PM and EV methods. The conclusion on the interest of the WSF method is presented in the end.

2. SUBSPACE FITTING FRAMEWORK

In this section, the signal model is presented and its covariance matrix is decomposed into its EVs. The significant eigenvalues correspond to signal subspace EVs, and the other eigenvalues correspond to the noise subspace. Then the WSF method is derived. Finally, we focus on the optimal weighted covariance matrix and calculate the PSD, which is believed to be more effective for Doppler frequency estimation.

A. Signal Model

The N samples of the lidar data consist of the signals of interest and the background noise, which assumes that the noise signal is additive and uncorrelated with the clean signal. Consider the following noise-corrupted Doppler lidar signal [16]:

$$\boldsymbol{x}(t) = \boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{1}$$

where $t = 1, 2, \ldots N$, N denotes the pulse number, $\boldsymbol{s}(t)$ is the signals at time t, and $\boldsymbol{n}(t)$ is the additive noise. In practical applications, the exact ensemble covariance matrix $\boldsymbol{R}$ is unknown. Nonetheless, a series of samples from $\boldsymbol{x}(t)$ are given, and an estimate of the covariance matrix is used instead of the exact one. The estimate sample covariance matrix $\hat{\boldsymbol{R}}$ is computed as follows:

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{x}(k)\boldsymbol{x}^H(k), \tag{2}$$

where $(\bullet)^H$ denotes the Hermitian transpose. Now, perform an eigenvalue decomposition of $\hat{\boldsymbol{R}}$:

$$\begin{aligned}\hat{\boldsymbol{R}} &= \sum_{k=1}^{M} \lambda_k \hat{\boldsymbol{e}}_k \hat{\boldsymbol{e}}_k^H \\ &= \hat{\boldsymbol{E}}_s \hat{\boldsymbol{\Lambda}}_s \hat{\boldsymbol{E}}_s^H + \hat{\sigma}^2 \hat{\boldsymbol{E}}_n \hat{\boldsymbol{E}}_n^H \\ &= \hat{\boldsymbol{E}}_s (\hat{\Lambda}_s - \hat{\sigma}^2)\hat{\boldsymbol{E}}_s^H + \hat{\sigma}^2 \boldsymbol{I},\end{aligned} \tag{3}$$

where M is sampling number of the CDL data in each range bin. $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p = \lambda_{p+1} = \cdots = \lambda_M = \hat{\sigma}^2$. The diagonal matrix $\hat{\boldsymbol{\Lambda}}_s$ contains the p largest eigenvalues in descending order, which is the rank of the signal covariance. The columns of $\hat{\boldsymbol{E}}_s$ are the corresponding EVs. Similarly, $\hat{\boldsymbol{E}}_n$ contains the M-p EVs that correspond to the eigenvalue σ^2. $\hat{\boldsymbol{E}}_s = [e_1 e_2 \cdots e_p]$, which are usually called the signal EVs, and the range space of $\hat{\boldsymbol{E}}_s$ is called the signal subspace. Its orthogonal complement, $\hat{\boldsymbol{E}}_n = [e_{p+1} \cdots e_M]$, is usually referred to as the noise subspace, which is spanned by the columns of $\hat{\boldsymbol{E}}_n$. Here, $\hat{\boldsymbol{\Lambda}}_s = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_p]$, and $\hat{\boldsymbol{\Lambda}}_n = \mathrm{diag}[\lambda_{p+1}, \cdots, \lambda_M] = \hat{\sigma}^2 \boldsymbol{I}$.

Determination of the dimension of the signal subspace p is crucial for subspace methods. Methods that accomplish this detection are discussed in [17–20]. In this paper, we concentrate on the modified Gerschgorin Disk Estimator (GDE) criterion [21,22] for more accurate eigenvalues.

B. WSF Method

As observed in Section A, the p dimensional signal subspace is confined to an M dimensional subspace, which is corresponding to the true signal parameters f. If the signal covariance $\boldsymbol{S}$ has full rank, $p = M$. However, for the general case, $p \leq M$.

A natural estimation criterion is to find the best least squares fit of the subspace, and the basic subspace fitting problem $\hat{f}$ can equivalently be expressed as [14]

$$\hat{f} = \arg\max_{f} \mathrm{Tr}\{\boldsymbol{P}_A(f)\hat{\boldsymbol{E}}_s \boldsymbol{W} \hat{\boldsymbol{E}}_s^H\}, \tag{4}$$

where $\boldsymbol{P_A}$ is the projection matrix, $\boldsymbol{P}_A(f) = \boldsymbol{A}(f)(\boldsymbol{A}(f)^* \boldsymbol{A}(f))^{-1}\boldsymbol{A}(f)^*$, $\boldsymbol{A}(f) = [1 e^{-i2\pi f} \cdots e^{-i(M-1)2\pi f}]^T (M \times 1)$ is the discrete Fourier transform kernel vector at frequency f, and $\boldsymbol{W}$ is a $p \times p$ positive definite weighting matrix.

Herein, (4), with this optimal choice of $\boldsymbol{W}$, is referred to as the WSF method, which affects the asymptotic properties of the estimation error. In [13], it is observed through numerical examples that the WSF method is asymptotically efficient for Gaussian signal waveforms, such as the PSD of the CDL echo data.

C. Weighting Matrix

The eigendecomposition of the estimate of the noise-free covariance matrix $\hat{\hat{\boldsymbol{R}}}$, which improves detection performance, is defined in a similar fashion as

$$\hat{\hat{\boldsymbol{R}}} = \hat{\boldsymbol{E}}_s \boldsymbol{W} \hat{\boldsymbol{E}}_s^H. \tag{5}$$

For $\boldsymbol{W} = \boldsymbol{I}$, this corresponds to a least squares solution of the nonlinear modeling equations proposed in [23]. Moreover, if $\boldsymbol{W}$ is set to be $\hat{\boldsymbol{\Lambda}}_s$, then $\hat{\hat{\boldsymbol{R}}}$ is a reduced-rank approximation of covariance matrix. In [24], the reduced-rank approximation is used to improve frequency estimation accuracy. As proved in [15], the weighting matrix that gives the lowest asymptotic estimation error variance, $\boldsymbol{W} = \hat{\boldsymbol{\Lambda}}^2 \hat{\boldsymbol{\Lambda}}_s^{-1}$, where $\hat{\boldsymbol{\Lambda}} = \hat{\boldsymbol{\Lambda}}_s - \hat{\sigma}^2 \boldsymbol{I}$, and $\hat{\sigma}^2$ is any consistent estimate of the noise variance, which can be expressed as

$$\hat{\sigma}^2 = \frac{1}{M-p} \sum_{l=p+1}^{M} \lambda_k. \tag{6}$$

Below, the WSF scheme is summarized as:

(a) Compute the sample covariance matrix $\hat{\boldsymbol{R}}$;
(b) Perform eigendecomposition on $\hat{\boldsymbol{R}}$;
(c) Use the GDE criterion to estimate the rank of $\hat{\boldsymbol{R}}$, say p, then obtain $\hat{\boldsymbol{E}}_s$, $\hat{\sigma}^2$, and $\hat{\boldsymbol{\Lambda}}_s$.
(d) Compute the estimate of the noise-free covariance matrix by $\hat{\hat{\boldsymbol{R}}} = \hat{\boldsymbol{E}}_s \boldsymbol{W}_{\text{Opt}} \hat{\boldsymbol{E}}_s^H$, where $\boldsymbol{W}_{\text{Opt}} = \hat{\hat{\boldsymbol{\Lambda}}}^2 \hat{\boldsymbol{\Lambda}}_s^{-1}$;
(e) Use the GDE criterion to estimate the rank of $\hat{\hat{\boldsymbol{R}}}$, say p', then obtain the signal EVs $\hat{\hat{\boldsymbol{E}}}_s$, the noise variance $\hat{\hat{\sigma}}^2$, and the signal eigenvalues $\hat{\hat{\boldsymbol{\Lambda}}}_s$.
(f) Estimate the PSD to obtain the peak frequency:

$$\boldsymbol{P}(f) = \sum_{l=1}^{p'} \hat{\hat{\boldsymbol{\Lambda}}}_s(l) \left| \hat{\hat{\boldsymbol{E}}}_s(l) \boldsymbol{A}(f) \right|^2. \tag{7}$$

Fourier transform can be computed for each $\hat{\hat{\boldsymbol{E}}}_s(l)$, and then the squared magnitudes can be weighted by the inverse of each signal eigenvalue of $\hat{\hat{\boldsymbol{R}}}$. Next, the above result can be summed to obtain the PSD. Afterwards, the signal parameter $\hat{f}$ can be estimated by locating the maximum of the PSD.

3. SIMULATION

A. Numerical Simulation of Signal

In this paper, heterodyne-detected signals are expressed as the summation of incoherent backscattered signals from small atmospheric ranges and detector noise. In general, heterodyne detected signals with intermediate frequencies and complex time samples are generated using a balanced photodetector and an analog-to-digital-converter (ADC) card. Consequently, the signals in the time domain are expressed as [25]

$$S(m,n,i) = \sqrt{\text{SNR}} \cdot \sqrt{\frac{2\sqrt{\ln 2}\, T_S}{\sqrt{\pi}\Delta t}} \cdot \sum_{\tau=-P}^{P} \left\{ x(\tau,n,i) \exp\left(j \left[\frac{4\pi\nu}{\lambda} (m-1) T_s \right] \right) \right\} \times \exp\left(-2\ln 2 \frac{(m - M/2 + \tau)^2 T_S^2}{\Delta t^2} \right) + N_{\text{oise}}(m,n,i), \tag{8}$$

where m denotes the sample number in the time gate; M is the total sample number in the gate, which determines the range resolution; n is the accumulation number; i is the iteration number in the Monte Carlo simulation; τ is the small atmospheric range number; Δt is the full width at half maximum (FWHM) of the transmitting pulse; T_S is the sampling interval; λ is the laser wavelength; and SNR denotes the wideband SNR, which is the SNR of the average signal power to the average noise power. Noise denotes the complex amplitude of the white Gaussian detector noise having normalized average power of one, and $\boldsymbol{x}(\tau, n, i)$ denotes the complex amplitude of the signal from each range and the complex statistically independent zero mean Gaussian variable with normalized average power of one. This statistical property expresses the speckle effect of the signals. $v(\tau, i)$ denotes the LOS wind velocity of each small range, which corresponds to the assumed wind field. The number of atmospheric layers considered in the simulation is $2P + 1$, and this value should be considerably greater than M, assuming that the SNR is constant in the gate in this model. The parameters listed in Table 1 are used in the simulation, and the generated time-domain signal is shown in Fig. 1.

B. Performance of the EV and WSF Methods

Different algorithms are used to exploit subspace properties of the covariance matrix, such as the EV and WSF methods. For simulation, the degree of discrepancy between the two approaches has been revealed in the strong and low SNR regime. A number of simulations are performed when SNR = −10 dB, and it could be a good or bad estimate for each simulation. Four representative simulations are shown in Figs. 2(a), 2(b), 2(c), and 2(d), respectively, which are arranged in the order of increasing intensity of uncorrelated receiver noise in the PSDs. The peak numbers are bigger with the increasing noise intensity, and the signal will be buried in the noise gradually. Consequently, in general, the peak numbers increase as the SNR decreases, and the peak numbers are random for each simulation in the same SNR. However, wind parameters such

Table 1. Parameters of the Simulation

Parameter	Description	Value
λ	Wavelength	1.5 μm
Δt	Pulse width	500 ns
TS	Sampling interval	1.8 ns
M	Sample number	256
$2P + 1$	Number of layers	512
Vr	Radial wind velocity	0 m/s
SNR	Signal-to-noise ratio	−10 dB, −20 dB, −30 dB
f_{AOM}	Intermediate frequency	55 MHz

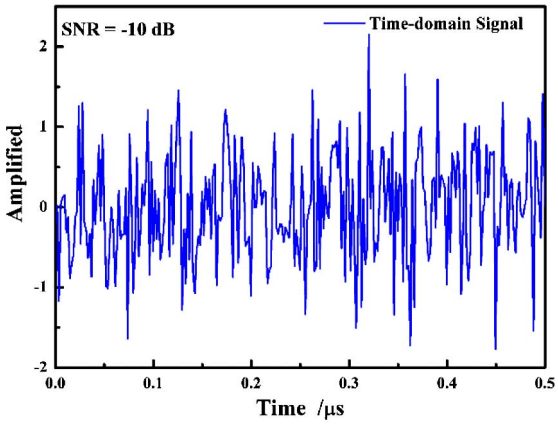

Fig. 1. Simulated time-domain signal of the CDL.

as Doppler shift (DS) and wind velocity are estimated by locating the maximum power peak while preventing ambiguity in the decision-making process. A subtle advantage of the WSF method shows up in the PSDs with the least disturbance compared with the other techniques. Consequently, in a strong SNR environment, the wind velocity estimators by PM, EV, and WSF methods are unbiased, and their performance becomes independent of the number of data samples that are being used for estimation.

In the regime of weak signal, however, one may need to implement a way of increasing the signal power before estimating the power spectra of lidar return. Often, an averaging process over multiple returns is utilized to improve the performance of wind measurement. When SNR = −20 dB, for these two cases, which are shown in Figs. 3(a) and 3(b), the WSF method provides an estimation of the mean velocity, which is consistent with the PM and EV spectrum. In Figs. 3(c) and 3(d), PM and EV methods perform poorly, while the WSF model maintains competitive.

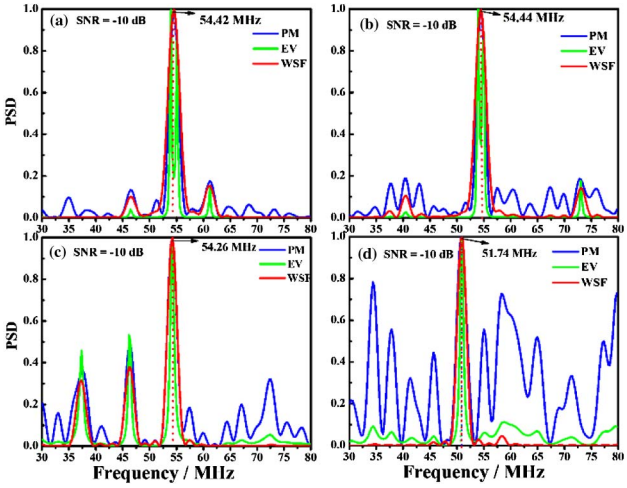

Fig. 2. Comparison of the PSD estimated by PM, EV, and WSF methods when SNR = −10 dB. Pulse number = 100, FFT size = 1024.

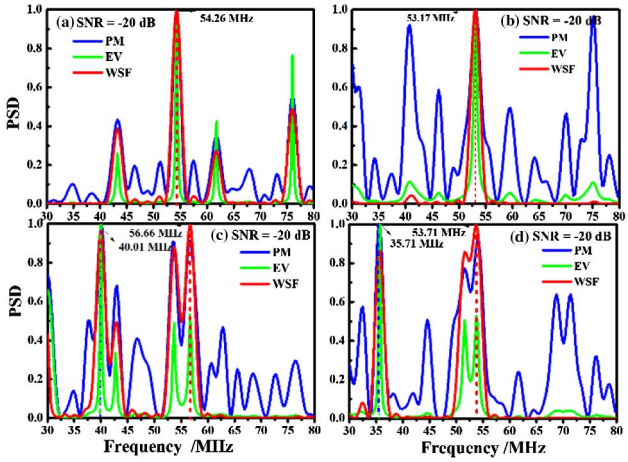

Fig. 3. Comparison of the PSDs estimated by PM, EV, and WSF methods when SNR = −20 dB. Pulse number = 1000, FFT size = 1024. (a) Estimated peak frequencies all are 54.26 MHz; (b) estimated peak frequencies all are 53.17 MHz; (c) peak frequencies estimated by the PM and EV methods are both 40.01 MHz, and 56.66 MHz by WSF method; and (d) peak frequencies estimated by the PM and EV methods are both 35.71 MHz, and 53.71 MHz by WSF method.

When SNR = −30 dB, robust extraction of the shift frequency is not possible due to the numerous peaks; the EV method exhibits several peaks inside that echo, and therefore a posterior analysis is necessary to select the maximum peak that is representative of the atmospheric contribution. As illustrated in Fig. 4(a), the DS estimated by the WSF method is reliable and consistent with the PM and EV methods. However, the EV method exhibits two significant peaks, one of which represents the "negative" frequency peak. In

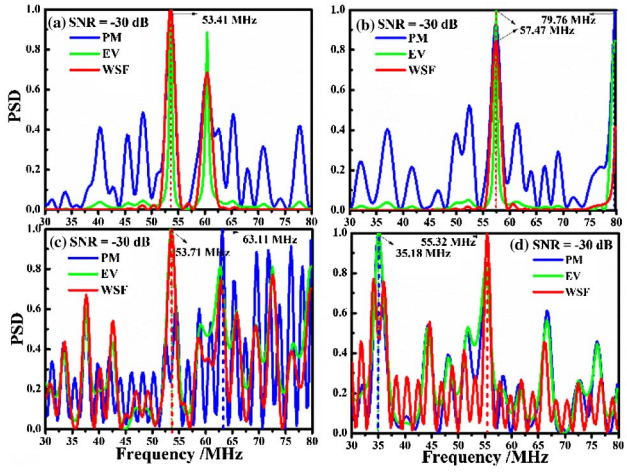

Fig. 4. Comparison of the PSDs estimated by PM, EV, and WSF methods when SNR = −30 dB. Pulse number = 1000, FFT size = 1024. (a) Estimated peak frequencies are 53.41 MHz; (b) peak frequencies estimated by EV and WSF methods are both 57.47 MHz, and 79.76 MHz by PM method; (c) peak frequencies estimated by the WSF and EV methods are both 53.71 MHz, and 63.11 MHz by PM method; and (d) peak frequencies estimated by the PM and EV methods are both 35.18 MHz, and 55.32 MHz by WSF method.

Figs. 4(b), 4(c), and 4(d), the DS estimated by the PM method corresponds to the "negative" peak, which is the erroneous DS estimate. However, correction of such an erroneous DS estimate can be made by the WSF method, which gives the reliable DS estimation.

C. Statistical Analysis

Statistical studies are carried out to show the performance of the PM, EV, and WSF methods in the reliability of wind velocity estimates for the CDL. The comparison of a histogram or PDF of 1000 estimates of wind velocity estimated by the three techniques is shown in Fig. 5. All three algorithms have similar performance when SNR = −10 dB. The estimates are well behaved and symmetrically distributed around the true value. The Gaussian fitting of the PDF indicated by the blue line approaches a Gaussian distribution. All the estimation methods have small error (<16%), but the WSF method has a small 11% negative bias. We use the mean, the median, and the standard deviation (SD) to adequately describe the PDF. The SD of the velocities estimated by WSF method is 4.23 m/s, which is the smallest among the three methods.

For the low SNR, there will be realizations in which the spectral feature of the signal is buried in the noise. An example of the histograms for 1000 estimates of the PM, EV, and WSF methods when SNR = −20 dB is shown in Fig. 6. The number of reliable estimators of the WSF method is the most among the three methods, and the spectral width of the Gaussian fitting of the PDF is becoming wider compared to that when SNR = −10 dB.

An example of the histograms for 1000 PM, EV, and the WSF estimates when SNR = −30 dB is displayed in Fig. 7. Both the green lines and the blue lines are the multi-peaks Gaussian fitting of the PDFs of the estimates. The predicted best-fit Gaussian model of the histogram for the PDFs of the estimates is indicated by the blue lines, and the Gaussian fitting of the "negative" estimates is indicated by the green lines. The peaks of the green lines are the sub-peaks, which are the erroneous peaks to estimate the velocity. The distributions of

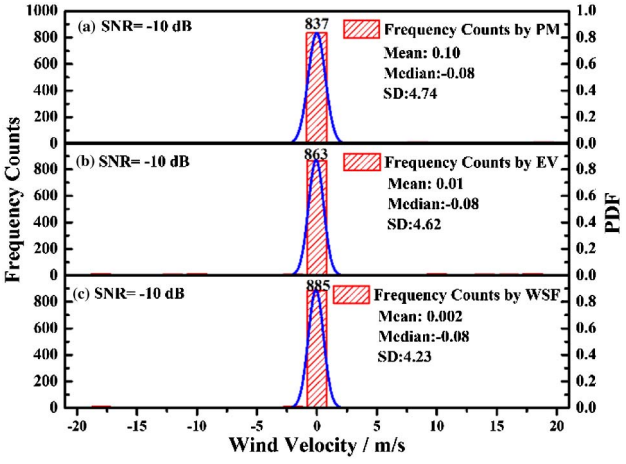

Fig. 5. Comparison of histograms from 1000 PM, EV, and WSF estimates of velocity when SNR = −10 dB. The best-fit Gaussian model for the PDF is indicated by the blue line. (a) PM; (b) EV; and (c) WSF.

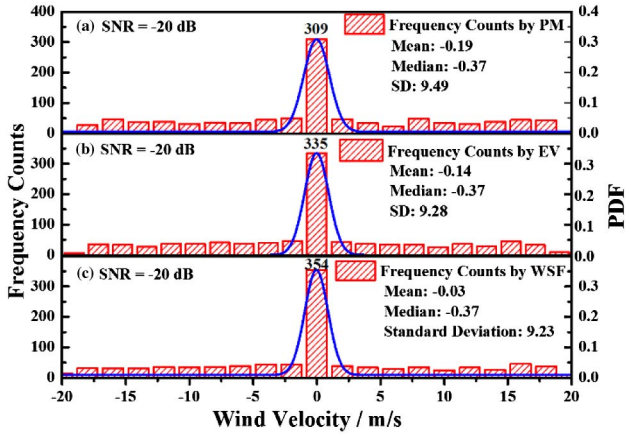

Fig. 6. Comparison of histograms from 1000 PM, EV, and WSF estimates of velocity when SNR = −20 dB. The best-fit Gaussian model for the PDF is indicated by the blue line. (a) PM; (b) EV; and (c) WSF.

1000 estimates of velocity when SNR = −30 dB are different from the PM, EV, and WSF methods. Consequently, the Gaussian fitting of the "negative" peaks for the PDFs of the 1000 estimates indicated by the green lines is different in the three methods. The predicted best-fit Gaussian model for the PDF of the wind velocities estimated by the WSF method has a narrower spectral width than those of PM and EV methods. Furthermore, the WSF method exhibits the least SD among the three methods.

D. Attention of the WSF Method

The key to estimate the number of signals based on eigenstructure is to determinate the boundary line between signal eigenvalues and noise eigenvalues [21]. The proposed GDE method achieves a better performance than Akaike information criterion (AIC) and minimum descriptive length (MDL) for estimating the number of signals. For the GDE approach,

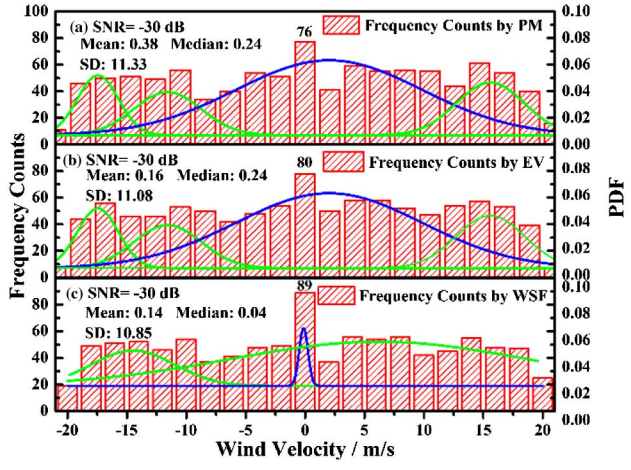

Fig. 7. Comparison of histograms from 1000 PM, EV, and WSF estimates of velocity when SNR = −30 dB. The best-fit Gaussian model for the PDF is indicated by the blue lines. The Gaussian fitting of the "negative" peaks for the PDF is indicated by the green lines. (a) PM; (b) EV; and (c) WSF.

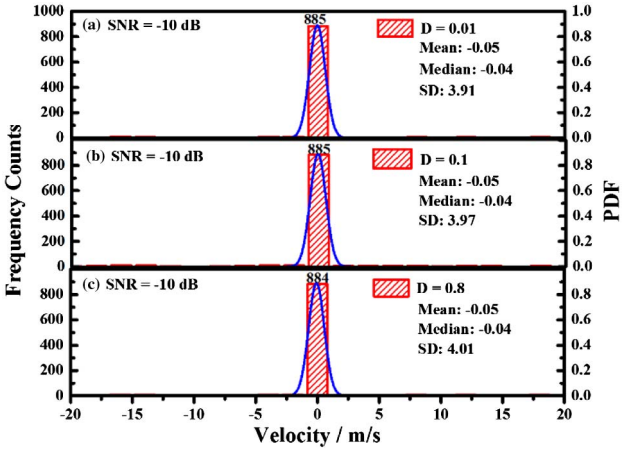

Fig. 8. Comparison of histograms from 1000 WSF estimates of velocity for different Ds when SNR = −10 dB. The best-fit Gaussian model for the PDF is indicated by the blue line. (a) $D = 0.01$; (b) $D = 0.1$; and (c) $D = 0.8$.

the adjustable factor D between 1 and 0, which becomes 0 while the pulse number becomes infinite, is used as the bound to decide the number of signal. To select D exactly is very important because an incorrect value would lead to an incorrect estimation. However, it is difficult to give an accurate value of D, because it is relative to the pulse number and sample number, but also to the SNR. The theoretic analysis indicates that it is better to select a little value. We can select more than one value of D to achieve the correct source number [21].

The comparison of different Ds is shown in Figs. 8 and 9. The adjustable factors are set to be 0.01, 0.1, and 0.8. The multi-peaks Gaussian fittings of the PDFs are performed at different Ds by means of peak-differentation-imitating analysis. Figure 8 shows that the value of D has little impact on the performance the WSF method when SNR = −10 dB. In Fig. 9, the best-fit Gaussian model for the PDFs is indicated by the blue lines, whose peaks are the principal peaks. The Gaussian fitting of the "negative" peaks is indicated by the green lines. Figure 9 shows that the number of the reliable estimates with $D = 0.01$ is the biggest, and the SD is the smallest.

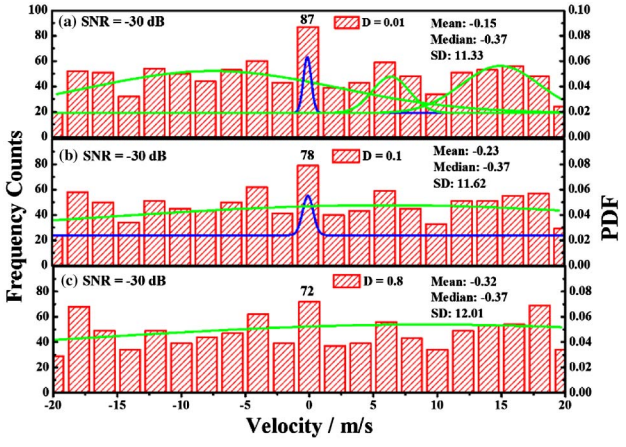

Fig. 9. Comparison of histograms from 1000 WSF estimates of wind velocity for different Ds when SNR = −30 dB. The best-fit Gaussian model for the PDF is indicated by the blue line. The green line is the Gaussian fitting of the "negative" peaks. (a) $D = 0.01$; (b) $D = 0.1$; and (c) $D = 0.8$.

The performance of velocity estimators is described by the velocity dispersion spectrum width. The spectrum width of the blue line with $D = 0.01$ is narrower than that with $D = 0.1$. Estimates with $D = 0.8$ are so poor that their PDF cannot be fitted by the Gaussian model. Therefore, the estimators for $D = 0.01$ exhibit the most reliable estimators, and the least SD though the "negative" peaks occurs when SNR = −30 dB.

4. EXPERIMENT RESULTS

A. Experimental System Description

The all-fiber CDL used in the experiment was located in the Shanghai Institute of Optics and Fine Mechanics. The system configuration of the CDL system is shown in Fig. 10 and consists of three components: laser transmitter, optical antenna, and signal detection and processing. All optical components are connected by a single-mode polarization-maintaining optical fiber. The laser transmitter consists of a laser seeder, acousto-optic modulator (AOM), and a fiber amplifier. The optical antenna, which consists of a polarization-insensitive optical circulator and a telescope, is separated from the laser transmitter with optical fiber. The amplified laser pulses are transmitted from port 1 to port 2 of the optical circulator. The backscattered signals are directed back through the same telescope optics, taking a path around the circulator and exiting through port 3. The seeder beam is split into two beams by a fiber coupler. One beam is directly modulated by AOM while the other is taken as the local oscillator (LO) of the CDL. Backscattered and LO signals are optically combined using an optical coupler and heterodyne-detected using a balanced photodetector. The electrical signals from the detector are acquired using an ADC card. Digital data are then streamed to a host computer for further processing. The main system parameters are listed in Table 2.

B. Implementation of the EV Method and the WSF Method

To reveal the performance of PM, EV, and WSF methods in inverting the wind profile, a detailed analysis was carried out on the data collected by the CDL on June 4, 2013, at various time intervals. The analysis includes mainly the performance of velocity estimators in obtaining the wind profiling up to the

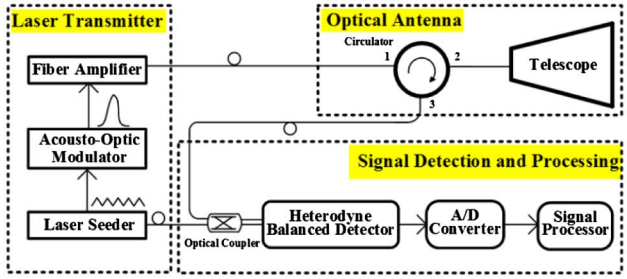

Fig. 10. Structure diagram of the CDL system.

Table 2. Specifications of the CDL

Parameter	Value
Wavelength	1.54 μm
Pulse energy	30 μJ (typical)
Pulse width	500 ns
Pulse repetition frequency	10 kHz
Intermediate frequency	55 MHz
Telescope aperture	100 mm
ADC sample rate	550 MHz
Sampling number for each bin/M	512

height of 5 km. The climate during the data acquisition was clear to cloudy, and the northwest wind ranged to levels 4–5 (wind speed 5.5–10.7 m/s). Each lidar return at $\langle 135, 0\rangle$ consists of 18,333 samples, which is segmented into 68 range bins with 50% overlap over adjacent range bins, where each range bin contains 512 samples. The first 1,024 samples are used to estimate the reference Doppler frequency. The last nominal number of samples in the wind data is assumed to be the noise, and its power spectrum is used to normalize the noise in the power spectrum of the entire wind data. In the data processing algorithm of the CDL, the power spectrum of each bin is adjusted for zero Doppler normalization and averaged over multiple returns. A total of 400 returns is used in this paper. Figure 11 shows the PSDs of the 10th, 20th, 30th, and 40th range bins using the PM, EV, and WSF methods. The estimated maximum value in the power is consistent in the three methods. As the range bin index increases, the SNR decreases, and the increasing noise gradually emerges from the signal. However, the PSD estimated by the WSF method exhibits the least noise fluctuations.

A comparison of PSDs estimated by the PM, EV, and WSF methods for the 53rd and 55th range bins of the CDL data is shown in Fig. 12, which is taken as an example to describe the

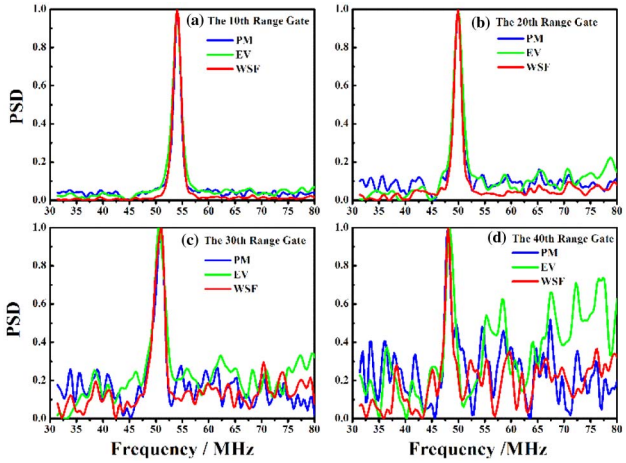

Fig. 11. PSDs of the four range bins at a look direction $\langle 135, 0\rangle$. The lidar returns were acquired on a sunny and clear day in June 2013. A total of 400 lidar returns was acquired, but only the wind data in the four range bins were used for illustration. The wind data were zero padded to the length of 2048 samples in PSD estimation via the PM, EV, and WSF methods. (a) 10th range bin; (b) 20th range bin; (c) 30th range bin; and (d) 40th range bin.

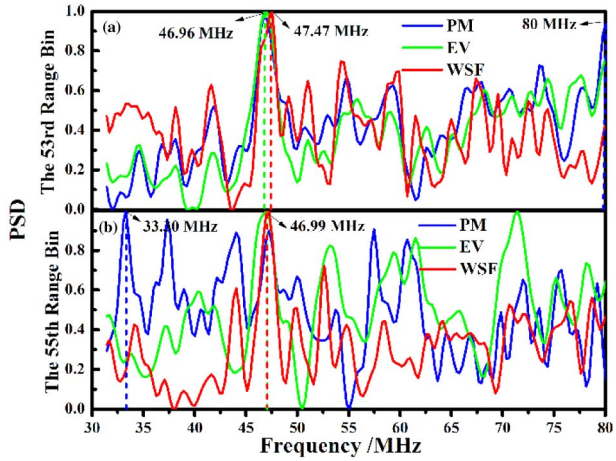

Fig. 12. Comparison of the PSDs for the 53rd and 55th range bins with the three methods. (a) 53rd range bin; and (b) 55th range bin.

performance of the three methods. In the 53rd range bin, the peak frequency of the PSD estimated by the WSF method is 47.47 MHz. The corresponding value of the EV method is 46.96 MHz, which is approximately the same as that of the WSF method, whereas that by the PM method is 80 MHz. A large deviation is yielded by the PM method compared with the EV and WSF methods. The same phenomenon may occur in the 55th range bin. Consequently, improper estimates return a congregated wrong-wind profile.

A vigorous demonstration of the behavior of the WSF method is illustrated in Fig. 13. A comparison of PSDs estimated by the PM, EV, and WSF methods for the 56th and, 64th range bins of the CDL data is taken as an example to describe the performance of the three methods in the weak SNR regime. In the 56th range bin, the peak frequency of the PSD estimated by the WSF method is 60.69 MHz, whereas those by the PM method and EV method are 76.80 MHz and 33.03 MHz, respectively. The same phenomenon occurs in the 64th range bin. A large deviation is yielded by the PM

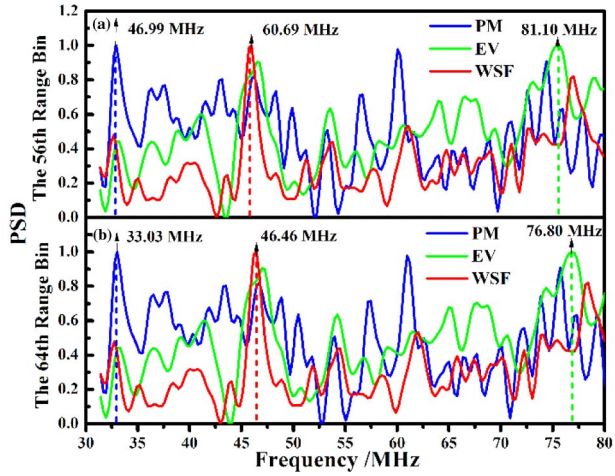

Fig. 13. Comparison of the PSDs for the 56th and 64th range bins with the three methods. (a) 56th range bin; and (b) 64th range bin.

and EV methods, which leads to an erroneous estimation, whereas the WSF method is opposite.

Panoramic displays of continuous observation of wind profiles are helpful in understanding wind trends and validating analyses. The results of wind profiling are compared in graphs to show the strength of each algorithm. Four groups of wind profiling of time-series data estimated by PM, EV, and WSF methods are illustrated in Fig. 14. An abrupt change in the graph indicates the fluctuating velocity, and such a change occurs more often, and more contrast appears as the detected range increases or as the SNR decreases. Good estimates will show gradual changes in either increasing or decreasing fashion [11]. In Fig. 14(a), the periodogram-based velocity begins to fluctuate starting about 3.5 km where the SNR is low, whereas that by the EV method fluctuates at 4 km, and by the WSF method maintains stable until 4.3 km. Such a steady velocity implies the better estimates, since the wind velocity tends to be decreasing gradually along the range. In Figs. 14(b), 14(c), and 14(d), wind profiling estimated by the three methods shows a similar performance. Therefore, the WSF method performs better because of its long-haul steady wind profile compared with the other two methods. The proposed method improves the detection range approximately up to 14.2% and 26.6% when compared to the EV method and the PM method, respectively. In order to reveal the performance of PM, EV, and WSF methods in processing the massive CDL data in continuous observation, the spatial and temporal distributions of wind velocity estimated by the EV and WSF methods will be compared with the PM method.

Figure 15 is a convincing demonstration of the behavior of the WSF method, which shows the spatial and temporal distributions of wind velocity for continuous signal acquisition in 20 min. As shown in Figs. 15(a), 15(b), and 15(c), the results from the WSF method are reliable in regions that suffer from an extremely low SNR. In regions in which the SNR is high, the WSF method returns results similar to the PM and EV methods. Gradually, changes are seen up to about 3 km,

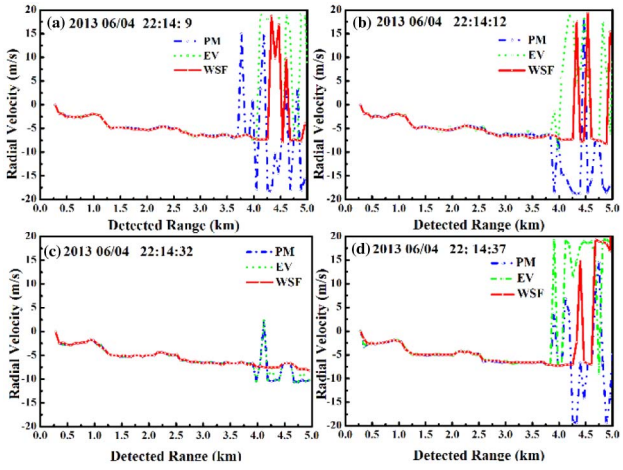

Fig. 14. Wind profiling estimated by PM, EV, and WSF methods for the CDL data on June 4, 2013. The WSF method exhibits a better performance with continuous wind velocity. (a) CDL data collected at 22:14:9; (b) CDL data collected at 22:14:12; (c) CDL data collected at 22:14:32; and (d) CDL data collected at 22:14:37.

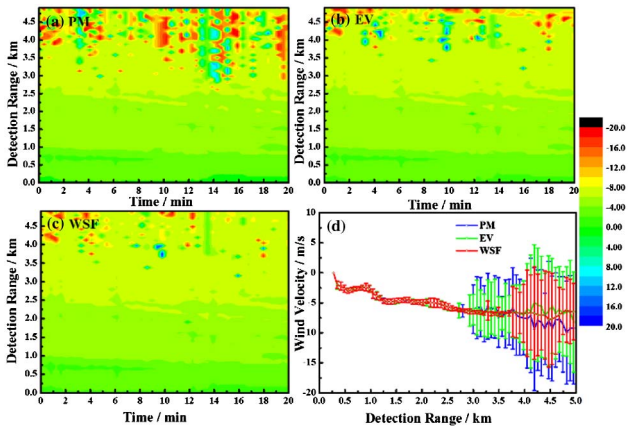

Fig. 15. Spatial and temporal distributions of wind velocity for continuous signal acquisition in 20 min. As the detection range increases, the SNR decreases, and the velocities estimated by the PM, EV, and WSF methods become less consistent. (a) Velocity estimation by the PM method, which has a large number of erroneous velocity estimations; (b) velocity estimation by the EV method, which has a small amount of erroneous velocity estimations; (c) velocity estimation by the WSF method, which has little erroneous velocity; and (d) comparison of velocity mean value and SD by PM, EV, and WSF methods.

and more sudden changes afterwards, yet less severe abrupt changes than the PM and EV methods. To compare the two methods more clearly, Fig. 15(d) shows the SD of velocities estimated by the WSF method is smallest when compared with the other two methods. Therefore, the WSF method improves wind velocity estimation and is superior to the PM and EV methods.

5. CONCLUSION

The WSF algorithm, by introducing an optimal weighting matrix, is presented in this paper. It is based on a parametric modelization of the covariance matrix of the atmospheric backscattered echoes, which is asymptotically efficient and achieves the CRB for Gaussian signals. The main focus in this paper is on the performance of the WSF method and the other subspace decomposition method, such as the EV method, for improving the accuracy of the wind velocity estimation.

The statistical performances of the PM, EV, and the WSF methods have been confirmed by simulations and actual data. The aerosol backscatter is stronger in the ranges closer to the ground than high altitudes. When the strength of the aerosol backscatter is strong in low altitudes, the periodogram-based DS estimates are of good quality, and so are the EV and WSF estimates. When the strength of the aerosol backscatter is weak in high altitudes, the PM algorithm suffers from the smoothing effects and the leakage problem, which hinder the accurate estimation of the frequency at maximum power. Also, the EV method cannot return reasonable signal locations that suffer from extremely low SNR. However, the atmospheric signals are distinctly identified from a noisy environment by the WSF method compared with the PM and EV methods. As the WSF method reduces the statistical uncertainties in the Doppler power spectrum, enhancement in the quality of the calculations of mean frequency leads to greater accuracy in

wind estimation. Depending on the analysis results of simulated spectra and real spectra, the proposed method improves the detection range approximately up to 14.2% and 26.6% when compared to the EV method and the PM method, respectively. Consequently, the WSF approach can be an alternate way to estimate the wind parameters accurately by distinctly identifying the peak frequencies of signals while suppressing erroneous peaks that are always present with conventional Fourier-transform-based frequency spectra.

Funding. National Natural Science Foundation of China (NSFC) (61505009).

Acknowledgment. We thank the Shanghai Institute of Optics and Fine Mechanics for the use of its equipment.

REFERENCES

1. R. T. Menzies and R. M. Hardesty, "Coherent Doppler lidar for measurements of wind fields," Proc. IEEE **77**, 449–462 (1989).
2. D. S. Zrnic, "Estimation of spectral moments for weather echoes," IEEE Trans. Geosci. Electron. **17**, 113–128 (1979).
3. P. R. Mahapatra and D. S. Zrnić, "Practical algorithms for mean velocity estimation in pulse Doppler weather radars using a small number of samples," IEEE Trans. Geosci. Remote Sens. **GE-21**, 491–501 (1983).
4. J. Makhoul, "Linear prediction: a tutorial review," Proc. IEEE **63**, 561–580 (1975).
5. J. Y. Beyon, G. J. Koch, and Z. Li, "Noise normalization and windowing functions for VALIDAR in wind parameter estimation," in *Defense and Security Symposium* (International Society for Optics and Photonics, 2006), pp. 621404–621411.
6. E. Boyer, M. Petitdidier, W. Corneil, C. Adnet, and P. Larzabal, "Application of model-based spectral analysis to wind-profiler radar observations," Ann. Geophys. **19**, 815–824 (2001).
7. H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," IEEE Signal Process. Mag. **13**, 67–94 (1996).
8. L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," Proc. IEEE **85**, 1195–1245 (1997).
9. A.-J. Van Der Veen, E. F. Deprettere, and A. L. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," Proc. IEEE **81**, 1277–1308 (1993).
10. R. Roy, A. Paulraj, and T. Kailath, "ESPRIT–A subspace rotation approach to estimation of parameters of cisoids in noise," IEEE Trans. Acoust. Speech Signal Process. **34**, 1340–1342 (1986).
11. J. Y. Beyon and G. J. Koch, "Wind profiling by a coherent Doppler lidar system VALIDAR with a subspace decomposition approach," in *Defense and Security Symposium* (International Society for Optics and Photonics, 2006), 623605–623608.
12. V. Sureshbabu, V. Anandan, T. Tsuda, J.-I. Furumoto, and S. Rao, "Denoising atmospheric radar signals using spectral-based subspace method applicable for PBS wind estimation," IEEE Trans. Geosci. Remote Sens. **51**, 3853–3861 (2013).
13. E. Boyer, M. Petitdidier, and P. Larzabal, "Stochastic Maximum Likelihood (SML) parametric estimation of overlapped Doppler echoes," Ann. Geophys. **22**, 3983–3993 (2004).
14. M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," IEEE Trans. Signal Process. **39**, 2436–2449 (1991).
15. M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," IEEE Trans. Signal Process. **39**, 1110–1121 (1991).
16. P. Stoica and T. Söderström, "Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies," IEEE Trans. Signal Process. **39**, 1836–1847 (1991).
17. C.-W. Ma and C.-C. Teng, "Detection of coherent signals using weighted subspace smoothing," IEEE Trans. Antennas Propag. **44**, 179–187 (1996).
18. Y. Wu, P. Guo, S. Chen, Y. Zhang, and H. Chen, "Performance of estimated Doppler velocity by maximum likelihood based on covariance matrix," Opt. Eng. **55**, 096112 (2016).
19. K. M. Wong, Q.-T. Zhang, J. P. Reilly, and P. Yip, "On information theoretic criteria for determining the number of signals in high resolution array processing," IEEE Trans. Acoust. Speech Signal Process. **38**, 1959–1971 (1990).
20. M. Wax and I. Ziskind, "Detection of the number of coherent signals by the MDL principle," IEEE Trans. Acoust. Speech Signal Process. **37**, 1190–1196 (1989).
21. Y. Zhigang, L. Xiaobo, and L. Deshu, "Study on the source number estimator using Gerschgorin radii," in *Fourth International Conference on Signal Processing Proceedings, 1998. ICSP'98* (IEEE, 1998), 152–155.
22. O. Caspary, P. Nus, and T. Cecchin, "The source number estimation based on gerschgorin radii," in *IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 1998), pp. 1993–1996.
23. J. A. Cadzow, "A high resolution direction-of-arrival algorithm for narrow-band coherent and incoherent sources," IEEE Trans. Acoust. Speech Signal Process. **36**, 965–979 (1988).
24. H. Krim and J. G. Proakis, "Smoothed eigenspace-based parameter estimation," Automatica **30**, 27–38 (1994).
25. S. Kameyama, T. Ando, K. Asaka, Y. Hirano, and S. Wadaka, "Performance of discrete-Fourier-transform-based velocity estimators for a wind-sensing coherent Doppler lidar system in the Kolmogorov turbulence regime," IEEE Trans. Geosci. Remote Sens. **47**, 3560–3569 (2009).