Sparsity and Smoothness Regularized Estimation of Power Spectral Density and Its Application to Weather Radar

Hiroki Kuroda[1], Daichi Kitahara[2], Eiichi Yoshikawa[3], Hiroshi Kikuchi[4], and Tomoo Ushio[2]

[1]College of Information Science and Engineering, Ritsumeikan University
[2]Division of Electrical, Electronic and Infocommunications Engineering, Osaka University
[3]Aeronautical Technology Directorate, Japan Aerospace Exploration Agency
[4]Center for Space Science and Radio Engineering, The University of Electro-Communications

Abstract In this paper, we propose a sparsity and smoothness regularized estimation of nonnegative power spectral densities (PSDs) of complex-valued random processes from mixtures of realizations. For this purpose, we design a model that jointly estimates frequency components of the realizations and the PSDs. In the proposed model, the PSDs are estimated by nonnegative variables, and thus the smoothness can be exploited without losing the convexity. Numerical experiments on the phased array weather radar show that the proposed approach achieves better estimation accuracy than the existing sparse estimation models combined with post-smoothing.

1 Introduction

Estimation of the power spectral density (PSD) of a random process is an important problem in science and engineering [1, 2]. In particular for weather radar applications, the PSD estimation is essential for analysis of weather phenomena because the PSD reflects the precipitation intensity and the doppler velocity distribution [3–7]. Our primal interest is in phased array weather radars (PAWRs) [8–15], which are developed to detect hazardous weather phenomena. For quick observation, the PAWR transmits a fan beam and then receives mixtures of realizations of random processes, i.e., backscattered signals from multiple elevation angles. Thus, the PSD estimation for the PAWR is much more challenging than the classical single source problem where realizations of a single random process are directly measured.

Separation of frequency components for each elevation angle, called *beamforming*, is a critical step for the subsequent PSD estimation. In many fields, sparsity-aware methods have achieved substantial improvements on estimation accuracy over the classical linear ones [15–21]. Spatial sparsity is exploited in [16, 17] by assuming that signals exist at a few angles. However, this assumption is not appropriate for the PAWR because targets such as clouds and raindrops exist in many angles [15]. In [18–20], isolated sparse frequency components, known as *line spectrum*, are estimated based on the ℓ_1 penalty. Although the line spectrum model (i.e., simple spare model) is not suitable for weather radar applications [3], it is demonstrated in [15] that a block-sparse model, which supposes that frequency components are clustered in a few blocks, is suitable for the PAWR due to the narrow-bandwidth of PSDs. The block structure adaptation technique of [21] further improves the estimation accuracy.

Meanwhile, for the single source problem, exploiting the smoothness of the PSD is crucial to reduce the estimation variance [1, 2]. This suggests that PSD smoothing is also important for the PAWR. Since the PSDs are estimated based on the squared magnitude of frequency components, a straightforward approach is to penalize difference between magnitude of frequency components in the beamforming. However, since this penalty is nonconvex [22], this approach has difficulty in computation of a globally optimal solution. While it is possible to apply existing smoothing techniques, e.g., those shown in [1, 2], as a post-processing, such a two-stage approach is sub-optimal because the smoothness is not exploited in the beamforming.

In this paper, we propose to exploit sparsity and smoothness for the estimation of PSDs from the mixtures of realizations by a convex model that simultaneously estimates frequency components and PSDs. Our major contribution is to leverage the nonnegative latent variable of the block-sparse model of [21], which is originally introduced to optimize the block structure, for the PSD estimation. Namely, we show that the nonnegative latent variable in fact corresponds to the square root of the PSD when the model of [21] is used for the block-sparse estimation of frequency components. Thanks to the nonnegativity of the latent variable, the smoothness can be exploited by the penalty for their differences without losing the convexity. Numerical simulations on the PAWR show that the proposed model achieves better estimation accuracy than the existing sparse estimation models combined with post-smoothing.

Notations: $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{C}$ respectively denote the sets of all real numbers, all nonnegative real numbers, and all complex numbers. We use $\imath \in \mathbb{C}$ to denote the imaginary unit, i.e., $\imath = \sqrt{-1}$. For matrices or vectors, we denote the simple transpose and the Hermitian transpose respectively by $(\cdot)^\top$ and $(\cdot)^{\mathrm{H}}$. We define the support of $\boldsymbol{x} \in \mathbb{C}^N$ by $\mathrm{supp}(\boldsymbol{x}) := \{n \in \{1, \ldots, N\} \,|\, x_n \neq 0\}$. The cardinality of a set $\mathcal{S}$ is denoted by $|\mathcal{S}|$. The ℓ_2 norm, the ℓ_1 norm, and the ℓ_0 pseudo-norm of $\boldsymbol{x} \in \mathbb{C}^N$ are respectively denoted by $\|\boldsymbol{x}\|_2 := \sqrt{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}}$, $\|\boldsymbol{x}\|_1 := \sum_{n=1}^{N} |x_n|$, and $\|\boldsymbol{x}\|_0 := |\mathrm{supp}(\boldsymbol{x})|$. The expectation operator is denoted by $E[\cdot]$.

2 Problem Formulation

We consider the estimation problem of power spectral densities (PSDs) of N random processes from mixtures of their realizations, which arises in phased array weather radars (PAWRs) [8–15]. Let the time-series

$$\bar{x}_{n,j}[\ell] \in \mathbb{C} \quad (\ell = 1, \dots, L)$$

be a realization of n-th discrete-time random process $X_n^\star[\ell]$ $(n = 1, \dots, N)$, where $j \in \{1, \dots, J\}$ is the trial index. For the PAWR, $X_n^\star[\ell]$ corresponds to the sum of backscattered signals in the angular interval $\left[\theta_n - \frac{\Delta\theta}{2}, \theta_n + \frac{\Delta\theta}{2}\right]$, where θ_n $(n = 1, \dots, N)$ are equally spaced angles with the spacing $\Delta\theta$. The PAWR observes mixtures of the realizations by a p-element uniform linear array:

$$\boldsymbol{y}_j[\ell] := \sum_{n=1}^{N} \boldsymbol{a}(\theta_n)\bar{x}_{n,j}[\ell] + \boldsymbol{\varepsilon}_j[\ell] \in \mathbb{C}^p \quad (\ell = 1, \dots, L) \tag{1}$$

for every $j \in \{1, \dots, J\}$, where $\boldsymbol{a}(\theta_n) \in \mathbb{C}^p$ is the *steering vector* [11,15] for the angle θ_n, and $\boldsymbol{\varepsilon}_j[\ell] \in \mathbb{C}^p$ is the white Gaussian noise. In particular, we suppose that $\bar{x}_{n,j}[\ell]$ $(j = 1, \dots, J)$ are realizations of a common random process $X_n^\star[\ell]$ (see Remark 1 for validity of such observations). The observation model (1) can be expressed in a single equation

$$\boldsymbol{y}_j := \sum_{n=1}^{N} \boldsymbol{A}_n \bar{\boldsymbol{x}}_{n,j} + \boldsymbol{\varepsilon}_j \in \mathbb{C}^{pL}, \tag{2}$$

for each $j \in \{1, \dots, J\}$, where

$$\begin{aligned} \boldsymbol{y}_j &:= (\boldsymbol{y}_j[1]^\top, \boldsymbol{y}_j[2]^\top, \dots, \boldsymbol{y}_j[L]^\top)^\top \in \mathbb{C}^{pL}, \\ \bar{\boldsymbol{x}}_{n,j} &:= (\bar{x}_{n,j}[1], \bar{x}_{n,j}[2], \dots, \bar{x}_{n,j}[L])^\top \in \mathbb{C}^L, \\ \boldsymbol{\varepsilon}_j &:= (\boldsymbol{\varepsilon}_j[1]^\top, \boldsymbol{\varepsilon}_j[2]^\top \dots, \boldsymbol{\varepsilon}_j[L]^\top)^\top \in \mathbb{C}^{pL}, \end{aligned}$$

and $\boldsymbol{A}_n \in \mathbb{C}^{pL \times L}$ is a block-diagonal matrix that contains L copies of $\boldsymbol{a}(\theta_n)$ on the diagonal blocks.

Since frequency components are important for the PSD estimation, we rewrite (2) to a observation model in terms of the discrete Fourier transform (DFT) coefficients

$$\bar{u}_{n,j}[k] := \frac{1}{\sqrt{L}} \sum_{\ell=1}^{L} \bar{x}_{n,j}[\ell] e^{-\imath 2\pi f_k (\ell-1)} \quad (k = 1, \dots, L),$$

where $f_k := (k - 1 - \frac{L}{2})/L$. With the DFT matrix $\boldsymbol{F} \in \mathbb{C}^{L \times L}$, $\bar{u}_{n,j}[k]$ $(k = 1, \dots, L)$ can be written as

$$(\bar{u}_{n,j}[1], \dots, \bar{u}_{n,j}[L])^\top =: \bar{\boldsymbol{u}}_{n,j} = \boldsymbol{F}\bar{\boldsymbol{x}}_{n,j} \in \mathbb{C}^L.$$

Thus, substituting $\bar{\boldsymbol{x}}_{n,j} = \boldsymbol{F}^{\mathrm{H}}\bar{\boldsymbol{u}}_{n,j}$ into (2), we have

$$\boldsymbol{y}_j := \sum_{n=1}^{N} \boldsymbol{A}_n \boldsymbol{F}^{\mathrm{H}} \bar{\boldsymbol{u}}_{n,j} + \boldsymbol{\varepsilon}_j \in \mathbb{C}^{pL}. \tag{3}$$

The goal of the PAWR is to estimate PSDs

$$S_n^\star(f_k) := \sum_{\ell=-\infty}^{\infty} R_n^\star[\ell] e^{-\imath 2\pi f_k \ell} \quad (k = 1, \dots, L),$$

of $X_n^\star[\ell]$ for each $n = 1, \dots, N$ from the observations in (3), where $R_n^\star[\ell] := E[X_n^\star[\tau + \ell]\mathrm{conj}(X_n^\star[\tau])]$ is the auto-correlation function, which does not depend on τ by assuming *second-order stationarity* [1] on $X_n^\star$.

Remark 1 (Tradeoff between L and J)**.** To obtain the observation model in (1) where $\bar{x}_{n,j}[\ell]$ $(j = 1, \dots, J)$ are realizations of a common random process, similarly to the single source case ($N = 1$) [1, 2], we split the whole observations into J subsets. Note that JL should be small enough for weather radar applications so that the statistics of the targets such as clouds and raindrops are (approximately) unchanged. Since L defines the frequency resolution, L should be kept large. Thus, for weather radar applications, typically $J < 10$ is used, and $J = 1$ is also of interest.

2.1 Major Challenge

The *periodogram*

$$|\bar{u}_{n,j}[k]|^2 \quad (k = 1, \dots, L)$$

is widely used as a estimate of the PSD because of its asymptotically unbiasedness

$$E[|\bar{u}_{n,j}[k]|^2] = S_n^\star(f_k) \quad (L \to \infty) \tag{4}$$

under mild conditions [1,2] (Note: the periodogram has to be estimated from the observations in (3) for the PAWR). However, in fact, the periodogram often exhibits erratic oscillation [1–3]. This phenomenon is theoretically proven, e.g., for [1,2] its variance is as large as the square of true PSD $S_n^\star(f_k)$ even when $L \to \infty$ [1,2]. In addition, $|\bar{u}_{n,j}[k]|^2$ and $|\bar{u}_{n,j}[k']|^2$ $(k \neq k')$ are uncorrelated under the same condition. To reduce the variance, a simple way is to exploit the situation that $|\bar{u}_{n,j}[k]|^2$ $(j = 1, \dots, J)$ are realizations of the common random variable, i.e., to take the ensemble average

$$\frac{1}{J} \sum_{j=1}^{J} |\bar{u}_{n,j}[k]|^2 \quad (k = 1, \dots, L). \tag{5}$$

However, since J is very small for PAWRs (see Remark 1), the ensemble average cannot reduce the variance sufficiently. Another approach is to exploit the smoothness of the PSD in the frequency domain, which is suitable for weather radar applications [3]. However, existing smoothing techniques [1, 2] are not directly applicable to the PAWR because these techniques suppose that the frequency components $\bar{u}_{n,j}[k]$ are known. Using smoothing techniques as a post-processing is sub-optimal because the smoothness of the PSDs cannot be exploited for the estimation of frequency components.

3 Proposed Model

To exploit sparsity and smoothness of PSDs, we design a convex model that jointly estimates frequency components and PSDs from the observations in (3). In Section 3.1, we apply the block-sparse model of [21] for the estimation of frequency components. Then, in Section 3.2, we show that its latent variable, which is originally introduced for the optimization of block structure, is in fact suitable for the estimation of sparse and smooth PSDs.

3.1 Block-Sparse Recovery of Frequency Components

We begin by designing a block-sparse penalty for frequency components for each $n \in \{1, \ldots, N\}$. As demonstrated in [15], the PSD $S_n^\star(f)$ is usually narrow-band for the PAWR, which implies that $\bar{\boldsymbol{u}}_{n,j}$ is block-sparse for each source $n \in \{1, \ldots, N\}$ and trial $j \in \{1, \ldots, J\}$ due to the relation (4). Moreover, an appropriate block partition is common for all trials $j = 1, \ldots, J$. Thus, we introduce the penalty for $\boldsymbol{u}_n := (\boldsymbol{u}_{n,j})_{j=1}^J$ based on the mixed ℓ_2/ℓ_1 norm:

$$\|\boldsymbol{u}_n\|_{2,1}^{(\mathcal{B}_m)_{m=1}^q} = \sum_{m=1}^{q} \sqrt{|\mathcal{B}_m|} \sqrt{\sum_{(k,j)\in\mathcal{B}_m\times\{1,\ldots,J\}} |u_{n,j}[k]|^2},$$

where $\mathcal{B}_m \subset \{1, \ldots, L\}$ $(m = 1, \ldots, q)$ is non-overlapping blocks. However, the appropriate block partition is unknown a priori because it depends on the unknown doppler velocity distribution [15]. Thus, following the approach of [21], we minimize the mixed ℓ_2/ℓ_1 norm over the partition of at most M blocks, i.e.,

$$\psi_M(\boldsymbol{u}_n) := \min_{q\in\{1,\ldots,M\}} \left[\min_{(\mathcal{B}_m)_{m=1}^q \in \mathcal{P}_q} \|\boldsymbol{u}_n\|_{2,1}^{(\mathcal{B}_m)_{m=1}^q} \right],$$

where $\mathcal{P}_q$ consists of all q block partitions of $\{1, \ldots, L\}$. Introducing the lower semicontinuous convex function $\phi\colon \mathbb{C}^J \times \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ defined by

$$\phi(\boldsymbol{v}, \sigma) := \begin{cases} \dfrac{\|\boldsymbol{v}\|^2}{2\sigma} + \dfrac{\sigma}{2}, & \text{if } \sigma > 0; \\ 0, & \text{if } \boldsymbol{v} = \boldsymbol{0} \text{ and } \sigma = 0; \\ \infty, & \text{otherwise,} \end{cases} \tag{6}$$

similarly to the derivation in [21, Section II], we have

$$\psi_M(\boldsymbol{u}_n) = \min_{\substack{\boldsymbol{\sigma}_n\in\mathbb{R}_+^L \\ \|\boldsymbol{D}\boldsymbol{\sigma}_n\|_0 \le M-1}} \sum_{k=1}^{L} \phi\left((u_{n,j}[k])_{j=1}^J, \sigma_n[k]\right), \tag{7}$$

which is an extension of the result in [21] for the partially fixed block partition, where $\boldsymbol{D} \in \mathbb{R}^{L\times L}$ is the first difference operator with periodic boundary condition[1)].

[1)]The periodic boundary condition is introduced due to possible aliasing, which is not a serious matter for weather radar applications unless the variation of doppler velocity is extremely large (see, e.g., [3]).

A tight convex relaxation can be obtained by replacing the ℓ_0 pseudo-norm in the constraint of (7) with the ℓ_1 norm. To simplify the tuning, we design the proposed penalty for $\boldsymbol{u} = (\boldsymbol{u}_n)_{n=1}^N$ as

$$\Psi_\alpha(\boldsymbol{u}) := \min_{\substack{\boldsymbol{\sigma}=(\boldsymbol{\sigma}_n)_{n=1}^N\in\mathbb{R}_+^{NL} \\ \sum_{n=1}^N \|\boldsymbol{D}\boldsymbol{\sigma}_n\|_1 \le \alpha}} \sum_{n=1}^{N}\sum_{k=1}^{L} \phi\left((u_{n,j}[k])_{j=1}^J, \sigma_n[k]\right),$$

where the tuning parameter $\alpha \in \mathbb{R}_+$ controls the total number of blocks.

We estimate the frequency components by regularizing the quadratic error for observations in (3), i.e.,

$$\operatorname*{minimize}_{\boldsymbol{u}\in\mathbb{C}^{JNL}} \frac{1}{2}\sum_{j=1}^{J} \left\|\boldsymbol{y}_j - \textstyle\sum_{n=1}^N \boldsymbol{A}_n \boldsymbol{F}^{\mathrm{H}} \boldsymbol{u}_{n,j}\right\|_2^2 + \lambda\Psi_\alpha(\boldsymbol{u}), \tag{8}$$

where $\lambda > 0$ is the regularization parameter. Substituting the definition of $\Psi_\alpha(\boldsymbol{u})$, we can solve (8) as

$$\left.\begin{aligned} &\operatorname*{minimize}_{\boldsymbol{u}\in\mathbb{C}^{JNL},\boldsymbol{\sigma}\in\mathbb{R}_+^{NL}} \frac{1}{2}\sum_{j=1}^{J} \left\|\boldsymbol{y}_j - \textstyle\sum_{n=1}^N \boldsymbol{A}_n \boldsymbol{F}^{\mathrm{H}} \boldsymbol{u}_{n,j}\right\|_2^2 \\ &\qquad + \lambda \sum_{n=1}^{N}\sum_{k=1}^{L} \phi\left((u_{n,j}[k])_{j=1}^J, \sigma_n[k]\right) \\ &\qquad \text{subject to } \sum_{n=1}^{N} \|\boldsymbol{D}\boldsymbol{\sigma}_n\|_1 \le \alpha \end{aligned}\right\}. \tag{9}$$

A globally optimal solution of (9) can be obtained by the combination of proximal splitting techniques [23–25] and interpretation of ϕ in (6) as a perspective function [26,27] with a slight reformulation similar to [21, Section III].

3.2 Exploiting Latent Variable for PSD Estimation

While it is possible to compute the periodogram as in (5) for the solution of (9) for the variable $\boldsymbol{u}$, say $\hat{\boldsymbol{u}}$, we show that the solution for the latent variable $\boldsymbol{\sigma}$, say $\hat{\boldsymbol{\sigma}}$, is more suitable for the estimation of sparse and smooth PSDs. In the proposed model (9), $\boldsymbol{\sigma}$ is contained in the part

$$\lambda \sum_{n=1}^{N}\sum_{k=1}^{K} \phi\left((u_{n,j}[k])_{j=1}^J, \sigma_n[k]\right) \tag{10}$$

and the constraint

$$\sum_{n=1}^{N} \|\boldsymbol{D}\boldsymbol{\sigma}_n\|_1 \le \alpha. \tag{11}$$

a) From [21, Lemma 1], the part (10) is minimized when $\sigma_n[k]$ is the square root of the averaged periodogram shown in (5), i.e., when

$$\sigma_n[k] = \sqrt{\frac{1}{J}\sum_{j=1}^{J} |\bar{u}_{n,j}[k]|^2}. \tag{12}$$

b) To satisfy the constraint (11), since $\|\boldsymbol{D}\boldsymbol{\sigma}_n\|_1$ penalizes the smoothness of $\boldsymbol{\sigma}_n$, $\boldsymbol{\sigma}_n$ should be smooth for each $n = 1 \ldots, N$.

Thus, balancing these terms makes $\hat{\boldsymbol{\sigma}}$ the square root of smoothed and averaged periodogram. Note that the block-sparsity is preserved in $\hat{\boldsymbol{\sigma}}$ because $\hat{\boldsymbol{u}}_{n,j}$ $(j = 1, \ldots, J)$ are regularized to have a common block-sparse support. Since the PSDs are expected to be smooth and block-sparse for the PAWR, the square of components of $\hat{\boldsymbol{\sigma}}$, i.e.,

$$\hat{S}_n(f_k) = |\hat{\sigma}_n[k]|^2 \tag{13}$$

is more suitable for the PSD estimation than the periodogram computed from $\hat{\boldsymbol{u}}_{n,j}$ $(j = 1, \ldots, J)$.

Intuitively, the proposed model can effectively estimates smooth and block-sparse PSD by the following mechanism. Since $\phi(\boldsymbol{v}, \sigma)$ is basically $\frac{\|\boldsymbol{v}\|^2}{2\sigma} + \frac{\sigma}{2}$ (see (6)), roughly speaking, the part (10) acts as

$$\lambda \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{j=1}^{J} \frac{|u_{n,j}[k]|^2}{2\sigma_n[k]} + \frac{\sigma_n[k]}{2}.$$

Due to the relation (4), $|u_{n,j}[k]|^2$ weighted by $\sigma_n[k]$ the estimate of PSD is an effective regularization. Refining $\boldsymbol{u}$ leads to a refined estimate of $\boldsymbol{\sigma}$ because $\boldsymbol{\sigma}$ is smoothed around the value in (12). Thanks to these interaction, $\hat{\boldsymbol{\sigma}}$ is expected to be a better estimate of the PSD than the sparse frequency estimation combined with the post-smoothing.

While $\|\boldsymbol{D}\boldsymbol{\sigma}_n\|_1$ is a good choice for the control of the block structure, more advanced smoothness penalties are expected to further enhance the estimation accuracy of the PSD. Thanks to the nonnegativity of $\boldsymbol{\sigma}_n$, the proposed formulation can incorporate many convex smoothness penalties without losing the convexity of the model. For instance, the high-order total variation [28, 29], which uses $\boldsymbol{D}^q$ $(q \geq 2)$ instead of $\boldsymbol{D}$, and the total generalized variation [30] can be incorporated into the proposed model and solved similarly to (9).

4 Simulation Results

For the PAWR, we compare the proposed approach against the existing sparse frequency component estimations: the ℓ_1 regularization and the mixed ℓ_2/ℓ_1 norm based model [15], which is the state-of-the-art for the PAWR. We also compare the proposed approach against the existing model combined with post-smoothing. For the post-smoothing, we employ the standard Daniell smoothing [1, 2], i.e., the moving average filter, where we use a small filter size 3 to avoid the frequency broadening.

We conduct numerical simulations on the PAWR in a setting similar to [11, 15]. The elevation angles θ_n $(n = 1, \ldots, N)$ are chosen uniformly from -15° to 30° degrees with $N = 110$. To objectively evaluate the models, we synthesize the (discrete-time) PSD $S_n^\star(f)$ for each elevation angle from a continuous-time Gaussian shaped PSD $G_n^\star(f) = \frac{P_n}{\sqrt{2\pi}\varsigma_n} e^{-\frac{(f-\mu_n)^2}{2\varsigma_n^2}}$, which is an appropriate model, e.g., when the atmospheric turbulence is dominant [3]. Namely, we set $S_n^\star(f) = \frac{1}{T}\sum_{m=-\infty}^{\infty} G_n^\star\left(\frac{f-m}{T}\right)$, where T is the pulse repetition time. We set the power P_n by the real reflection intensity measured by the PAWR at Osaka University on March 30, 2014. We define the mean Doppler frequency μ_n by the certain sine curve used in [15]. The Doppler velocity width is chosen uniformly from $[1, 3]$ [m/s], and then converted to the width ς_n in the frequency domain. Note that this setting is more realistic than that of [15] where the Doppler velocity width is simply fixed to 2 [m/s]. We define $X_n^\star[\ell]$ as the Gaussian process that has the specified PSD $S_n^\star(f)$. We generate $\bar{x}_{n,j}[\ell]$ $(\ell = 1, \ldots, L)$ as a realization of $X_n^\star[\ell]$. The measurements $\boldsymbol{y}_j$ is given by (2), where $\boldsymbol{\varepsilon}_j$ is generated as the white Gaussian noise of the standard deviation $\sqrt{2.5}$. We set the parameters as $L = 32$, $J = 1$, $p = 128$, and $T = 0.4$ [ms].

Table 1: Comparison of the NMAE of $S_n^\star(f_k)$, where the result is averaged over 100 independent trials. Values shown in parenthesis are the NMAEs without post-smoothing for the existing methods.

ℓ_1	ℓ_2/ℓ_1	Proposed
0.7607 (0.8789)	0.7301 (0.7388)	**0.6262**

We compare the models in terms of the normalized mean absolute error (NMAE)

$$\frac{\sum_{n=1}^{N}\sum_{k=1}^{L}\left|S_n^\star(f_k) - \hat{S}_n(f_k)\right|}{\sum_{n=1}^{N}\sum_{k=1}^{L} S_n^\star(f_k)}$$

averaged over 100 independent trails. For the proposed approach, we compute $\hat{S}_n(f_k)$ by (13) with $\hat{\boldsymbol{\sigma}}$ the solution of (9). For the existing approaches, $\hat{S}_n(f_k)$ is defined by (5) with the estimated frequency components, and then processed by the post-smoothing. The regularization parameter λ, α for the proposed model, and the block size for the mixed ℓ_2/ℓ_1 norm based model [15], are tuned independently for each model to obtain the best accuracy. From Table 1, we see that the proposed approach achieves better estimation accuracy the existing models combined with the post-smoothing.

Acknowledgment: This work was supported in part by JSPS Grants-in-Aid 21K17827 and 21H01592.

References

[1] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Prentice-Hall, 2005.

[2] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer, 2009.

[3] V. N. Bringi and V. Chandrasekar, *Polarimetric Doppler Weather Radar: Principles and Applications*, Cambridge University Press, 2001.

[4] D. S. Zrnic and R. J. Doviak, "Velocity spectra of vortices scanned with a pulse-doppler radar," *J. Appl. Meteorol.*, vol. 14, no. 8, pp. 1531–1539, 1975.

[5] L. H. Janssen and G. A. Van Der Spek, “The shape of doppler spectra from precipitation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-21, no. 2, pp. 208–219, 1985.

[6] Y. Wang, T.-Y. Yu, M. Yeary, A. Shapiro, S. Nemati, M. Foster, D. L. Andra, and M. Jain, “Tornado detection using a neuro–fuzzy system to integrate shear and spectral signatures,” *J. Atmos. Ocean. Technol.*, vol. 25, no. 7, pp. 1136–1148, 2008.

[7] T.-Y. Yu, R. R. Rondinel, and R. D. Palmer, “Investigation of non-Gaussian doppler spectra observed by weather radar in a tornadic supercell,” *J. Atmos. Ocean. Technol.*, vol. 26, no. 3, pp. 444–461, 2009.

[8] D. S. Zrnic, J. F. Kimpel, D. E. Forsyth, A. Shapiro, G. Crain, R. Ferek, J. Heimmer, W. Benner, F. T. J. McNellis, and R. J. Vogt, “Agile-beam phased array radar for weather observations,” *Bull. Am. Meteorol. Soc.*, vol. 88, no. 11, pp. 1753–1766, 2007.

[9] D. McLaughlin, D. Pepyne, V. Chandrasekar, B. Philips, J. Kurose, M. Zink, K. Droegemeier, S. Cruz-Pol, F. Junyent, J. Brotzge, D. Westbrook, N. Bharadwaj, Y. Wang, E. Lyons, K. Hondl, Y. Liu, E. Knapp, M. Xue, A. Hopf, K. Kloesel, A. DeFonzo, P. Kollias, K. Brewster, R. Contreras, B. Dolan, T. Djaferis, E. Insanic, S. Frasier, and F. Carr, “Short-wavelength technology and the potential for distributed networks of small radar systems,” *Bull. Am. Meteorol. Soc.*, vol. 90, no. 12, pp. 1797–1818, 2009.

[10] B. Isom, R. Palmer, R. Kelley, J. Meier, D. Bodine, M. Yeary, B.-L. Cheong, Y. Zhang, T.-Y. Yu, and M. I. Biggerstaff, “The atmospheric imaging radar: Simultaneous volumetric observations using a phased array weather radar,” *J. Atmos. Ocean. Technol.*, vol. 30, no. 4, pp. 655–675, Apr. 2013.

[11] E. Yoshikawa, T. Ushio, Z. Kawasaki, S. Yoshida, T. Morimoto, F. Mizutani, and M. Wada, “MMSE beam forming on fast-scanning phased array weather radar,” *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 3077–3088, May 2013.

[12] D. S. Zrnic, V. M. Melnikov, R. J. Doviak, and R. Palmer, “Scanning strategy for the multifunction phased-array radar to satisfy aviation and meteorological needs,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1204–1208, 2015.

[13] F. Mizutani, T. Ushio, E. Yoshikawa, S. Shimamura, H. Kikuchi, M. Wada, S. Satoh, and T. Iguchi, “Fast-scanning phased-array weather radar with angular imaging technique,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2664–2673, May 2018.

[14] E. Yoshikawa, T. Ushio, and H. Kikuchi, “A study of comb beam transmission on phased array weather radars,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6346–6356, 2021.

[15] D. Kitahara, H. Kuroda, A. Hirabayashi, E. Yoshikawa, H. Kikuchi, and T. Ushio, “Nonlinear beamforming based on group-sparsities of periodograms for phased array weather radar,” *IEEE Trans. Geosci. Remote Sens.*, 2022.

[16] P. Stoica, P. Babu, and J. Li, “SPICE: A sparse covariance-based estimation method for array processing,” *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, 2011.

[17] L. Wang, L. Zhao, G. Bi, C. Wan, L. Zhang, and H. Zhang, “Novel wideband DOA estimation based on sparse Bayesian learning with Dirichlet process priors,” *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 275–289, 2016.

[18] P. Stoica, P. Babu, and J. Li, “New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data,” *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35–47, 2011.

[19] C. R. Rojas, D. Katselis, and H. Hjalmarsson, “A note on the SPICE method,” *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4545–4551, 2013.

[20] J. Fang, F. Wang, Y. Shen, H. Li, and R. S. Blum, “Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach,” *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4649–4662, 2016.

[21] H. Kuroda and D. Kitahara, “Block-sparse recovery with optimal block partition,” *IEEE Trans. Signal Process.*, vol. 70, pp. 1506–1520, 2022.

[22] P. Ciuciu, J. Idier, and J.-F. Giovannelli, “Regularized estimation of mixed spectra using a circular Gibbs-Markov model,” *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2202–2213, 2001.

[23] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *J. Math. Imaging Vis.*, vol. 40, no. 1, pp. 120–145, 2011.

[24] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation,” *Comput. Math. Appl.*, vol. 2, no. 1, pp. 17–40, 1976.

[25] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, “Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists,” *SIAM Rev.*, 2022.

[26] P. L. Combettes, “Perspective functions: Properties, constructions, and examples,” *Set-Valued and Variational Analysis*, vol. 26, no. 2, pp. 247–264, 2018.

[27] P. L. Combettes and C. L. Müller, “Perspective maximum likelihood-type estimation via proximal decomposition,” *Electron. J. Statist.*, vol. 14, no. 1, pp. 207–238, 2020.

[28] T. Chan, A. Marquina, and P. Mulet, “High-order total variation-based image restoration,” *SIAM J. Sci. Comput.*, vol. 22, no. 2, pp. 503–516, 2000.

[29] G. D. Maso, I. Fonseca, G. Leoni, and M. Morini, “A higher order model for image restoration: The one-dimensional case,” *SIAM J. Math. Anal.*, vol. 40, no. 6, pp. 2351–2391, 2009.

[30] K. Bredies, K. Kunisch, and T. Pock, “Total generalized variation,” *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 492–526, 2010.