

Nonlinear Beamforming Based on Group-Sparsities of Periodograms for Phased Array Weather Radar

Daichi Kitahara, *Member, IEEE*, Hiroki Kuroda, *Member, IEEE*, Akira Hirabayashi, *Member, IEEE*,
Eiichi Yoshikawa, *Member, IEEE*, Hiroshi Kikuchi, *Member, IEEE*,
and Tomoo Ushio, *Senior Member, IEEE*

Abstract—We propose *nonlinear* beamforming for phased array weather radars (PAWRs). Conventional beamforming is *linear* in the sense that a backscattered signal arriving from each elevation is reconstructed by a weighted sum of received signals, which can be seen as a linear transform for the received signals. For *distributed targets* such as raindrops, however, the number of scatterers is significantly large, differently from the case of *point targets* that are standard targets in array signal processing. Thus, the spatial resolution of the conventional linear beamforming is limited. To improve the spatial resolution, we exploit two characteristics of a *periodogram* of each backscattered signal from the distributed targets. The periodogram is a series of the powers of the discrete Fourier transform (DFT) coefficients of each backscattered signal and utilized as a nonparametric estimate of the *power spectral density*. Since each power spectral density is proportional to the Doppler frequency distribution, 1) major components of the periodogram are concentrated in the vicinity of the *mean Doppler frequency*, and 2) frequency indices of the major components are similar between adjacent elevations. These are expressed as *group-sparsities* of the DFT coefficient matrix of the backscattered signals, and we propose to reconstruct the signals through convex optimization exploiting the group-sparsities. We consider two optimization problems. One problem roughly evaluates the group-sparsities and is relatively easy to solve. The other evaluates the group-sparsities more accurately but requires more time to solve. Both problems are solved with the *alternating direction method of multipliers* including *nonlinear mappings*. Simulations using synthetic and real-world PAWR data show that the proposed method dramatically improves the spatial resolution.

Index Terms—Beamforming, convex optimization, distributed target, group-sparsity, periodogram, phased array weather radar.

Manuscript received March 1, 2021; revised September 26, 2021 and January 7, 2022; accepted February 19, 2022. Date of publication February 24, 2022; date of current version March 31, 2022. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP-19K20361 and Grant JP21H01592. (*Corresponding author: Daichi Kitahara.*)

Daichi Kitahara, Hiroki Kuroda, and Akira Hirabayashi are with the College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan (e-mail: d-kita@media.ritsumeiki.ac.jp; kuroda@media.ritsumeiki.ac.jp; akirahrb@media.ritsumeiki.ac.jp).

Eiichi Yoshikawa is with the Aeronautical Technology Directorate, Japan Aerospace Exploration Agency, Mitaka, Tokyo 181-0015, Japan (e-mail: yoshikawa.eiichi@jaxa.jp).

Hiroshi Kikuchi is with the Center for Space Science and Radio Engineering, The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan (e-mail: hkikuchi@uec.ac.jp).

Tomoo Ushio is with the Division of Electrical, Electronic and Infocommunications Engineering, Osaka University, Suita, Osaka 565-0871, Japan (e-mail: ushio@eei.eng.osaka-u.ac.jp).

Digital Object Identifier 10.1109/TGRS.2022.3154118

I. INTRODUCTION

PHASED array weather radars (PAWRs) [1]–[6] have been developed to quickly detect hazardous weather phenomena such as a thunderstorm with heavy rain. Classical parabolic Doppler weather radars [7], [8] transmit a pencil beam and receive backscattered signals within a *narrow* range of elevation angles. By repeatedly transmitting and receiving the signals at several elevation angles with the mechanical vertical scan, the classical Doppler radar can observe the weather in the sky. On the other hand, PAWRs transmit a fan beam and receive backscattered signals within a *wide* range of elevation angles by an array antenna. Then, the backscattered signals within the *narrow* ranges are reconstructed from the received signals of the array antenna by digital *beamforming* [9]–[20]. This is the key technology in the PAWR because it gets rid of the mechanical vertical scan and hence the temporal resolution is significantly improved in the weather observation. For example, the PAWR developed at Osaka University performs a volume scan of 60 kilometers radius with 600 range, 300 azimuth, and 110 elevation points (i.e., with a resolution of about $100\text{ [m]} \times 1.2\text{ [deg]} \times 0.82\text{ [deg]}$) in 30 seconds by one rotation in the azimuth direction [5], [21], while the Doppler radar performs a similar scan with freely-selected 10 to 20 elevation points of about 1.2 [deg] beamwidth in 5 to 10 minutes by rotations equal to the number of the elevations [8] (see [22] for comparison of these radars).

The weather observation accuracy in the PAWR depends on beamforming. Major methods [9]–[19] reconstruct the signal arriving from each elevation as a complex weighted sum of the received signals. In particular, Capon (CP) beamforming [12]–[17], also known as the minimum variance distortionless response (MVDR) beamforming [18], is widely used because it can adaptively reduce the influence of so-called *sidelobes* if a sufficient number of beams (pulses) are transmitted. However, for quick weather observations, the number of pulses should be as small as possible. To meet this demand, Yoshikawa *et al.* [19] proposed the minimum mean square error (MMSE) beamforming by modifying the direction-of-arrival (DOA) estimation method of Blunt *et al.* in [23] to a beamforming method imposed with a directional gain constraint. In MMSE beamforming, differently from CP beamforming, the covariance matrix of the received signals is iteratively updated, and hence the backscattered signal arriving from each elevation is reconstructed robustly even if the number of the pulses is small.

Beamforming methods were developed originally for observations of *point targets* in array signal processing, while targets of the PAWR are *distributed targets* such as raindrops. In this case, the number of scatterers is significantly large, i.e., much greater than the number of antenna elements. As a result, the spatial resolution of the above *linear* methods [9]–[19] is limited, and fine fluctuations of the *average signal powers* corresponding to precipitation profile have not been captured.

To overcome the limitation of the *linear* methodology, this article proposes a *nonlinear* beamforming method. We formulate the beamforming as an ill-conditioned inverse problem. To solve it, we exploit properties of *periodograms* of the backscattered signals from the distributed targets. The periodogram is a series of the powers of the discrete Fourier transform (DFT) coefficients of each backscattered signal and often used as a nonparametric estimator for the *power spectral density* of the backscattered signal. Since the power spectral density is proportional to the Doppler frequency distribution [7], 1) dominant components of each periodogram are concentrated in the vicinity of the mean Doppler frequency, and 2) frequency indices of the dominant components are similar between the periodograms of adjacent elevation angles. We express these two properties as *group-sparsities* of the DFT coefficient matrix of the backscattered signals and propose to reconstruct the DFT coefficient matrix by minimizing a cost function that consists of a data-fidelity term and two group ℓ_1 -norms evaluating the group-sparsities. We consider two optimization problems that compute the group ℓ_1 -norms in different ways. One roughly evaluates the group-sparsities and is relatively easy to solve, while the other evaluates the group-sparsities more accurately but requires more time to solve. Both problems are solved with the *alternating direction method of multipliers (ADMM)* [24]–[27]. Since nonlinear mappings called the *proximity operators* are used in the ADMM iterations, the proposed beamforming is a *nonlinear* transform for the received signals. Numerical simulations using synthetic and real-world PAWR data show that the proposed nonlinear beamforming greatly improves the spatial resolution in comparison with the linear beamforming.

There are several studies that use certain sparsities in various domains of different applications to improve the resolution of radars. For example, the *sparsity of point targets in the range-velocity domain* is already used for Doppler radars [28]–[31] and phased array radars [32], [33]. In [34], [35], by supposing that there exist multiple point targets of the same velocity at different ranges, the *velocity-wise group-sparsity in the range-velocity domain* is used for Doppler radars. In [36]–[38], the *sparsity of point targets in the elevation axis* is used for phased array radars. For synthetic aperture radar imaging, to obtain a high-resolution image in the *range-azimuth domain*, the *sparsity of the image or its discrete cosine transform coefficients* [39]–[41] and *group-sparsities due to the continuity of objects* [42]–[44] are used. In [28]–[44], the resolution is improved via convex or nonconvex optimization based mainly on the (group) ℓ_1 -norm or ℓ_0 -pseudonorm that evaluates each sparsity. In this article, *group-sparsities of distributed targets in the elevation-velocity domain* are newly used for phased array radars, and we obtain high-resolution beamforming results as the solutions to convex optimization problems based on two group ℓ_1 -norms.

Note that in [34], [35] the velocity-wise group ℓ_1 -norm (or ℓ_0 -pseudonorm) can be used from the assumption on the velocity, while in [42]–[44] complicated nonconvex regularization terms based on hierarchical probability models are used to evaluate group-sparsities whose group partitions are unknown. To avoid the group partition problem, as with *latent group lasso* [45]–[48], this article uses *small overlapping groups* that enable us to evaluate the group-sparsities by convex regularization terms.

The rest of this article is organized as follows. In Section II, after explaining mathematical notation, we introduce the basic principles of the classical Doppler weather radar, and then we formulate the observation model in the PAWR. In Section III, we describe three conventional linear beamforming [11], [12], [19]. Section IV presents the proposed nonlinear beamforming. First, we derive the data-fidelity term in the frequency domain. Second, we explain the group ℓ_1 -norms evaluating two group-sparsities of periodograms in Section IV-A and Section IV-B. Third, we propose two convex optimization problems including their ADMM iterations in Section IV-C and Section IV-D. Section V shows the effectivity of the proposed beamforming by simulations, and finally Section VI concludes this article.

II. PRELIMINARIES

Let \mathbb{R} , \mathbb{R}_+ , and \mathbb{C} be the sets of all real numbers, nonnegative real numbers, and complex numbers, respectively. We use $j \in \mathbb{C}$ to denote the imaginary unit, i.e., $j := \sqrt{-1}$. For any complex number $x \in \mathbb{C}$, \bar{x} denotes its complex conjugate, and $|x| := \sqrt{x\bar{x}}$ denotes its absolute value. We write vectors and matrices by boldface small and capital letters, respectively. We use $\mathbf{I}_n \in \{0, 1\}^{n \times n}$ to denote the identity matrix of order n . The transpose and Hermitian transpose of a vector or a matrix are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively. The ℓ_2 -norm (or the Euclidean norm) of a vector $\mathbf{x} := (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$ is defined by $\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^n |x_i|^2}$. A weighted group ℓ_1 -norm¹ of \mathbf{x} is defined, with a non-overlapping partition $G := (\mathcal{G}_i)_{i=1}^{n_G}$ and a nonnegative vector $\boldsymbol{\xi} := (\xi_1, \xi_2, \dots, \xi_{n_G})^T \in \mathbb{R}_+^{n_G}$, by $\|\mathbf{x}\|_{1, \boldsymbol{\xi}}^G := \sum_{i=1}^{n_G} \xi_i \|\mathbf{x}_{\mathcal{G}_i}\|_2$, where $\mathcal{G}_i (\neq \emptyset)$ is an index set for the i th group s.t. $\bigcup_{i=1}^{n_G} \mathcal{G}_i = \{1, 2, \dots, n\}$ and $\mathcal{G}_{i_1} \cap \mathcal{G}_{i_2} = \emptyset$ ($i_1 \neq i_2$), and $\mathbf{x}_{\mathcal{G}} := (x_i)_{i \in \mathcal{G}}$ is the subvector extracted from \mathbf{x} with \mathcal{G} . If $\mathbf{x}_{\mathcal{G}_i} \approx \mathbf{0}$ for many groups \mathcal{G}_i , \mathbf{x} is called *group-sparse*² [47]. The Frobenius norm of a matrix $\mathbf{X} := (x_{i_1, i_2}) \in \mathbb{C}^{n_1 \times n_2}$ is defined by $\|\mathbf{X}\|_F := \sqrt{\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} |x_{i_1, i_2}|^2}$. We use $E[\cdot]$ to denote the expected value of some random variable. Table I shows constants used in this article for radar settings, where the left column is used for both Doppler weather radar and PAWR, and the right column is used for only PAWR.

A. Basic Principles of Weather Observation by Doppler Radar

Classical parabolic Doppler weather radars transmit pulses for a specific elevation angle and receive backscattered signals generated by the incidence of the transmitted signals to targets

¹We also define a weighted group ℓ_1 -norm of a matrix \mathbf{X} in a similar way.

²We also use the word “group-sparse” for matrices in a similar sense. If all \mathcal{G}_i consist only of consecutive indices, \mathbf{x} is also called *block-sparse* [49], [50].

TABLE I
CONSTANTS USED IN THIS ARTICLE FOR RADAR SETTINGS

Both Doppler and PAWR		Only PAWR	
c	speed of light	d	antenna element spacing
λ	carrier wavelength	N	number of antenna elements
r	target range (distance from radar to target)	θ_{\min} θ_{\max}	bottom/top of angular interval where fan beam is transmitted
T	pulse repetition time (sampling period)	M	number of subintervals dividing $[\theta_{\min}, \theta_{\max}]$ for beamforming
L	number of pulses (number of samples)	K	number of point targets or subintervals where scatterers exist

(clouds and raindrops). For a fixed range r [m], the l th discrete-time sample $\hat{x}_l \in \mathbb{C}$ of the received signal is expressed as

$$\hat{x}_l := x_l + \varepsilon_l := x\left((l-1)T + \frac{2r}{c}\right) + \varepsilon_l \quad (l = 1, 2, \dots, L), \quad (1)$$

where c [m/s] is the speed of light, T [s] is the pulse repetition time satisfying $T \geq \frac{2r}{c}$, L is the number of the pulses, $x_l \in \mathbb{C}$ is the sample of the backscattered signal $x(t)$ generated by the pulse transmitted at $t = (l-1)T$, and $\varepsilon_l \in \mathbb{C}$ is additive white Gaussian noise of variance $E[|\varepsilon_l|^2] = \sigma_\varepsilon^2$. Normally, the total observation time LT is very short³ and we can consider $x(t)$ to be one realization of a continuous-time random process $X(t)$ that is *wide-sense stationary*,⁴ *ergodic*,⁵ zero-mean and zero-pseudoautocorrelation, i.e., $\mu_x := E[X(t)] = 0$ and $Q_x(\tau) := E[X(t+\tau)X(t)] = 0$. Moreover, its autocorrelation function $R_x(\tau) := E[X(t+\tau)\overline{X(t)}]$ satisfies $\int_{-\infty}^{\infty} |\tau R_x(\tau)| d\tau < \infty$.

Strictly speaking, $x(t)$ is expressed as the sum of individual backscattered signals in the scattering resolution volume [7]:

$$x(t) = \sum_s a_s(t) e^{j(2\pi f_s t + \phi_s)}, \quad (2)$$

where $a_s(t) \geq 0$ is the scattering amplitude of the s th particle, f_s [Hz] is called the Doppler frequency, which is independently generated from a *common probability density function* $q(f)$, and ϕ_s [rad] is called the phase shift, which is uniformly distributed from $[-\pi, \pi)$ under the assumption that precipitation profile is uniform within the scattering resolution volume. The *power spectral density* of the backscattered signal $x(t)$ is given by the Fourier transform of the autocorrelation function $R_x(\tau)$:

$$S(f) := \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f \tau} d\tau. \quad (3)$$

³For example, if the maximum range is $r = 60$ [km] and the speed of light is $c = 300$ [km/ms], the pulse repetition time in (1) should be set to $T = 0.4$ [ms] or more. In this situation, when the number of pulses is $L = 20$, the observation time is about $LT = 8$ [ms]. Raindrops fall at several meters per second, and thus they move only several centimeters during the observation.

⁴Let $X(t)$ be a continuous-time random process. $X(t)$ is called *wide-sense stationary* or *weak-sense stationary* if its mean and (pseudo)autocorrelation do not vary by shifts in the time t , i.e., $\forall t E[X(t)] = \int_{\mathbb{C}} xq(x|t) dx = \mu_x$ and $\forall t \forall \tau E[X(t+\tau)\overline{X(t)}] = \iint_{\mathbb{C}^2} x_1 \overline{x_2} q(x_1, x_2 | t + \tau, t) dx_1 dx_2 = R_x(\tau)$ and $E[X(t+\tau)X(t)] = \iint_{\mathbb{C}^2} x_1 x_2 q(x_1, x_2 | t + \tau, t) dx_1 dx_2 = Q_x(\tau)$, where $q(x|t)$ is the probability density function of $x(t)$ and $q(x_1, x_2 | t + \tau, t)$ is the joint probability density function of $x_1 := x(t + \tau)$ and $x_2 := x(t)$.

⁵A wide-sense stationary process $X(t)$ is called *ergodic* if any realization $x(t)$ of the random process $X(t)$ satisfies $\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt = \mu_x$ and $\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t + \tau) \overline{x(t)} dt = R_x(\tau)$ by the mean-square convergence.

Suppose that the amplitude $a_s(t)$ and the Doppler frequency f_s are independent, i.e., $\forall t \forall \tau E[\sum_s a_s(t + \tau) a_s(t)] \approx P$ holds, during the observation [7], where $P := R_x(0) = E[|x_l|^2]$ is the *average signal power*. Then, $R_x(\tau) \approx PE[e^{j2\pi f_s \tau}]$ holds, and the power spectral density $S(f)$ in (3) can be expressed as

$$S(f) \approx PE[\delta(f - f_s)] = Pq(f) = \frac{\lambda P}{2} q_v\left(\frac{\lambda f}{2}\right). \quad (4)$$

In (4), δ is the delta function, λ [m] is the carrier wavelength, and $q_v : \mathbb{R} \rightarrow \mathbb{R}_+$ is a common probability density function for each Doppler velocity $v_s = \frac{\lambda f_s}{2}$ [m/s], where the direction of approach to the radar is defined as the positive direction of the Doppler velocity. From (4), we can see that the power spectral density $S(f)$ is almost proportional to the Doppler frequency distribution $q(f)$ [7]. On the frequency f , we customarily consider $\frac{S(f)}{P}$ as its probability density $q(f)$, and we call its mean

$$\mu_f := \int_{-\infty}^{\infty} f \frac{S(f)}{P} df \approx \int_{-\infty}^{\infty} f q(f) df \quad (5)$$

the *mean Doppler frequency* and its standard deviation

$$\sigma_f := \sqrt{\int_{-\infty}^{\infty} (f - \mu_f)^2 \frac{S(f)}{P} df} \approx \sqrt{\int_{-\infty}^{\infty} (f - \mu_f)^2 q(f) df} \quad (6)$$

the *Doppler frequency width*. Similarly, we call $\mu_v = \frac{\lambda \mu_f}{2}$ the *mean Doppler velocity* and call $\sigma_v = \frac{\lambda \sigma_f}{2}$ the *Doppler velocity width*. The main purpose of the Doppler weather radars is to estimate P , μ_f (or μ_v) and σ_f (or σ_v) from the measurements \hat{x}_l in (1), since these parameters correspond to the amount of rainfall, the mean wind speed and the variation of wind speeds.

As shown in (4), $S(f)$ has information about P and $q(f)$. However, the backscattered signal $x(t)$ is observed in discrete-time, and we have to consider the power spectral density of x_l that is given by the discrete-time Fourier transform of $R_x(lT)$:

$$S^{(d)}(f) := \sum_{l=-\infty}^{\infty} R_x(lT) e^{-j2\pi f lT} = \frac{1}{T} \sum_{n=-\infty}^{\infty} S\left(f - \frac{n}{T}\right). \quad (7)$$

If $S(f) = 0$ for $|f| \geq \frac{1}{2T}$, $S(f)$ can be perfectly reconstructed from $S^{(d)}(f)$. Furthermore, we have to estimate $S^{(d)}(f)$ in (7) from finite samples x_l ($l = 1, 2, \dots, L$), where we assume L is *even* and $\varepsilon_l = 0$ in (1) for simplicity but the discussion can be extended to noisy samples \hat{x}_l with *odd* L . To estimate $S^{(d)}(f)$, we compute the normalized DFT coefficients $u_k \in \mathbb{C}$ of x_l by

$$u_k := \frac{1}{\sqrt{L}} \sum_{l=1}^L x_l e^{-j \frac{2\pi(k-1-L/2)(l-1)}{L}} \quad (k = 1, 2, \dots, L). \quad (8)$$

Then, we compute a *periodogram* as a series of the powers of the normalized DFT coefficients u_k :

$$|u_k|^2 = \frac{1}{L} \left| \sum_{l=1}^L x_l e^{-j \frac{2\pi(k-1-L/2)(l-1)}{L}} \right|^2 \quad (k = 1, 2, \dots, L). \quad (9)$$

The expected value of the periodogram $(|u_k|^2)_{k=1}^L$ is given by

$$\begin{aligned} E[|u_k|^2] &= \frac{1}{L} E \left[\sum_{l_1=1}^L \sum_{l_2=1}^L x_{l_1} \bar{x}_{l_2} e^{-j \frac{2\pi(k-1-L/2)(l_1-l_2)}{L}} \right] \\ &= \frac{1}{L} \sum_{l_1=1}^L \sum_{l_2=1}^L E[x_{l_1} \bar{x}_{l_2}] e^{-j \frac{2\pi(k-1-L/2)(l_1-l_2)}{L}} \\ &= \frac{1}{L} \sum_{l_1=1}^L \sum_{l_2=1}^L R_x((l_1-l_2)T) e^{-j \frac{2\pi(k-1-L/2)(l_1-l_2)}{L}} \\ &= \sum_{l=-(L-1)}^{L-1} \left(1 - \frac{|l|}{L}\right) R_x(lT) e^{-j \frac{2\pi(k-1-L/2)l}{L}}. \quad (10) \end{aligned}$$

From (7), (10), and $\sum_{l=-\infty}^{\infty} |lR_x(lT)| < \infty$, we have

$$\lim_{L \rightarrow \infty} E[|u_k|^2] = S^{(d)} \left(\frac{k-1-L/2}{LT} \right). \quad (11)$$

Therefore, the periodogram is often used as a nonparametric estimator for $S^{(d)}(f)$. Moreover, the parameters P , μ_f and σ_f can be estimated from the periodogram $(|u_k|^2)_{k=1}^L$, e.g., by

$$\left. \begin{aligned} \hat{P} &= \frac{1}{L} \sum_{l=1}^L |x_l|^2 = \frac{1}{L} \sum_{k=1}^L |u_k|^2 \\ \hat{\mu}_f &= \frac{1}{\hat{P}L} \sum_{k=k_{\text{peak}}-L/2}^{k_{\text{peak}}+L/2-1} \frac{k-1-L/2}{LT} |u_{k_{\text{mod}}L}|^2 \\ \hat{\sigma}_f &= \sqrt{\frac{1}{\hat{P}L} \sum_{k=k_{\text{peak}}-L/2}^{k_{\text{peak}}+L/2-1} \left(\frac{k-1-L/2}{LT} - \hat{\mu}_f \right)^2 |u_{k_{\text{mod}}L}|^2} \end{aligned} \right\} \quad (12)$$

in [51], where $k_{\text{peak}} := \arg\max_k |u_k|$, $k_{\text{mod}}L := k - \lfloor \frac{k-1}{L} \rfloor L$, and $\lfloor \cdot \rfloor$ is the floor function. In (12), $\hat{\mu}_f$ and $\hat{\sigma}_f$ are calculated by unfolding the frequency f in (7) wrapped below the Nyquist frequency $\frac{1}{2T}$. In most cases, the Doppler frequency width σ_f is smaller than the sampling frequency $\frac{1}{T}$, and thus the Doppler frequency close to the Nyquist frequency is properly measured.

Note that if pulse compression by frequency modulation is utilized as in [52], [53], then we treat the received signal after matched filtering with the carrier frequency shift as \hat{x}_l in (1). Differently from the basic case of transmitting truncated sinusoidal pulses, the pulse compression causes some dependence between the received signals \hat{x}_l of adjacent range bins. In this article, we suppose that the influence of this dependence is not dominant, and hence \hat{x}_l at each range is processed separately.

B. Observation Model in Phased Array Weather Radar

First of all, we show the observation model for K point targets. Let a PAWR have an N -element uniform linear array with the inter-element spacing d [m]. A plane wave signal scattered from the κ th point target hits on the array antenna at an unknown angle $\theta_\kappa^* \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, where $\theta_1^* < \theta_2^* < \dots < \theta_K^*$ [rad]. The l th time sample $\mathbf{y}_l \in \mathbb{C}^N$ of the received signal is given by

$$\mathbf{y}_l = \sum_{\kappa=1}^K x_{\kappa,l}^* \mathbf{a}(\theta_\kappa^*) + \boldsymbol{\varepsilon}_l \quad (l = 1, 2, \dots, L), \quad (13)$$

where $x_{\kappa,l}^* \in \mathbb{C}$ is the l th time sample of the κ th plane wave signal, $\mathbf{a}(\theta_\kappa^*) \in \mathbb{C}^N$ is the so-called *steering vector* defined by

$$\mathbf{a}(\theta) := \left(1, e^{-j \frac{2\pi d \sin \theta}{\lambda}}, e^{-j \frac{4\pi d \sin \theta}{\lambda}}, \dots, e^{-j \frac{2(N-1)\pi d \sin \theta}{\lambda}} \right)^\top, \quad (14)$$

and $\boldsymbol{\varepsilon}_l \in \mathbb{C}^N$ is additive white Gaussian noise, independent of the signals $x_{\kappa,l}^*$, of covariance matrix $\mathbf{R}_\boldsymbol{\varepsilon} := E[\boldsymbol{\varepsilon}_l \boldsymbol{\varepsilon}_l^H] = \sigma_\boldsymbol{\varepsilon}^2 \mathbf{I}_N$.

On the other hand, our targets such as raindrops are called *distributed targets* [19], that are supposed to exist continuously (strictly speaking, a sufficient number of raindrops exist within the antenna beamwidth). In this case, the number K of scatterers (raindrops) is too large and the observation model in (13) is difficult to use. Instead, we approximate the observation model by dividing the whole angular interval $[\theta_{\min}, \theta_{\max}] \subset [-\frac{\pi}{2}, \frac{\pi}{2}]$ into non-overlapping subintervals $[\theta_m - \frac{\Delta\theta}{2}, \theta_m + \frac{\Delta\theta}{2}]$ ($m = 1, 2, \dots, M$), where the number M of the subintervals is *user-defined*, $\Delta\theta := \frac{\theta_{\max} - \theta_{\min}}{M}$, and $\theta_m := \theta_{\min} + (m - \frac{1}{2})\Delta\theta$. As a result, we use the following discretized approximation model

$$\mathbf{y}_l = \sum_{m=1}^M x_{m,l} \mathbf{a}_m + \boldsymbol{\varepsilon}_l = \mathbf{A} \mathbf{x}_l + \boldsymbol{\varepsilon}_l \quad (l = 1, 2, \dots, L), \quad (15)$$

where $x_{m,l} \in \mathbb{C}$ is the l th sample of the sum of backscattered plane wave signals in the m th subinterval $[\theta_m - \frac{\Delta\theta}{2}, \theta_m + \frac{\Delta\theta}{2}]$, $\mathbf{x}_l := (x_{1,l}, x_{2,l}, \dots, x_{M,l})^\top \in \mathbb{C}^M$, $\mathbf{a}_m := \mathbf{a}(\theta_m) \in \mathbb{C}^N$, and $\mathbf{A} := [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_M] \in \mathbb{C}^{N \times M}$. Note that the signal x_l in (1) corresponds to $x_{m,l}$ in (15) for a specific elevation angle θ_m , and the noise vector $\boldsymbol{\varepsilon}_l$ in (15) includes model errors due to the discretization. These model errors become smaller as M increases, i.e., as $\Delta\theta$ decreases, but it is difficult to predict the accurate amount of the model errors. Since standard Doppler weather radars perform observations with about 1 [deg] beamwidth, we consider that the model errors have little effect on weather observations if $\Delta\theta \leq 1$ [deg]. Therefore, we suppose that $\Delta\theta \leq 1$ [deg] and $\boldsymbol{\varepsilon}_l$ is still additive white Gaussian noise independent of \mathbf{x}_l . In the PAWR system, θ_m denotes the m th elevation angle, and the average signal power vector

$$\begin{aligned} \mathbf{p} &:= (P_1, P_2, \dots, P_M)^\top \\ &:= (E[|x_{1,l}|^2], E[|x_{2,l}|^2], \dots, E[|x_{M,l}|^2])^\top \in \mathbb{R}_+^M \quad (16) \end{aligned}$$

corresponds to precipitation profile along the elevation axis.

Even if pulse compression is used in the PAWR, we can treat the received signal after matched filtering with the carrier frequency shift as \mathbf{y}_l in (15). In this case, \mathbf{y}_l and \mathbf{x}_l contain some dependence in the range direction. As in Section II-A, we suppose the influence of this dependence is small, and the received signal \mathbf{y}_l at each range is processed separately in the following.

III. CONVENTIONAL METHODS: LINEAR BEAMFORMING

The beamforming is a reconstruction problem of \mathbf{x}_l in (15). Major beamforming methods [9]–[19] estimate \mathbf{x}_l by multiplying complex weights $\mathbf{w}_m \in \mathbb{C}^N$ ($m = 1, 2, \dots, M$) and \mathbf{y}_l as

$$\begin{aligned} \hat{\mathbf{x}}_l &:= (\hat{x}_{1,l}, \hat{x}_{2,l}, \dots, \hat{x}_{M,l})^\top \\ &:= (\mathbf{w}_1^H \mathbf{y}_l, \mathbf{w}_2^H \mathbf{y}_l, \dots, \mathbf{w}_M^H \mathbf{y}_l)^\top = \mathbf{W}^H \mathbf{y}_l. \quad (17) \end{aligned}$$

In this article, we call the methodology based on (17) the *linear beamforming*. Note that the least squares (LS) method

$$\hat{\mathbf{x}}_{\text{LS},l} := \mathbf{W}_{\text{LS}}^{\text{H}} \mathbf{y}_l := \mathbf{A}^\dagger \mathbf{y}_l \quad (18)$$

does not necessarily work well, even if $N \geq M$, since \mathbf{A} is *ill-conditioned* when $\Delta\theta$ is smaller than the antenna beamwidth that is determined by the antenna size, where $\mathbf{A}^\dagger \in \mathbb{C}^{M \times N}$ is the Moore–Penrose pseudoinverse of \mathbf{A} . In the following, we explain three linear methods, *Fourier (FR) beamforming* [11], *Capon (CP) beamforming* [12], and *MMSE beamforming* [19].

A. FR Beamforming

FR beamforming [11] is the most basic linear method and its complex weight vector is defined by

$$\mathbf{w}_{\text{FR},m} := \frac{\mathbf{a}_m}{N} \quad (19)$$

independently of \mathbf{y}_l . The vector $\mathbf{w}_{\text{FR},m}$ is equal to a *matched filter* that maximizes the signal-to-noise ratio $\frac{E[|x_{m,l} \mathbf{w}_{\text{FR},m}^{\text{H}} \mathbf{a}_m|^2]}{E[|\mathbf{w}_{\text{FR},m}^{\text{H}} \boldsymbol{\varepsilon}_l|^2]}$. If $N \geq M$ and there exists some positive integer n satisfying $\forall m \frac{d}{\lambda} (\sin \theta_{m+1} - \sin \theta_m) \approx \frac{n}{N}$, then FR beamforming can reconstruct the signal x_l in the manner of the inverse DFT, but this ideal condition is usually not met. Actually, in the typical case, the average signal power P_m is often overestimated [19]. To explain this fact, we derive the observation model as in (13) from (15) by redefining $K (\leq M)$ as the number of the sub-intervals $[\theta_m - \frac{\Delta\theta}{2}, \theta_m + \frac{\Delta\theta}{2}]$ in which scatterers (raindrops) exist, and θ_κ^* as the center of such a subinterval. Then, from

$$\hat{\mathbf{x}}_{\text{FR},m,l} := \frac{\mathbf{a}_m^{\text{H}}}{N} \mathbf{y}_l = x_{m,l} + \frac{1}{N} \sum_{\theta_\kappa^* \neq \theta_m} x_{\kappa,l}^* \mathbf{a}_m^{\text{H}} \mathbf{a}(\theta_\kappa^*) + \frac{\mathbf{a}_m^{\text{H}} \boldsymbol{\varepsilon}_l}{N}, \quad (20)$$

the signal power $P_m = E[|x_{m,l}|^2]$ is overestimated by (12):

$$\hat{P}_{\text{FR},m} = \frac{1}{L} \sum_{l=1}^L |\hat{\mathbf{x}}_{\text{FR},m,l}|^2 \gg E[|x_{m,l}|^2] + \frac{\sigma_\varepsilon^2}{N} \quad (21)$$

for many elevation angles θ_m due to the influence of *sidelobes*, i.e., the influence of $\theta_\kappa^* (\neq \theta_m)$ satisfying $|\mathbf{a}_m^{\text{H}} \mathbf{a}(\theta_\kappa^*)| \not\approx 0$.

B. CP Beamforming

CP beamforming [12], that is also known as MVDR beamforming [18], is a data-dependent method, differently from FR beamforming. To avoid the overestimation, CP beamforming minimizes the signal power $\frac{1}{L} \sum_{l=1}^L |\hat{x}_{m,l}|^2 = \mathbf{w}_m^{\text{H}} \hat{\mathbf{R}}_y \mathbf{w}_m$ under the condition $\mathbf{w}_m^{\text{H}} \mathbf{a}_m = 1$, where $\hat{\mathbf{R}}_y := \frac{1}{L} \sum_{l=1}^L \mathbf{y}_l \mathbf{y}_l^{\text{H}} \in \mathbb{C}^{N \times N}$ denotes the sample covariance matrix of the zero-mean random vector \mathbf{y}_l . The weight vector $\mathbf{w}_{\text{CP},m}$ is defined as the solution to the following optimization problem

$$\underset{\mathbf{w}_m}{\text{minimize}} \quad \mathbf{w}_m^{\text{H}} \hat{\mathbf{R}}_y \mathbf{w}_m \quad \text{subject to} \quad \mathbf{w}_m^{\text{H}} \mathbf{a}_m = 1, \quad (22)$$

and given by

$$\mathbf{w}_{\text{CP},m} := \frac{\hat{\mathbf{R}}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^{\text{H}} \hat{\mathbf{R}}_y^{-1} \mathbf{a}_m} \quad (23)$$

if $L \geq N$ (strictly speaking, if $\text{rank}(\hat{\mathbf{R}}_y) = N$). In particular, if $N \geq K + 1$ and L is sufficiently large, then we have

$$\hat{P}_{\text{CP},m} := \frac{1}{L} \sum_{l=1}^L |\hat{x}_{\text{CP},m,l}|^2 \approx E[|x_{m,l}|^2] + \sigma_\varepsilon^2 \|\mathbf{w}_{\text{CP},m}\|_2^2 \quad (24)$$

for all m since $\forall \theta_\kappa^* \neq \theta_m \quad |\mathbf{w}_{\text{CP},m}^{\text{H}} \mathbf{a}(\theta_\kappa^*)| \approx 0$ holds. However, if L is not so large, the average signal powers P_m are underestimated [19]. Moreover if $L < N$, $\hat{\mathbf{R}}_y^{-1}$ cannot be computed.

C. MMSE Beamforming

The MMSE method is often used in communications [54]–[60] to recover x_l by (17), and it solves the following problem

$$\underset{\mathbf{w}_m}{\text{minimize}} \quad E[|x_{m,l} - \mathbf{w}_m^{\text{H}} \mathbf{y}_l|^2]. \quad (25)$$

The optimal solution to this problem is given by

$$\mathbf{W}_{\text{op}} := [\mathbf{w}_{\text{op},1} \quad \mathbf{w}_{\text{op},2} \quad \cdots \quad \mathbf{w}_{\text{op},M}] := \mathbf{R}_y^{-1} \mathbf{A} \mathbf{R}_x, \quad (26)$$

where x_l or its covariance matrix $\mathbf{R}_x := E[x_l x_l^{\text{H}}] \in \mathbb{C}^{M \times M}$ is *known*⁶ in the communications scenario. In the latter case, if the matrix \mathbf{A} is also *known*, then the covariance matrix $\mathbf{R}_y := E[\mathbf{y}_l \mathbf{y}_l^{\text{H}}] \in \mathbb{C}^{N \times N}$ of \mathbf{y}_l is given by $\mathbf{R}_y = \mathbf{A} \mathbf{R}_x \mathbf{A}^{\text{H}} + \sigma_\varepsilon^2 \mathbf{I}_N$.

Blunt *et al.* [23] extended the MMSE method in (26) to the DOA estimation for signals $x_{\kappa,l}^*$ from point targets, where each angle θ_κ^* in (13) is estimated as a discretized angle θ_m in (15). x_l and \mathbf{R}_x are *unknown* in this scenario, but we can suppose $E[x_{m_1,l} \bar{x}_{m_2,l}] = E[x_{m_1,l}] E[\bar{x}_{m_2,l}] = 0$ ($m_1 \neq m_2$). Hence, \mathbf{R}_x is a diagonal matrix $\mathbf{R}_x = \text{diag}(\mathbf{p})$ and approximated by $\hat{\mathbf{R}}_x \odot \mathbf{I}_M := (\frac{1}{L} \sum_{l=1}^L x_l x_l^{\text{H}}) \odot \mathbf{I}_M$, where \odot is the Hadamard product. The signal vector x_l is estimated, from the initial estimate $\hat{x}_{\text{DOA},l}^{(0)} = \hat{\mathbf{x}}_{\text{FR},l} = \mathbf{A}^{\text{H}} \mathbf{y}_l / N$, by iteratively computing

$$\begin{cases} \mathbf{R}_x^{(i)} = \left(\frac{1}{L} \sum_{l=1}^L \hat{x}_{\text{DOA},l}^{(i)} \hat{x}_{\text{DOA},l}^{(i)\text{H}} \right) \odot \mathbf{I}_M \\ \mathbf{W}_{\text{DOA}}^{(i+1)} = (\mathbf{A} \mathbf{R}_x^{(i)} \mathbf{A}^{\text{H}} + \sigma_\varepsilon^2 \mathbf{I}_N)^{-1} \mathbf{A} \mathbf{R}_x^{(i)} \\ \hat{x}_{\text{DOA},l}^{(i+1)} = \mathbf{W}_{\text{DOA}}^{(i+1)\text{H}} \mathbf{y}_l \quad (l = 1, 2, \dots, L) \end{cases} \quad (27)$$

for $i \geq 0$ until $\hat{x}_{\text{DOA},l}$ converges. The signal power P_m is estimated by $\hat{P}_{\text{DOA},m} = \frac{1}{L} \sum_{l=1}^L |\hat{x}_{\text{DOA},m,l}|^2$ ($m = 1, 2, \dots, M$), and each angle θ_κ^* is estimated as a peak position of $\hat{P}_{\text{DOA},m}$.

Differently from the DOA estimation where we detect peak positions of P_m , the estimation accuracy of x_l and P_m is important in the beamforming. In this point of view, large peaks of $\hat{P}_{\text{DOA},m}$ are stable even for small L while small peaks are unstable for dense scatterers with small L . To improve the accuracy, Yoshikawa *et al.* [19] proposed MMSE beamforming by adding a directional gain constraint to the method in (27) as in CP beamforming in (23) and the adaptive pulse compression in [61]. MMSE beamforming considers the following problem

$$\underset{\mathbf{w}_m}{\text{minimize}} \quad E[|x_{m,l} - \mathbf{w}_m^{\text{H}} \mathbf{y}_l|^2] \quad \text{subject to} \quad \mathbf{w}_m^{\text{H}} \mathbf{a}_m = 1, \quad (28)$$

⁶In the former case, the observation matrix \mathbf{A} is *unknown*.

and the exact optimal solution to this problem is given by

$$\mathbf{w}_{\text{MMSE},m} := \frac{\mathbf{R}_y^{-1} \mathbf{a}_m}{\mathbf{a}_m^H \mathbf{R}_y^{-1} \mathbf{a}_m}. \quad (29)$$

Based on (27) and (29), the signal vector \mathbf{x}_l is estimated, from the initial estimate $\hat{\mathbf{x}}_{\text{MMSE},l}^{(0)} = \hat{\mathbf{x}}_{\text{FR},l}$, by iteratively computing

$$\begin{cases} \mathbf{R}_x^{(i)} = \left(\frac{1}{L} \sum_{l=1}^L \hat{\mathbf{x}}_{\text{MMSE},l}^{(i)} \hat{\mathbf{x}}_{\text{MMSE},l}^{(i)H} \right) \odot \mathbf{I}_M \\ \mathbf{R}_y^{(i)} = \mathbf{A} \mathbf{R}_x^{(i)} \mathbf{A}^H + \sigma_\varepsilon^2 \mathbf{I}_N \\ \mathbf{w}_{\text{MMSE},m}^{(i+1)} = \frac{\mathbf{R}_y^{(i)-1} \mathbf{a}_m}{\mathbf{a}_m^H \mathbf{R}_y^{(i)-1} \mathbf{a}_m} \quad (m = 1, 2, \dots, M) \\ \hat{\mathbf{x}}_{\text{MMSE},l}^{(i+1)} = \mathbf{W}_{\text{MMSE}}^{(i+1)H} \mathbf{y}_l \quad (l = 1, 2, \dots, L) \end{cases} \quad (30)$$

for $i \geq 0$ until $\hat{\mathbf{x}}_{\text{MMSE},l}$ converges. Differently from CP beamforming, \mathbf{R}_y and \mathbf{x}_l can be stably estimated even for small L .

IV. NONLINEAR BEAMFORMING VIA CONVEX OPTIMIZATION BASED ON GROUP-SPARSITIES

An essential problem for the linear beamforming in (17) is that when the number K of subintervals including scatterers is close to N or larger than N , the estimation accuracy degrades, i.e., fine fluctuations of the average signal powers P_m in (16) cannot be captured. This is because each weight vector $\mathbf{w}_m \in \mathbb{C}^N$ can only create at most $N-1$ null directions and the influence of sidelobes occurs. In this article, to eliminate the sidelobes and improve the spatial resolution, we consider the use of certain *sparsities* in optimization problems as in [28]–[44].

In radar applications, the sparsity has been used mainly for point targets since the signal to be estimated can itself be regarded as sparse [28]–[38]. On the other hand, in this article, we show that the signal from distributed targets (raindrops) can be regarded as *group-sparse in the elevation-velocity domain*, but its appropriate *non-overlapping group partition* is unknown in advance. Instead of finding the non-overlapping groups, we use *small overlapping groups* as with *latent group lasso* [45]–[48], and evaluate the group-sparsities by convex regularizers.

In this section, we propose a *nonlinear beamforming* method based on *convex optimization*. First, we gather \mathbf{x}_l and \mathbf{y}_l into $\mathbf{X} := [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_L] = [\tilde{\mathbf{x}}_1 \ \tilde{\mathbf{x}}_2 \ \dots \ \tilde{\mathbf{x}}_L]^T \in \mathbb{C}^{M \times L}$ and $\mathbf{Y} := [\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \ \mathbf{y}_L] \in \mathbb{C}^{N \times L}$, respectively, where $\tilde{\mathbf{x}}_m := (x_{m,1}, x_{m,2}, \dots, x_{m,L})^T \in \mathbb{C}^L$. Then the beamforming can be translated into a reconstruction problem of \mathbf{X} from \mathbf{Y} , and the data-fidelity for the measurements \mathbf{y}_l in (15) is evaluated by

$$\sum_{l=1}^L \|\mathbf{y}_l - \mathbf{A} \mathbf{x}_l\|_2^2 = \|\mathbf{Y} - \mathbf{A} \mathbf{X}\|_{\text{F}}^2 = \|\mathbf{Y}^T - \mathbf{X}^T \mathbf{A}^T\|_{\text{F}}^2. \quad (31)$$

For each discrete-time backscattered signal $\tilde{\mathbf{x}}_m$ at θ_m , define its normalized DFT coefficients $u_{k,m} \in \mathbb{C}$ ($k = 1, 2, \dots, L$) as in (8). Let $\mathbf{F} := \frac{1}{\sqrt{L}} [\mathbf{f}_{-\frac{L}{2}} \ \mathbf{f}_{-\frac{L}{2}+1} \ \dots \ \mathbf{f}_{\frac{L}{2}-1}]^T \in \mathbb{C}^{L \times L}$ be the normalized discrete Fourier transform matrix, where $\mathbf{f}_i := (1, e^{-j\frac{2\pi i}{L}}, e^{-j\frac{4\pi i}{L}}, \dots, e^{-j\frac{2(L-1)\pi i}{L}})^T \in \mathbb{C}^L$. By using \mathbf{F} , the DFT coefficient vector $\mathbf{u}_m := (u_{1,m}, u_{2,m}, \dots, u_{L,m})^T \in \mathbb{C}^L$

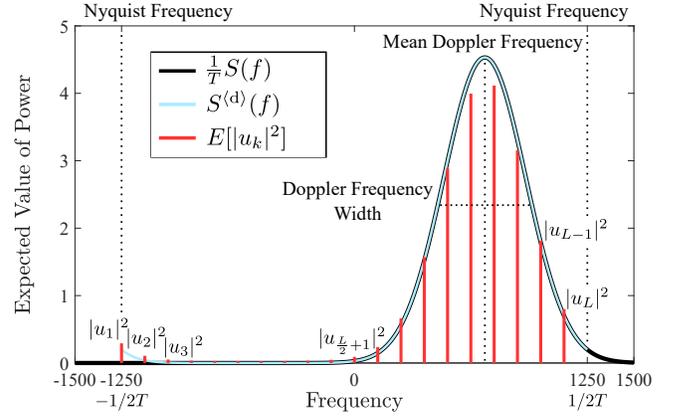


Fig. 1. Relation between the power spectral density $S(f)$ of $x(t)$, the power spectral density $S^{(d)}(f)$ of x_l , and the periodogram $(|u_k|^2)_{k=1}^L$, where the Gaussian function in (34) is used as the Doppler frequency distribution with $P = 1$, $\mu_f = 700$ [Hz], $\sigma_f = 220$ [Hz], $T = 4 \times 10^{-4}$ [s], and $L = 20$. If we set the carrier wavelength to $\lambda = 3.18 \times 10^{-2}$ [m], then $\sigma_v \approx 3.5$ [m/s]. Black, blue, and red lines depict $\frac{1}{T}S(f)$, $S^{(d)}(f)$, and $E[|u_k|^2]$, respectively.

is expressed as $\mathbf{u}_m = \mathbf{F} \tilde{\mathbf{x}}_m$ ($m = 1, 2, \dots, M$). Furthermore, we define the DFT coefficient matrix $\mathbf{U} = (u_{k,m}) \in \mathbb{C}^{L \times M}$ by

$$\mathbf{U} := [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_M] = \mathbf{F} \mathbf{X}^T. \quad (32)$$

The reconstruction of \mathbf{X} and that of \mathbf{U} are essentially equivalent due to the unitarity of \mathbf{F} , and the weather parameters can be computed from \mathbf{U} by (12). Thus in this article, we consider the beamforming as a *reconstruction problem of \mathbf{U} from \mathbf{Y}* . We have $\mathbf{X}^T = \mathbf{F}^H \mathbf{U}$, and (31) can be transformed into

$$\|\mathbf{Y}^T - \mathbf{X}^T \mathbf{A}^T\|_{\text{F}}^2 = \|\mathbf{Y}^T - \mathbf{F}^H \mathbf{U} \mathbf{A}^T\|_{\text{F}}^2 = \|\mathbf{F} \mathbf{Y}^T - \mathbf{U} \mathbf{A}^T\|_{\text{F}}^2 \quad (33)$$

from the unitarity of \mathbf{F} . In the following, after describing two *group-sparsities* of \mathbf{U} , we estimate \mathbf{U} by solving convex optimization problems based on (33) and the two group-sparsities (see Footnote 2 for the meaning of the word “group-sparse”).

A. Narrow Bandwidth of Each Power Spectral Density

As mentioned in Section II-A, the power spectral density of the backscattered signal is almost proportional to the Doppler frequency distribution. In addition, the Doppler frequency distribution and the Doppler velocity distribution are essentially the same (see (4)). Thus, unless the Doppler velocity width σ_v is extremely large, the bandwidth of the power spectral density, similar to the Doppler frequency width σ_f , is narrow compared to the sampling frequency $\frac{1}{T}$ [Hz]. In particular, when the influence of the atmospheric turbulence is large,⁷ the power spectral density can be modeled by a Gaussian function [7], [62], [63]:

$$S(f) = \frac{P}{\sqrt{2\pi}\sigma_f} e^{-\frac{(f-\mu_f)^2}{2\sigma_f^2}} \approx Pq(f). \quad (34)$$

In this model, the autocorrelation function is exactly given by

$$R_x(\tau) = \int_{-\infty}^{\infty} S(f) e^{j2\pi f\tau} df = P e^{j2\pi\mu_f\tau} e^{-2\pi^2\sigma_f^2\tau^2}. \quad (35)$$

⁷See, e.g., [62] for a detailed relation between the atmospheric turbulence intensity and the Doppler velocity distribution at each elevation angle.

Figure 1 illustrates the relation between $S(f)$ in (3), $S^{(d)}(f)$ in (7), and the periodogram $(|u_k|^2)_{k=1}^L$ in (9) using (34) and (35). As shown in Fig. 1, the frequency indices k of large $E[|u_k|^2]$ are gathered in the vicinity of the mean Doppler frequency μ_f .

Since $E[|u_k|^2] \approx 0$ implies $u_k \approx 0$ and the above discussion holds for every elevation θ_m , each DFT coefficient vector \mathbf{u}_m will be *group-sparse* by an appropriate *non-overlapping group partition*, but the appropriate partition is *unknown* in advance. Instead, we divide \mathbf{u}_m into L *overlapping groups* of size b_f :

$$\left. \begin{aligned} \mathbf{z}_{1,m}^{(1)} &:= (u_{1,m}, u_{2,m}, \dots, u_{b_f,m})^T \in \mathbb{C}^{b_f} \\ \mathbf{z}_{2,m}^{(1)} &:= (u_{2,m}, u_{3,m}, \dots, u_{b_f+1,m})^T \in \mathbb{C}^{b_f} \\ &\vdots \\ \mathbf{z}_{L-b_f+1,m}^{(1)} &:= (u_{L-b_f+1,m}, u_{L-b_f+2,m}, \dots, u_{L,m})^T \in \mathbb{C}^{b_f} \\ \mathbf{z}_{L-b_f+2,m}^{(1)} &:= (u_{L-b_f+2,m}, \dots, u_{L,m}, u_{1,m})^T \in \mathbb{C}^{b_f} \\ &\vdots \\ \mathbf{z}_{L,m}^{(1)} &:= (u_{L,m}, u_{1,m}, \dots, u_{b_f-1,m})^T \in \mathbb{C}^{b_f} \end{aligned} \right\} \quad (36)$$

under the periodic boundary condition. Note that we connect $u_{L,m}$ and $u_{1,m}$ in $\mathbf{z}_{k,m}^{(1)}$ ($k = L - b_f + 2, L - b_f + 3, \dots, L$) since the frequency components of the discrete-time signal $\tilde{\mathbf{x}}_m$ are wrapped below the Nyquist frequency as shown in Fig. 1. By arranging all $\mathbf{z}_{k,m}^{(1)}$ ($k = 1, 2, \dots, L$ and $m = 1, 2, \dots, M$), we define a matrix

$$\mathbf{Z}_1 := \begin{bmatrix} \mathbf{z}_{1,1}^{(1)} & \mathbf{z}_{1,2}^{(1)} & \cdots & \mathbf{z}_{1,M}^{(1)} \\ \mathbf{z}_{2,1}^{(1)} & \mathbf{z}_{2,2}^{(1)} & \cdots & \mathbf{z}_{2,M}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{z}_{L,1}^{(1)} & \mathbf{z}_{L,2}^{(1)} & \cdots & \mathbf{z}_{L,M}^{(1)} \end{bmatrix} \in \mathbb{C}^{b_f L \times M}. \quad (37)$$

Then, $\mathbf{z}_{k,m}^{(1)} \approx \mathbf{0}$ holds for many k and m . Therefore, the matrix \mathbf{Z}_1 will be *group-sparse* by a *non-overlapping partition* $G_1 := (\mathcal{G}_{k,m}^{(1)})$ s.t. $\mathcal{G}_{k,m}^{(1)} := \{(b_f(k-1) + i, m)\}_{i=1}^{b_f}$. Moreover, there is a simple matrix $\mathbf{B}_f \in \{0, 1\}^{b_f L \times L}$ satisfying $\mathbf{B}_f \mathbf{U} = \mathbf{Z}_1$. This property can be evaluated by a weighted group ℓ_1 -norm

$$\|\mathbf{Z}_1\|_{1, \Xi_1}^{G_1} := \sum_{k=1}^L \sum_{m=1}^M \xi_{k,m}^{(1)} \|\mathbf{z}_{k,m}^{(1)}\|_2, \quad (38)$$

where $\Xi_1 := (\xi_{k,m}^{(1)}) \in \mathbb{R}_+^{L \times M}$ is a nonnegative weight matrix.

B. Similarity between the Adjacent Power Spectral Densities

In the previous section, we focused on the narrow bandwidth of the power spectral density for each elevation angle θ_m . In this section, we explain the similarity between the power spectral densities at adjacent elevation angles θ_m and θ_{m+1} . When $\Delta\theta$ is relatively small, i.e., M is set to a relatively large value to acquire the high-resolution precipitation profile, the mean Doppler velocity μ_v and the Doppler velocity width σ_v at θ_m and those at θ_{m+1} are expected to be similar. As a result, the power spectral density at θ_m and that at θ_{m+1} are similar.

Such a similarity also applies to the adjacent periodograms $(|u_{k,m}|^2)_{k=1}^L$ and $(|u_{k,m+1}|^2)_{k=1}^L$. For a fixed frequency index k , if $E[|u_{k,m}|^2]$ is large, then $E[|u_{k,m+1}|^2]$ will be also large. If $E[|u_{k,m}|^2] \approx 0$ holds, then $E[|u_{k,m+1}|^2] \approx 0$ will also hold. As a result, a vector $\tilde{\mathbf{u}}_k := (u_{k,1}, u_{k,2}, \dots, u_{k,M})^T \in \mathbb{C}^M$ will be *group-sparse* by an appropriate *non-overlapping group partition*, but it is unknown in advance. Instead, as in the previous section, we divide $\tilde{\mathbf{u}}_k$ into M *overlapping groups* of size b_e :

$$\left. \begin{aligned} \mathbf{z}_{1,k}^{(2)} &:= (u_{k,1}, u_{k,2}, \dots, u_{k,b_e})^T \in \mathbb{C}^{b_e} \\ \mathbf{z}_{2,k}^{(2)} &:= (u_{k,2}, u_{k,3}, \dots, u_{k,b_e+1})^T \in \mathbb{C}^{b_e} \\ &\vdots \\ \mathbf{z}_{M-b_e+1,k}^{(2)} &:= (u_{k,M-b_e+1}, u_{k,M-b_e+2}, \dots, u_{k,M})^T \in \mathbb{C}^{b_e} \\ \mathbf{z}_{M-b_e+2,k}^{(2)} &:= (u_{k,M-b_e+2}, \dots, u_{k,M}, u_{k,1})^T \in \mathbb{C}^{b_e} \\ &\vdots \\ \mathbf{z}_{M,k}^{(2)} &:= (u_{k,M}, u_{k,1}, \dots, u_{k,b_e-1})^T \in \mathbb{C}^{b_e} \end{aligned} \right\} \quad (39)$$

under the periodic boundary condition. By arranging all $\mathbf{z}_{m,k}^{(2)}$ ($m = 1, 2, \dots, M$ and $k = 1, 2, \dots, L$), we define a matrix

$$\mathbf{Z}_2 := \begin{bmatrix} \mathbf{z}_{1,1}^{(2)} & \mathbf{z}_{1,2}^{(2)} & \cdots & \mathbf{z}_{1,L}^{(2)} \\ \mathbf{z}_{2,1}^{(2)} & \mathbf{z}_{2,2}^{(2)} & \cdots & \mathbf{z}_{2,L}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{z}_{M,1}^{(2)} & \mathbf{z}_{M,2}^{(2)} & \cdots & \mathbf{z}_{M,L}^{(2)} \end{bmatrix} \in \mathbb{C}^{b_e M \times L}. \quad (40)$$

Then $\mathbf{z}_{m,k}^{(2)} \approx \mathbf{0}$ holds for many m and k . Therefore, the matrix \mathbf{Z}_2 will be *group-sparse* by a *non-overlapping partition* $G_2 := (\mathcal{G}_{m,k}^{(2)})$ s.t. $\mathcal{G}_{m,k}^{(2)} := \{(b_e(m-1) + i, k)\}_{i=1}^{b_e}$, and there exists a simple matrix $\mathbf{B}_e \in \{0, 1\}^{b_e M \times M}$ satisfying $\mathbf{B}_e \mathbf{U}^T = \mathbf{Z}_2$. This property can be evaluated by a weighted group ℓ_1 -norm

$$\|\mathbf{Z}_2\|_{1, \Xi_2}^{G_2} := \sum_{m=1}^{M-b_e+1} \sum_{k=1}^L \xi_{m,k}^{(2)} \|\mathbf{z}_{m,k}^{(2)}\|_2, \quad (41)$$

where $\Xi_2 := (\xi_{m,k}^{(2)}) \in \mathbb{R}_+^{M \times L}$ is a nonnegative weight matrix s.t. $\xi_{m,k}^{(2)} := 0$ for $m = M - b_e + 2, M - b_e + 3, \dots, M$. Note that $E[|u_{k,M}|^2]$ and $E[|u_{k,1}|^2]$ are not necessarily similar, and hence we define the weights $\xi_{m,k}^{(2)}$ as zeros for the vectors $\mathbf{z}_{m,k}^{(2)}$ including both $u_{k,M}$ and $u_{k,1}$ not to evaluate these vectors. In addition, if we know in advance that there is almost no signal at particular elevations or frequencies, we can use this information by increasing only the particular weights $\xi_{k,m}^{(1)}$ and $\xi_{m,k}^{(2)}$.

Remark 1 (Difference from Our Previous Work [20]): In our previous work [20], we considered the beamforming as a reconstruction problem of \mathbf{X} , and used the same group ℓ_1 -norm $\|\mathbf{B}_f \mathbf{F} \mathbf{X}^T\|_{1, \Xi_1}^{G_1} = \|\mathbf{B}_f \mathbf{U}\|_{1, \Xi_1}^{G_1} = \|\mathbf{Z}_1\|_{1, \Xi_1}^{G_1}$ as in (38) for the narrow bandwidth of each power spectral density. On the other hand, we did not use the similarity of the adjacent power spectral densities but used the continuity of precipitation profile, i.e., we assumed that if $\tilde{\mathbf{x}}_m \approx \mathbf{0}$, then it is highly possible that

$\tilde{\mathbf{x}}_{m+1} \approx \mathbf{0}$. This was evaluated by a weighted group ℓ_1 -norm

$$\begin{aligned} \|\mathbf{B}_e \mathbf{X}\|_{1,\xi_0}^{G_0} &= \sum_{m=1}^{M-b_e+1} \xi_m^{(0)} \|(\tilde{\mathbf{x}}_m^T, \tilde{\mathbf{x}}_{m+1}^T, \dots, \tilde{\mathbf{x}}_{m+b_e-1}^T)^T\|_2 \\ &= \sum_{m=1}^{M-b_e+1} \xi_m^{(0)} \|(\mathbf{u}_m^T, \mathbf{u}_{m+1}^T, \dots, \mathbf{u}_{m+b_e-1}^T)^T\|_2 \\ &= \sum_{m=1}^{M-b_e+1} \xi_m^{(0)} \|(\mathbf{z}_{m,1}^{(2)T}, \mathbf{z}_{m,2}^{(2)T}, \dots, \mathbf{z}_{m,L}^{(2)T})^T\|_2 = \|\mathbf{Z}_2\|_{1,\xi_0}^{G_0}, \end{aligned} \quad (42)$$

where $G_0 := (\mathcal{G}_m^{(0)})$ s.t. $\mathcal{G}_m^{(0)} := \{(b_e(m-1) + i, k)\}_{k=1,2,\dots,L}^{i=1,2,\dots,b_e}$ and $\xi_0 := (\xi_m^{(0)}) \in \mathbb{R}_+^M$ is some vector s.t. $\xi_m^{(0)} = 0$ for $m = M - b_e + 2, M - b_e + 3, \dots, M$. Therefore, the current work evaluates the group-sparsity of \mathbf{Z}_2 in detail for each frequency by (41), while the previous work roughly evaluated it by collecting all frequency components. In addition, if $\|\mathbf{Z}_1\|_{1,\Xi_1}^{G_1}$ and $\|\mathbf{Z}_2\|_{1,\Xi_2}^{G_2}$ are small, then $\|\mathbf{Z}_2\|_{1,\xi_0}^{G_0}$ will be automatically small. Thus, we do not use our previous cost $\|\mathbf{Z}_2\|_{1,\xi_0}^{G_0}$ in this article.

C. Proposed Nonlinear Beamforming (Formulation I)

Based on the data-fidelity term in (33) and the two weighted group ℓ_1 -norms (38) and (41), we reconstruct \mathbf{U} from \mathbf{Y} by solving a convex optimization problem

$$\begin{aligned} &\underset{\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{F}\mathbf{Y}^T - \mathbf{U}\mathbf{A}^T\|_F^2 + \|\mathbf{Z}_1\|_{1,\Xi_1}^{G_1} + \|\mathbf{Z}_2\|_{1,\Xi_2}^{G_2} \\ &\text{subject to} \quad \mathbf{Z}_1 = \mathbf{B}_f \mathbf{U} \text{ and } \mathbf{Z}_2 = \mathbf{B}_e \mathbf{U}^T, \end{aligned} \quad (43)$$

with the use of ADMM [24]–[27] (see Appendix for details of ADMM). By defining two convex functions g and h in (78) as

$$g(\mathbf{Z}_1, \mathbf{Z}_2) := \|\mathbf{Z}_1\|_{1,\Xi_1}^{G_1} + \|\mathbf{Z}_2\|_{1,\Xi_2}^{G_2} \quad (44)$$

and

$$h(\mathbf{U}) := \frac{1}{2} \|\mathbf{F}\mathbf{Y}^T - \mathbf{U}\mathbf{A}^T\|_F^2, \quad (45)$$

the problem in (43) is expressed as an ADMM-form:

$$\begin{aligned} &\underset{\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2}{\text{minimize}} \quad g(\mathbf{Z}_1, \mathbf{Z}_2) + h(\mathbf{U}) \\ &\text{subject to} \quad (\mathbf{Z}_1, \mathbf{Z}_2) = \mathcal{L}(\mathbf{U}) := (\mathbf{B}_f \mathbf{U}, \mathbf{B}_e \circ \mathcal{T}(\mathbf{U})), \end{aligned} \quad (46)$$

where \mathcal{T} is the transpose operator, \circ denotes the composition of two mappings, and $\mathcal{L} : \mathbb{C}^{L \times M} \rightarrow \mathbb{C}^{b_f L \times M} \times \mathbb{C}^{b_e M \times L}$ is a linear mapping that maps \mathbf{U} to $(\mathbf{Z}_1, \mathbf{Z}_2)$ with \mathbf{B}_f and \mathbf{B}_e .

The ADMM iterations for (46) are given as follows. On the first line in (79), since the function $g(\mathbf{Z}_1, \mathbf{Z}_2)$ is divided into two parts $\|\mathbf{Z}_1\|_{1,\Xi_1}^{G_1}$ and $\|\mathbf{Z}_2\|_{1,\Xi_2}^{G_2}$, \mathbf{Z}_1 and \mathbf{Z}_2 are updated by

$$\left. \begin{aligned} \mathbf{Z}_1^{(i)} &= \text{prox}_{\gamma \|\cdot\|_{1,\Xi_1}^{G_1}} (\mathbf{B}_f \mathbf{U}^{(i)} + \mathbf{V}_1^{(i)}) \\ \mathbf{Z}_2^{(i)} &= \text{prox}_{\gamma \|\cdot\|_{1,\Xi_2}^{G_2}} (\mathbf{B}_e \mathbf{U}^{(i)T} + \mathbf{V}_2^{(i)}) \end{aligned} \right\} \quad (47)$$

with the use of the *proximity operators* of the group ℓ_1 -norms in (81), where $\gamma > 0$ and $\mathbf{V}_1 \in \mathbb{C}^{b_f L \times M}$ and $\mathbf{V}_2 \in \mathbb{C}^{b_e M \times L}$ are dual variables. Then, on the second line in (79), we have

$$\left. \begin{aligned} \tilde{\mathbf{V}}_1^{(i)} &= \mathbf{V}_1^{(i)} + \mathbf{B}_f \mathbf{U}^{(i)} - \mathbf{Z}_1^{(i)} \\ \tilde{\mathbf{V}}_2^{(i)} &= \mathbf{V}_2^{(i)} + \mathbf{B}_e \mathbf{U}^{(i)T} - \mathbf{Z}_2^{(i)} \end{aligned} \right\}. \quad (48)$$

On the third line in (79), \mathbf{U} is updated as the unique solution to a least squares problem, and the solution $\mathbf{U}^{(i+1)}$ has to satisfy

$$\begin{aligned} &\gamma \mathbf{U}^{(i+1)} \mathbf{A}^T \bar{\mathbf{A}} + \mathcal{L}^* \circ \mathcal{L}(\mathbf{U}^{(i+1)}) \\ &= \gamma \mathbf{F}\mathbf{Y}^T \bar{\mathbf{A}} + \mathcal{L}^*(\mathbf{Z}_1^{(i)} - \tilde{\mathbf{V}}_1^{(i)}, \mathbf{Z}_2^{(i)} - \tilde{\mathbf{V}}_2^{(i)}), \end{aligned} \quad (49)$$

with the adjoint operator $\mathcal{L}^* : \mathbb{C}^{b_f L \times M} \times \mathbb{C}^{b_e M \times L} \rightarrow \mathbb{C}^{L \times M}$

$$\mathcal{L}^*(\mathbf{Z}_1, \mathbf{Z}_2) := \mathbf{B}_f^T \mathbf{Z}_1 + \mathcal{T}(\mathbf{B}_e^T \mathbf{Z}_2) = \mathbf{B}_f^T \mathbf{Z}_1 + \mathbf{Z}_2^T \mathbf{B}_e. \quad (50)$$

Moreover, the composite mapping $\mathcal{L}^* \circ \mathcal{L}$ can be computed by

$$\begin{aligned} \mathcal{L}^* \circ \mathcal{L}(\mathbf{U}) &= \mathcal{L}^*(\mathbf{B}_f \mathbf{U}, \mathbf{B}_e \mathbf{U}^T) \\ &= \mathbf{B}_f^T \mathbf{B}_f \mathbf{U} + \mathcal{T}(\mathbf{B}_e^T \mathbf{B}_e \mathbf{U}^T) \\ &= b_f \mathbf{I}_L \mathbf{U} + \mathcal{T}(b_e \mathbf{I}_M \mathbf{U}^T) = (b_f + b_e) \mathbf{U}. \end{aligned} \quad (51)$$

By substituting (51) into (49), we have

$$\begin{aligned} &\mathbf{U}^{(i+1)} (\gamma \mathbf{A}^T \bar{\mathbf{A}} + (b_f + b_e) \mathbf{I}_M) \\ &= \gamma \mathbf{F}\mathbf{Y}^T \bar{\mathbf{A}} + \mathbf{B}_f^T (\mathbf{Z}_1^{(i)} - \tilde{\mathbf{V}}_1^{(i)}) + (\mathbf{Z}_2^{(i)} - \tilde{\mathbf{V}}_2^{(i)})^T \mathbf{B}_e, \end{aligned} \quad (52)$$

and $\mathbf{U}^{(i+1)}$ is obtained by applying $(\gamma \mathbf{A}^T \bar{\mathbf{A}} + (b_f + b_e) \mathbf{I}_M)^{-1}$ to (52) from the right side. Finally, on the fourth line in (79), \mathbf{V}_1 and \mathbf{V}_2 are updated, with the use of $\rho^{(i+1)} \in [0, 2]$, by

$$\left. \begin{aligned} \mathbf{V}_1^{(i+1)} &= \tilde{\mathbf{V}}_1^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{B}_f \mathbf{U}^{(i+1)} - \mathbf{Z}_1^{(i)}) \\ \mathbf{V}_2^{(i+1)} &= \tilde{\mathbf{V}}_2^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{B}_e \mathbf{U}^{(i+1)T} - \mathbf{Z}_2^{(i)}) \end{aligned} \right\}. \quad (53)$$

By repeating (47), (48), (52) and (53) until a convergence condition is met, the solution to (46) is obtained as an estimate $\hat{\mathbf{U}}$.

As described in [27], convergence to the optimal solution is easily guaranteed by setting $\rho^{(i+1)}$ as a constant $\rho \in [1, 2)$. If $\rho^{(i+1)} = 1$ for all i , we can replace $\tilde{\mathbf{V}}_1^{(i)}$ and $\tilde{\mathbf{V}}_2^{(i)}$ with $\mathbf{V}_1^{(i+1)}$ and $\mathbf{V}_2^{(i+1)}$, respectively. If $\rho^{(i+1)} > 1$ for all i , we can accelerate the convergence compared to the case of $\rho^{(i+1)} = 1$.

D. Proposed Nonlinear Beamforming (Formulation II)

In (43), the matrices \mathbf{Z}_1 and \mathbf{Z}_2 are constructed from \mathbf{U} by applying \mathbf{B}_f and $\mathbf{B}_e \circ \mathcal{T}$, respectively, i.e., each component of \mathbf{U} is copied to b_f groups in \mathbf{Z}_1 and b_e groups in \mathbf{Z}_2 . Then, we compute the ℓ_2 -norm of each group in (38) and (41). Since the energy of each $u_{k,m}$ is spread over multiple groups, evaluation values for the group-sparsities of \mathbf{U} will be slightly inaccurate.

In the theoretical field of compressed sensing, a formulation called *latent group lasso* has been proposed to evaluate group-sparsities [45]–[48], where overlapping groups have been used in a slightly different way from Formulation I. To evaluate the group-sparsities of \mathbf{U} more accurately than (43), we consider another optimization problem based on the latent group lasso formulation in which the relation between \mathbf{U} and $(\mathbf{Z}_1, \mathbf{Z}_2)$ is reversed. Specifically, we also propose to solve a problem

$$\begin{aligned} &\underset{\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{F}\mathbf{Y}^T - \mathbf{U}\mathbf{A}^T\|_F^2 + \|\mathbf{Z}_1\|_{1,\Xi_1}^{G_1} + \|\mathbf{Z}_2\|_{1,\Xi_2}^{G_2} \\ &\text{subject to} \quad \mathbf{B}_f^T \mathbf{Z}_1 = \mathbf{U} \text{ and } \mathbf{B}_e^T \mathbf{Z}_2 = \mathbf{U}^T, \end{aligned} \quad (54)$$

where the weight matrix $\Xi_2 = (\xi_{m,k}^{(2)})$ satisfies $\xi_{m,k}^{(2)} = \infty$ for $m = M - b_e + 2, M - b_e + 3, \dots, M$, i.e., $\mathbf{z}_{m,k}^{(2)} = \mathbf{0}$ has to hold for $m = M - b_e + 2, M - b_e + 3, \dots, M$. Differently

from the problem in (43), \mathbf{U} and \mathbf{U}^T in (54) are constructed by applying \mathbf{B}_f^T and \mathbf{B}_e^T to \mathbf{Z}_1 and \mathbf{Z}_2 , respectively. Such a latent group lasso formulation can concentrate the energy of each $u_{k,m}$ on a few groups in \mathbf{Z}_1 and \mathbf{Z}_2 , although it requires more time to solve. Thus, the number of almost-zero vectors $\mathbf{z}_{k,m}^{(1)} \approx \mathbf{0}$ and $\mathbf{z}_{m,k}^{(2)} \approx \mathbf{0}$ increases, and the group-sparsity with *fixed-size overlapping groups* can be accurately evaluated [47].

To solve the problem in (54) with ADMM, we define five auxiliary variables $\mathbf{\Lambda}_1 \in \mathbb{C}^{b_f L \times M}$, $\mathbf{\Lambda}_2 \in \mathbb{C}^{b_e M \times L}$, and $\mathbf{\Gamma}_n \in \mathbb{C}^{L \times M}$ ($n = 1, 2, 3$). Then, define functions g and h in (78) as

$$g(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) := \|\mathbf{\Lambda}_1\|_{1, \Xi_1}^{G_1} + \|\mathbf{\Lambda}_2\|_{1, \Xi_2}^{G_2} + \iota(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) \quad (55)$$

and

$$h(\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2) := \frac{1}{2} \|\mathbf{F}\mathbf{Y}^T - \mathbf{U}\mathbf{A}^T\|_{\mathbf{F}}^2, \quad (56)$$

with the use of an *indicator function* ι s.t. $\iota(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) := 0$ if $\mathbf{\Gamma}_1 = \mathbf{\Gamma}_2 = \mathbf{\Gamma}_3$ and $\iota(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) := \infty$ otherwise. Moreover, by defining an linear mapping \mathcal{L} in (78) as

$$\mathcal{L}(\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2) := (\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{U}, \mathbf{B}_f^T \mathbf{Z}_1, \mathcal{T}(\mathbf{B}_e^T \mathbf{Z}_2)), \quad (57)$$

the problem in (54) is expressed as an ADMM-form:

$$\begin{aligned} & \underset{\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3}{\text{minimize}} && g(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) + h(\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2) \\ & \text{subject to} && (\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) = \mathcal{L}(\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2). \end{aligned} \quad (58)$$

Note that although the function $h(\mathbf{U}, \mathbf{Z}_1, \mathbf{Z}_2)$ actually depends only on \mathbf{U} , we added \mathbf{Z}_1 and \mathbf{Z}_2 as arguments to make it easier to understand that the problem in (58) as the ADMM-form.

The ADMM iterations for (58) are given as follows. On the first line in (79), since $g(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3)$ is divided into three parts $\|\mathbf{\Lambda}_1\|_{1, \Xi_1}^{G_1}$, $\|\mathbf{\Lambda}_2\|_{1, \Xi_2}^{G_2}$, and $\iota(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3)$, the variables $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are updated, with the use of some $\gamma > 0$, by

$$\left. \begin{aligned} \mathbf{\Lambda}_1^{(i)} &= \text{prox}_{\gamma \|\cdot\|_{1, \Xi_1}^{G_1}} (\mathbf{Z}_1^{(i)} + \mathbf{V}_1^{(i)}) \\ \mathbf{\Lambda}_2^{(i)} &= \text{prox}_{\gamma \|\cdot\|_{1, \Xi_2}^{G_2}} (\mathbf{Z}_2^{(i)} + \mathbf{V}_2^{(i)}) \end{aligned} \right\}, \quad (59)$$

and $\mathbf{\Gamma}_n$ ($n = 1, 2, 3$) are updated, with the use of the proximity operators of the indicator function in (82), by

$$\mathbf{\Gamma}_n^{(i)} = \frac{1}{3} (\mathbf{U}^{(i)} + \mathbf{B}_f^T \mathbf{Z}_1^{(i)} + \mathbf{Z}_2^{(i)T} \mathbf{B}_e + \mathbf{V}_3^{(i)} + \mathbf{V}_4^{(i)} + \mathbf{V}_5^{(i)}), \quad (60)$$

where $\mathbf{V}_1 \in \mathbb{C}^{b_f L \times M}$, $\mathbf{V}_2 \in \mathbb{C}^{b_e M \times L}$, and $\mathbf{V}_n \in \mathbb{C}^{L \times M}$ ($n = 3, 4, 5$) are dual variables. On the second line in (79), we have

$$\left. \begin{aligned} \tilde{\mathbf{V}}_1^{(i)} &= \mathbf{V}_1^{(i)} + \mathbf{Z}_1^{(i)} - \mathbf{\Lambda}_1^{(i)} \\ \tilde{\mathbf{V}}_2^{(i)} &= \mathbf{V}_2^{(i)} + \mathbf{Z}_2^{(i)} - \mathbf{\Lambda}_2^{(i)} \\ \tilde{\mathbf{V}}_3^{(i)} &= \mathbf{V}_3^{(i)} + \mathbf{U}^{(i)} - \mathbf{\Gamma}_1^{(i)} \\ \tilde{\mathbf{V}}_4^{(i)} &= \mathbf{V}_4^{(i)} + \mathbf{B}_f^T \mathbf{Z}_1^{(i)} - \mathbf{\Gamma}_2^{(i)} \\ \tilde{\mathbf{V}}_5^{(i)} &= \mathbf{V}_5^{(i)} + \mathbf{Z}_2^{(i)T} \mathbf{B}_e - \mathbf{\Gamma}_3^{(i)} \end{aligned} \right\}. \quad (61)$$

On the third line in (79), \mathbf{U} , \mathbf{Z}_1 , and \mathbf{Z}_2 can be updated as the solutions to three different least squares problems since the linear mapping \mathcal{L} does not mix \mathbf{U} , \mathbf{Z}_1 and \mathbf{Z}_2 . Specifically, from

$$\mathcal{L}^*(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \mathbf{\Gamma}_3) := (\mathbf{\Gamma}_1, \mathbf{\Lambda}_1 + \mathbf{B}_f \mathbf{\Gamma}_2, \mathbf{\Lambda}_2 + \mathbf{B}_e \mathbf{\Gamma}_3^T), \quad (62)$$

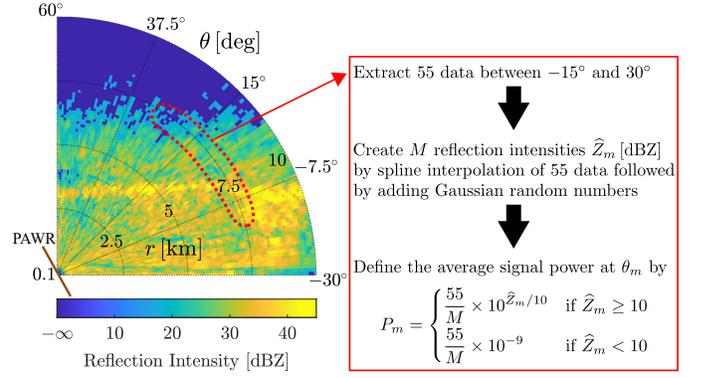


Fig. 2. Reflection intensities \hat{Z}_m observed by the PAWR at Osaka University on March 30, 2014, and a flowchart on the average signal powers P_m in our simulations. These intensities were given from the results of FR beamforming.

\mathbf{U} is updated by

$$\mathbf{U}^{(i+1)} = (\gamma \mathbf{F}\mathbf{Y}^T \bar{\mathbf{A}} + \mathbf{\Gamma}_1^{(i)} - \tilde{\mathbf{V}}_3^{(i)}) (\gamma \mathbf{A}^T \bar{\mathbf{A}} + \mathbf{I}_M)^{-1}. \quad (63)$$

\mathbf{Z}_1 and \mathbf{Z}_2 are respectively updated by

$$\mathbf{Z}_1^{(i+1)} = (\mathbf{I}_{b_f L} + \mathbf{B}_f \mathbf{B}_f^T)^{-1} (\mathbf{\Lambda}_1^{(i)} - \tilde{\mathbf{V}}_1^{(i)} + \mathbf{B}_f (\mathbf{\Gamma}_2^{(i)} - \tilde{\mathbf{V}}_4^{(i)})), \quad (64)$$

and

$$\mathbf{Z}_2^{(i+1)} = (\mathbf{I}_{b_e M} + \mathbf{B}_e \mathbf{B}_e^T)^{-1} (\mathbf{\Lambda}_2^{(i)} - \tilde{\mathbf{V}}_2^{(i)} + \mathbf{B}_e (\mathbf{\Gamma}_3^{(i)} - \tilde{\mathbf{V}}_5^{(i)})^T). \quad (65)$$

Note that, if we use the *matrix inversion lemma* [64], [65], then $(\mathbf{I}_{b_f L} + \mathbf{B}_f \mathbf{B}_f^T)^{-1}$ in (64) can be computed by

$$\begin{aligned} (\mathbf{I}_{b_f L} + \mathbf{B}_f \mathbf{B}_f^T)^{-1} &= \mathbf{I}_{b_f L} - \mathbf{B}_f (\mathbf{I}_L + \mathbf{B}_f^T \mathbf{B}_f)^{-1} \mathbf{B}_f^T \\ &= \mathbf{I}_{b_f L} - \frac{1}{b_f + 1} \mathbf{B}_f \mathbf{B}_f^T, \end{aligned} \quad (66)$$

and $(\mathbf{I}_{b_e M} + \mathbf{B}_e \mathbf{B}_e^T)^{-1}$ in (65) can also be computed similarly. Finally, on the fourth line in (79), \mathbf{V}_1 , \mathbf{V}_2 , \mathbf{V}_3 , \mathbf{V}_4 and \mathbf{V}_5 are updated, with the use of $\rho^{(i+1)} \in [0, 2]$, by

$$\left. \begin{aligned} \mathbf{V}_1^{(i+1)} &= \tilde{\mathbf{V}}_1^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{Z}_1^{(i+1)} - \mathbf{\Lambda}_1^{(i)}) \\ \mathbf{V}_2^{(i+1)} &= \tilde{\mathbf{V}}_2^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{Z}_2^{(i+1)} - \mathbf{\Lambda}_2^{(i)}) \\ \mathbf{V}_3^{(i+1)} &= \tilde{\mathbf{V}}_3^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{U}^{(i+1)} - \mathbf{\Gamma}_1^{(i)}) \\ \mathbf{V}_4^{(i+1)} &= \tilde{\mathbf{V}}_4^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{B}_f^T \mathbf{Z}_1^{(i+1)} - \mathbf{\Gamma}_2^{(i)}) \\ \mathbf{V}_5^{(i+1)} &= \tilde{\mathbf{V}}_5^{(i)} + (\rho^{(i+1)} - 1) (\mathbf{Z}_2^{(i+1)T} \mathbf{B}_e - \mathbf{\Gamma}_3^{(i)}) \end{aligned} \right\}. \quad (67)$$

By repeating (59)–(61), (63)–(65) and (67) until a convergence condition is met, the solution to (58) is obtained as an estimate $\hat{\mathbf{U}}$. From the ADMM iterations in Formulations I and II, we see that the problem in (54) is more difficult than that in (43).

V. NUMERICAL SIMULATIONS

A. Settings for Synthetic Data

To show the effectiveness of the proposed nonlinear beamforming, first we conducted simulations using synthetic data based on the *reflection intensities*, in Fig. 2, observed by the PAWR at Osaka University on March 30, 2014. The reflection

intensity \widehat{Z} [dBZ] indicates the scattering factor per unit volume considering the radar range equation and the permittivity of raindrops, and it is computed from the signal power P as

$$\widehat{Z} = 10 \log_{10} \left(\frac{r^4 P}{\nu \Omega} \right), \quad (68)$$

where ν is a constant determined from the characteristics of the radar and the permittivity of raindrops, and Ω is the size of the scattering resolution volume at range r . Note that in the beamforming, Ω is *inversely proportional* to the number M of sub-intervals. In Fig. 2, there exist 110 reflection intensities at each range between the whole elevations of -30 [deg] and 60 [deg], and they were computed from the results of FR beamforming. This PAWR system considers that if the value of the reflection intensity is less than 10 [dBZ], there is no signal at that point.

We conducted the following simulations based on [66]. By setting $r = 7.5$ [km], $\theta_{\min} = -15$ [deg] and $\theta_{\max} = 30$ [deg], we extracted $110 \frac{(\theta_{\max} - \theta_{\min})}{\pi/2} = 55$ data samples from Fig. 2. The target reflection intensity \widehat{Z}_m [dBZ] at each θ_m was generated by cubic spline interpolation of these 55 samples followed by adding Gaussian random numbers of variance 9. Note that the random numbers were needed to create the high-resolution precipitation profile. We supposed that $\frac{r^4}{\nu \Omega} = 1$ for simplicity when $M = 55$. From (68) and the relation between Ω and M , we defined the average signal power P_m at each θ_m by

$$P_m = \begin{cases} \frac{55}{M} \times 10^{\widehat{Z}_m/10} & \text{if } \widehat{Z}_m \geq 10, \\ \frac{55}{M} \times 10^{-9} & \text{if } \widehat{Z}_m < 10. \end{cases} \quad (69)$$

We used $\lambda = 31.8$ [mm], $d = 16.5$ [mm], and $T = 0.4$ [ms]. As a result, the Nyquist frequency was $\frac{1}{2T} = 1250$ [Hz] and the Nyquist Doppler velocity was $\frac{\lambda}{4T} \approx 19.9$ [m/s]. We used the Gaussian function in (34) as the power spectral density $S_m(f)$ at each θ_m , and we defined the mean Doppler frequency by

$$\mu_{f_m} = -\frac{2}{\lambda} \left(20 \sin \left(\frac{2\pi m}{M} \right) + n_m \right) \text{ [Hz]}, \quad (70)$$

where $n_m \in \mathbb{R}$ is a Gaussian random number of variance 1. The Doppler frequency width at each θ_m was simply fixed to $\sigma_{f_m} = \frac{4}{\lambda} \approx 126$ [Hz] which corresponds to $\sigma_{v_m} = 2$ [m/s].

The exact autocorrelation function $R_{x_m}(\tau)$ of the backscattered signal $x_m(t)$ at each θ_m can be computed as in (35), and the covariance matrix of each random vector $\tilde{\mathbf{x}}_m$ is given by

$$\begin{aligned} \mathbf{R}_m &:= E[\tilde{\mathbf{x}}_m \tilde{\mathbf{x}}_m^H] \\ &= \begin{bmatrix} R_{x_m}(0) & R_{x_m}(T) & \cdots & R_{x_m}((L-1)T) \\ R_{x_m}(T) & R_{x_m}(0) & \cdots & R_{x_m}((L-2)T) \\ \vdots & \vdots & \ddots & \vdots \\ R_{x_m}((L-1)T) & R_{x_m}((L-2)T) & \cdots & R_{x_m}(0) \end{bmatrix}. \end{aligned} \quad (71)$$

Based on [7], [67], each target signal $\tilde{\mathbf{x}}_m$ was generated⁸ from

⁸In this model, the variance $V[|u_{k,m}|^2] := E[(|u_{k,m}|^2 - E[|u_{k,m}|^2])^2]$ of the periodogram satisfies $V[|u_{k,m}|^2] = (E[|u_{k,m}|^2])^2$, which leads to

$$\lim_{L \rightarrow \infty} V[|u_{k,m}|^2] = \left(S_m^{(d)} \left(\frac{k-1-L/2}{LT} \right) \right)^2, \quad (72)$$

i.e., the variance does not converge to zero when L approaches infinity [68].

a circularly-symmetric complex Gaussian distribution

$$\tilde{q}_m(\tilde{\mathbf{x}}) = \frac{1}{\pi^L \det(\mathbf{R}_m)} e^{-\tilde{\mathbf{x}}^H \mathbf{R}_m^{-1} \tilde{\mathbf{x}}}. \quad (73)$$

In (15), the number of antenna elements was set to $N = 128$, and the variance of the noise ε_l was set to $\sigma_\varepsilon^2 = 2.5$.

We compared the proposed nonlinear beamforming, NL-I in (43) and NL-II in (54), to LS in (18),⁹ FR in (19), CP in (23), and MMSE beamforming in (30) for $M = 110, 160$ and $L = 20, 60, 256$. Note that CP beamforming in (23) cannot be used for $L = 20, 60$ since $L < N$ holds. When $M = 110$, the number of subintervals in which signals exist¹⁰ was $K = 93 < N$. On the other hand, when $M = 166$, that was $K = 142 > N$, and hence it is difficult to estimate \mathbf{X} or \mathbf{U} . Moreover, in both cases, the condition number of $\mathbf{A}^H \mathbf{A}$, denoted by $\text{cn}(\mathbf{A}^H \mathbf{A})$ in Table II, is extremely large, and the beamforming problem based on the observation model in (15) is very ill-conditioned.

The weights of NL-I in (43) were set to $\xi_{k,m}^{(1)} = 1.5 \frac{N}{\sqrt{b_f}} \times 10^{-3}$ and $\xi_{m,k}^{(2)} = 2.5 \frac{N}{\sqrt{b_e}} \times 10^{-4}$, the group size b_f was set to $b_f = 3$ for $L = 20$, $b_f = 8$ for $L = 60$, and $b_f = 12$ for $L = 256$, and the group size b_e was set to $b_e = 5$ for $M = 110$ and $b_e = 7$ for $M = 166$. The weights of NL-II in (54) were set to $\xi_{k,m}^{(1)} = 6.5N\sqrt{b_f} \times 10^{-4}$ and $\xi_{m,k}^{(2)} = 7.5N\sqrt{b_e} \times 10^{-5}$, the group size b_f was set to $b_f = 5$ for $L = 20$, $b_f = 13$ for $L = 60$, and $b_f = 21$ for $L = 256$, and the group size b_e was set to $b_e = 6$ for $M = 110$ and $b_e = 7$ for $M = 166$.

We evaluated the estimate $\widehat{\mathbf{X}}$ or $\widehat{\mathbf{U}} = \mathbf{F} \widehat{\mathbf{X}}^T$ of each method by the following normalized mean square error (NMSE).

- NMSE for $x_{m,l}$ or $u_{k,m}$

$$\begin{aligned} \frac{\|\mathbf{X} - \widehat{\mathbf{X}}\|_F^2}{\|\mathbf{X}\|_F^2} &= \frac{\sum_{m=1}^M \sum_{l=1}^L |x_{m,l} - \widehat{x}_{m,l}|^2}{\sum_{m=1}^M \sum_{l=1}^L |x_{m,l}|^2} \\ &= \frac{\|\mathbf{U} - \widehat{\mathbf{U}}\|_F^2}{\|\mathbf{U}\|_F^2} = \frac{\sum_{k=1}^L \sum_{m=1}^M |u_{k,m} - \widehat{u}_{k,m}|^2}{\sum_{k=1}^L \sum_{m=1}^M |u_{k,m}|^2}. \end{aligned} \quad (74)$$

Moreover, we evaluated the estimated periodograms ($|\widehat{u}_{k,m}|^2$) by the following normalized mean absolute errors (NMAEs).

- NMAE for $|u_{k,m}|^2$

$$\frac{\sum_{k=1}^L \sum_{m=1}^M \left| |u_{k,m}|^2 - |\widehat{u}_{k,m}|^2 \right|}{\sum_{k=1}^L \sum_{m=1}^M |u_{k,m}|^2}. \quad (75)$$

- NMAE for $E[|u_{k,m}|^2]$

$$\frac{\sum_{k=1}^L \sum_{m=1}^M \left| E[|u_{k,m}|^2] - |\widehat{u}_{k,m}|^2 \right|}{\sum_{k=1}^L \sum_{m=1}^M E[|u_{k,m}|^2]}. \quad (76)$$

- NMAE for $S_m^{(d)} \left(\frac{k-1-L/2}{LT} \right)$

$$\frac{\sum_{k=1}^L \sum_{m=1}^M \left| S_m^{(d)} \left(\frac{k-1-L/2}{LT} \right) - |\widehat{u}_{k,m}|^2 \right|}{\sum_{k=1}^L \sum_{m=1}^M S_m^{(d)} \left(\frac{k-1-L/2}{LT} \right)}. \quad (77)$$

NMAE in (75) compares the estimate ($|\widehat{u}_{k,m}|^2$) with true periodograms ($|u_{k,m}|^2$) in (9), as NMSE in (74) compares the estimate $\widehat{\mathbf{U}} = (\widehat{u}_{k,m})$ with true DFT coefficients $\mathbf{U} = (u_{k,m})$ in

⁹To avoid the numerical instability in the computation of the pseudoinverse \mathbf{A}^\dagger , we truncated the singular values of \mathbf{A} that are smaller than 0.005.

¹⁰We defined K as the number of the indices m s.t. $\widehat{Z}_m \geq 10$.

TABLE II
AVERAGES OF THE NORMALIZED ERRORS OF THE ESTIMATED SIGNALS AND THE ESTIMATED PERIODOGRAMS OF EACH METHOD IN 100 TRIALS

Simulation Settings			Evaluation Metric \ Method	LS	FR	CP	MMSE	NL-I	NL-II
$N = 128$ $T = 0.4$ [ms] $c = 300$ [km/ms] $\lambda = 31.8$ [mm] $d = 16.5$ [mm] $r = 7.5$ [km] $\theta_{\min} = -15$ [deg] $\theta_{\max} = 30$ [deg]	$M = 110$ $K = 93$ $\Delta\theta \approx 0.41$ [deg] $\text{cn}(\mathbf{A}^H \mathbf{A}) > 10^{30}$ (we used the same power spectrum $S_m(f)$ for all $L = 20, 60, 256$)	$L = 20$	NMSE for $u_{k,m}$	0.6377	1.1060	—	0.4189	0.2608	0.2556
			NMAE for $ u_{k,m} ^2$	0.9231	1.3744	—	0.6372	0.5066	0.4931
			NMAE for $E[u_{k,m} ^2]$	0.9185	1.6742	—	0.9726	0.7223	0.7143
			NMAE for $S_m^{(d)}(\frac{k-1-L/2}{LT})$	0.9546	1.6975	—	1.0039	0.7428	0.7366
	$L = 60$	NMSE for $u_{k,m}$	0.6324	1.1057	—	0.4271	0.2328	0.2265	
		NMAE for $ u_{k,m} ^2$	0.9263	1.3621	—	0.6549	0.4787	0.4669	
		NMAE for $E[u_{k,m} ^2]$	0.9222	1.6341	—	1.0041	0.6845	0.6808	
		NMAE for $S_m^{(d)}(\frac{k-1-L/2}{LT})$	0.9338	1.6407	—	1.0131	0.6899	0.6866	
	$L = 256$	NMSE for $u_{k,m}$	0.6306	1.1042	0.7538	0.4183	0.2188	0.2143	
		NMAE for $ u_{k,m} ^2$	0.9275	1.3465	0.9558	0.6447	0.4611	0.4501	
		NMAE for $E[u_{k,m} ^2]$	0.9259	1.6690	0.9225	1.0344	0.6809	0.6668	
		NMAE for $S_m^{(d)}(\frac{k-1-L/2}{LT})$	0.9290	1.6706	0.9297	1.0368	0.6821	0.6680	
$M = 166$ $K = 142$ $\Delta\theta \approx 0.27$ [deg] $\text{cn}(\mathbf{A}^H \mathbf{A}) = \infty$ (we used the same power spectrum $S_m(f)$ for all $L = 20, 60, 256$)	$L = 20$	NMSE for $u_{k,m}$	0.7324	2.1567	—	1.4603	0.6012	0.6078	
		NMAE for $ u_{k,m} ^2$	0.8670	2.4415	—	1.6916	0.7816	0.7813	
		NMAE for $E[u_{k,m} ^2]$	0.8184	2.4541	—	1.7487	0.7859	0.7828	
		NMAE for $S_m^{(d)}(\frac{k-1-L/2}{LT})$	0.8639	2.4797	—	1.7845	0.8069	0.8060	
	$L = 60$	NMSE for $u_{k,m}$	0.7261	2.1812	—	1.5514	0.5694	0.5664	
		NMAE for $ u_{k,m} ^2$	0.8733	2.4994	—	1.8467	0.7600	0.7556	
		NMAE for $E[u_{k,m} ^2]$	0.8292	2.5310	—	1.9364	0.7487	0.7454	
		NMAE for $S_m^{(d)}(\frac{k-1-L/2}{LT})$	0.8473	2.5390	—	1.9474	0.7559	0.7530	
	$L = 256$	NMSE for $u_{k,m}$	0.7281	2.1718	1.2436	1.5854	0.5601	0.5581	
		NMAE for $ u_{k,m} ^2$	0.8814	2.4785	1.2371	1.8901	0.7560	0.7515	
		NMAE for $E[u_{k,m} ^2]$	0.8372	2.5225	1.1516	1.9895	0.7354	0.7435	
		NMAE for $S_m^{(d)}(\frac{k-1-L/2}{LT})$	0.8422	2.5245	1.1585	1.9921	0.7370	0.7451	

(8). Although \mathbf{U} and $(|u_{k,m}|^2)$ are the targets to be estimated directly, they are realizations of random variables, and the expected values $E[|u_{k,m}|^2]$ in (10) or the power spectral densities $S_m^{(d)}(\frac{k-1-L/2}{LT})$ in (7) have more accurate information on P_m , μ_{f_m} and σ_{f_m} . Hence, in (76) and (77), we compare $(|\hat{u}_{k,m}|^2)$ with the two more ideal values $E[|u_{k,m}|^2]$ and $S_m^{(d)}(\frac{k-1-L/2}{LT})$.

B. Simulation Results for Synthetic Data

Table II summarizes the average, for each situation and each method, of the normalized errors in 100 trials, where the power spectral density $S_m(f)$ at each θ_m was fixed throughout all the 100 trials. From Table II, we see that, when $M = 110$, the proposed methods, NL-I and NL-II, reduced NMSE from LS by $100 \times \frac{0.6377-0.2608}{0.6377} \approx 59\%$ and $100 \times \frac{0.6377-0.2556}{0.6377} \approx 60\%$ for $L = 20$, $100 \times \frac{0.6324-0.2328}{0.6324} \approx 63\%$ and $100 \times \frac{0.6324-0.2265}{0.6324} \approx 64\%$ for $L = 60$, and $100 \times \frac{0.6306-0.2188}{0.6306} \approx 65\%$ and $100 \times \frac{0.6306-0.2143}{0.6306} \approx 66\%$ for $L = 256$. We write this as “NL-I and NL-II reduced NMSE by 59–66% from LS when $M = 110$.” We also see that NL-I and NL-II reduced NMSE by 76–81% from FR, 71–72% from CP, and 38–49% from MMSE when $M = 110$, while by 17–23% from LS, 72–74% from FR, 55% from CP, and 58–65% from MMSE when $M = 166$. Thus, the proposed methods greatly improved the estimation accuracy of \mathbf{U} . In addition, we find that MMSE beamforming estimated \mathbf{X} with relatively high accuracy among the linear methods when $M = 110$ while the estimation accuracy significantly degraded due to the occurrence of sidelobes when $M = 166$ and became lower than CP beamforming when $L = 256$. This implies that

MMSE beamforming is affected by the initial value $\hat{\mathbf{x}}_{\text{MMSE},L}^{(0)}$, which may lead to more overestimation than CP beamforming.

On the estimated periodograms $(|\hat{u}_{k,m}|^2)$, from Table II and the same calculations as in the previous paragraph, we see that NL-I and NL-II reduced NMAE for $|u_{k,m}|^2$ by 45–51% from LS, 63–67% from FR, 52–53% from CP, and 20–30% from MMSE when $M = 110$, while by 10–15% from LS, 68–70% from FR, 39% from CP, and 54–60% from MMSE when $M = 166$. Further, NL-I and NL-II reduced NMAEs for $E[|u_{k,m}|^2]$ and $S_m^{(d)}(\frac{k-1-L/2}{LT})$ by 21–28% from LS, 56–60% from FR, 26–28% from CP, and 26–36% from MMSE when $M = 110$, while by 4–12% from LS, 67–71% from FR, 35–36% from CP, and 55–63% from MMSE when $M = 166$. Thus, the proposed methods greatly improved the accuracy of the periodograms.

On the number of pulses, three linear methods, LS, FR, and MMSE, did not necessarily reduce NMSE and NMAEs, even if L was changed from 20 to 60 and 256. On the other hand, the proposed methods, NL-I and NL-II, could reduce NMSE and NMAEs by the increase of L . This is because, as L increases, $E[|u_{k,m}|^2]$ approaches $S_m^{(d)}(\frac{k-1-L/2}{LT})$ and hence \mathbf{U} becomes more group-sparse, which is ideal for the proposed methods.

Figures 3, 4, 6 and 7 show examples of beamforming results for $(M, L) = (110, 20), (110, 60), (166, 20), (166, 60)$, where (a) shows the power spectral densities $S_m^{(d)}(\frac{k-1-L/2}{LT})$ in (7), (b) shows the expected values $E[|u_{k,m}|^2]$ of periodograms in (10), (c) shows the true periodograms $(|u_{k,m}|^2)$ in (9) to be estimated, and (d), (e), (f), (g) and (h) show the estimated periodograms $(|\hat{u}_{k,m}|^2)$ by LS, FR, MMSE, NL-I and NL-II beamforming, respectively. Figures 5 and 8 show examples of beam-

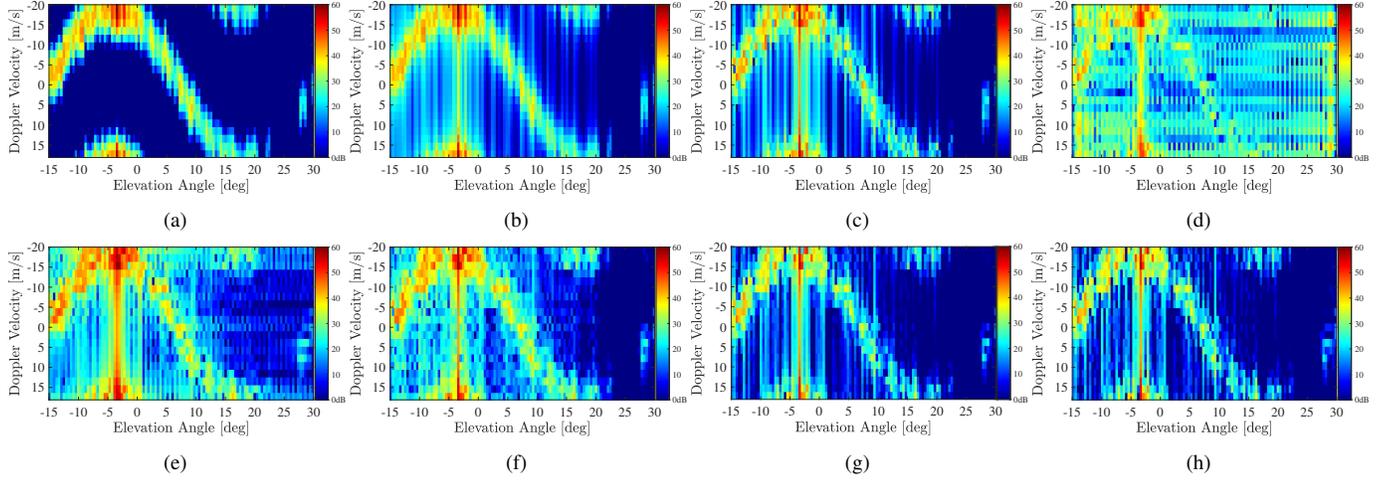


Fig. 3. Comparison of the power spectral densities, true periodograms, and each beamforming result in case of $M = 110$ and $L = 20$: (a) $S_m^{(d)}(\frac{k-1-L/2}{LT})$, (b) $E[|u_{k,m}|^2]$, (c) $|u_{k,m}|^2$, (d) $|\hat{u}_{k,m}|^2$ by LS, (e) $|\hat{u}_{k,m}|^2$ by FR, (f) $|\hat{u}_{k,m}|^2$ by MMSE, (g) $|\hat{u}_{k,m}|^2$ by NL-I, and (h) $|\hat{u}_{k,m}|^2$ by NL-II beamforming.

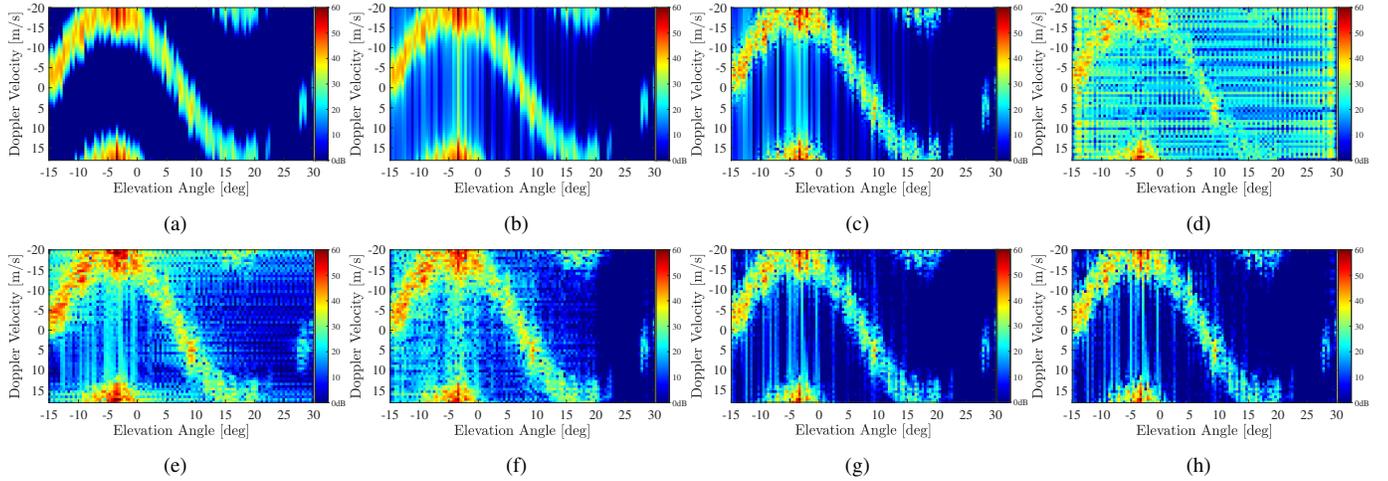


Fig. 4. Comparison of the power spectral densities, true periodograms, and each beamforming result in case of $M = 110$ and $L = 60$: (a) $S_m^{(d)}(\frac{k-1-L/2}{LT})$, (b) $E[|u_{k,m}|^2]$, (c) $|u_{k,m}|^2$, (d) $|\hat{u}_{k,m}|^2$ by LS, (e) $|\hat{u}_{k,m}|^2$ by FR, (f) $|\hat{u}_{k,m}|^2$ by MMSE, (g) $|\hat{u}_{k,m}|^2$ by NL-I, and (h) $|\hat{u}_{k,m}|^2$ by NL-II beamforming.

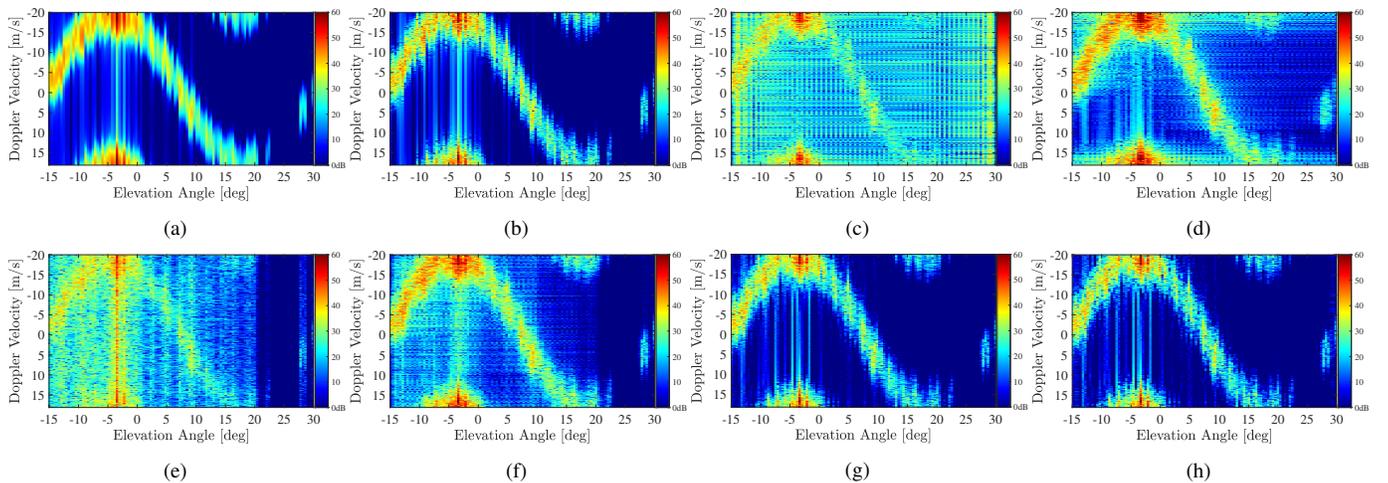


Fig. 5. Comparison of true periodograms and each beamforming result in case of $M = 110$ and $L = 256$: (a) $E[|u_{k,m}|^2]$, (b) $|u_{k,m}|^2$, (c) $|\hat{u}_{k,m}|^2$ by LS, (d) $|\hat{u}_{k,m}|^2$ by FR, (e) $|\hat{u}_{k,m}|^2$ by CP, (f) $|\hat{u}_{k,m}|^2$ by MMSE, (g) $|\hat{u}_{k,m}|^2$ by NL-I, and (h) $|\hat{u}_{k,m}|^2$ by NL-II beamforming, where we omit the figure of the power spectral densities $S_m^{(d)}(\frac{k-1-L/2}{LT})$ since it is almost the same as Fig. 4(a), and the result of CP beamforming is added instead since $L > N$ holds.

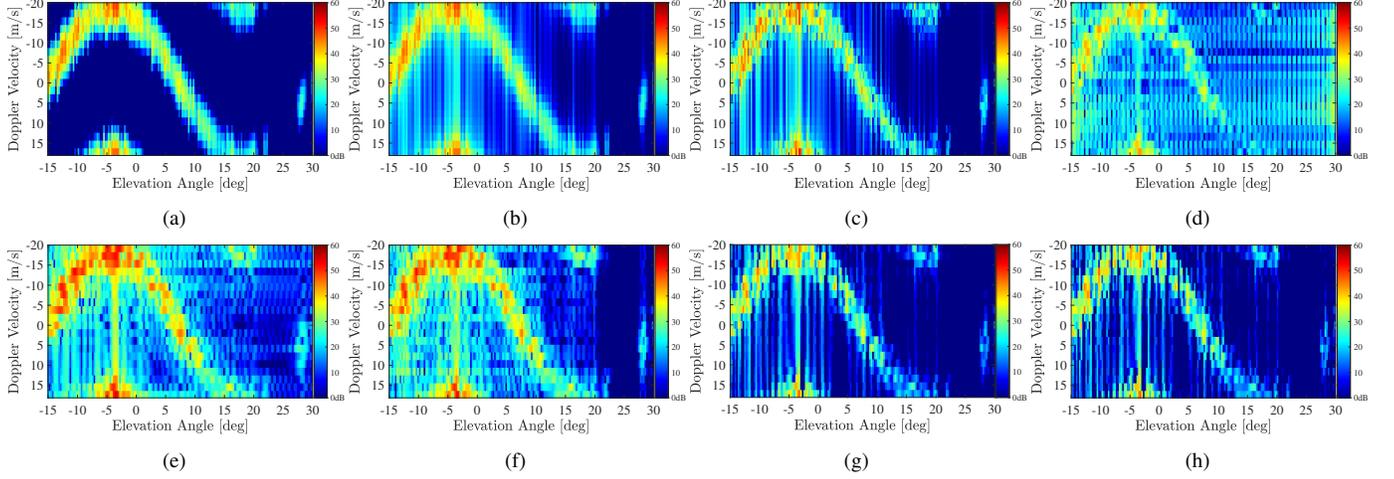


Fig. 6. Comparison of the power spectral densities, true periodograms, and each beamforming result in case of $M = 166$ and $L = 20$: (a) $S_m^{(d)}(\frac{k-1-L/2}{LT})$, (b) $E[|u_{k,m}|^2]$, (c) $|u_{k,m}|^2$, (d) $|\hat{u}_{k,m}|^2$ by LS, (e) $|\hat{u}_{k,m}|^2$ by FR, (f) $|\hat{u}_{k,m}|^2$ by MMSE, (g) $|\hat{u}_{k,m}|^2$ by NL-I, and (h) $|\hat{u}_{k,m}|^2$ by NL-II beamforming.

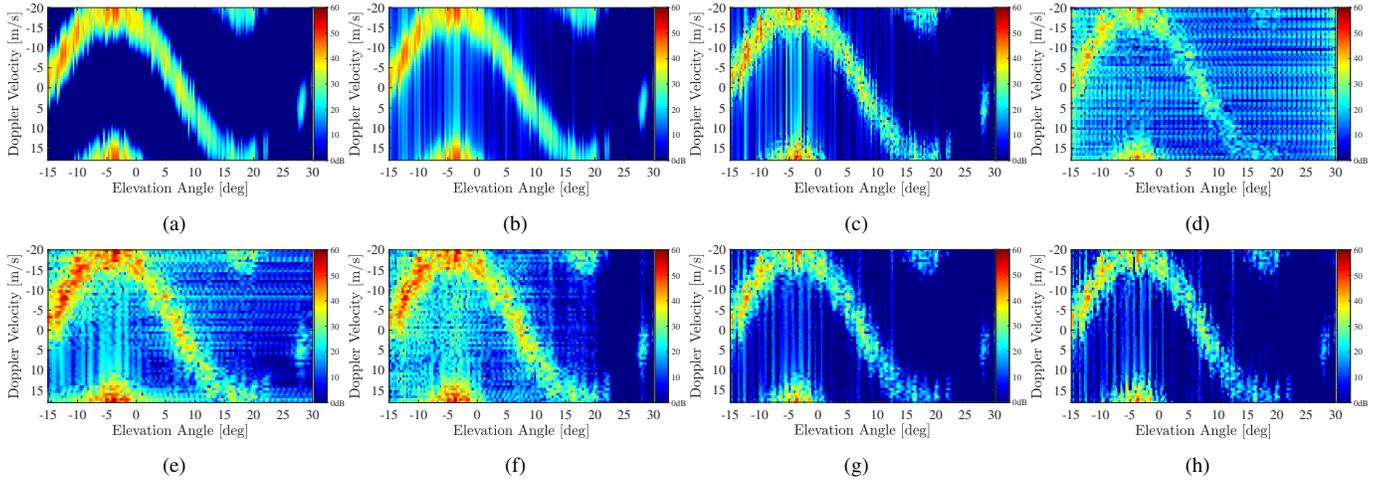


Fig. 7. Comparison of the power spectral densities, true periodograms, and each beamforming result in case of $M = 166$ and $L = 60$: (a) $S_m^{(d)}(\frac{k-1-L/2}{LT})$, (b) $E[|u_{k,m}|^2]$, (c) $|u_{k,m}|^2$, (d) $|\hat{u}_{k,m}|^2$ by LS, (e) $|\hat{u}_{k,m}|^2$ by FR, (f) $|\hat{u}_{k,m}|^2$ by MMSE, (g) $|\hat{u}_{k,m}|^2$ by NL-I, and (h) $|\hat{u}_{k,m}|^2$ by NL-II beamforming.

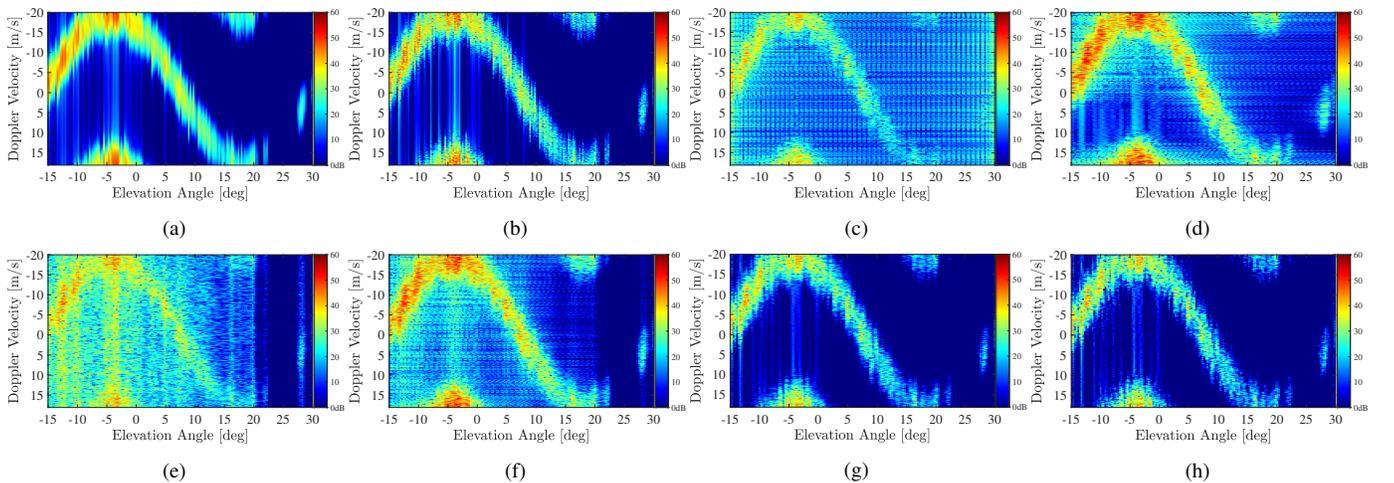


Fig. 8. Comparison of true periodograms and each beamforming result in case of $M = 166$ and $L = 256$: (a) $E[|u_{k,m}|^2]$, (b) $|u_{k,m}|^2$, (c) $|\hat{u}_{k,m}|^2$ by LS, (d) $|\hat{u}_{k,m}|^2$ by FR, (e) $|\hat{u}_{k,m}|^2$ by CP, (f) $|\hat{u}_{k,m}|^2$ by MMSE, (g) $|\hat{u}_{k,m}|^2$ by NL-I, and (h) $|\hat{u}_{k,m}|^2$ by NL-II beamforming, where we omit the figure of the power spectral densities $S_m^{(d)}(\frac{k-1-L/2}{LT})$ since it is almost the same as Fig. 7(a), and the result of CP beamforming is added instead since $L > N$ holds.

forming results for $(M, L) = (110, 256), (166, 256)$, where we omit the figures of $S_m^{(d)}(\frac{k-1-L/2}{LT})$ because they are almost the same as Fig. 4(a) and Fig. 7(a), and instead, those of $(|\hat{u}_{k,m}|^2)$ by CP beamforming are inserted into (e). In these figures, we consider that there is no signal at points shown in blue. From Figs. 3(a)–3(c), 4(a)–4(c), 6(a)–6(c) and 7(a)–7(c), we can see that $E[|u_{k,m}|^2]$ in (b) approached $S_m^{(d)}(\frac{k-1-L/2}{LT})$ in (a), and the target periodograms $(|u_{k,m}|^2)$ in (c) became more group-sparse as the number L of pulses increased from 20 to 60. In addition, From Figs. 5(a), 5(b), 8(a) and 8(b), we also see that $E[|u_{k,m}|^2]$ further approached $S_m^{(d)}(\frac{k-1-L/2}{LT})$ when $L = 256$ as shown in (11), and $(|u_{k,m}|^2)$ became further group-sparse. In the following, we judge that the closer the figure of the estimated periodograms $(|\hat{u}_{k,m}|^2)$ is to that of the target values $(|u_{k,m}|^2)$ and those of the two more ideal values $E[|u_{k,m}|^2]$ and $S_m^{(d)}(\frac{k-1-L/2}{LT})$, the better the beamforming result is.

When $M = 110$, from Figs. 3(d), 4(d) and 5(c), LS beamforming diffused the energy of the backscattered signals to all elevations, and these results cannot be used for the weather observation because they cannot detect the elevations with no or few raindrops. From Figs. 3(e), 4(e) and 5(d), FR beamforming overestimated the signal power due to sidelobes as shown in (21), which is noticeable particularly from 23 [deg] to 27 [deg] because the signal power should be almost zero in this interval. From Fig. 5(e), CP beamforming could find that there is no signal from 23 [deg] to 27 [deg] but diffused the energy of the backscattered signals to all frequencies at the other elevations. From Figs. 3(f), 4(f) and 5(f), MMSE beamforming estimated the periodograms with relatively high accuracy, but the energy of the backscattered signals was still diffused mainly along the elevation axis. On the other hand, from Figs. 3(g), 3(h), 4(g), 4(h), 5(g) and 5(h), we can confirm that the proposed methods, NL-I and NL-II beamforming, could obtain the most accurate periodograms with very little energy diffusion. Indeed, the no-signal points of $(|u_{k,m}|^2)$ and $E[|u_{k,m}|^2]$ shown in blue and narrow vertical lines in light blue were well reproduced. The results of NL-I and NL-II beamforming were very similar, but NL-II beamforming could obtain the slightly higher-resolution periodograms because the color transition is a little bit finer in Figs. 3(h), 4(h) and 5(h) than in Figs. 3(g), 4(g) and 5(g).

When $M = 166$, each signal power P_m decreases compared to the case of $M = 110$ from (69). Thus, not only the number of variables to be estimated increases, but also the signal-to-noise ratio decreases. From Figs. 6(d), 7(d) and 8(c), LS beamforming did not cause overestimation but diffused the energy of the backscattered signals to all elevations. From Figs. 6(e), 6(f), 7(e), 7(f) and 8(d)–8(f), FR, CP, and MMSE beamforming caused some overestimation due to the inevitable occurrence of sidelobes, while features of the estimated periodograms were inherited from the case of $M = 110$. The results of FR beamforming in Figs. 6(e), 7(e) and 8(d) were greatly overestimated, and those of MMSE beamforming in Figs. 6(f), 7(f) and 8(f) were affected by the initial estimate $\hat{\mathbf{x}}_{\text{MMSE},l}^{(0)} = \hat{\mathbf{x}}_{\text{FR},l}$, which led to more overestimation than CP beamforming in Fig. 8(e) when $L = 256$. In this severe situation, the proposed methods, NL-I and NL-II beamforming, obtained the most accurate periodograms without the overestimation and the energy diffusion as shown in Figs. 6(g), 6(h), 7(g), 7(h), 8(g) and 8(h). As in

TABLE III
AVERAGE COMPUTATION TIME [s] OF EACH METHOD IN 100 TRIALS

(M, L)	LS	FR	CP	MMSE	NL-I	NL-II
(110, 20)	0.00010	0.00011	—	0.0368	1.304	8.91
(110, 60)	0.00013	0.00019	—	0.0416	3.989	16.80
(110, 256)	0.00048	0.00060	0.0171	0.0967	17.372	127.73
(166, 20)	0.00012	0.00013	—	0.0526	2.348	11.13
(166, 60)	0.00020	0.00024	—	0.0680	5.706	34.68
(166, 256)	0.00077	0.00081	0.0204	0.1983	19.911	261.59

the case of $M = 110$, the no-signal points shown in blue and narrow vertical lines in light blue were well reproduced. Thus, we see that the proposed nonlinear beamforming greatly improved the spatial resolution compared to the linear methods.

In both cases of $M = 110$ and $M = 166$, the proposed nonlinear methods greatly improved the spatial resolution, but two slight drawbacks were also found. One is that there was slight underestimation due to the soft-thresholding-like computation in (81), i.e., the power of the estimated signals by the proposed methods tended to be a little smaller than the true signal power. The other is that the vertical line in the figure of periodograms was sometimes estimated at an elevation next to the correct one due to the large condition number of $\mathbf{A}^H \mathbf{A}$. In addition, when the number L of pulses was small, there was a gap between the target periodogram values $(|u_{k,m}|^2)$ (or their expected values $E[|u_{k,m}|^2]$) and the power spectral densities $S_m^{(d)}(\frac{k-1-L/2}{LT})$, and hence the original group-sparsities of $S_m^{(d)}(\frac{k-1-L/2}{LT})$ have not yet been fully exploited. We plan to address these issues in future work, e.g., by further improving the regularization term.

Next, we compared the computation time of each method. Note that the time of the proposed beamforming depends on γ , $\rho^{(i+1)}$, and the convergence condition. We fixed $\rho^{(i+1)} = 1.9$ for both NL-I and NL-II beamforming and all (M, L) , γ was manually set between 100 and 2000 for each method and each (M, L) to make the convergence as fast as possible, and we stopped the ADMM iterations when $\|\mathbf{U}^{(i+1)} - \mathbf{U}^{(i)}\|_F \leq \frac{L}{256}$ or $i = 1000$ was satisfied. Table III shows the average computation time [s] of each method in 100 trials, where the same simulations as in Table II were done by MATLAB R2020a on Surface Book (Windows 10, Intel Core i7-6600U, 2.60 GHz, 16 GB). From Table III, we can see that the nonlinear methods require much more computation time than the linear methods because they compute the exact solutions to complicated optimization problems with iterations. Comparing NL-I and NL-II beamforming, NL-II requires 4–13 times longer time, although it indicates slightly higher accuracy than NL-I in Table II. Thus NL-I has a better balance between the accuracy and the time. If we want high-resolution results in a short computation time, it is better to use NL-I beamforming. If it is acceptable to take a longer time to obtain more accurate results, we can use NL-II beamforming that outputs slightly higher-resolution results, as argued in theoretical papers [45]–[48] on compressed sensing.

C. Results for Real-World Data

We finally applied the proposed method, NL-I beamforming, to real-world data observed by the PAWR at Osaka University on August 6, 2015, and verified its result. Figure 9 shows the

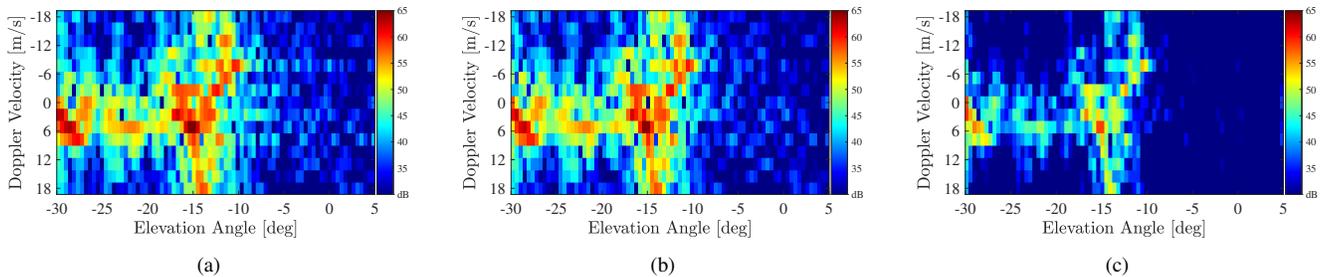


Fig. 9. Beamforming results for real-world data with $(M, L) = (80, 15)$: (a) $|\hat{u}_{k,m}|^2$ by FR, (b) $|\hat{u}_{k,m}|^2$ by MMSE, and (c) $|\hat{u}_{k,m}|^2$ by NL-I beamforming.

results of (a) FR, (b) MMSE, and (c) NL-I beamforming. Main PAWR parameters were set to $N = 100$, $L = 15$, $r = 14$ [km], $\theta_{\min} = -30$ [deg], $\theta_{\max} = 5$ [deg], and $M = 80$ (i.e., $\Delta\theta = 0.4375$ [deg]), and the other parameters had the same values as in Table II. Since pulse compression was used, we treated the received signal after matched filtering as \mathbf{y}_l in (15). Note that the results of LS, CP, and NL-II beamforming are not shown because the result of LS was too bad, CP cannot be used when $L < N$, and the result of NL-II was similar to that of NL-I.

True signal values are unknown in this situation,¹¹ and thus we check the effectiveness of the proposed method from features of the estimated periodograms in Fig. 9. The results of FR and MMSE beamforming in Figs. 9(a) and 9(b) are similar, but we can confirm from the elevations of -7 [deg] to 5 [deg] that the number of light blue points in Fig. 9(b) is smaller than in Fig. 9(a). This is because the influence of sidelobes could be suppressed by MMSE beamforming. However, both figures are blurred, and the high-resolution precipitation profile cannot be acquired. On the other hand, the result of NL-I beamforming in Fig. 9(c), is group-sparse, i.e., regions with almost no signal shown in blue are clearly visible. The energy diffusion along the elevation axis also seems to be suppressed compared to the linear methods in Figs. 9(a) and 9(b), and hence we judge that the proposed method, NL-I beamforming, acquired the higher-resolution precipitation profile than the linear beamforming.

Since the above real-world data contained much more noise than the synthetic data and was affected by the pulse compression, we had to set the weights $\xi_{k,m}^{(1)}$ and $\xi_{m,k}^{(2)}$ in (43) to large values to obtain the result shown in Fig. 9(c). As a result, it is highly possible that some underestimation occurs in Fig. 9(c) due to the soft-thresholding-like computation in (81). To acquire the high-resolution precipitation profile with less underestimation, we need to make periodograms more group-sparse. To achieve this, the use of window functions and the adaptive pulse compression in [69] can be considered as future work.

VI. CONCLUSION

In this article, we proposed a nonlinear beamforming method for PAWRs. Differently from a standard radar which observes point targets, the PAWR receives a lot of backscattered signals from distributed targets, and the spatial resolution of the linear beamforming methods for the distributed targets is limited due

to the lack of the number of null directions. To obtain the high-resolution precipitation profile, we treated beamforming as an inverse problem and solved it by using two characteristics of periodograms of the backscattered signals. One is the narrow bandwidth of each periodogram, and the other is the similarity between the adjacent periodograms. Both characteristics were expressed as group-sparsities of the DFT coefficient matrix of the backscattered signals. We proposed to reconstruct the DFT coefficient matrix by minimizing a convex cost function which consists of one data-fidelity term and two group ℓ_1 -norms that evaluate the group-sparsities, where two slightly different formulations, NL-I and NL-II beamforming, were considered. We computed the exact minimizer of the convex cost, with ADMM including the nonlinear proximity operators, as the result of the proposed beamforming. Existing radar applications have used certain sparsities of point targets, while this study, to the best of the authors' knowledge, is the first radar application using *group-sparsities of distributed targets in the elevation-velocity domain*. Although an appropriate *non-overlapping group partition* was unknown, we could evaluate each group-sparsity by using *small overlapping groups* as with latent group lasso.

To show the effectiveness of the proposed nonlinear beamforming, we first conducted simulations for synthetic data. The proposed beamforming greatly improved the estimation accuracy of both backscattered signals and periodograms compared to major linear beamforming methods. The linear methods did not necessarily reduce the normalized errors even if the number of pulses increased, while the proposed beamforming could reduce the normalized errors by the increase of the pulses. In addition, some overestimation occurred in the linear methods when the number of subintervals increased, while the proposed beamforming did not cause the overestimation. Then, from the figures of the estimated periodograms, we confirmed that no-signal regions and narrow vertical lines were well reproduced without the energy diffusion, i.e., the proposed beamforming dramatically improved the resolution of precipitation profile.

Next, we compared the average computation time of each beamforming method. The proposed beamforming took much more computation time than the linear methods, since it solves complicated optimization problems with iterations. Therefore, the proposed beamforming is useful as an *offline algorithm* to acquire the high-resolution precipitation profile, but it is difficult to directly use it in real-time processing at present. On the two formulations, NL-I and NL-II beamforming, NL-II could obtain the slightly higher-resolution precipitation profile with slightly higher accuracy than NL-I, but it required 4–13 times

¹¹If we try to observe the true signals by a parabolic Doppler weather radar, the antenna size has to be very large and the speed of the vertical scan has to be extremely fast. In practice, it is difficult to develop such a Doppler radar.

longer computation time. Although the latent group lasso formulation as in NL-II has been proposed in the theoretical field of compressed sensing, NL-I had a better balance between the accuracy and the time than NL-II at least for the beamforming, which can be considered a novel finding in the theoretical field.

Finally, we applied each beamforming method to real-world data including pulse compression. From the figures of the estimated periodograms, we confirmed that no-signal regions were well reproduced without the energy diffusion by the proposed beamforming, as in the results for the synthetic data. Thus, the proposed beamforming improved the spatial resolution for the real-world data. Although fine fluctuations of the periodograms were captured by the proposed beamforming, probably some underestimation was caused by increasing the weight parameters $\xi_{k,m}^{(1)}$ and $\xi_{m,k}^{(2)}$ to suppress large noise and the influence of the pulse compression. To acquire the high-resolution results with less underestimation for such real-world data, we need to make periodograms more group-sparse. For this purpose, we plan to integrate window functions and the adaptive pulse compression in [69] into the proposed beamforming as future work.

APPENDIX

ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Let us consider the following convex optimization problem

$$\underset{\mathbf{x} \in \mathcal{X}, \mathbf{z} \in \mathcal{Z}}{\text{minimize}} \quad g(\mathbf{z}) + h(\mathbf{x}) \quad \text{subject to} \quad \mathbf{z} = \mathcal{L}(\mathbf{x}), \quad (78)$$

where \mathcal{X} and \mathcal{Z} are finite-dimensional real Hilbert spaces with the standard inner products, $\mathcal{L} : \mathcal{X} \rightarrow \mathcal{Z}$ is a linear mapping, and $g : \mathcal{Z} \rightarrow \mathbb{R} \cup \{\infty\}$ and $h : \mathcal{X} \rightarrow \mathbb{R} \cup \{\infty\}$ are *proper*, *lower semicontinuous*, and *convex* functions.¹² ADMM [24]–[27] solves the problem in (78), together with the corresponding dual problem, from any $(\mathbf{x}^{(0)}, \mathbf{v}^{(0)}) \in \mathcal{X} \times \mathcal{Z}$ by iterating

$$\begin{cases} \mathbf{z}^{(i)} = \text{prox}_{\gamma g}(\mathcal{L}(\mathbf{x}^{(i)}) + \mathbf{v}^{(i)}) \\ \tilde{\mathbf{v}}^{(i)} = \mathbf{v}^{(i)} + \mathcal{L}(\mathbf{x}^{(i)}) - \mathbf{z}^{(i)} \\ \mathbf{x}^{(i+1)} = \underset{\mathbf{x} \in \mathcal{X}}{\text{argmin}} \quad \gamma h(\mathbf{x}) + \frac{1}{2} \|\mathcal{L}(\mathbf{x}) - \mathbf{z}^{(i)} + \tilde{\mathbf{v}}^{(i)}\|_{\mathcal{X}}^2 \\ \mathbf{v}^{(i+1)} = \tilde{\mathbf{v}}^{(i)} + (\rho^{(i+1)} - 1)(\mathcal{L}(\mathbf{x}^{(i+1)}) - \mathbf{z}^{(i)}) \end{cases} \quad (79)$$

for $i \geq 0$. In (79), $\gamma > 0$, $\mathbf{v} \in \mathcal{Z}$ and $\tilde{\mathbf{v}} \in \mathcal{Z}$ are dual variables, $\|\cdot\|_{\mathcal{X}}$ is the Euclidean norm of \mathcal{X} induced by the standard inner product, $\text{prox}_{\gamma g} : \mathcal{Z} \rightarrow \mathcal{Z}$ is the *proximity operator* defined by

$$\text{prox}_{\gamma g}(\zeta) := \underset{\mathbf{z} \in \mathcal{Z}}{\text{argmin}} \quad \gamma g(\mathbf{z}) + \frac{1}{2} \|\zeta - \mathbf{z}\|_{\mathcal{Z}}^2, \quad (80)$$

and $\rho^{(i+1)} \in [0, 2]$ satisfies $\sum_{i=0}^{\infty} \rho^{(i+1)}(2 - \rho^{(i+1)}) = \infty$. In particular, if $\rho^{(i+1)} = 1$ for all i , which is the most commonly used setting, then we can replace $\tilde{\mathbf{v}}^{(i)}$ with $\mathbf{v}^{(i+1)}$ and remove the fourth line in (79). The sequence $(\mathbf{x}^{(i)}, \mathbf{z}^{(i)})_{i=0}^{\infty}$ defined by the iterations in (79) converges to a solution to (78). Moreover, the sequences $(\mathbf{v}^{(i)}/\gamma)_{i=0}^{\infty}$ and $(\tilde{\mathbf{v}}^{(i)}/\gamma)_{i=0}^{\infty}$ both converge to a

¹²A function $h : \mathcal{X} \rightarrow \mathbb{R} \cup \{\infty\}$ is called *proper*, *lower semicontinuous*, and *convex* if $\text{dom}(h) := \{\mathbf{x} \in \mathcal{X} \mid h(\mathbf{x}) < \infty\}$ is nonempty, $\text{lev}_{\leq \alpha}(h) := \{\mathbf{x} \in \mathcal{X} \mid h(\mathbf{x}) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$, and $h(\beta \mathbf{x}_1 + (1 - \beta)\mathbf{x}_2) \leq \beta h(\mathbf{x}_1) + (1 - \beta)h(\mathbf{x}_2)$ for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ and all $\beta \in (0, 1)$, respectively.

solution to the dual problem. As described in [27], the over-relaxation setting, i.e., $\rho^{(i+1)} > 1$ for all i , can accelerate the convergence, compared to the standard setting s.t. $\rho^{(i+1)} = 1$.

In the following, we introduce the proximity operators used in this article. When g is a weighted group ℓ_1 -norm with a non-overlapping group partition $G := (G_i)_{i=1}^{n_G}$ and a weight vector $\xi := (\xi_1, \xi_2, \dots, \xi_{n_G})^T \in \mathbb{R}_+^{n_G}$, from $\gamma g(\mathbf{z}) + \frac{1}{2} \|\zeta - \mathbf{z}\|_{\mathcal{Z}}^2 = \sum_{i=1}^{n_G} (\gamma \xi_i \|\mathbf{z}_{G_i}\|_2 + \frac{1}{2} \|\zeta_{G_i} - \mathbf{z}_{G_i}\|_2^2)$, the computation of $\text{prox}_{\gamma g}$ is divided into those of $\text{prox}_{\gamma \xi_i \|\cdot\|_2}$. Specifically, $\text{prox}_{\gamma g}(\zeta) = \text{prox}_{\gamma \|\cdot\|_{\xi}}(\zeta)$ can be computed for each subvector ζ_{G_i} by

$$\text{prox}_{\gamma \xi_i \|\cdot\|_2}(\zeta_{G_i}) = \begin{cases} \frac{\|\zeta_{G_i}\|_2 - \gamma \xi_i}{\|\zeta_{G_i}\|_2} \zeta_{G_i} & \text{if } \|\zeta_{G_i}\|_2 > \gamma \xi_i, \\ \mathbf{0} & \text{if } \|\zeta_{G_i}\|_2 \leq \gamma \xi_i. \end{cases} \quad (81)$$

Let $\zeta_i \in \mathcal{Z}$ ($i = 1, 2, \dots, n$) be n variables of the same size. Let ι be the *indicator function* of the consensus set [70] for the n -tuple $(\zeta_1, \zeta_2, \dots, \zeta_n)$, i.e., $\iota(\zeta_1, \zeta_2, \dots, \zeta_n) := 0$ if $\zeta_1 = \zeta_2 = \dots = \zeta_n$ and $\iota(\zeta_1, \zeta_2, \dots, \zeta_n) := \infty$ otherwise. The proximity operator of ι can be easily computed by

$$\text{prox}_{\gamma \iota}(\zeta_1, \zeta_2, \dots, \zeta_n) = \frac{1}{n} \left(\sum_{i=1}^n \zeta_i, \sum_{i=1}^n \zeta_i, \dots, \sum_{i=1}^n \zeta_i \right). \quad (82)$$

As can be seen from (82), the proximity operator of the indicator function is equal to the projection onto the consensus set and does not depend on the value of γ differently from (81).

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their many constructive comments which were very helpful to improve the contents and the readability of this article.

REFERENCES

- [1] D. S. Zrnic, J. F. Kimpel, D. E. Forsyth, A. Shapiro, G. Crain, R. Ferek, J. Heimmer, W. Benner, 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, Nov. 2007.
- [2] 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, Dec. 2009.
- [3] 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.
- [4] 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, Jun. 2015.
- [5] 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.
- [6] 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, Aug. 2021.

- [7] V. N. Bringi and V. Chandrasekar, *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge, UK: Cambridge University Press, 2001.
- [8] M. Wada, J. Horikomi, and F. Mizutani, "Development of solid-state weather radar," in *Proc. IEEE Radar Conf.*, Pasadena, CA, USA, May 2009, 4 pages.
- [9] E. Kudeki and R. F. Woodman, "A poststatistics steering technique for MST radar applications," *Radio Sci.*, vol. 25, no. 4, pp. 591–594, Jul./Aug. 1990.
- [10] R. D. Palmer, R. F. Woodman, S. Fukao, T. Tsuda, and S. Kato, "Three-antenna poststatistic steering using the MU radar," *Radio Sci.*, vol. 25, no. 6, pp. 1105–1110, Nov./Dec. 1990.
- [11] E. Kudeki and F. Sürücü, "Radar interferometric imaging of field-aligned plasma irregularities in the equatorial electrojet," *Geophys. Res. Lett.*, vol. 18, no. 1, pp. 41–44, Jan. 1991.
- [12] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.
- [13] R. D. Palmer, S. Gopalani, T. Y. Yu, and S. Fukao, "Coherent radar imaging using Capon's method," *Radio Sci.*, vol. 33, no. 6, pp. 1585–1598, Nov./Dec. 1998.
- [14] R. D. Palmer, T. Y. Yu, and P. B. Chilson, "Range imaging using frequency diversity," *Radio Sci.*, vol. 34, no. 6, pp. 1485–1496, Nov./Dec. 1999.
- [15] H. Luce, M. Yamamoto, S. Fukao, D. Helal, and M. Crochet, "A frequency domain radar interferometric imaging (FII) technique based on high-resolution methods," *J. Atmos. Solar-Terr. Phys.*, vol. 63, nos. 2–3, pp. 221–234, Jan. 2001.
- [16] L. Cheong, M. W. Hoffman, and R. D. Palmer, "Efficient atmospheric simulation for high-resolution radar imaging applications," *J. Atmos. Ocean. Technol.*, vol. 21, no. 2, pp. 374–378, Feb. 2004.
- [17] B. L. Cheong, M. W. Hoffman, R. D. Palmer, S. J. Frasier, and F. J. López-Dekker, "Pulse pair beamforming and the effects of reflectivity field variations on imaging radars," *Radio Sci.*, vol. 39, no. 3, 13 pages, Jun. 2004.
- [18] L. Yang, M. R. McKay, R. Couillet, "High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1933–1947, Apr. 2018.
- [19] 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.
- [20] D. Kitahara, M. Nakahara, A. Hirabayashi, E. Yoshikawa, H. Kikuchi, and T. Ushio, "Nonlinear beamforming via convex optimization for phased array weather radar," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annual Summit Conf. (APSIPA ASC)*, Honolulu, HI, USA, Nov. 2018, pp. 1831–1835.
- [21] R. Kawami, D. Kitahara, A. Hirabayashi, E. Yoshikawa, H. Kikuchi, and T. Ushio, "Three-dimensional data compression and fast high-quality reconstruction for phased array weather radar," *IEEE Trans. Electron. Inf. Syst.*, vol. 140, no. 1, pp. 40–48, Jan. 2020.
- [22] M. Wada, N. Anraku, H. Yamauchi, A. Adachi, U. Tomoo, and S. Satoh, "Development of advanced radar technologies for weather application," in *Proc. Eur. Conf. Radar Meteorol. Hydrol. (ERAD)*, Garmisch-Partenkirchen, Germany, Sep. 2014, 7 pages.
- [23] S. D. Blunt, T. Chan, and K. Gerlach, "Robust DOA estimation: The reiterative superresolution (RISR) algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 332–346, Jan. 2011.
- [24] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite elements approximations," *Comput. Math. Appl.*, vol. 2, no. 1, pp. 17–40, 1976.
- [25] R. Glowinski, "On alternating direction methods of multipliers: A historical perspective," in *Modeling, Simulation and Optimization for Science and Technology*, W. Fitzgibbon, Y. A. Kuznetsov, P. Neittaanmäki, and O. Pironneau, Eds. Dordrecht, The Netherlands: Springer Netherlands, 2014, pp. 59–82.
- [26] J. Eckstein and W. Yao, "Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives," *Pacific J. Optim.*, vol. 11, no. 4, pp. 619–644, Oct. 2015.
- [27] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, "Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists," *preprint arXiv:1912.00137*, 65 pages, 2021.
- [28] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2275–2284, Jun. 2009.
- [29] R. Volz and S. Close, "A compressed sensing approach to observing distributed radar targets," in *Proc. 30th URSI Gen. Assem. Sci. Symp. (GASS)*, Istanbul, Turkey, Aug. 2011, 4 pages.
- [30] M. A. Hadi, S. Alshebeili, K. Jamil, and F. E. A. El-Samie, "Compressive sensing applied to radar systems: An overview," *Signal Image Video Process.*, vol. 9, no. S1, pp. 25–39, Dec. 2015.
- [31] T. Huang, Y. Liu, X. Xu, Y. C. Eldar, and X. Wang, "Analysis of frequency agile radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 66, no. 23, pp. 6228–6240, Dec. 2018.
- [32] T. Huang, Y. Liu, D. Ma, N. Shlezinger, and Y. C. Eldar, "Multi-carrier agile phased array radar," in *Proc. IEEE Int. Conf. Signal Inf. Data Process. (ICSIDP)*, Chongqing, China, Dec. 2019, 6 pages.
- [33] T. Huang, N. Shlezinger, X. Xu, D. Ma, Y. Liu, and Y. C. Eldar, "Multi-carrier agile phased array radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 5706–5721, 2020.
- [34] L. Wang, T. Huang, and Y. Liu, "Theoretical analysis for extended target recovery in randomized stepped frequency radars," *preprint arXiv:1908.02929*, 12 pages, 2019.
- [35] Y. Li, T. Huang, X. Xu, Y. Liu, L. Wang, and Y. C. Eldar, "Phase transitions in frequency agile radar using compressed sensing," *IEEE Trans. Signal Process.*, vol. 69, pp. 4801–4818, 2021.
- [36] M. Rossi, A. M. Haimovich, and Y. C. Eldar, "Spatial compressive sensing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 419–430, Jan. 2014.
- [37] Y. Han and J. Wang, "Adaptive beamforming based on compressed sensing with smoothed ℓ_0 norm," *Int. J. Antennas. Propag.*, vol. 2015, 10 pages, Mar. 2015.
- [38] K. V. Mishra, I. Kahane, A. Kaufmann, and Y. C. Eldar, "High spatial resolution radar using thinned arrays," in *Proc. IEEE Radar Conf.*, Seattle, WA, USA, May 2017, pp. 1119–1124.
- [39] J. Yang, J. Thompson, X. Huang, T. Jin, and Z. Zhou, "Segmented reconstruction for compressed sensing SAR imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4214–4225, Jul. 2013.
- [40] J. Yang, T. Jin, and X. Huang, "Compressed sensing radar imaging with magnitude sparse representation," *IEEE Access*, vol. 7, pp. 29722–29733, 2019.
- [41] J. Yang, T. Jin, C. Xiao, and X. Huang, "Compressed sensing radar imaging: Fundamentals, challenges, and advances," *Sensors*, vol. 19, no. 14, 19 pages, Jul. 2019.
- [42] L. Wang, L. Zhao, G. Bi, C. Wan, and L. Yang, "Enhanced ISAR imaging by exploiting the continuity of the target scene," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5736–5750, Sep. 2014.
- [43] H. Duan, L. Zhang, J. Fang, L. Huang, and H. Li, "Pattern-coupled sparse Bayesian learning for inverse synthetic aperture radar imaging," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1995–1999, Nov. 2015.
- [44] J. Fang, L. Zhang, and H. Li, "Two-dimensional pattern-coupled sparse Bayesian learning via generalized approximate message passing," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2920–2930, Jun. 2016.
- [45] L. Jacob, G. Obozinski, and J. P. Vert, "Group lasso with overlap and graph lasso," in *Proc. 26th Annual Int. Conf. Mach. Learn. (ICML)*, Montreal, Canada, 2009, pp. 433–440.
- [46] G. Obozinski, L. Jacob, and J. P. Vert, "Group lasso with overlaps: The latent group lasso approach," *preprint arXiv:1110.0413*, 60 pages, 2011.
- [47] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Found. Trends Mach. Learn.*, vol. 4, no. 1, pp. 1–106, Jan. 2012.
- [48] S. Villa, L. Rosasco, S. Mosci, and A. Verri, "Proximal methods for the latent group lasso penalty," *Comput. Optim. Appl.*, vol. 58, no. 2, pp. 381–407, Jun. 2014.
- [49] H. Kuroda, D. Kitahara, and A. Hirabayashi, "A convex penalty for block-sparse signals with unknown structures," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Toronto, Canada, Jun. 2021, pp. 5410–5414.
- [50] H. Kuroda and D. Kitahara, "Block-sparse recovery with optimal block partition," *IEEE Trans. Signal Process.*, vol. 70, pp. 1506–1520, 2022.
- [51] D. S. Zrnic, "Estimation of spectral moments for weather echoes," *IEEE Trans. Geosci. Electron.*, vol. GE-17, no. 4, pp. 113–128, Oct. 1979.
- [52] A. S. Mudukutore, V. Chandrasekar, and R. J. Keeler, "Pulse compression for weather radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 1, pp. 125–142, Jan. 1998.
- [53] N. Anraku, M. Wada, H. Yamauchi, and A. Adachi, "Solid-state pulse compression radars in Japan," in *Proc. Conf. Radar Meteorol.*, Breckenridge, CO, USA, Sep. 2013, 12 pages.
- [54] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proc. IEEE*, vol. 55, no. 12, pp. 2143–2159, Dec. 1967.
- [55] R. L. Riegler and R. T. Compton, "An adaptive array for interference rejection," *Proc. IEEE*, vol. 61, no. 6, pp. 748–758, Jun. 1973.
- [56] P. Mosen, "MMSE equalization of interference on fading diversity channels," *IEEE Trans. Commun.*, vol. COM-32, no. 1, pp. 5–12, Jan. 1984.

- [57] S. M. Kay, *Fundamentals of Statistical Processing, Volume I: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice Hall, 1993.
- [58] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Mag.*, vol. 14, no. 6, pp. 49–83, Nov. 1997.
- [59] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 858–871, May 1997.
- [60] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge University Press, 1998.
- [61] T. Higgins, S. D. Blunt, and K. Gerlach, "Gain-constrained adaptive pulse compression via an MVDR framework," in *Proc. IEEE Radar Conf.*, Pasadena, CA, USA, May 2009, 6 pages.
- [62] F. J. Yanovsky, H. W. J. Russchenberg, and C. M. H. Unal, "Retrieval of information about turbulence in rain by using Doppler-polarimetric radar," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 2, pp. 444–450, Feb. 2005.
- [63] S. Fukao and K. Hamazu, *Radar for Meteorological and Atmospheric Observations*. Tokyo, Japan: Springer Japan, 2014.
- [64] W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, Jun. 1989.
- [65] G. H. Golub, and C. F. V. Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: Johns Hopkins University Press, 2013.
- [66] E. Yoshikawa, N. Takizawa, H. Kikuchi, T. Mega, and T. Ushio, "An estimator for weather radar Doppler power spectrum via minimum mean square error," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 16 pages, 2022.
- [67] D. Middleton, *An Introduction to Statistical Communication Theory*. New York, NY, USA: McGraw-Hill, 1960.
- [68] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2009.
- [69] H. Kikuchi, E. Yoshikawa, T. Ushio, F. Mizutani, and M. Wada, "Adaptive pulse compression technique for X-band phased array weather radar," *IEEE Geosci. Remote. Sens. Lett.*, vol. 14, no. 10, pp. 1810–1814, Oct. 2017.
- [70] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Mach. Learn.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.



Daichi Kitahara (Member, IEEE) received the B.E. degree in computer science, and the M.E. and Ph.D. degrees in communications and computer engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 2012, 2014 and 2017, respectively.

Since 2017, he has been an Assistant Professor with the College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan. His research interests are in signal processing and its applications, inverse problem, optimization theory, and multivariate spline theory.

Dr. Kitahara received the Young Researcher Award from the Institute of Electronics, Information and Communication Engineers (IEICE) Technical Group on Signal Processing in 2013, the Society of Instrument and Control Engineers (SICE) Technical Committee on Remote Sensing in 2015, and the IEEE Computational Intelligence Society Japan Chapter in 2016. He also received the IEEE Signal Processing Society (SPS) Japan Student Journal Paper Award from the IEEE SPS Tokyo Joint Chapter in 2016 and the Yasujiro Niwa Outstanding Paper Award from the Tokyo Denki University in 2018. From 2014 to 2017, he was a recipient of the Research Fellowship of the Japan Society for the Promotion of Science (JSPS).



Hiroki Kuroda (Member, IEEE) received the B.E. degree in computer science, and the M.E. and Ph.D. degrees in information and communications engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 2013, 2015 and 2019, respectively.

From April 2019 to March 2020, he was a Post-Doctoral Researcher with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan. Since April 2020, he has been an Assistant Professor with the College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan. His research interests are in signal processing and its applications, sparse modeling, and optimization theory.



Akira Hirabayashi (Member, IEEE) received the D.E. degree in computer science from the Tokyo Institute of Technology, Tokyo, Japan in 1999.

Since 2013, he has been a Professor with the College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan. His research interests include compressed sensing, sampling theory, and signal and image processing.

Prof. Hirabayashi is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).



Eiichi Yoshikawa (Member, IEEE) received the B.E. degree in aerospace engineering from Osaka Prefecture University, Sakai, Japan, in 2005, and the M.E. and Ph.D. degrees from Osaka University, Suita, Japan, in 2008 and 2011, respectively.

In 2011, he was a Post-Doctoral Researcher with Osaka University and Colorado State University, Fort Collins, CO, USA, and a Post-Doctoral Fellow of the Research Fellowship for Young Scientists sponsored by the Japan Society for the Promotion of Science (JSPS). In 2012, he joined the Japan Aerospace Exploration Agency (JAXA), Mitaka, Tokyo, Japan, where he is currently an Associate Senior Researcher. He is also a Visiting Associate Professor with Tokyo Metropolitan University, Hino, Tokyo, Japan. His research interests include weather radar remote sensing, radar-based analyses, and applications for general and aviation weather.



Hiroshi Kikuchi (Member, IEEE) received the B.S. degree from the Department of Engineering, Doshisha University, Kyoto, Japan, in 2008, and the M.S. and Ph.D. degrees from the Division of Electrical, Electronic, and Information Engineering, Osaka University, Suita, Japan, in 2010 and 2013, respectively.

He joined the Division of Electrical, Electronic, and Information Engineering, Osaka University as a Specially Appointed Researcher in 2013. In 2017, he was a Research Assistant Professor at Tokyo Metropolitan University, Hino, Tokyo, Japan. In 2018, he joined The University of Electro-Communications, where he is currently an Assistant Professor. His research specialties are remote sensing for atmospheric electricity with space-borne platforms, weather radar remote sensing, and development of radar systems.



Tomoo Ushio (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Osaka University, Suita, Japan, in 1993, 1995, and 1998, respectively.

He was with the Global Hydrology and Climate Center, Huntsville, AL, USA, as a Post-Doctoral Researcher from 1998 to 2000. In 2000, he joined the Department of Aerospace Engineering, Osaka Prefecture University, Sakai, Japan, as an Assistant Professor. In 2006, he joined the Division of Electrical, Electronic and Information Engineering, Osaka University, as an Associate Professor, where he has been a Professor since 2019. In 2017, he joined the Department of Aeronautics and Astronautics, Tokyo Metropolitan University, Hino, Tokyo, Japan, as a Professor. His research specialties are radar-based remote sensing, passive and active remote sensing of atmosphere from space-borne platforms, and atmospheric electricity.