Nonlinear Beamforming via Convex Optimization for Phased Array Weather Radar

Daichi Kitahara*, Maya Nakahara*, Akira Hirabayashi*, Eiichi Yoshikawa†, Hiroshi Kikuchi‡, and Tomoo Ushio‡
*College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan
†Aeronautical Technology Directorate, Japan Aerospace Exploration Agency, Mitaka, Tokyo, Japan
‡Department of Aerospace Engineering, Tokyo Metropolitan University, Hino, Tokyo, Japan

***Abstract*—In this paper, we propose a *nonlinear* beamforming method for a phased array weather radar (PAWR). Conventional beamforming methods are *linear* in the sense that a signal arriving from each elevation is reconstructed by computing a weighted sum of the received signals. For *distributed targets* such as raindrops, however, the number of backscattered signals is very large differently from the case for *point targets*. Thus, the spatial resolution of the linear methods is limited. To improve the spatial resolution, we focus on two characteristics of the signals from the distributed targets. One is the continuity of the reflection intensity in the time and spatial domains. The other is the narrow bandwidth in the frequency domain. These can be expressed as group-sparsity of certain matrices, and we reconstruct the signals by solving a convex optimization problem based on the group-sparsity. Simulations using real PAWR data show that the proposed method captures fine variation of precipitation profile with high precision.**

I. Introduction

Phased array weather radar (PAWR) [1], [2] has been developed to rapidly detect hazardous weather phenomena such as a thunderstorm with heavy rain. A classical parabolic radar transmits a pencil beam and receives backscattered signals within a *narrow* range of elevation angles. On the other hand, a PAWR transmits a fan beam and receives backscattered signals within a *wide* range of elevation angles simultaneously by an antenna array. Then, the backscattered signals within the *narrow* ranges are reconstructed from the received signals of the antenna array by digital *beamforming* [3]–[5]. This is the key technology in the PAWR because it gets rid of the mechanical vertical scan and hence the temporal resolution can be drastically improved in weather observation. Indeed, the PAWR developed at Osaka University [1] observes the weather in a hemisphere of a radius 60 kilometers in 30 seconds, while the classical parabolic radar requires 5 to 10 minutes for a similar observation [6].

Major beamforming methods [3]–[5] reconstruct the signal arriving from each elevation as a (complex) weighted sum of the received signals. In particular, Capon beamforming [4] is a famous method that can adaptively reduce the influence of sidelobes if a sufficient number of pulses are transmitted. For fast weather observation, however, the number of pulses should be as small as possible. To deal with such a situation, the minimum mean square error (MMSE) beamforming [5] was proposed. In this method, differently from Capon's method, the sample covariance matrix of the received signals is not used, and hence the signal arriving from each elevation is robustly reconstructed even if the number of pulses is small. Such beamforming methods were developed originally for observation of *point targets*, but targets of the PAWR are *distributed targets* such as raindrops. In this case, the number of backscattered signals is very large, and the spatial resolution of the above *linear* methods [3]–[5] is limited, i.e., fine variation of the reflection intensity corresponding to precipitation profile is not captured.

To overcome the limitation of the *linear* methodology, this paper proposes a *nonlinear* beamforming method. We formulate the beamforming as an ill-conditioned inverse problem. To solve it, we utilize two properties of signals from distributed targets. One is the continuity of the reflection intensity in the time and spatial domains. The other is the narrow bandwidth in the frequency domain. These properties can be expressed as group-sparsity of certain two matrices, and we reconstruct the signals by minimizing a cost function that consists of the data-fidelity term and two group ℓ_1-norms. The optimal solution is effectively computed with the alternating direction method of multipliers (ADMM) [7]. Numerical experiments based on real PAWR data show the effectiveness of the proposed nonlinear beamforming in comparison with the linear methods [3]–[5].

This work was supported by JSPS KAKENHI Grant Number 17H02069.

II. Preliminaries

Let $\mathbb{R}$ and $\mathbb{C}$ be the sets of all real numbers and complex numbers, respectively. We use $j \in \mathbb{C}$ to denote the imaginary unit, i.e., $j = \sqrt{-1}$. For any $x \in \mathbb{C}$, $\bar{x}$ denotes its complex conjugate, and $|x| := \sqrt{x\bar{x}}$ denotes its absolute value. We write vectors with lowercase boldface letters and matrices with capital letters. We use $I_n \in \mathbb{R}^{n\times n}$ to denote the identity matrix of order n. The transpose and the Hermitian transpose of vectors or matrices are expressed as $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$, respectively. The ℓ_2-norm (or the Euclidean norm) of $\boldsymbol{x} := (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{C}^n$ is defined as $\|\boldsymbol{x}\|_2 := \sqrt{\sum_{i=1}^{n} |x_i|^2}$, and a group ℓ_1-norm with non-overlapping groups is defined as $\|\boldsymbol{x}\|_1^G := \sum_{i=1}^{n_G} \|\boldsymbol{x}_{\mathcal{G}_i}\|_2$, where $\boldsymbol{x}_{\mathcal{G}_i}$ $(i = 1, 2, \ldots, n_G)$ are sub-vectors of $\boldsymbol{x}$ divided by $\{\mathcal{G}_i\}_{i=1}^{n_G}$ s.t. $\bigcup_{i=1}^{n_G} \mathcal{G}_i = \{1, 2, \ldots, n\}$ and $\mathcal{G}_i \cap \mathcal{G}_{i'} = \varnothing$ $(i \neq i')$. We use $E[\cdot]$ to denote the expected values of random variables.

A. Signal Model

First of all, we give a signal model for observation of K *point targets*. Let a PAWR have an N-element uniform linear array with the inter-element spacing $d\,[\mathrm{m}]$. A plane wave signal scattered from the kth point target impinges on the antenna array at an angle $\theta_k^\star \in [\theta_{\min}, \theta_{\max}]$ $(\theta_1^\star < \theta_2^\star < \cdots < \theta_K^\star\,[\mathrm{rad}])$. Then,

the lth time sample of the received signal $\boldsymbol{y}_l \in \mathbb{C}^N$ is given by

$$\boldsymbol{y}_l = \sum_{k=1}^{K} x^\star_{k,l}\boldsymbol{a}(\theta^\star_k) + \boldsymbol{v}_l \qquad (l = 1, 2, \ldots, L), \tag{1}$$

where $x^\star_{k,l} \in \mathbb{C}$ is the lth sample of the kth plane wave signal s.t. $E[x^\star_{k,l}] = 0$, $\boldsymbol{a}(\theta^\star_k) \in \mathbb{C}^N$ is the steering vector defined by

$$\boldsymbol{a}(\theta) := \left(1, e^{-j\frac{2\pi d \sin\theta}{\lambda}}, e^{-j\frac{4\pi d \sin\theta}{\lambda}}, \ldots, e^{-j\frac{2(N-1)\pi d \sin\theta}{\lambda}}\right)^{\mathrm{T}}$$

with the carrier wavelength $\lambda\,[\mathrm{m}]$, and $\boldsymbol{v}_l \in \mathbb{C}^N$ is white Gaussian noise whose covariance matrix is $R_v := E[\boldsymbol{v}_l\boldsymbol{v}_l^{\mathrm{H}}] = \sigma_v^2 I_N$.

On the other hand, our targets such as raindrops are called *distributed targets*, which are supposed to exist continuously (strictly speaking, a sufficient number of raindrops exist within the antenna beamwidth). Let us observe the distributed targets while dividing the whole angular interval $[\theta_{\min}, \theta_{\max}]$ into M sub-intervals $[\theta_m - \frac{\Delta\theta}{2}, \theta_m + \frac{\Delta\theta}{2}]$, where $\Delta\theta := \frac{\theta_{\max}-\theta_{\min}}{M}$ and $\theta_m := \theta_{\min} + (m - \frac{1}{2})\Delta\theta$ $(m = 1, 2, \ldots, M)$. Therefore, instead of (1), we use the following signal model

$$\boldsymbol{y}_l = \sum_{m=1}^{M} x_{m,l}\boldsymbol{s}_m + \boldsymbol{v}_l = S\boldsymbol{x}_l + \boldsymbol{v}_l, \tag{2}$$

where $x_{m,l} \in \mathbb{C}$ s.t. $E[x_{m,l}] = 0$ is the lth sample of the sum of plane wave signals in the sub-interval $[\theta_m - \frac{\Delta\theta}{2}, \theta_m + \frac{\Delta\theta}{2}]$, $\boldsymbol{x}_l := (x_{1,l}, x_{2,l}, \ldots, x_{M,l})^{\mathrm{T}} \in \mathbb{C}^M$, $\boldsymbol{s}_m := \boldsymbol{a}(\theta_m) \in \mathbb{C}^N$, and $S := (\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_M) \in \mathbb{C}^{N\times M}$. Furthermore, we can derive (1) from (2) by redefining K $(\leq M)$ as the number of the sub-intervals $[\theta_m - \frac{\Delta\theta}{2}, \theta_m + \frac{\Delta\theta}{2}]$ where plane wave signals exist, and $\theta^\star_k$ as the center of such a sub-interval. In PAWR systems, θ_m means the mth elevation angle, and the reflection intensity

$$\boldsymbol{p} := \left(E\left[|x_{1,l}|^2\right], E\left[|x_{2,l}|^2\right], \ldots, E\left[|x_{M,l}|^2\right]\right)^{\mathrm{T}} \in \mathbb{R}^M$$

corresponds to precipitation profile in the elevation angles.

B. Linear Beamforming

Beamforming is an estimation problem of $\boldsymbol{x}_l$ from $\boldsymbol{y}_l$ in (2). Major beamforming methods [3]–[5] estimate $\boldsymbol{x}_l$ by multiplying complex weights $\boldsymbol{w}_m \in \mathbb{C}^N$ $(m = 1, 2, \ldots, M)$ and $\boldsymbol{y}_l$ as

$$\begin{aligned}\widehat{\boldsymbol{x}}_l &:= (\widehat{x}_{1,l}, \widehat{x}_{2,l}, \ldots, \widehat{x}_{M,l})^{\mathrm{T}} \\ &:= (\boldsymbol{w}_1^{\mathrm{H}}\boldsymbol{y}_l, \boldsymbol{w}_2^{\mathrm{H}}\boldsymbol{y}_l, \ldots, \boldsymbol{w}_M^{\mathrm{H}}\boldsymbol{y}_l)^{\mathrm{T}} = W\boldsymbol{y}_l.\end{aligned} \tag{3}$$

In this paper, the methods based on (3) are called the *linear beamforming*. Note that the least squares (LS) method

$$\widehat{\boldsymbol{x}}_{\mathrm{LS},l} := W_{\mathrm{LS}}\,\boldsymbol{y}_l := S^\dagger \boldsymbol{y}_l \tag{4}$$

does not necessarily work well, even if $N \geq M$, because S is ill-conditioned when $\Delta\theta$ is smaller than the antenna beamwidth that is determined by the antenna size, where $S^\dagger \in \mathbb{C}^{M\times N}$ is the Moore-Penrose pseudoinverse of S. In the following, we introduce three linear methods, *Fourier (FR) beamforming* [3], *Capon (CP) beamforming* [4], and *MMSE beamforming* [5].

1) FR Beamforming: FR beamforming [3] is the most basic method and its complex weight vector is defined by

$$\boldsymbol{w}_{\mathrm{FR},m} := \frac{\boldsymbol{s}_m}{N} \tag{5}$$

independently of $\boldsymbol{y}_l$. The weight vector $\boldsymbol{w}_{\mathrm{FR},m}$ is a matched filter that maximizes the signal-to-noise ratio $\frac{E[|x_{m,l}\boldsymbol{w}_m^{\mathrm{H}}\boldsymbol{s}_m|^2]}{E[|\boldsymbol{w}_m^{\mathrm{H}}\boldsymbol{v}_l|^2]}$. However, from

$$\widehat{x}_{\mathrm{FR},m,l} := \frac{\boldsymbol{s}_m^{\mathrm{H}}}{N}\boldsymbol{y}_l = x_{m,l} + \frac{1}{N}\sum_{\theta^\star_k \neq \theta_m} x^\star_{k,l}\boldsymbol{s}_m^{\mathrm{H}}\boldsymbol{a}(\theta^\star_k) + \frac{1}{N}\boldsymbol{s}_m^{\mathrm{H}}\boldsymbol{v}_l,$$

the precipitation profile is overestimated for many elevation angles θ_m since

$$\widehat{p}_{\mathrm{FR},m} := \frac{1}{L}\sum_{l=1}^{L}|\widehat{x}_{\mathrm{FR},m,l}|^2 \gg E\left[|x_{m,l}|^2\right] + \frac{\sigma_v^2}{N}$$

often holds for m satisfying $\exists\theta^\star_k \neq \theta_m$ $|\boldsymbol{s}_m^{\mathrm{H}}\boldsymbol{a}(\theta^\star_k)| \not\approx 0$.

2) CP Beamforming: CP beamforming [4] is an adaptive method and minimizes $\frac{1}{L}\sum_{l=1}^{L}|\widehat{x}_{m,l}|^2 = \boldsymbol{w}_m^{\mathrm{H}}\widehat{R}_y\boldsymbol{w}_m$ under the condition $\boldsymbol{w}_m^{\mathrm{H}}\boldsymbol{s}_m = 1$ to avoid the above overestimation, where $\widehat{R}_y := \frac{1}{L}\sum_{l=1}^{L}\boldsymbol{y}_l\boldsymbol{y}_l^{\mathrm{H}} \in \mathbb{C}^{N\times N}$ is the sample covariance matrix of the zero-mean random variable $\boldsymbol{y}_l$. The weight vector $\boldsymbol{w}_{\mathrm{CP},m}$ is defined as the solution of the optimization problem

$$\underset{\boldsymbol{w}_m}{\text{minimize}}\ \boldsymbol{w}_m^{\mathrm{H}}\widehat{R}_y\boldsymbol{w}_m \qquad \text{subject to } \boldsymbol{w}_m^{\mathrm{H}}\boldsymbol{s}_m = 1$$

by

$$\boldsymbol{w}_{\mathrm{CP},m} := \frac{\widehat{R}_y^{-1}\boldsymbol{s}_m}{\boldsymbol{s}_m^{\mathrm{H}}\widehat{R}_y^{-1}\boldsymbol{s}_m} \tag{6}$$

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

$$\widehat{p}_{\mathrm{CP},m} := \frac{1}{L}\sum_{l=1}^{L}|\widehat{x}_{\mathrm{CR},m,l}|^2 \approx E\left[|x_{m,l}|^2\right] + \sigma_v^2\|\boldsymbol{w}_{\mathrm{CP},m}\|_2^2$$

for all m since $\forall\theta^\star_k \neq \theta_m$ $|\boldsymbol{w}_{\mathrm{CP},m}^{\mathrm{H}}\boldsymbol{a}(\theta^\star_k)| \approx 0$ holds. However, if L is not large, then the precipitation profile is often underestimated [5]. Moreover, if $L < N$, $\widehat{R}_y^{-1}$ cannot be computed.

3) MMSE Beamforming: MMSE beamforming [5] was developed to improve the estimation accuracy in case of small L. This method approximately solves the optimization problem

$$\underset{\boldsymbol{w}_m}{\text{minimize}}\ E\left[|x_{m,l} - \boldsymbol{w}_m^{\mathrm{H}}\boldsymbol{y}_l|^2\right] \qquad \text{subject to } \boldsymbol{w}_m^{\mathrm{H}}\boldsymbol{s}_m = 1.$$

Assuming $E[x_{m,l}\bar{x}_{m',l}] = E[x_{m,l}]E[\bar{x}_{m',l}] = 0$ if $m \neq m'$, the exact solution of the above optimization problem is

$$\boldsymbol{w}_{\mathrm{MMSE},m} := \frac{R_y^{-1}\boldsymbol{s}_m}{\boldsymbol{s}_m^{\mathrm{H}}R_y^{-1}\boldsymbol{s}_m}, \tag{7}$$

where $R_y := E[\boldsymbol{y}_l\boldsymbol{y}_l^{\mathrm{H}}]$ is the covariance matrix of $\boldsymbol{y}_l$. By using the covariance matrix $R_x := E[\boldsymbol{x}_l\boldsymbol{x}_l^{\mathrm{H}}]$ of $\boldsymbol{x}_l$, R_y is expressed as $R_y = SR_xS^{\mathrm{H}} + \sigma_v^2 I_N$. Moreover, $R_x = \mathrm{diag}(\boldsymbol{p})$ can be approximated by $\widehat{R}_x \odot I_M := (\frac{1}{L}\sum_{l=1}^{L}\boldsymbol{x}_l\boldsymbol{x}_l^{\mathrm{H}}) \odot I_M$, where $\odot$ denotes the Hadamard product. As a result, the weight vector $\boldsymbol{w}_{\mathrm{MMSE},m}$ in (7) is approximated, from the initial estimate $\widehat{\boldsymbol{x}}^{(0)}_{\mathrm{MMSE},l} = \widehat{\boldsymbol{x}}_{\mathrm{FR},l} = S^{\mathrm{H}}\boldsymbol{y}_l/N$, by iteratively computing

$$\left|\begin{aligned} R_x^{(i)} &= \left(\frac{1}{L}\sum_{l=1}^{L}\widehat{\boldsymbol{x}}^{(i)}_{\mathrm{MMSE},l}\widehat{\boldsymbol{x}}^{(i)\mathrm{H}}_{\mathrm{MMSE},l}\right) \odot I_M \\ R_y^{(i)} &= SR_x^{(i)}S^{\mathrm{H}} + \sigma_v^2 I_N \\ \boldsymbol{w}^{(i+1)}_{\mathrm{MMSE},m} &= \frac{R_y^{(i)-1}\boldsymbol{s}_m}{\boldsymbol{s}_m^{\mathrm{H}}R_y^{(i)-1}\boldsymbol{s}_m} \qquad (m = 1, 2, \ldots, M) \\ \widehat{\boldsymbol{x}}^{(i+1)}_{\mathrm{MMSE},l} &= W^{(i+1)}_{\mathrm{MMSE}}\,\boldsymbol{y}_l \qquad (l = 1, 2, \ldots, L)\end{aligned}\right. \tag{8}$$

for $i \geq 0$ until $\delta^{(i+1)} = \frac{1}{M}\sum_{m=1}^{M} \frac{\sum_{l=1}^{L} |\widehat{x}^{(i+1)}_{\mathrm{MMSE},m,l} - \widehat{x}^{(i)}_{\mathrm{MMSE},m,l}|^2}{\sum_{l=1}^{L} |\widehat{x}^{(i)}_{\mathrm{MMSE},m,l}|^2}$ becomes sufficiently small. In this method, even for small L, R_y can be stably estimated. However, if K is close to N or larger than N, then the estimation accuracy degrades, i.e., fine variation of the precipitation profile cannot be captured.

III. Nonlinear Beamforming via Convex Optimization Based on Double Group-Sparsity

In this section, we propose a *nonlinear beamforming* method based on convex optimization. First, we gather $\boldsymbol{x}_l$ and $\boldsymbol{y}_l$ into $X := (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L) = (\widetilde{\boldsymbol{x}}_1, \widetilde{\boldsymbol{x}}_2, \ldots, \widetilde{\boldsymbol{x}}_M)^{\mathrm{T}} \in \mathbb{C}^{M\times L}$ and $Y := (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L) \in \mathbb{C}^{N\times L}$, respectively, where $\widetilde{\boldsymbol{x}}_m := (x_{m,1}, x_{m,2}, \ldots, x_{m,L})^{\mathrm{T}} \in \mathbb{C}^L$. Then the beamforming can be translated into an estimation problem of X from Y, and the data fidelity in (2) can be evaluated by the Frobenius norm as

$$\|Y - SX\|_{\mathrm{F}}^2 := \sum_{l=1}^{L} \|\boldsymbol{y}_l - S\boldsymbol{x}_l\|_2^2. \tag{9}$$

In the following, after describing two characteristics on $\widetilde{\boldsymbol{x}}_m$ as group-sparsity of certain matrices, we solve a convex optimization problem based on (9) and the double group-sparsity.

A. Continuity of Precipitation Profile in Elevation Angles

Many PAWR systems employ contiguous pair sampling [8]. In such a system, the pulse repetition time (PRT) is designed by $T_{\mathrm{PRT}} := \frac{2r_{\max}}{c}$, where $r_{\max}$ is the maximum range to be observed and c is the speed of light. For example, the PAWR system developed at Osaka University observes the weather in a hemisphere of a radius 60 kilometers [1] and hence $T_{\mathrm{PRT}} \approx 0.0004\,[\mathrm{s}]$. If L is not large, the total observation time LT_{PRT} is sufficiently short to consider $\theta_k^\star$ to be constant for the time index $l = 1, 2, \ldots, L$. Therefore, $\widetilde{\boldsymbol{x}}_m$ becomes a dense vector if $\theta_m = \theta_k^\star$ for some k, and $\widetilde{\boldsymbol{x}}_m = \boldsymbol{0}$ otherwise. Furthermore, from the continuity of precipitation profile, if $\widetilde{\boldsymbol{x}}_m = \boldsymbol{0}$, it is highly possible that $\widetilde{\boldsymbol{x}}_{m-1}$ and $\widetilde{\boldsymbol{x}}_{m+1}$ are also $\boldsymbol{0}$. This property is expressed as group-sparsity of X and evaluated by

$$\|X\|_1^{G_1} := \sum_{i=1}^{M/q} \left\|(\widetilde{\boldsymbol{x}}_{(i-1)q+1}^{\mathrm{T}}, \widetilde{\boldsymbol{x}}_{(i-1)q+2}^{\mathrm{T}}, \ldots, \widetilde{\boldsymbol{x}}_{iq}^{\mathrm{T}})^{\mathrm{T}}\right\|_2 \tag{10}$$

with the use of a factor $q \geq 1$ of M.

B. Narrow Bandwidth of Signals from Distributed Scatterers

The power spectral density of the backscattered signal from the distributed target, such as raindrops, fog droplets and cloud droplets, can be modeled by a Gaussian function according to the central limit theorem [9], [10]. The first raw moment of the normalized power spectral density, i.e., the center of the above Gaussian function, is called the *mean Doppler frequency* (or the *mean Doppler shift*). The square root of the second central moment, i.e., the standard deviation of the Gaussian function, is called the *Doppler frequency spectrum width* (see Fig. 1).

Define the normalized discrete Fourier transform matrix by $F := \frac{1}{\sqrt{L}}(\boldsymbol{f}_0, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_{L-1}) \in \mathbb{C}^{L\times L}$, where

$$\boldsymbol{f}_i := \left(1, e^{-j\frac{2\pi i}{L}}, e^{-j\frac{4\pi i}{L}}, \ldots, e^{-j\frac{2(L-1)\pi i}{L}}\right)^{\mathrm{T}} \in \mathbb{C}^L.$$

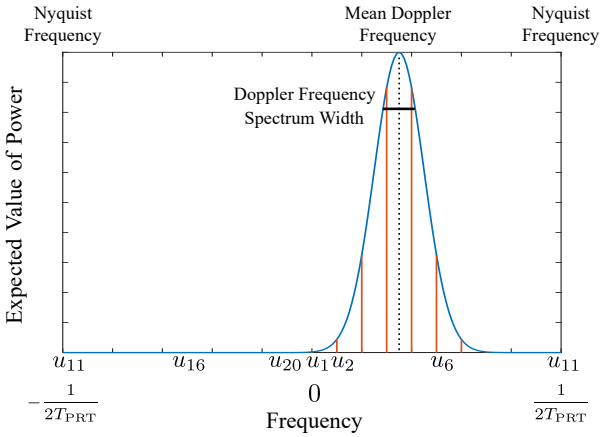

Fig. 1. Power spectral density model of backscattered signals from distributed targets. Blue and red lines respectively depict the power spectral density and $E[|u_{m,l}|^2]$ ($l = 1, 2, \ldots, L := 20$). In the graph, m is omitted for simplicity.

Then the vector $\boldsymbol{u}_m := F\widetilde{\boldsymbol{x}}_m = (u_{m,1}, u_{m,2}, \ldots, u_{m,L})^{\mathrm{T}} \in \mathbb{C}^L$ is group-sparse, i.e., the indices l having large $|u_{m,l}|$ concentrate in the vicinity of the mean Doppler frequency as shown in Fig. 1. However, we cannot specify the center and the width of such a group because the mean Doppler frequency and the spectrum width are different for each elevation angle θ_m. Alternatively, we divide $\boldsymbol{u}_m$ into L overlapping blocks of size b:

$$\begin{cases} \boldsymbol{b}_{m,1} := (u_{m,1}, u_{m,2}, \ldots, u_{m,b})^{\mathrm{T}} \in \mathbb{C}^b \\ \boldsymbol{b}_{m,2} := (u_{m,2}, u_{m,3}, \ldots, u_{m,b+1})^{\mathrm{T}} \in \mathbb{C}^b \\ \qquad\vdots \\ \boldsymbol{b}_{m,L-b+1} := (u_{m,L-b+1}, u_{m,L-b+2}, \ldots, u_{m,L})^{\mathrm{T}} \in \mathbb{C}^b \\ \boldsymbol{b}_{m,L-b+2} := (u_{m,L-b+2}, \ldots, u_{m,L}, u_{m,1})^{\mathrm{T}} \in \mathbb{C}^b \\ \qquad\vdots \\ \boldsymbol{b}_{m,L} := (u_{m,L}, u_{m,1}, \ldots, u_{m,b-1})^{\mathrm{T}} \in \mathbb{C}^b \end{cases}$$

under the periodic boundary condition. As a result, if we define a matrix $B \in \mathbb{R}^{bL\times L}$ satisfying

$$BF\widetilde{\boldsymbol{x}}_m = B\boldsymbol{u}_m = (\boldsymbol{b}_{m,1}^{\mathrm{T}}, \boldsymbol{b}_{m,2}^{\mathrm{T}}, \ldots, \boldsymbol{b}_{m,L}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{C}^{bL},$$

the matrix BFX^{T} becomes group-sparse without overlapping, and this property can be evaluated by a group ℓ_1-norm

$$\|BFX^{\mathrm{T}}\|_1^{G_2} := \sum_{m=1}^{M}\sum_{l=1}^{L} \|\boldsymbol{b}_{m,l}\|_2. \tag{11}$$

C. The Proposed Nonlinear Beamforming

On the basis of (9), (10), and (11), we estimate X by solving a convex optimization problem

$$\underset{X}{\text{minimize}}\ \frac{1}{2}\|Y - SX\|_{\mathrm{F}}^2 + \nu_1\|X\|_1^{G_1} + \nu_2\|BFX^{\mathrm{T}}\|_1^{G_2} \tag{12}$$

with the use of ADMM [7] (see Appendix), where $\nu_1 > 0$ and $\nu_2 > 0$. The problem in (12) is expressed as an ADMM-form

$$\underset{X\in\mathcal{X}, Z\in\mathcal{Z}}{\text{minimize}}\ \frac{1}{2}\|Y - SX\|_{\mathrm{F}}^2 + \nu_1\|Z_1\|_1^{G_1} + \nu_2\|Z_2\|_1^{G_2}$$
$$\text{subject to } Z := \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \mathcal{L}(X) := \begin{bmatrix} I_M \\ BF\circ\mathcal{T} \end{bmatrix}(X), \tag{13}$$

where $\mathcal{X} := \mathbb{C}^{M\times L}$, $\mathcal{Z} := \mathbb{C}^{M\times L}\times\mathbb{C}^{bL\times M}$, $\mathcal{T}$ is the transpose operator, $\circ$ denotes the composition of mappings, and convex functions f and g in (19) are respectively defined by $f(X) := \frac{1}{2}\|Y - SX\|_{\mathrm{F}}^2$ and $g(Z) := \nu_1\|Z_1\|_1^{G_1} + \nu_2\|Z_2\|_1^{G_2}$.

On the first line in (20), since X is updated as the solution of a least squares problem, the solution $X^{(i+1)}$ satisfies

$$\left(S^{\mathrm{H}}S + \frac{1}{\gamma}\mathcal{L}^*\circ\mathcal{L}\right)(X^{(i+1)}) = S^{\mathrm{H}}Y + \frac{1}{\gamma}\mathcal{L}^*(Z^{(i)} - D^{(i)}), \tag{14}$$

where $\mathcal{L}^* : \mathcal{Z}\to\mathcal{X}$ is the adjoint operator of $\mathcal{L}$ and defined by $\mathcal{L}^*(Z) := I_M(Z_1) + \mathcal{T}\circ F^{\mathrm{H}}B^{\mathrm{T}}(Z_2) = Z_1 + Z_2^{\mathrm{T}}BF^{\mathrm{H}}$. Moreover, the composite mapping $\mathcal{L}^*\circ\mathcal{L}$ is expressed as

$$\begin{aligned}\mathcal{L}^*\circ\mathcal{L} &= \begin{bmatrix} I_M & \mathcal{T}\circ F^{\mathrm{H}}B^{\mathrm{T}}\end{bmatrix}\begin{bmatrix} I_M \\ BF\circ\mathcal{T}\end{bmatrix}\\ &= I_M + \mathcal{T}\circ F^{\mathrm{H}}B^{\mathrm{T}}BF\circ\mathcal{T}\\ &= I_M + \mathcal{T}\circ F^{\mathrm{H}}(bI_L)F\circ\mathcal{T}\\ &= I_M + \mathcal{T}\circ bI_L\circ\mathcal{T} = (1+b)\,I_M. \end{aligned}\tag{15}$$

By substituting (15) into (14), $X^{(i+1)}$ is computed by

$$\begin{aligned}X^{(i+1)} = &\left(S^{\mathrm{H}}S + \frac{1+b}{\gamma}I_M\right)^{-1}\\ &\cdot\left(S^{\mathrm{H}}Y + \frac{1}{\gamma}\left(Z_1^{(i)} - D_1^{(i)} + \left(Z_2^{(i)\mathrm{T}} - D_2^{(i)\mathrm{T}}\right)BF^{\mathrm{H}}\right)\right).\end{aligned}\tag{16}$$

Next, on the second line in (20), since the computation of $g(Z)$ is divided into $\nu_1\|Z_1\|_1^{G_1}$ and $\nu_2\|Z_2\|_1^{G_2}$, $Z^{(i+1)}$ is computed by

$$\begin{cases} Z_1^{(i+1)} = \mathrm{prox}_{\gamma\nu_1\|\cdot\|_1^{G_1}}\left(X^{(i+1)} + D_1^{(i)}\right)\\ Z_2^{(i+1)} = \mathrm{prox}_{\gamma\nu_2\|\cdot\|_1^{G_2}}\left(BFX^{(i+1)\mathrm{T}} + D_2^{(i)}\right)\end{cases}\tag{17}$$

with the use of the proximity operators of the group ℓ_1-norms in (21). Finally, on the third line in (20), $D^{(i+1)}$ is computed by

$$\begin{cases} D_1^{(i+1)} = D_1^{(i)} + X^{(i+1)} - Z_1^{(i+1)}\\ D_2^{(i+1)} = D_2^{(i)} + BFX^{(i+1)\mathrm{T}} - Z_2^{(i+1)}\end{cases}\tag{18}$$

and the optimal solution of the problem in (13) is obtained by repeating (16)–(18) until a convergence condition is satisfied.

IV. Numerical Experiments

To show the effectiveness of the proposed nonlinear beamforming, we conducted simulations based on the real reflection intensity, observed by the PAWR at Osaka University, in Fig. 2. At the range $r = 7.5\,[\mathrm{km}]$, we picked out 55 samples between $\theta_{\min} = -15°\,[\mathrm{deg}]$ and $\theta_{\max} = 30°\,[\mathrm{deg}]$. The true reflection intensity $\boldsymbol{p}$ was created by cubic spline interpolation of these 55 samples followed by adding Gaussian random numbers. We set $\lambda = 0.0318\,[\mathrm{m}]$, $d = 0.0165\,[\mathrm{m}]$, and $T_{\mathrm{PRT}} = 0.0004\,[\mathrm{s}]$. Random signals $\widetilde{\boldsymbol{x}}_m$ were generated in the frequency domain so that $E[|u_{m,l}|^2]$ would follow a Gaussian distribution, which was wrapped into $[-\frac{1}{2T_{\mathrm{PRT}}}, \frac{1}{2T_{\mathrm{PRT}}}]$, on the basis of [9]. The mean Doppler frequency for each elevation angle was generated from a uniform distribution $\mathcal{U}(-\frac{1}{2T_{\mathrm{PRT}}}, \frac{1}{2T_{\mathrm{PRT}}})$, and the Doppler frequency spectrum width was simply fixed to $\sigma = 125.7\,[\mathrm{Hz}]$. The standard deviation of $\boldsymbol{v}_l$ was set to $\sigma_v = \sqrt{5}$.

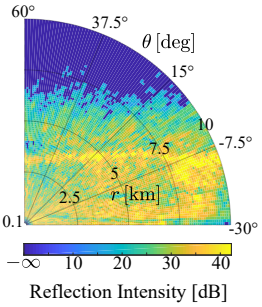

Fig. 2. Real PAWR data.

TABLE I
Mean of the Normalized Errors [%] of Each Beamformer in 10 Trials

Method	$\Delta\theta = 45°/110$ $K = 93 < N$ $L = 128,\ 20$	$\Delta\theta = 45°/160$ $K = 136 > N$ $L = 128,\ 20$
LS	73.64, 74.39	93.77, 92.17
FR	104.2, 104.1	142.5, 143.3
CP	99.85, ——	100.1, ——
MMSE	77.56, 74.24	115.2, 111.9
Proposed	**20.04**, **24.63**	**43.29**, **47.35**

For[1] $N = 128$ and $M = 110, 160$, we compared the proposed method with LS in (4),[2] FR in (5), CP in (6), and MMSE in (8), in cases of $L = 128$ and $L = 20$. The parameters in (12) were set to $\nu_1 = 0.0025\frac{N\sqrt{qL}}{M}$ and $\nu_2 = 0.25\frac{N}{M\sqrt{b}}$. The group sizes in $\|\cdot\|_1^{G_1}$ and $\|\cdot\|_1^{G_2}$ were set to $q = 5$ for $M = 110$, $q = 8$ for $M = 160$, $b = 3$ for $L = 20$, and $b = 10$ for $L = 128$.

Table I summarizes the average, for each situation, of the normalized errors $100\,\|\widehat{X} - X\|_{\mathrm{F}}/\|X\|_{\mathrm{F}}\,[\%]$ in 10 trials, where the results of CP are empty for $L = 20$ because $\widehat{R}_y^{-1}$ cannot be computed. From Table I, we can see that the proposed method reduced the normalized errors by 70–80 [%] when $M = 110$ and 50–70 [%] when $M = 160$. Figures 3, 4, 5, and 6 depict the reflection intensity $E[|x_{m,l}|^2]$ and estimates $\frac{1}{L}\sum_{l=1}^{L}|\widehat{x}_{m,l}|^2$ for $(M, L) = (110, 128), (110, 20), (160, 128), (160, 20)$. When $M = 110$, the proposed method accurately reconstructed the true reflection intensity in case of $L = 128$. LS and FR failed in estimation particularly from 23° [deg] to 27° [deg] because the reflection intensity should be 0 in this interval. CP severely underestimated the reflection intensity in the whole angles, and MMSE reconstructed only a rough shape of the reflection intensity. Even in case of $L = 20$, the proposed method achieved the highest estimation accuracy though the reflection intensity was underestimated at several angles. When $M = 160$, the proposed method could reconstruct the true reflection intensity at many angles in case of $L = 128$. In case of $L = 20$, although the proposed method achieved the highest estimation accuracy, the estimated reflection intensity became 0 in several angles.

V. Conclusion

In this paper, we have proposed a nonlinear beamforming method for a PAWR. Differently from a radar that observes point targets, the PAWR receives a lot of backscattered signals from distributed targets, and the performance of linear beamforming methods degrades. We regarded the beamforming as an inverse problem and estimated the backscattered signals by solving an optimization problem based on the group-sparsity. Numerical experiments showed that, compared with the linear beamforming, the proposed method can reconstruct precipitation profile with higher accuracy from a few time samples.

[1]When $M = 110$, the number of the sub-intervals, where signals exist, is $K = 93 < N$. On the other hand, when $M = 160$, the number of the sub-intervals is $K = 136 > N$, and hence it is very difficult to estimate X.

[2]To avoid the numerical instability in the computation of $S^{\dagger}$, we truncated the singular values of S that are smaller than 0.005.

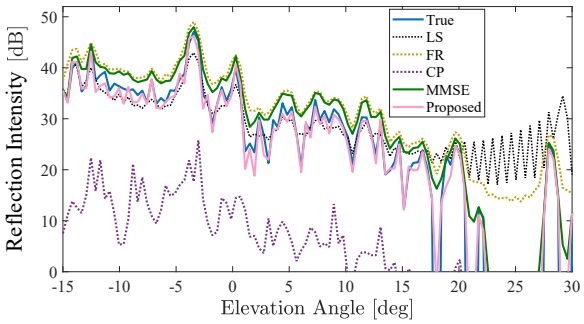

Fig. 3. Estimates of the reflection intensity for $(M, L) = (110, 128)$.

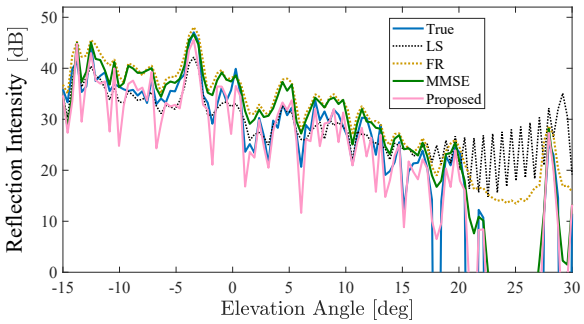

Fig. 4. Estimates of the reflection intensity for $(M, L) = (110, 20)$.

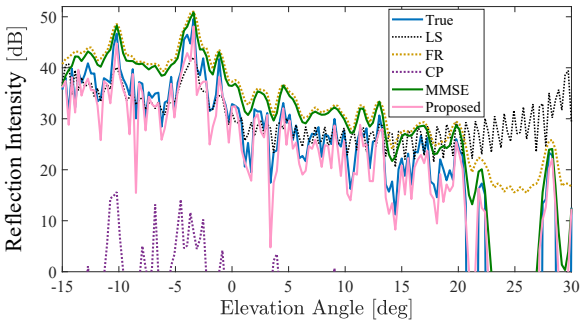

Fig. 5. Estimates of the reflection intensity for $(M, L) = (160, 128)$.

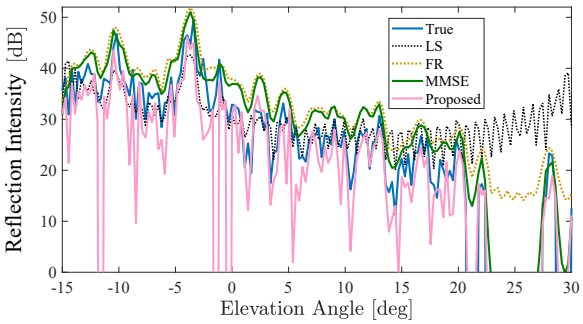

Fig. 6. Estimates of the reflection intensity for $(M, L) = (160, 20)$.

APPENDIX
ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Let us consider the following convex optimization problem:

$$\underset{\boldsymbol{x}\in\mathcal{X},\,\boldsymbol{z}\in\mathcal{Z}}{\text{minimize}}\ f(\boldsymbol{x}) + g(\boldsymbol{z}) \qquad \text{subject to } \boldsymbol{z} = \mathcal{L}(\boldsymbol{x}), \tag{19}$$

where $\mathcal{X}$ and $\mathcal{Z}$ are finite-dimensional Hilbert spaces with the standard inner products, $\mathcal{L} : \mathcal{X} \to \mathcal{Z}$ is a linear mapping, and functions $f : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ and $g : \mathcal{Z} \to \mathbb{R} \cup \{\infty\}$ are proper, lower semicontinuous, and convex.[3] The alternating direction method of multipliers (ADMM) [7] solves the problem in (19) by iteratively computing, from any $(\boldsymbol{z}^{(0)}, \boldsymbol{d}^{(0)}) \in \mathcal{Z} \times \mathcal{Z}$,

$$\begin{cases} \boldsymbol{x}^{(i+1)} = \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmin}}\ f(\boldsymbol{x}) + \dfrac{1}{2\gamma}\|\boldsymbol{z}^{(i)} - \mathcal{L}(\boldsymbol{x}) - \boldsymbol{d}^{(i)}\|_{\mathcal{Z}}^2 \\ \boldsymbol{z}^{(i+1)} = \operatorname{prox}_{\gamma g}(\mathcal{L}(\boldsymbol{x}^{(i+1)}) + \boldsymbol{d}^{(i)}) \\ \boldsymbol{d}^{(i+1)} = \boldsymbol{d}^{(i)} + \mathcal{L}(\boldsymbol{x}^{(i+1)}) - \boldsymbol{z}^{(i+1)} \end{cases} \tag{20}$$

for $i \geq 0$, where $\gamma > 0$, $\|\cdot\|_{\mathcal{Z}}$ is the Euclidean norm introduced by the standard inner product in $\mathcal{Z}$, and $\operatorname{prox}_{\gamma g} : \mathcal{Z} \to \mathcal{Z}$ is the *proximity operator* defined by

$$\operatorname{prox}_{\gamma g}(\boldsymbol{y}) := \underset{\boldsymbol{z}\in\mathcal{Z}}{\operatorname{argmin}}\ g(\boldsymbol{z}) + \frac{1}{2\gamma}\|\boldsymbol{z} - \boldsymbol{y}\|_{\mathcal{Z}}^2.$$

If g is a group ℓ_1-norm with non-overlapping groups $\{\mathcal{G}_i\}_{i=1}^{n_G}$, from $g(\boldsymbol{z}) + \frac{1}{2\gamma}\|\boldsymbol{z} - \boldsymbol{y}\|_{\mathcal{Z}}^2 = \sum_{i=1}^{n_G}(\|\boldsymbol{z}_{\mathcal{G}_i}\|_2 + \frac{1}{2\gamma}\|\boldsymbol{z}_{\mathcal{G}_i} - \boldsymbol{y}_{\mathcal{G}_i}\|_2^2)$, the computation of $\operatorname{prox}_{\gamma g}$ is divided into those of $\operatorname{prox}_{\gamma\|\cdot\|_2}$.

[3] A function $f : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ is called proper, lower semicontinuous, and convex if $\operatorname{dom}(f) := \{\boldsymbol{x} \in \mathcal{X} \,|\, f(\boldsymbol{x}) < \infty\} \neq \varnothing$, $\operatorname{lev}_{\leq\alpha}(f) := \{\boldsymbol{x} \in \mathcal{X} \,|\, f(\boldsymbol{x}) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$, and $f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ and all $\lambda \in (0, 1)$, respectively.

Therefore, $\operatorname{prox}_{\gamma\|\cdot\|_1^G}(\boldsymbol{y})$ can be computed for each group by

$$\operatorname{prox}_{\gamma\|\cdot\|_2}(\boldsymbol{y}_{\mathcal{G}_i}) = \begin{cases} \dfrac{\|\boldsymbol{y}_{\mathcal{G}_i}\|_2 - \gamma}{\|\boldsymbol{y}_{\mathcal{G}_i}\|_2}\,\boldsymbol{y}_{\mathcal{G}_i} & \text{if } \|\boldsymbol{y}_{\mathcal{G}_i}\|_2 > \gamma, \\ \boldsymbol{0} & \text{if } \|\boldsymbol{y}_{\mathcal{G}_i}\|_2 \leq \gamma. \end{cases} \tag{21}$$

REFERENCES

[1] 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 Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2664–2673, May 2018.

[2] 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," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 4, pp. 655–675, Apr. 2013.

[3] E. Kudeki and F. Sürücü, "Radar interferometric imaging of field-aligned plasma irregularities in the equatorial electrojet," *Geophysical Research Letters*, vol. 18, no. 1, pp. 41–44, Jan. 1991.

[4] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[5] 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 Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 3077–3088, May 2013.

[6] M. Wada, J. Horikomi, and F. Mizutani, "Development of solid-state weather radar," in *Proceedings of IEEE Radar Conference*, 2009, 4 pages.

[7] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite elements approximations," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[8] T. Y. Yu and M. B. Orescanin, "Beam multiplexing using the phased-array weather radar," *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 4, pp. 616–626, Apr. 2007.

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

[10] S. Fukao and K. Hamazu, *Radar for Meteorological and Atmospheric Observations*. Springer Japan, 2014.