Compressive Radar Imaging of Stationary Indoor Targets With Low-Rank Plus Jointly Sparse and Total Variation Regularizations

Van Ha Tang, Abdesselam Bouzerdoum, *Senior Member, IEEE*, and Son Lam Phung, *Senior Member, IEEE*

***Abstract*— This paper addresses the problem of wall clutter mitigation and image reconstruction for through-wall radar imaging (TWRI) of stationary targets by seeking a model that incorporates low-rank (LR), joint sparsity (JS), and total variation (TV) regularizers. The motivation of the proposed model is that LR regularizer captures the low-dimensional structure of wall clutter; JS guarantees a small fraction of target occupancy and the similarity of sparsity profile among channel images; TV regularizer promotes the spatial continuity of target regions and mitigates background noise. The task of wall clutter mitigation and target image reconstruction is formulated as an optimization problem comprising LR, JS, and TV regularization terms. To handle this problem efficiently, an iterative algorithm based on the forward-backward proximal gradient splitting technique is introduced, which captures wall clutter and yields target images simultaneously. Extensive experiments are conducted on real radar data under compressive sensing scenarios. The results show that the proposed model enhances target localization and clutter mitigation even when radar measurements are significantly reduced.**

***Index Terms*— Through-the-wall radar imaging, wall clutter mitigation, compressed sensing, regularized optimization, low-rank matrix recovery, sparse signal reconstruction, proximal gradient technique.**

I. Introduction

In RECENT years, there has been a growing interest in through-wall radar (TWR) imaging technology, which aims to capture targets situated behind walls or inside enclosed structures. The ability to sense through walls is useful for numerous applications, e.g., detecting hidden hostages in police operations, tracking concealed hostile forces in military missions, and locating buried victims in environmental disasters [1]–[3]. In these applications, it is vital to provide a high-resolution image of the indoor targets. This task, however, is challenging due to several factors, including strong wall clutter and prolonged data collection. To suppress wall reflections and alleviate the burden of data collection, the task of wall clutter mitigation and target image reconstruction needs to be tacked in the context of compressive sensing (CS) [4]–[6].

Early TWRI techniques perform target image formation assuming that strong wall clutter has been entirely removed using measurements collected from an empty scene devoid of targets. The background measurements are then subtracted from the measurements collected from a scene populated with targets, and finally, the target image is formed using backprojection (BP) techniques, such as delay-and-sum (DS) beamforming [7]–[9]. The BP techniques are suitable for the full sensing mode where all antennas and frequencies are used. This full sensing operation, however, leads to a burden on data collection and becomes ineffective in many applications. These issues can be addressed by using the CS technique. Most CS-based TWRI methods have used the sparsity constraint for single-channel image formation [10]–[13] and multi-polarization image reconstruction [14], [15]. It has been shown that employing sparse regularization improves image reconstruction even when the measurements are reduced. However, the sparsity-based methods become ineffective in the presence of wall clutter due to the dominance of wall returns over target reflections. Consequently, the sparse image recovery detects only pixels associated with the wall clutter.

In the presence of wall clutter, several TWRI approaches have employed the sparsity penalty for wall clutter mitigation and image formation [16]–[19]. In [16], a sparse representation (SR) was used for range profile estimation through Fourier basis, followed by image formation through ℓ_1 convex relaxation. In [17] and [18], the sparsity penalty was imposed on a representation of radar signals in a combined Fourier and wavelet bases for signal recovery and wall clutter removal. In [19], SR was applied in a two-stage processing: wall clutter estimation with the discrete prolate spheroidal sequence (DPSS) basis and image reconstruction with ℓ_1 minimization. The multistage sparse-regularized processing, however, can suffer from uncertainty and suboptimality. This issue has been addressed by the studies presented in [20], [21],

Manuscript received April 15, 2019; revised January 5, 2020 and February 5, 2020; accepted February 7, 2020. Date of current version February 26, 2020. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.01-2017.307. The work of Abdesselam Bouzerdoum and Son Lam Phung was supported by a grant from the Australian Research Council (ARC). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jocelyn Chanussot. *(Corresponding author: Van Ha Tang.)*

Van Ha Tang is with the Faculty of Information Technology, Le Quy Don Technical University, Hanoi 10000, Vietnam (e-mail: hatv@lqdtu.edu.vn).

Abdesselam Bouzerdoum is with the Information and Computing Technology Division, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar, and also with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: a.bouzerdoum@uow.edu.au).

Son Lam Phung is with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: phung@uow.edu.au).

Digital Object Identifier 10.1109/TIP.2020.2973819

where wall signals and target returns are separated based on a robust principal component analysis (RPCA) model. The RPCA model comprised LR and sparse regularizers imposed directly on the received signals. By processing the wall return and the target signal jointly, RPCA has been shown to be suitable for TWRI. However, this model does not exploit the characteristics of TWR signals fully and could be ineffective with background clutter and artifacts, which typically appear for non-homogeneous walls and in multi-channel radar imaging.

This paper introduces an imaging model that comprises low-rank, jointly sparse, and total variation regularizations for wall clutter mitigation and target image reconstruction in compressive TWRI. Instead of imposing the LR directly in the signal domain, we employ a low-rank representation technique with an analysis prior (AP) to model the low-dimensional structure of wall clutter expansion in a transform domain. The joint sparsity penalty ensures target pixels occupy only a small fraction of the total image pixels. In addition, JS enforces the same sparsity profile among channel images for multi-channel imaging. Furthermore, because in the formed image targets occupy regions of connected sets of pixels and outliers typically appear as isolated points, we incorporate a TV regularizer to promote the continuity of the target regions and suppress the artifacts. As a result, the TV regularization further benefits the sparse representation of the target images. It is worth noting that TV has been introduced in [22] and widely used for noise removal, local spatial consistency, and structure preservation in signal/image recovery [23]–[27]. In this paper, the original sparse-derivative TV is extended to a joint derivative-sparsity prior to maintain the same sparsity profile among channel images in the gradient domain.

The task of wall clutter mitigation and target image formation is formulated as a low-rank analysis prior-jointly sparse total variation (LRAP-JSTV) regularized least squares (LS) minimization problem. Solving this minimization problem is challenging due to the non-smooth nature of the regularization terms. To deal with this issue, this paper introduces an iterative algorithm using the proximal forward-backward splitting (PFBS) technique [28]–[30] and its accelerated variants [31]–[33]. The iterative minimization involves the forward gradient evaluation of the LS term and backward proximal operators for regularization terms. As the PFBS technique allows the splitting of regularization terms, the proposed algorithm is computationally efficient, especially when proximal operators of the regularizers have closed-form expressions. In the proposed model, LRAP and JS regularizer proximals can be evaluated efficiently using shrinkage/thresholding techniques. The evaluation of TV proximal, however, is complicated and typically requires a nested optimization algorithm. To maintain computational efficiency, we use the inexact proximal technique and approximate the TV proximal as an average of several simpler counterparts that have closed-form solutions. As a result, wall clutter mitigation and target image reconstruction are performed efficiently through shrinkage techniques.

The main contributions of this study can be highlighted as follows.

1) A new model is proposed to simultaneously solve three important TWRI problems: wall clutter mitigation, target image reconstruction, and background artifact suppression. The proposed model can be applied to both single and multi-channel radar imaging, and it is capable of handling the indoor radar imaging problem in the CS context, thereby enabling fast data collection and efficient data storage.
2) This paper introduces an imaging model that comprises LR, JS, and TV regularizations for compressive TWRI. In the proposed model, instead of enforcing LR directly on the signal domain, we propose an LR representation technique with an analysis prior to model the low-dimensional structure of wall clutter expansion in a transform domain. Furthermore, the sparsity is extended to JS to guarantee a small fraction of the target pixels and capture the correlation among channel images. The original TV prior is generalized to a joint derivative-sparsity TV regularizer and incorporated into the proposed model to ensure continuity of the target regions, suppress the artifacts, and maintain the same sparsity profile among channel images in the gradient domain.
3) This paper also develops an efficient algorithm to solve the joint LRAP-JSTV regularized LS optimization problem based on PFBS technique. The key idea is to decompose the composite problem into subproblems that can be handled more efficiently. In the proposed algorithm, the proximal operations of LRAP and JS regularization terms are evaluated efficiently through their convex relaxation of nuclear-norm and $\ell_{2,1}$-norm, respectively. The proximal evaluation of TV term generally requires a nested optimization algorithm, thereby increasing the computational complexity. This paper addresses this issue by developing an inexact proximal technique that approximates the TV proximal as an average of several simpler counterparts having closed-form expressions. This new strategy makes the overall algorithm computational efficient. Furthermore, an acceleration approach is incorporated into the forward gradient evaluation, leading to a fast convergence rate for the proposed algorithm.
4) The proposed LRAP-JSTV model is evaluated using several real radar datasets under different CS operations. Through extensive experimental validation, comparison, and analysis, we find that the proposed approach enhances indoor target localization significantly, even when the measurements are significantly reduced. In compressive TWR sensing, the proposed method outperforms the existing CS-based indoor radar imaging models in terms of target-to-clutter ratio.

The remainder of the paper is organized as follows. Section II introduces the TWR signal model. Section III describes the proposed LRAP-JSTV regularization model and an iterative algorithm for wall clutter mitigation and image reconstruction. Section IV presents the experimental results and analysis. Finally, Section V gives the concluding remarks.

II. TWR Signal Model

This section briefly presents the signal model of a monostatic stepped-frequency TWR used to image targets situated behind a wall. Let us consider the TWR sensing operations where the indoor targets are imaged by L channels: $L = 1$ for single-channel TWRI, and $L > 1$ for multi-polarization TWRI. A transceiver is placed at several scan positions parallel to the wall to synthesize a horizontal N-element linear antenna array. Each antenna transceives a stepped-frequency signal comprising M frequencies, equally spaced over the sensing bandwidth, to image the indoor scene. Suppose that the scene contains P targets placed behind the wall. Let $z_l(m,n)$ denote the mth frequency radar signal received by the nth antenna through the lth channel. The signal $z_l(m,n)$ can be modeled as a superposition of the wall reflection $z_l^w(m,n)$ (including the wall reverberations), target return $z_l^t(m,n)$, and noise $v_l(m,n)$:

$$z_l(m,n) = z_l^w(m,n) + z_l^t(m,n) + v_l(m,n). \tag{1}$$

The wall component $z_l^w(m,n)$ can be expressed as [16], [34]

$$z_l^w(m,n) = \sum_{r=1}^{R} \sigma_w a_r e^{-j2\pi f_m \tau_{n,w}^r}, \tag{2}$$

where σ_w is the reflectivity of the wall, R is the number of wall reverberations, a_r is the path loss factor of the rth wall return, and $\tau_{n,w}^r$ is the propagation delay of the rth wall reverberation. The target return is modeled as the superposition of all the target reflections [19], [35]:

$$z_l^t(m,n) = \sum_{p=1}^{P} \sigma_p^l e^{-j2\pi f_m \tau_{n,p}}, \tag{3}$$

where σ_p^l is the reflectivity of the pth target at the lth channel, and $\tau_{n,p}$ is the round-trip travel time of the signal from the nth antenna location to the pth target. By stacking M frequency measurements collected along the nth antenna, $\mathbf{z}_{l,n} = [z_l(1,n), \ldots, z_l(M,n)]^T$, the signal model in (1) can be represented in vector-form:

$$\mathbf{z}_{l,n} = \mathbf{z}_{l,n}^w + \mathbf{z}_{l,n}^t + \boldsymbol{\upsilon}_{l,n}. \tag{4}$$

Arranging the N vectors $\mathbf{z}_{l,n}$, for $n = 1, \ldots, N$, as columns of matrix $\mathbf{Z}_l \in \mathbb{C}^{M\times N}$, we have the following matrix form:

$$\mathbf{Z}_l = [\mathbf{z}_{l,1}, \ldots, \mathbf{z}_{l,N}] = \mathbf{Z}_l^w + \mathbf{Z}_l^t + \boldsymbol{\Upsilon}_l. \tag{5}$$

For image formation, it is essential to formulate a linear model that relates the target component $\mathbf{Z}_l^t$ to a target image. In doing so, the target space is partitioned into a rectangular grid comprising Q pixels along the crossrange and downrange. Let $s_l(q)$ denote a weighted indicator function representing the pth target reflectivity at the lth channel:

$$s_l(q) = \begin{cases} \sigma_p^l, & \text{if the } q\text{th pixel includes the } p\text{th target;} \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

From (3) and (4), the target signal at the nth antenna of the lth polarization, $\mathbf{z}_{l,n}^t$, can be related to the target image $\mathbf{s}_l = [s_l(1), \ldots, s_l(Q)]^T$ as

$$\mathbf{z}_{l,n}^t = \mathbf{A}_n \, \mathbf{s}_l, \tag{7}$$

where $\mathbf{A}_n \in \mathbb{C}^{M\times Q}$ with the (m,q)th entry given by $A_n(m,q) = \exp(-j2\pi f_m \tau_{n,q})$. Here $\tau_{n,q}$ is the focusing delay between the nth antenna and the qth pixel, considering the penetration delay of signal through the wall [7], [9]. Stacking all measurements collected from the N antennas yields

$$\mathbf{z}_l^t = \text{vec}(\mathbf{Z}_l^t) = \mathbf{A}\, \mathbf{s}_l, \tag{8}$$

where $\mathbf{z}_l^t = [(\mathbf{z}_{l,1}^t)^T, \ldots, (\mathbf{z}_{l,N}^t)^T]^T$, $\mathbf{A} = [\mathbf{A}_1^T, \ldots, \mathbf{A}_N^T]^T$, and hereafter vec$(\cdot)$ denotes the vectorization operator forming a composite column vector by stacking the columns of a matrix in lexicographic order.

From (8), the lth channel image $\mathbf{s}_l$ can be recovered from the target signal $\mathbf{z}_l^t$ using DS beamforming or sparsity-driven techniques. DS beamforming reconstructs an image $\mathbf{s}_l$ by premultiplying the target signal $\mathbf{z}_l^t$ with the adjoint operator $\mathbf{A}^H$:

$$\mathbf{s}_l = \mathbf{A}^H \, \mathbf{z}_l^t. \tag{9}$$

The DS beamformed image typically contains sidelobes and background noise that affect target localization and resolution. To overcome this issue, the sparsity-based model has been proposed in [10]–[12] which yields an image $\mathbf{s}_l$ by solving the following ℓ_1 minimization problem,

$$\min_{\mathbf{s}_l} \|\mathbf{s}_l\|_1 \quad \text{subject to} \|\mathbf{z}_l^t - \mathbf{A}\, \mathbf{s}_l\|_2^2 \;\leq\; \epsilon, \tag{10}$$

where ϵ is a noise bound. An efficient approach considers solving this problem in Lagrangian form:

$$\mathbf{s}_l = \arg\min_{\mathbf{s}_l} \left\{ \frac{1}{2} \|\mathbf{z}_l^t - \mathbf{A}\, \mathbf{s}_l\|_2^2 + \lambda \, \|\mathbf{s}_l\|_1 \right\}, \tag{11}$$

where λ is a positive parameter used to trade off between the error and the penalty terms. Standard convex analysis theory proves that Problems (10) and (11) are equivalent [36], provided that ϵ and λ obey certain relationships. The precise relationship between a constrained optimization problem and its Lagrangian version is discussed in [30], [37], and [38].

Note that in practice the target signal $\mathbf{z}_l^t$ used for image formation in Eqs. (9) and (11) with DS beamforming or ℓ_1-minimization is unavailable. Instead, we have only the radar signal $\mathbf{z}_l$, which is the target signal corrupted by the wall component $\mathbf{z}_l^w$ and noise $\boldsymbol{\upsilon}_l$. Thus, prior to image formation, the target component needs to be segregated from the wall interferences. If full measurements or same frequencies are available along antennas, wall clutter mitigation techniques, such as spatial filtering [39] or subspace projection [40]–[42], can be applied to $\mathbf{Z}_l$ in (5) to estimate $\mathbf{z}_l^t$. In general compressed sensing TWR, however, only a subset of frequency samples is acquired, which may vary from one spatial position to another. To address this issue, several CS-based imaging techniques have been considered previously, where the missing measurements are first estimated, followed by wall clutter mitigation applied to the estimated data measurements. The issue with these CS approaches is that the image formation depends on the accuracy of the signal recovery and wall clutter mitigation operations. The next section describes the proposed LRAP-JSTV regularized model for simultaneous wall clutter mitigation and target image reconstruction.

III. LRAP-JSTV Regularized TWRI

In this section, we first describe the formulation of the proposed LRAP-JSTV regularized LS problem in Subsection III-A. We then introduce an iterative algorithm for solving the optimization problem in Subsection III-B, and provide an analysis of the computational complexity of the proposed algorithm in Subsection III-C.

A. Problem Formulation

The signal model presented in Eqs. (4) and (5) assumes a full set of measurements collected at all N antennas using all M frequencies. For fast data acquisition and efficient data storage, we consider the problem in the CS operation where the radar system uses a reduced number of antennas and frequencies for imaging targets. Suppose only K ($K < N$) randomly selected antennas are used for data collection. Let $k \in \{1, \ldots, N\}$ be the index of a selected antenna. Instead of recording a full of M measurements, the kth antenna acquires only J ($J \ll M$) frequency samples. Let $\boldsymbol{\varphi}_k \in \mathbb{R}^{J \times M}$ denote a sensing matrix in which each row has only one non-zero element (equal to 1) indicating the selected frequency at the kth antenna. The relation between the compressed measurement vector $\mathbf{y}_{l,k}$ with J entries and full data vector $\mathbf{z}_{l,k}$ can be expressed as

$$\mathbf{y}_{l,k} = \boldsymbol{\varphi}_k \, \mathbf{z}_{l,k}. \tag{12}$$

Arranging K measurement vectors $\mathbf{y}_{l,k}$, for $k = 1, \ldots, K$ into a column vector $\mathbf{y}_l = [\mathbf{y}_{l,1}^T, \ldots, \mathbf{y}_{l,K}^T]^T$, $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1^T, \ldots, \boldsymbol{\varphi}_K^T]^T$, it follows from Eqs. (5) and (8) that

$$\mathbf{y}_l = \boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_l) = \boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_l^w + \mathbf{Z}_l^w + \boldsymbol{\Upsilon}_l). \tag{13}$$

Note that matrix $\mathbf{Z}_l$ can be obtained from vector $\mathbf{y}_l$ as $\mathbf{Z}_l = \mathrm{mat}(\boldsymbol{\Phi}^\dagger \mathbf{y}_l)$, where mat denotes the operator reshaping a column vector of MK elements into an $M \times K$ matrix, and † denotes the pseudoinverse operator.

Given the measurement vector $\mathbf{y}_l$, a low-rank matrix $\mathbf{Z}_l^w$ carrying wall returns and a sparse matrix $\mathbf{Z}_l^t$ containing target signals can be obtained by solving the following problem:

$$\min_{\mathbf{Z}_l^w, \mathbf{Z}_l^t} \left\| \mathbf{Z}_l^w \right\|_* + \lambda \left\| \mathbf{Z}_l^t \right\|_1 \quad \text{subject to } \left\| \mathbf{y}_l - \boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_l^w + \mathbf{Z}_l^t) \right\|_2^2 \le \epsilon. \tag{14}$$

Here, $\left\| \mathbf{Z}_l^w \right\|_*$ denotes the nuclear-norm defined as the sum of the singular values of the matrix, $\left\| \mathbf{Z}_l^w \right\|_* = \sum_{v=1}^{V} \lambda_v(\mathbf{Z}_l^w)$ with $\lambda_v(\mathbf{Z}_l^w)$ being the vth largest singular value of matrix $\mathbf{Z}_l^w$ of rank at most V, $\left\| \mathbf{Z}_l^t \right\|_1$ is the ℓ_1-norm defined as the sum of absolute entries, $\|\mathbf{Z}_l^t\|_1 = \sum_m \sum_k |Z_l^t(m,k)|$, λ is a regularization parameter reflecting a trade off between the low-rank and sparse terms, and ϵ is a noise bound. The resulting target signal $\mathbf{Z}_l^t$ can be used to form a target image [20], [21]. The formulation in (14) can be considered as RPCA model since it involves the estimates of LR and sparse components. In this paper, we aim to integrate the image formation stage into the optimization model, upgrade LR to LRAP, sparsity to JS, and incorporate a TV regularizer to enforce the continuity property of the target occupancy and stabilize the solution.

The clutter mitigation and target image reconstruction can be performed in an optimization framework by exploiting the relation between target signal and image as well as processing multi-channel signals simultaneously. It follows from (8) and (13) that

$$\mathbf{y}_l = \boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_l) = \boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_l^w) + \boldsymbol{\Phi} \, \mathbf{A} \, \mathbf{s}_l + \boldsymbol{\Phi} \, \mathrm{vec}(\boldsymbol{\Upsilon}_l). \tag{15}$$

Now, we arrange the measurement sets along L channels into a matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L]$, the wall reflections into a matrix $\mathbf{Z}^w = [\mathbf{Z}_1^w, \ldots, \mathbf{Z}_L^w]$, the multi-channel images $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_L]$. Defining an operator $\mathcal{A}(\mathbf{Z}^w) = [\boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_1^w), \ldots, \boldsymbol{\Phi} \, \mathrm{vec}(\mathbf{Z}_L^w)]$, and $\boldsymbol{\Psi} = \boldsymbol{\Phi} \, \mathbf{A}$, the wall clutter matrix $\mathbf{Z}^w$ and target image $\mathbf{S}$ can be estimated by solving the following optimization problem:

$$\min_{\mathbf{Z}^w, \mathbf{S}} \left\| \boldsymbol{\Theta} \, \mathbf{Z}^w \right\|_* + \lambda \, \|\mathbf{S}\|_{2,1} + \gamma \, \|\mathbf{S}\|_{\mathrm{TV}} \quad \text{subject to } \left\| \mathbf{Y} - [\mathcal{A}(\mathbf{Z}^w) + \boldsymbol{\Psi} \, \mathbf{S}] \right\|_F^2 \le \epsilon. \tag{16}$$

It is worth noting that in (16), nuclear-norm regularization is not directly imposed on $\mathbf{Z}^w$, but on a representation of $\mathbf{Z}^w$ through the low-rank analysis operator $\boldsymbol{\Theta} \in \mathbb{C}^{H \times M}$. In this analysis model, the resulting matrix $\mathbf{R} = \boldsymbol{\Theta} \, \mathbf{Z}^w$ is expected to be low-rank. The mixed $\ell_{2,1}$-norm is used, instead of ℓ_1-norm, to promote the joint sparsity among the L channel images, defined as $\|\mathbf{S}\|_{2,1} = \sum_{q=1}^{Q} \|\mathbf{s}_q\|_2$ with $\mathbf{s}_q$ being the qth row of $\mathbf{S}$. The joint sparsity structure of the L images stored in the columns of $\mathbf{S}$ is because they represent the targets of the same scene. The term $\|\mathbf{S}\|_{\mathrm{TV}}$ denotes a variant of TV defined as $\|\mathbf{S}\|_{\mathrm{TV}} = \|\mathbf{D} \, \mathbf{S}\|_{2,1}$. Here, the gradient matrix $\mathbf{D}$ is applied to each image $\mathbf{s}_l$ yielding a vector containing the magnitude of the discrete gradient of $\mathbf{s}_l$ at all pixels:

$$[\mathbf{D}\mathbf{s}_l]_q = |\mathbf{s}_l(q) - \mathbf{s}_l(q+1)|. \tag{17}$$

This sparse-derivative promoting penalty is known as the anisotropic TV, which enforces local spatial consistency, removes noise, and preserves edges in image recovery [23], [25], [27]. Here, this sparse-derivative is extended to joint derivative-sparsity to enforce the same sparsity profile among channel images in the gradient domain. To solve Problem (16) efficiently, we cast this problem into its Lagrangian form:

$$\min_{\mathbf{Z}^w, \mathbf{S}} \; \{ f(\mathbf{Z}^w, \mathbf{S}) = \frac{1}{2} \left\| \mathbf{Y} - [\mathcal{A}(\mathbf{Z}^w) + \boldsymbol{\Psi} \, \mathbf{S}] \right\|_F^2 + \beta \left\| \boldsymbol{\Theta} \, \mathbf{Z}^w \right\|_* + \lambda \, \|\mathbf{S}\|_{2,1} + \gamma \, \|\mathbf{D} \, \mathbf{S}\|_{2,1} \}, \tag{18}$$

where β is a positive parameter controlling the strength of the LR term. Minimizing $f(\mathbf{Z}^w, \mathbf{S})$ produces the wall clutter matrix $\mathbf{Z}^w$ and multi-channel images $\mathbf{S}$ jointly.

B. Iterative Algorithm

This subsection introduces an iterative algorithm to solve the LRAP-JSTV regularized LS problem in (18). Although this minimization problem is convex, it is complicated to solve it directly due to the non-smoothness of the regularization terms. To handle this issue, we develop an algorithm based on the PFBS technique. Before presenting the algorithm, let

us consider a generic case of minimizing a composite objective function:

$$\min_{\mathbf{X}}\{f(\mathbf{X}) = g(\mathbf{X}) + h(\mathbf{X})\}, \tag{19}$$

where $g(\mathbf{X})$ is convex and differentiable with a C-Lipschitz continuous gradient ∇g, and $h(\mathbf{X})$ is convex but not necessary smooth. Problem (19) is minimized by an iterative scheme that involves a forward gradient evaluation of $g(\mathbf{X})$ and a backward proximal operator of $h(\mathbf{X})$. Let $\mathbf{X}_t$ denote an estimate of the solution at the tth iteration. The next estimate is obtained by

$$\mathbf{X}_{t+1} = \underbrace{\text{prox}_{\mu_t h}}_{\text{backward step}} \underbrace{(\mathbf{X}_t - \mu_t \nabla g(\mathbf{X}_t))}_{\text{forward step}}. \tag{20}$$

Here, the stepsize μ_t satisfies $0 < \mu_t \leq 1/C$ to ensure convergence, and the proximal operator is defined as

$$\text{prox}_{\mu_t h}(\mathbf{Z}) = \arg\min_{\mathbf{X}}\{\frac{1}{2}\|\mathbf{Z} - \mathbf{X}\|_F^2 + \mu_t\, h(\mathbf{X})\}. \tag{21}$$

In other words, let $\mathbf{Z}_t$ denote the result of the gradient step evaluated using the current estimate $\mathbf{X}_t$,

$$\mathbf{Z}_t = \mathbf{X}_t - \mu_t \nabla g(\mathbf{X}_t), \tag{22}$$

the next estimate of the solution is obtained by proximal evaluation

$$\mathbf{X}_{t+1} = \text{prox}_{\mu_t h}(\mathbf{Z}_t). \tag{23}$$

As the forward gradient is simple, the proximal evaluation is the main computation cost. Hence, this algorithm is computational-efficient if the proximal operator has a closed-form solution. For example, if $h(\mathbf{X})$ is the ℓ_1-norm, i.e., $h(\mathbf{X}) = \|\mathbf{X}\|_1$, its proximal is equivalent to the soft-thresholding operator, and the overall estimation scheme reduces to an iterative shrinkage/thresholding algorithm (ISTA) [37], [43]–[45]. In theory, the PFBS technique or its instance ISTA achieves a global convergence rate of $\mathcal{O}(1/t)$. Accelerated PFBS or its instance of fast iterative shrinkage/thresholding algorithm (FISTA) can obtain the convergence rate of $\mathcal{O}(1/t^2)$ and remain the computational cost [28], [31], [33]. The acceleration is mainly achieved by modifying the forward gradient step. In this step, instead of exploiting only $\mathbf{X}_t$, accelerated PFBS uses an auxiliary variable $\mathbf{U}_t$ obtained as a linear combination of $\mathbf{X}_t$ and $\mathbf{X}_{t-1}$, which leads to the following scheme:

$$\mathbf{U}_t = \mathbf{X}_t + \frac{q_{t-1} - 1}{q_t}(\mathbf{X}_t - \mathbf{X}_{t-1}) \tag{24}$$

$$\mathbf{Z}_t = \mathbf{U}_t - \mu_t \nabla g(\mathbf{U}_t), \tag{25}$$

$$\mathbf{X}_{t+1} = \text{prox}_{\mu_t h}(\mathbf{Z}_t), \tag{26}$$

$$q_{t+1} = \left(1 + \sqrt{1 + 4q_t^2}\right)/2. \tag{27}$$

We use the accelerated computation scheme in (24)–(27) to solve the LRAP-JSTV regularized LS TWRI problem in (18). Consider the variable $\mathbf{X} = \mathcal{A}(\mathbf{Z}^w) + \mathbf{\Psi}\mathbf{S}$, $g(\mathbf{X})$ is the quadratic term, and $h(\mathbf{X})$ is the sum of regularization terms. Defining an adjoint operator $\mathcal{A}^*(\mathbf{Y}) = [\text{mat}(\mathbf{\Phi}^\dagger\, \mathbf{y}_1), \ldots, \text{mat}(\mathbf{\Phi}^\dagger\, \mathbf{y}_L)]$, and noting that $\mathbf{Z}^w$ and $\mathbf{S}$ are separable, the gradient evaluation involves

$$\mathbf{U}_t^w = \mathbf{Z}_t^w + \frac{q_{t-1} - 1}{q_t}(\mathbf{Z}_t^w - \mathbf{Z}_{t-1}^w), \tag{28}$$

$$\mathbf{U}_t^s = \mathbf{S}_t + \frac{q_{t-1} - 1}{q_t}\,(\mathbf{S}_t - \mathbf{S}_{t-1}), \tag{29}$$

$$\mathbf{Z}_t = \mathbf{U}_t^w + \mathcal{A}^*(\mathbf{\Psi}\,\mathbf{U}_t^s) - \mathcal{A}^*(\mathcal{A}(\mathbf{U}_t^w) + \mathbf{\Psi}\,\mathbf{U}_t^s - \mathbf{Y}). \tag{30}$$

In Eq. (30), the stepsize μ_t is omitted because it is set to $\mu_t = 1/C$ with $C = \lambda_{\max}(\mathbf{\Phi}^T\mathbf{\Phi}) = 1$. The proximal operator of $h(\mathbf{X})$ is separated into the evaluation for $h_1(\mathbf{Z}^w) = \beta\|\mathbf{\Theta}\mathbf{Z}^w\|_*$ and $h_2(\mathbf{S}) = \lambda\,\|\mathbf{S}\|_{2,1} + \gamma\,\|\mathbf{D}\,\mathbf{S}\|_{2,1}$. The next estimate $(\mathbf{Z}_{t+1}^w, \mathbf{S}_{t+1})$ is obtained by

$$\mathbf{Z}_{t+1}^w = \text{prox}_{h_1}(\mathbf{Z}_t - \mathcal{A}^*(\mathbf{\Psi}\,\mathbf{S}_t)), \tag{31}$$

$$\mathbf{S}_{t+1} = \arg\min_{\mathbf{S}}\left\{\frac{1}{2}\|\mathcal{A}(\mathbf{Z}_t - \mathbf{Z}_{t+1}^w) - \mathbf{\Psi}\,\mathbf{S}\|_F^2 + h_2(\mathbf{S})\right\}. \tag{32}$$

The separation in proximal evaluation is similar to the alternating technique in which the resultant estimate of $\mathbf{Z}_{t+1}^w$ is used for the next estimate of $\mathbf{S}_{t+1}$. Subproblem (32) can be solved by applying the gradient computation for the quadratic term and evaluating the proximal of $h_2(\mathbf{S})$, yielding

$$\mathbf{B}_t = \mathbf{U}_t^s - \kappa_t \mathbf{\Psi}^H\left[\mathbf{\Psi}\,\mathbf{U}_t^s - \mathcal{A}(\mathbf{Z}_t - \mathbf{Z}_{t+1}^w)\right], \tag{33}$$

$$\mathbf{S}_{t+1} = \text{prox}_{\kappa_t h_2}(\mathbf{B}_t), \tag{34}$$

where the stepsize κ_t is set to $\kappa_t = 1/\lambda_{\max}(\mathbf{\Psi}^H\mathbf{\Psi})$ for fast convergence.

The task now is to evaluate the proximal operators of the two convex functions $h_1(\mathbf{Z}^w)$ in (31) and $h_2(\mathbf{S})$ in (34). Function $h_1(\mathbf{Z}^w)$ is the nuclear-norm with the analysis operator $\mathbf{\Theta}$, and thus its proximal operator has a closed-form expression through the singular value soft-thresholding (SVT):

$$\mathbf{Z}_{t+1}^w = \mathbf{\Theta}^\dagger \mathcal{D}(\mathbf{\Theta}[\mathbf{Z}_t - \mathcal{A}^*(\mathbf{\Psi}\,\mathbf{S}_t)], \beta). \tag{35}$$

In (35), the SVT operator $\mathcal{D}(\mathbf{Z}, \tau)$ is a nonlinear function which applies a soft-thresholding at level τ to the singular values of the input matrix $\mathbf{Z}$:

$$\mathcal{D}(\mathbf{Z}, \tau) = \mathbf{U}\,\mathcal{T}(\mathbf{\Lambda}, \tau)\,\mathbf{V}^H, \tag{36}$$

where $\mathbf{Z} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{V}^H$ is the singular value decomposition of $\mathbf{Z}$, and $\mathcal{T}(\mathbf{\Lambda}, \tau)$ is the component-wise shrinkage function,

$$\mathcal{T}(x, \tau) = \text{sgn}(x)\max(|x| - \tau, 0) = \frac{x}{|x|}\max(|x| - \tau, 0). \tag{37}$$

The proximal operator of $h_2(\mathbf{S})$ consists of an $\ell_{2,1}$ term and a TV term. For single channel imaging ($L = 1$), this proximal evaluation is referred as *fused lasso signal appropriator* (FLSA) [46]. Here it can be regarded as an extension of FLSA to the multiple measurement vector (MMV) problem ($L > 1$). Let $\mathbf{S}_{t+1}^{\text{tv}}$ denote the resultant proximal evaluation of the TV regularizer $h_{\text{tv}}(\mathbf{S}) = \gamma\,\|\mathbf{D}\,\mathbf{S}\|_{2,1}$,

$$\mathbf{S}_{t+1}^{\mathbf{tv}} = \text{prox}_{\kappa_t h_{\text{tv}}}(\mathbf{B}_t),$$
$$= \arg\min_{\mathbf{S}}\left\{\frac{1}{2}\,\|\mathbf{B}_t - \mathbf{S}\|_F^2 + \kappa_t\, h_{\text{tv}}(\mathbf{S})\right\}. \tag{38}$$

Using $\mathbf{S}_{t+1}^{\text{tv}}$, the proximal operator of $h_2(\mathbf{S})$ is reduced to the proximal evaluation of the $\ell_{2,1}$ function, which is equivalent to the row-shrinkage operator $\mathcal{R}(\mathbf{X}, \tau)$:

$$\mathbf{S}_{t+1} = \text{prox}_{\kappa_t \lambda \|\mathbf{S}\|_{2,1}}(\mathbf{S}_{t+1}^{\text{tv}}) = \mathcal{R}(\mathbf{S}_{t+1}^{\text{tv}}, \kappa_t \lambda). \quad (39)$$

In (39), the row-shrinkage $\mathcal{R}(\mathbf{X}, \tau)$ is a nonlinear function that applies the soft-thresholding operator at level τ to the ℓ_2-norm of the rows of the input matrix $\mathbf{X}$:

$$\mathcal{R}(\mathbf{X}, \tau) = \frac{\mathbf{x}_q}{\|\mathbf{x}_q\|_2} \max(\|\mathbf{x}_q\|_2 - \tau, 0), \quad q = 1, \ldots, Q, \quad (40)$$

where $\mathbf{x}_q$ denotes the qth row of matrix $\mathbf{X}$.

The remaining task is to evaluate the TV proximal or equivalently solve the TV-minimization problem in (38), which typically requires a nested optimization algorithm [24], [26]. This leads to computation complexity and makes the overall algorithm inefficient. To overcome this issue, we use the inexact proximal operator and approximate the exact TV proximal as an average of several operators that have closed-form expressions. This aim is achieved by considering the relation between TV regularization and cycle spinning with Haar wavelets [47]. The Haar wavelet transform applying to the qth pixel of image $\mathbf{s}$ yields coefficients,

$$c_q = \frac{1}{\sqrt{2}}(s_{2q-1} + s_{2q}), \quad d_q = \frac{1}{\sqrt{2}}(s_{2q-1} - s_{2q}), \quad (41)$$

where $q = 1, \ldots, Q/2$, assuming that Q is even. The vectors $\mathbf{c}$ and $\mathbf{d}$ are known as coarse and detail coefficients, respectively. Denoising is performed by applying a shrinkage operator to the detail coefficient component $\mathbf{d}$, which plays the role as the derivative vector. From Eqs. (17) and (41), TV regularization is deemed to be equivalent to Haar wavelet shrinkage (HWS). The studies in [27], [48], [49] have shown that HWS with cycle spinning is equivalent to a single iteration of TV denoising. The HWS with cycle spinning works by first denoising the wavelet expansion of shifted versions of the image and then applying inverse circulant shift operator to the denoised image; the results for a range of shifts are averaged to obtain the final image.

Here, the HWS with cycle spinning is exploited for efficient TV proximal approximation. Let $\mathbf{W}$ denote the orthogonal wavelet transform (e.g., Haar wavelet basis) and $\mathcal{S}_g(\mathbf{s})$ be the circulant shift operator of vector $\mathbf{s}$ by g, $\mathcal{S}_g(s_q) = s_{(q+g) \bmod Q}$. This operator is unitary and hence invertible $\mathcal{S}_g^{-1}(\cdot) = \mathcal{S}_{-g}(\cdot)$. The TV regularizer can be expressed as the following sum:

$$h_{\text{tv}}(\mathbf{S}) = \gamma \sqrt{2} \sum_{g=1}^{G} \|\mathbf{W}\, \mathcal{S}_g(\mathbf{S})\|_{2,1}. \quad (42)$$

Note that a 1D Haar representation of a 2D image can be computed by applying the transform either horizontally or vertically. Hence, we need four ($G = 4$) orthogonal 1D Haar transforms—two for each dimension of the gradient. Here the circulant operator $\mathcal{S}_g(\mathbf{S})$ is applied to each channel image stored as a column in matrix $\mathbf{S}$. Because both analysis operators $\mathbf{W}$ and $\mathcal{S}_g$ are invertible, Problem (38) admits the following closed-form solution,

$$\mathbf{S}_{t+1}^{\text{tv}} = \frac{1}{G} \sum_{g=1}^{G} \mathbf{W}^T\, \mathcal{S}_g^{-1}[\mathcal{R}(\mathbf{W}\, \mathcal{S}_g(\mathbf{B}_t), \sqrt{2}\, G\, \kappa_t\, \gamma)]. \quad (43)$$

TABLE I

ALGORITHM 1: ACCELERATED GRADIENT AND PROXIMAL EVALUATIONS OF LRAP-JSTV REGULARIZERS FOR WALL CLUTTER MITIGATION AND TARGET IMAGE RECONSTRUCTION

1) Initialize $\mathbf{Z}_1^w = \mathbf{Z}_0^w \leftarrow \mathcal{A}^*(\mathbf{Y})$, $\mathbf{S}_1 = \mathbf{S}_0 \leftarrow \mathbf{0}$, $t \leftarrow 1$, $q_1 = q_0 \leftarrow 1$.

2) Perform gradient evaluation using Eqs. (28)–(30):

$$\begin{aligned}
&\mathbf{U}_t^w \leftarrow \mathbf{Z}_t^w + \tfrac{q_{t-1}-1}{q_t}(\mathbf{Z}_t^w - \mathbf{Z}_{t-1}^w),\\
&\mathbf{U}_t^s \leftarrow \mathbf{S}_t + \tfrac{q_{t-1}-1}{q_t}(\mathbf{S}_t - \mathbf{S}_{t-1}),\\
&\mathbf{Z}_t \leftarrow \mathbf{U}_t^w + \mathcal{A}^*(\mathbf{\Psi}\, \mathbf{U}_t^s) - \mathcal{A}^*(\mathcal{A}(\mathbf{U}_t^w) + \mathbf{\Psi}\, \mathbf{U}_t^s - \mathbf{Y}).
\end{aligned}$$

3) Estimate wall component using (35):

$$\mathbf{Z}_{t+1}^w \leftarrow \mathbf{\Theta}^{\dagger} \mathcal{D}(\mathbf{\Theta}[\mathbf{Z}_t - \mathcal{A}^*(\mathbf{\Psi}\, \mathbf{S}_t)], \beta).$$

4) Estimate the target images using Eqs. (33), (43), and (39):

$$\begin{aligned}
&\mathbf{B}_t \leftarrow \mathbf{U}_t^s - \kappa_t \mathbf{\Psi}^H \left[\mathbf{\Psi}\, \mathbf{U}_t^s - \mathcal{A}(\mathbf{Z}_t - \mathbf{Z}_{t+1}^w)\right],\\
&\mathbf{S}_{t+1}^{\text{tv}} \leftarrow \tfrac{1}{G} \textstyle\sum_{g=1}^{G} \mathbf{W}^T\, \mathcal{S}_g^{-1}[\mathcal{R}(\mathbf{W}\, \mathcal{S}_g(\mathbf{B}_t), \sqrt{2}\, G\, \kappa_t\, \gamma)].\\
&\mathbf{S}_{t+1} \leftarrow \mathcal{R}(\mathbf{S}_{t+1}^{\text{tv}}, \kappa_t \lambda).
\end{aligned}$$

5) Evaluate the cost function $f(\mathbf{Z}_{t+1}^w, \mathbf{S}_{t+1})$ using (18).

If $\frac{|f(\mathbf{Z}_{t+1}^w, \mathbf{S}_{t+1}) - f(\mathbf{Z}_t^w, \mathbf{S}_t)|}{|f(\mathbf{Z}_t^w, \mathbf{S}_t)|} < \text{tol}$, then terminate the algorithm,

otherwise update $q_{t+1} \leftarrow \frac{1 + \sqrt{4q_t^2 + 1}}{2}$, $t \leftarrow t+1$, and go to Step 2.

In summary, the iterative steps of the algorithm for solving Problem (18) are provided in Table I. The algorithm takes an input set of the data matrix $\mathbf{Y}$, the parameters β, γ, λ, and a predefined tolerance tol. The regularization parameters β, γ, and λ are problem-dependent and need to be tuned appropriately. Further discussion on selecting suitable parameters is given in the experimental section IV-A. The algorithm performs two major tasks: wall-component estimation (Step 3) and image reconstruction (Step 4). It is observed that the wall-component is estimated by applying soft-thresholding on the singular values of the data matrix after subtracting the target component. The target images are obtained by solving a joint $\ell_{2,1}$ and TV regularized LS minimization problem in which the measurement matrix is obtained from the data matrix after removing the estimated wall component. The algorithm is computationally efficient because (1) it incorporates an acceleration gradient evaluation (Step 2) and (2) it evaluates the proximals of LRAP, JS, and TV through operators that have closed-form expressions. The algorithm stops when it converges to a local optimum. In this implementation, the algorithm terminates if the relative change of the objective function is negligible (see Step 5). After that, the sparse columns in $\mathbf{S}$ are rearranged into 2D maps representing the formed multi-channel images of the behind-the-wall targets.

C. Complexity and Convergence Analysis of LRAP-JSTV Algorithm

This subsection presents the computational analysis of the proposed LRAP-JSTV algorithm. This algorithm has two major stages: forward gradient and backward proximal evaluation. Because the computation of forward gradient is simple, the evaluations of the proximal operators are the most time-consuming steps and constitute the computational complexity of Algorithm 1. Let $\mathbf{Y} \in \mathbb{C}^{KJ\times L}$, $\mathbf{\Psi} \in \mathbb{C}^{KJ\times Q}$, $\mathbf{Z}^w \in \mathbb{C}^{M\times KL}$, $\mathbf{\Theta} \in \mathbb{C}^{H\times M}$, and $\mathbf{S} \in \mathbb{C}^{Q\times L}$, where $Q \gg H \gg M \gg N \gg K \gg L$. The proximity evaluation of the LRAP for wall-component estimation in (Step 3) involves SVD operation, which has the computational complexity of $\mathcal{O}(H\ K^2\ L^2)$. The time complexity of the row-shrinkage operator in Step 4 is $\mathcal{O}(L\ Q)$. Thus, the overall computational complexity of the LRAP-JSTV algorithm is $\mathcal{O}(t(H\ K^2\ L^2 + L\ Q))$, where t is the number of iterations.

The convergence of the proposed algorithm can be analyzed based on the generic framework of PFBS and its variant (accelerated PFBS). In theory, PFBS has a global convergence rate of $\mathcal{O}(1/t)$, and accelerated PFBS has a convergence rate of $\mathcal{O}(1/t^2)$ [28], [31], [33]. In this paper, the proposed LRAP-JSTV is developed using the accelerated PFBS scheme in Eqs. (24)–(27). The convergence of this computation scheme depends on that of the proximal operations. Here, the LRAP-JSTV involves the evaluation of three proximity operators: LRAP, JS, and TV. The LRAP operator is evaluated through SVT, whose convergence has been proven in [50]. The convergence of the proximal evaluations of JS and TV is based on the row-shrinkage operator, whose convergence proof is given in [30]. Subsection IV-C.5 provides experimental results which illustrate the convergence of the proposed LRAP-JSTV algorithm, with and without the acceleration feature.

D. Estimation of Regularization Parameters

The LRAP-JSTV algorithm has three regularization parameters β, λ, and γ to control the importance of the LR, JS, and TV regularization terms, respectively. Using large regularization parameters suppresses background clutter but also removes target pixels, while small regularization parameters yield imaging results with heavy clutter. Thus, the regularization parameters need to be tuned appropriately. A cross-validation grid search can be used to seek the suitable regularization parameters. For large search area, however, this technique is time-consuming. A different approach is to use Bayesian optimization that has been shown to be more effective for the selection of regularization parameters than grid search and random search techniques [51], [52]. Due to its efficiency for the global optimization of functions with expensive evaluations, Bayesian optimization with Gaussian processes has been used for tuning the hyperparameters of several optimization models, including convolutional neural networks [53], deep generative models [54], and radar signal detection [55]. In this paper, Bayesian optimization with Gaussian processes (GP) is employed for finding the optimal regularization parameters of the LRAP-JSTV model.

Let $f(\boldsymbol{\omega})$ be the objective function that yields the quality score of the target image $\mathbf{S}$, that is the target-to-clutter ratio resulted from the LRAP-JSTV using the set of hyperparameters $\boldsymbol{\omega} = [\beta,\ \lambda,\ \gamma] \in \mathbf{\Omega}$, where $\mathbf{\Omega}$ is the bounded hyperparameter space. Bayesian optimization finds the optimal hyperparameters by solving the following maximization,

$$\hat{\boldsymbol{\omega}} = \arg\max_{\boldsymbol{\omega}\in\mathbf{\Omega}} f(\boldsymbol{\omega}). \tag{44}$$

Bayesian optimization is fundamentally a sequential approach to solving Problem (44). In particular, it first defines a prior function $p(f)$ and constructs an acquisition function $a : \mathbf{\Omega} \to \mathbb{R}^+$ to determine which point in $\mathbf{\Omega}$ used for the next evaluation. Then, the sequential optimization involves the following three steps: (i) select the new hyperparameter $\boldsymbol{\omega}_{t+1}$ by optimizing acquisition function $\boldsymbol{\omega}_{t+1} = \arg\max_{\boldsymbol{\omega}\in\mathbf{\Omega}} a(\boldsymbol{\omega})$; (ii) query the objective function to obtain $y_{t+1} \approx f(\boldsymbol{\omega}_{t+1})$, and augment the resulting point $(\boldsymbol{\omega}_{t+1},\ y_{t+1})$ to the set of observed data $\mathcal{D}_{t+1} = \{\boldsymbol{\omega}_j,\ y_j\}_{j=1}^{t+1}$; and (iii) update the statistical model $p(f|\mathcal{D}_{t+1})$ and $a(f|\mathcal{D}_{t+1})$. A Gaussian process is the most common choice for $p(f)$ due to its flexibility and tractability in terms of conditioning and inference. By GP, $p(f)$ is completely defined by its mean function $m(\boldsymbol{\omega})$ and covariance function $c(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$ that can be derived in a closed form. Furthermore, if the prior mean function is zero, then the Gaussian process is fully determined by the covariance function. For the estimation of hyperparameters, the covariance function is typically defined using the automatic relevance determination (ARD) Matérn 5/2 kernel [56]

$$c_{M52}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) = \theta\left(1 + \sqrt{5}d(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) + \frac{5}{3}d(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)^2\right) \times \exp(-\sqrt{5}d(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)). \tag{45}$$

Here, $d(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$ denotes the Mahalanobis distance, and θ is the characteristic length scale, which determines how far apart the input $\boldsymbol{\omega}$ can be for the response value to be uncorrelated.

Bayesian optimization can select several acquisition functions, such as probability of improvement [57], expected improvement [58], or upper confidence bound [59]. In this paper, we define the acquisition function using the expected improvement given by

$$J(\boldsymbol{\omega}) = \max(0,\ f(\boldsymbol{\omega}) - f(\hat{\boldsymbol{\omega}})), \tag{46}$$

where $\hat{\boldsymbol{\omega}}$ is the current best hyperparameters at iteration t, $\hat{\boldsymbol{\omega}} = \arg\max_{\boldsymbol{\omega}\in\mathbf{\Omega}_{1:t}} f(\boldsymbol{\omega})$. The acquisition function is defined on the expected value of $J(\boldsymbol{\omega})$ as

$$\arg\max_{\boldsymbol{\omega}} \mathbb{E}(J(\boldsymbol{\omega})|\mathcal{D}_t), \tag{47}$$

where the closed form of $\mathbb{E}(J(\boldsymbol{\omega}))$ can be expressed as [58]

$$\mathbb{E}(J(\boldsymbol{\omega})) = \begin{cases} (m(\boldsymbol{\omega}) - f(\hat{\boldsymbol{\omega}}))\eta(z) + s(\boldsymbol{\omega})\kappa(z), & \text{if } s(\boldsymbol{\omega}) > 0; \\ 0, & \text{if } s(\boldsymbol{\omega}) = 0. \end{cases} \tag{48}$$

Here, $z = (m(\boldsymbol{\omega}) - f(\hat{\boldsymbol{\omega}}))/s(\boldsymbol{\omega})$, $s(\boldsymbol{\omega})$ is the standard deviation function of the GP, and $\eta(\cdot)$ and $\kappa(\cdot)$ denote the cumulative distribution function and the probability density function of a standard normal distribution, respectively.

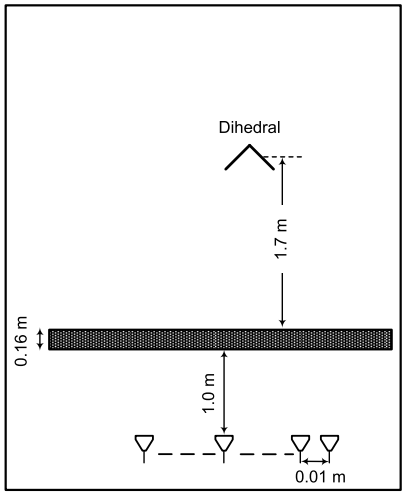

Fig. 1. Layout of the scene with a wooden wall and one dihedral target for TWR data acquisition.

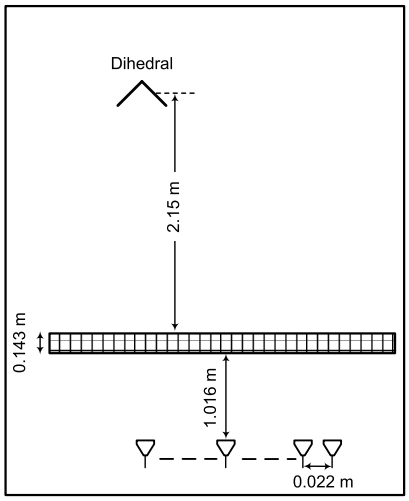

Fig. 2. Layout of the scene with a concrete wall and one dihedral target for TWR data acquisition.

IV. Experimental Results

In this section, we present the experimental results obtained using real radar datasets collected from the Imaging Laboratory at the University of Wollongong, Australia and the Radar Imaging Laboratory at Villanova University, USA. Subsection IV-A describes the radar systems, the scenes, and the parameters used in the experiments. Subsection IV-B gives the performance evaluation and metrics, and Subsection IV-C presents imaging results, performance analysis, and comparison for several TWRI methods under different sensing scenarios.

A. Experimental Methods

Real multi-channel radar datasets were acquired from two real scenes which consist of different targets and types of walls. The first scene contains one dihedral target placed behind a wooden wall, whereas the second scene comprises one target behind a homogeneous concrete wall. The first scene, shown in Fig. 1, is illuminated by a stepped-frequency synthetic aperture radar system. The aperture array has 81 elements, with an inter-element spacing of 0.01 m. The stepped-frequency signal comprises 801 frequencies, equispaced over 2 GHz bandwidth centered at 2.0 GHz with 2.5 MHz frequency step. The radar system was placed at a standoff distance of 1.0 m from the wall made up of 0.16 m thick wooden boards.

The second scene, shown in Fig. 2, contains a 0.4 m high by 0.3 m wide dihedral, placed on a turntable made of two 1.2 m x 2.4 m sheets of 0.013 m thick plywood. A full polarization stepped-frequency radar system was used for data collection. Datasets for co- and cross-polarizations were collected by placing the radar system in front of a concrete wall of thickness 0.15 m. A transceiver is moved along the wall to synthesize a 57-element linear antenna array. At each antenna, a set of 801 monochromatic signals, covering a frequency range of [0.7–3.1] GHz, is transmitted to illuminate the scene behind the wall. Both scenes partitioned into 100×100 pixels have the downrange and crossrange extending from 0 to 3 m, and -1.5 to 1.5 m, respectively.

The proposed LRAP-JSTV model has three hyperparameters β, λ, and γ, that can be determined using the Bayesian optimization technique. The Gaussian process Bayesian optimization requires the initialization of the search boundary of the parameters. The parameter β is used in SVT to estimate the low-rank matrix, see Step 3 in Algorithm 1. Setting β to a very large value, e.g., $\beta = 0.99\|\mathcal{A}^*(\mathbf{Y})\|_2$, leads to the solution of $\mathbf{Z}^\omega$ being rank 1, whereas choosing a small value, e.g., $\beta = 10^{-3}\|\mathcal{A}^*(\mathbf{Y})\|_2$ makes the algorithm converge very slowly. The values $10^{-3}\|\mathcal{A}^*(\mathbf{Y})\|_2$ and $0.99\|\mathcal{A}^*(\mathbf{Y})\|_2$ can be regarded as the lower and upper bounds for β, respectively. Here, in the experiments, the search boundary for β was set to $[10^{-2}\beta_{\max},\ 0.5\beta_{\max}]$ where $\beta_{\max} = \|\mathcal{A}^*(\mathbf{Y})\|_2$.

The parameter λ guarantees the sparsity level of the multi-channel image $\mathbf{S}$. For λ greater than $\lambda_{\max} = \max\{\|(\mathbf{\Psi}^H\mathbf{y}_l)\|_\infty\}_{l=1}^L$, the unique solution to Problem (18) for $\mathbf{S}$ is the zero matrix. Here, the boundary values of λ were set to $[10^{-3}\lambda_{\max},\ 10^{-1}\lambda_{\max}]$. Similarly, the search interval for γ was set to $[10^{-2}\gamma_{\max},\ 10^{-1}\gamma_{\max}]$ where $\gamma_{\max} = \max\{\|\mathbf{W}\,(\mathbf{\Psi}^H\mathbf{y}_l)\|_\infty\}_{l=1}^L$. The algorithm converges if the relative change of the objective function is smaller than $\text{tol} = 10^{-4}$ (see Step 5 in Algorithm 1). The analysis operator $\mathbf{\Theta}$ is constructed with the discrete prolate spheroidal sequence (DPSS) basis. The reason is DPSS is more suitable to represent TWR signals than the Fourier or wavelet counterparts [19], [35].

B. Performance Evaluation and Metrics

To quantify the capability of clutter mitigation and target reconstruction for different approaches, the *target-to-clutter ratio* (TCR) performance measure is used. Let I denote the formed image. The TCR of image I is defined as the ratio between the average power of the target region and the average power of the clutter region (in dB):

$$\text{TCR} = 10\log_{10}\left(\frac{P_{\text{target}}}{P_{\text{clutter}}}\right). \tag{49}$$

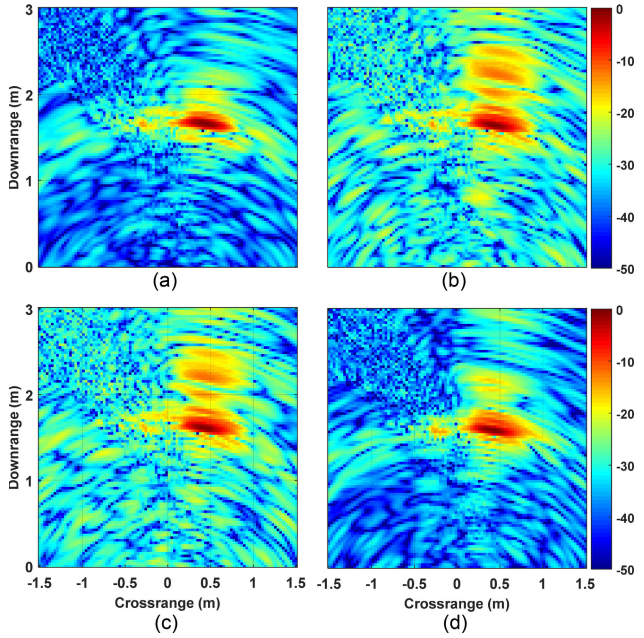

Fig. 3. Multi-polarization images of the wooden-wall dihedral scene formed with DS beamforming using measurements after background subtraction: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

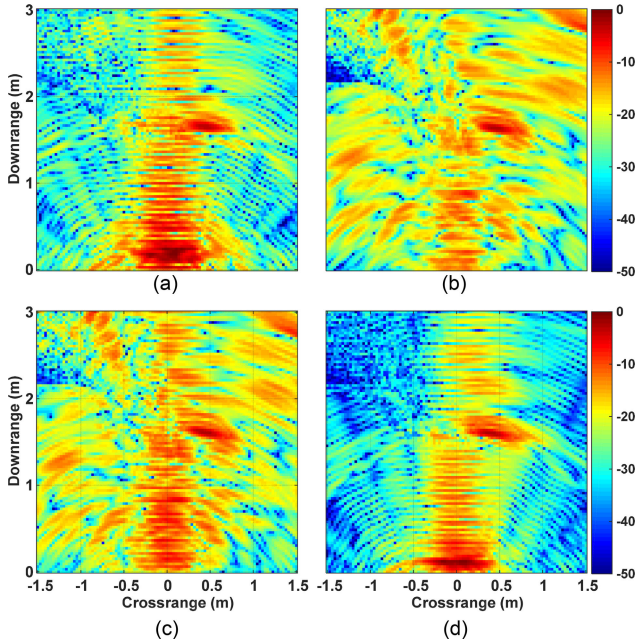

Fig. 4. Multi-polarization images of the wooden-wall dihedral scene formed with DS beamforming using radar measurements without background subtraction: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

The target region consists of pixels that belong to the target responses, while the clutter region comprises pixels that indicate wall reflections, multi-path interferences, and other image artifacts, such as sidelobe clutter. For a region R (R = target, clutter) in the image I, the average power of R is defined as

$$P_{\mathrm{R}} = \frac{1}{N_{\mathrm{R}}} \sum_{(x,y)\in \mathrm{R}} |I(x,y)|^2, \tag{50}$$

where N_{R} is the number of pixels in region R.

To form the ground-truth pixels belonging to the target region, the target image is first obtained by DS beamforming and background subtraction (BS) using the full measurements. The BS is performed by using the measurements collected from a reference scene, which contains the wall but no targets. The background signals are then subtracted from those of the actual scene, consisting of the front wall and the targets. Finally, DS beamforming is applied to the residual measurements to form an image of the indoor targets. Figs. 3 and 4 show the images of the wooden-wall scene formed with DS beamforming with and without clutter mitigation, respectively. Fig. 5(a) shows the target image formed by using the additive fusion of channel images in Fig. 3. Fig. 5(b) presents the target region (target mask) obtained by thresholding the target image shown in Fig. 5(a) using a threshold of constant false alarm rate (CFAR) detector. Here, the cell-averaging CFAR with (FAR = 5%) is used to determine the threshold value. Fig. 5(c) shows the clutter image, on the other hand, obtained by additive fusion of the DS beamforming images without background subtraction. Fig. 5(d) presents the corresponding clutter mask, which indicates pixels that belong to the wall clutter, sidelobes, and other background noise, but excludes the target pixels. A similar strategy is applied to the second concrete-wall scene data, which yields target

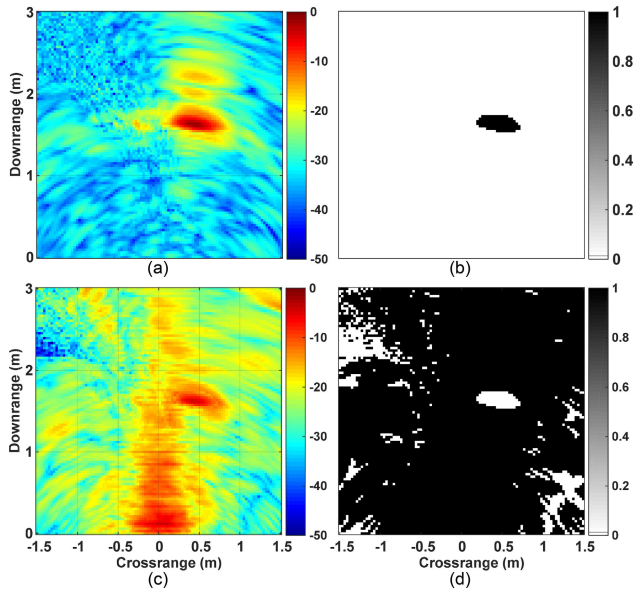

Fig. 5. Target and clutter regions of the dihedral scene with a wooden wall: (a) target image formed with additive fusion of DS beamforming images after background subtraction, (b) target mask obtained by thresholding the target image, (c) clutter image formed with addtive fusion of DS beamforming images without clutter mitigation, and (d) clutter mask obtained by thresholding the clutter image.

and clutter images shown in Figs. 6 and 7. Fig. 8 shows the ground-truth target and clutter regions of the second real dihedral scene, which are used for computing the performance measure TCR of the images reconstructed by different imaging approaches.

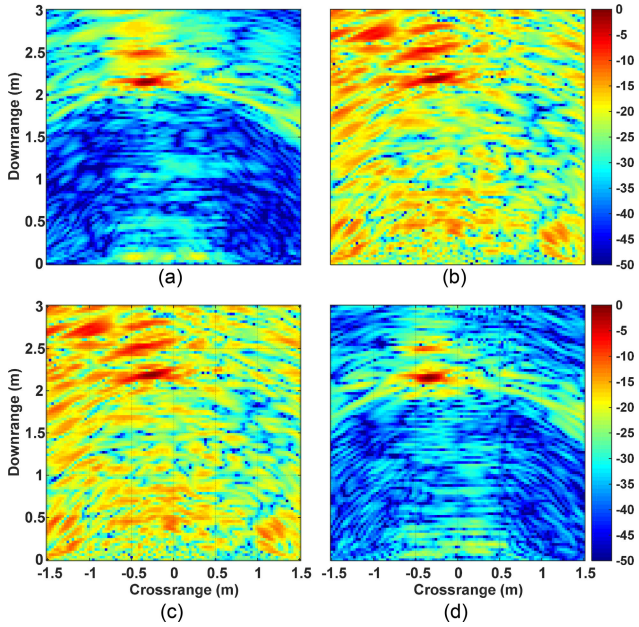

Fig. 6. Multi-polarization images of the concrete-wall dihedral scene formed with DS beamforming using measurements after background subtraction: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

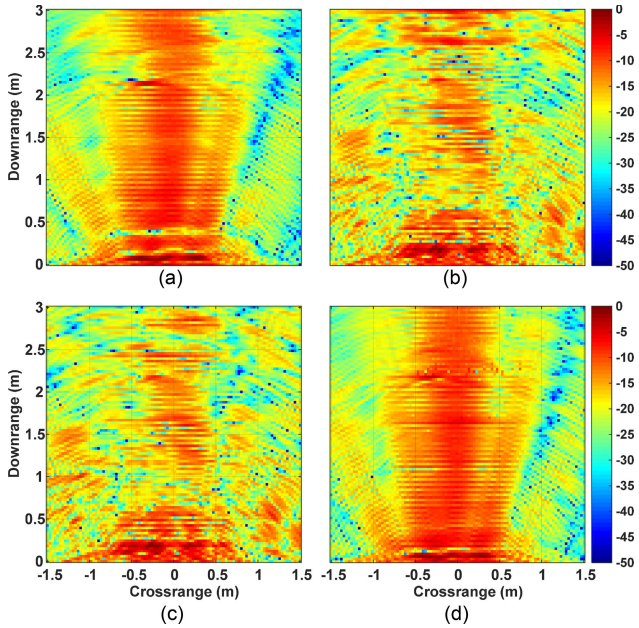

Fig. 7. Multi-polarization images of the concrete-wall dihedral scene formed with DS beamforming using measurements without background subtraction: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

C. Experimental Results and Analysis

This subsection presents the results, analysis, comparison, and performance evaluation for the proposed LRAP-JSTV approach and other imaging methods using several real radar datasets.

1) Clutter Mitigation & Image Reconstruction with Different Subset Frequencies for Reduced Antennas: This experiment aims to evaluate the performance of the LRAP-JSTV approach under generic CS operations where both antennas and frequencies are reduced. For the first wooden-wall dihedral scene, the reduced dataset was generated by randomly selecting $J = 160$ out of $M = 801$ (20%) of the total frequencies at half the antenna locations ($K = 41$ out of $N = 81$) selected randomly. Therefore, $L = 4$ datasets considered in this experiment collectively represent only 10% of the full measurements. The input measurement matrix $\mathbf{Y}$ is of size $KJ \times L = 6{,}560 \times 4$. The dictionary $\mathbf{\Psi}$ is overcomplete with the size of $KJ \times Q = 6{,}560 \times 10{,}000$. Using $\mathbf{Y}$, $\mathbf{\Psi}$, and the selected parameters, Fig. 9 shows the output target image $\mathbf{S}$ estimated by the LRAP-JSTV model. It can be observed that the target is reconstructed and its structure is preserved as target pixels are connected spatially. Furthermore, clutter is suppressed significantly for all channel images. For the second concrete-wall dihedral scene, the reduced dataset was generated in a similar protocol: selecting only half the antenna locations and using 40% of total frequencies at each chosen antenna (collectively representing of 20% of total measurements). Fig. 10 presents the target images reconstructed by LRAP-JSTV approach. The target is localized, and clutter is considerably suppressed.

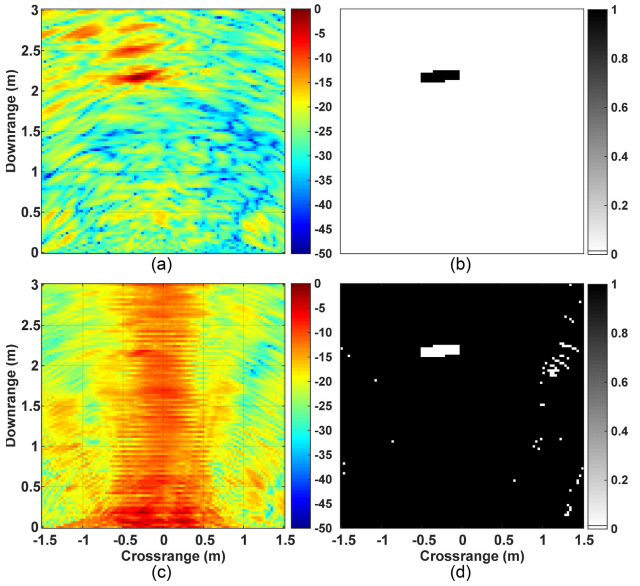

Fig. 8. Target and clutter regions of the dihedral scene with a concrete wall: (a) target image formed with additive fusion of DS beamforming images after background subtraction, (b) target mask obtained by thresholding the target image, (c) clutter image formed with additive fusion of DS beamforming images without clutter mitigation, and (d) clutter mask obtained by thresholding the clutter image.

2) Comparison With Different CS-Based Clutter Mitigation & Image Reconstruction Methods: In this experiment, the proposed LRAP-JSTV approach is compared with five other clutter mitigation and image reconstruction methods in the CS context: (1) Extended RPCA [35], (2) RPCA [20], [21], (3) two-stage CS using subspace projection (SP) and the ℓ_1-norm minimization [16], [42], (4) two-stage CS using spatial filtering (SF) and the ℓ_1-norm minimization [16], [39], and (5) direct ℓ_1 minimization [10], [11]. The extended RPCA (ERPCA) model uses LR and JS representations, i.e., LR-JS, for wall clutter mitigation and target

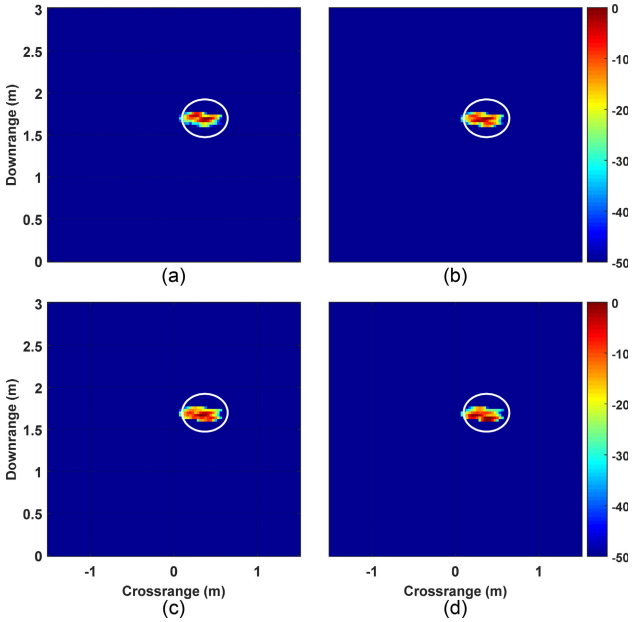

Fig. 9. Multi-polarization images of the wooden-wall dihedral scene reconstructed by the proposed LRAP-JSTV approach using 20% frequencies and 50% antennas; this reduced dataset collectively represents only 10% of the full measurements: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

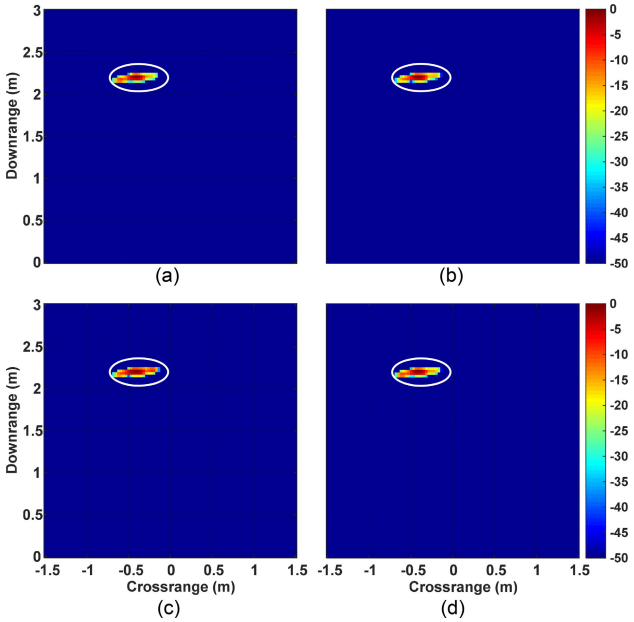

Fig. 10. Multi-polarization images of the concrete-wall dihedral scene reconstructed by the proposed LRAP-JSTV approach using 40% frequencies and 50% antennas; this reduced dataset collectively represents only 20% of the full measurements: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

image formation. Thus, it is equivalent to the proposed LRAP-JSTV model without the analysis prior and the TV regularization term. The RPCA regularizes the wall clutter with LR, but imposes sparsity on the target image (LR-S). While ERPCA and RPCA process wall and target signals simultaneously, the two-stage CS methods mitigate wall clutter

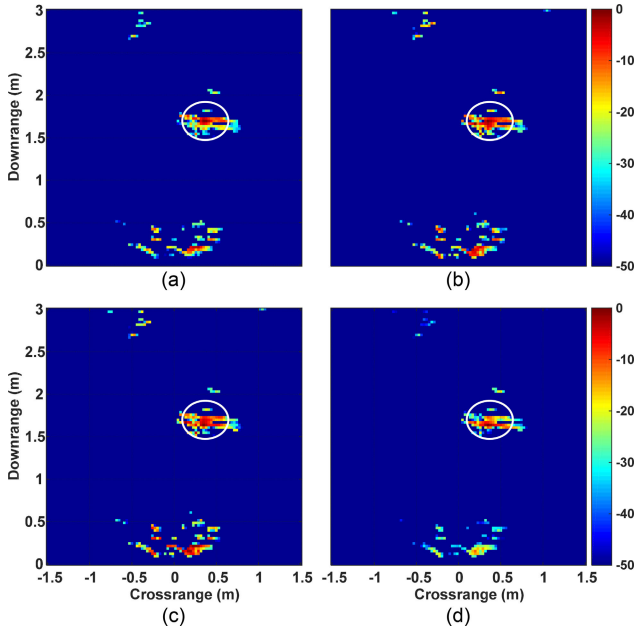

Fig. 11. Multi-polarization images of the wooden-wall dihedral scene reconstructed by the ERPCA (LR-JS) model using 20% frequencies and 50% antennas: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

and reconstruct the target image independently. Note that in the two-stage CS approaches, prior to wall clutter mitigation, missing measurements need to be recovered since the subset of frequencies varies along the reduced set of antennas.

Using the 10% reduced datasets collected from the wooden-wall dihedral scene, Fig. 11 depicts the multi-channel image $\mathbf{S}$ reconstructed by the ERPCA model. The target is detected, but the target pixels are not connected well. Moreover, the level of background clutter suppression is not as high as that of the LRAP-JSTV approach. Fig. 12 shows the target images reconstructed by the RPCA model. Although strong wall clutter is mitigated, the sparsity constraint is not effective for cross-polarization channels. Fig. 13 presents the ℓ_1-norm-based polarimetric images formed after applying SP to the recovered radar signals. The dominant wall clutter is suppressed, but the recovered target is contaminated with heavy background clutter, especially in the cross-polarization images. Fig. 14 shows that the quality of image reconstruction degrades even further when using the SF (instead of SP) for clutter mitigation. Fig. 15 shows the images obtained when applying the CS-based sparse image reconstruction directly, without clutter mitigation. Clearly, the wall clutter dominates the target of interest.

Table II shows the TCR values of the images reconstructed by the different imaging techniques. The TCR values are averaged over all the channels. It can be seen that the proposed LRAP-JSTV approach significantly improves image quality and yields the highest TCR value (31.61 dB). A similar experimental protocol is used for evaluating the different clutter mitigation and image reconstruction methods with the 20% reduced datasets acquired from the concrete-wall dihedral scene. Table III lists the performances of the different approaches in terms of TCR. LRAP-JSTV achieves 30.57 dB,

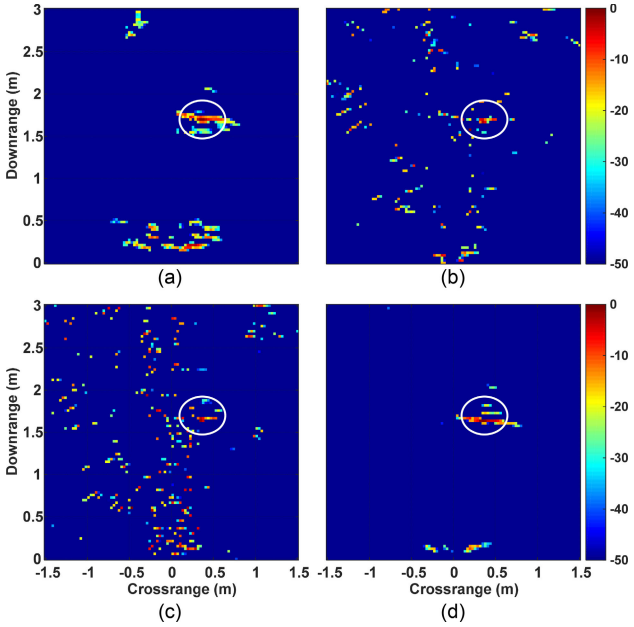

Fig. 12. Multi-polarization images of the wooden-wall dihedral scene reconstructed by the RPCA (LR-S) model using 20% frequencies and 50% antennas: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

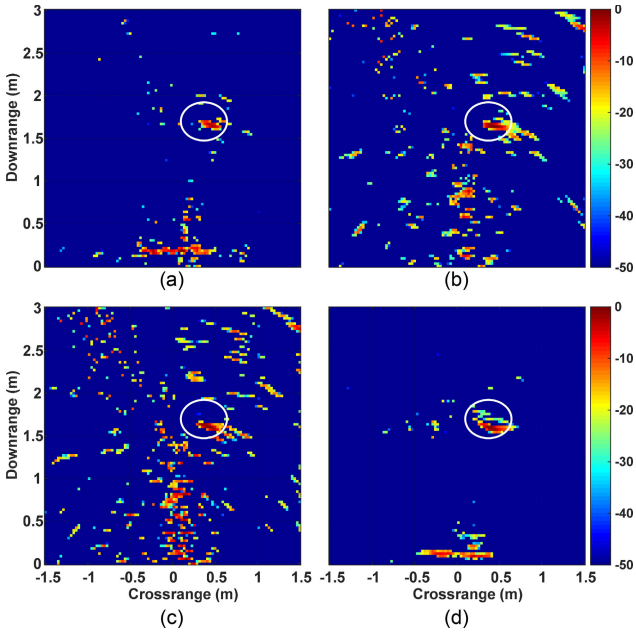

Fig. 14. Multi-polarization images of the wooden-wall target scene reconstructed by the two-stage wall clutter mitigation with spatial filtering and sparse image reconstruction using 20% frequencies and 50% antennas: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

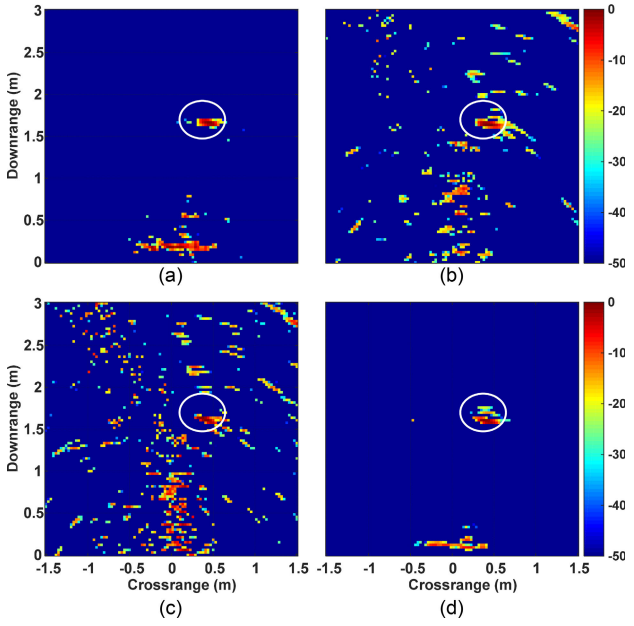

Fig. 13. Multi-polarization images of the wooden-wall target scene reconstructed by the two-stage wall clutter mitigation with subspace projection and sparse image reconstruction using 20% frequencies and 50% antennas: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

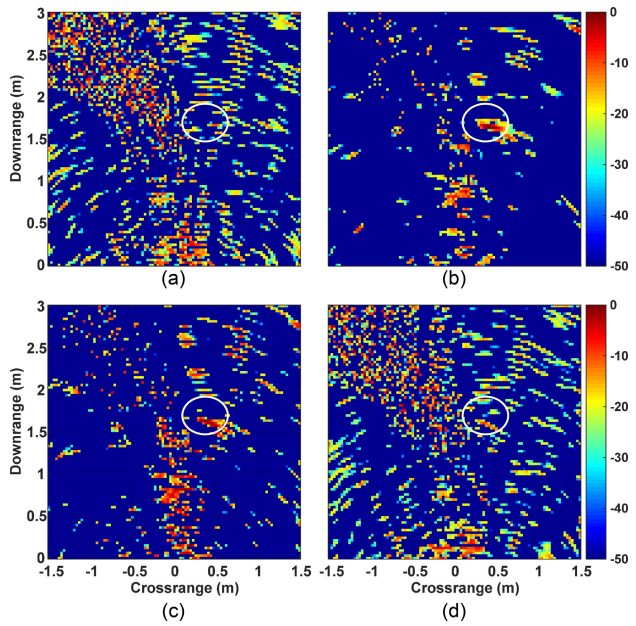

Fig. 15. Multi-polarization images of the wooden-wall target scene reconstructed by applying direct ℓ_1-norm minimization with 20% frequencies and 50% antennas: (a) HH channel, (b) HV channel, (c) VH channel, and (d) VV channel.

followed by ERPCA with a TCR of 21.83 dB and RPCA with a TCR of 17.08 dB.

3) Effects of Data Reduction on Performance: This subsection evaluates the performance of the proposed LRAP-JSTV model under different levels of compressive data measurements. These datasets were generated from the concrete-wall dihedral scene by randomly selecting only 40% of total frequencies at a randomly selected antenna subset. The number of selected antennas is varied from 14 (25%) to 28 (50%) of the total antennas. Therefore, the datasets considered in this experiment collectively represent only 10% to 20% of the total measurements. For each dataset, the rank of the wall component $\mathbf{Z}^w$, the number of iterations, TCR values of the formed images $\mathbf{S}$, and the relative change of the objective function are recorded after convergence. Table IV presents the obtained values for all datasets, averaged over 50 trials for the

TABLE II

AVERAGE TARGET-TO-CLUTTER RATIO (TCR) OVER ALL THE POLARIZATION IMAGES FORMED BY DIFFERENT CLUTTER MITIGATION AND IMAGE RECONSTRUCTION METHODS, USING 20% FREQUENCIES AND 50% ANTENNAS OF THE TOTAL MEASUREMENTS COLLECTED FROM THE WOODEN-WALL DIHEDRAL SCENE

Clutter mitigation & image reconstruction methods	TCR (dB)
Proposed LRAP-JSTV	**31.61**
ERPCA (LR-JS) [35]	22.01
RPCA (LR-S) [20, 21]	18.42
Two-stage subspace projection & ℓ_1 minimization [42, 16]	15.62
Two-stage spatial filtering & ℓ_1 minimization [39, 16]	14.79
Direct ℓ_1 minimization [10, 11]	4.25

TABLE III

AVERAGE TARGET-TO-CLUTTER RATIO (TCR) OVER ALL THE POLARIZATION IMAGES FORMED BY DIFFERENT CLUTTER MITIGATION AND IMAGE RECONSTRUCTION METHODS, USING 40% FREQUENCIES AND 50% ANTENNAS OF THE TOTAL MEASUREMENTS COLLECTED FROM THE CONCRETE-WALL DIHEDRAL SCENE

Clutter mitigation & image reconstruction methods	TCR (dB)
Proposed LRAP-JSTV	**30.57**
ERPCA (LR-JS) [35]	21.83
RPCA (LR-S) [20, 21]	17.08
Two-stage subspace projection & ℓ_1 minimization [42, 16]	10.18
Two-stage spatial filtering & ℓ_1 minimization [39, 16]	7.61
Direct ℓ_1 minimization [10, 11]	0.64

TABLE IV

PERFORMANCE MEASURES OF THE LRAP-JSTV ALGORITHM FOR DIFFERENT REDUCED DATASETS OBTAINED FROM THE CONCRETE-WALL DIHEDRAL SCENE

Dataset	Rank($\mathbf{Z}^w$)	TCR (dB)	Iterations	$\frac{\lvert f(\mathbf{Z}^w_{t+1},\mathbf{S}_{t+1})-f(\mathbf{Z}^w_t,\mathbf{S}_t)\rvert}{\lvert f(\mathbf{Z}^w_t,\mathbf{S}_t)\rvert}$
10%	7	24.48	66	5.11×10^{-5}
12%	6	26.61	64	6.37×10^{-5}
14%	6	28.68	59	8.33×10^{-5}
16%	6	29.01	57	8.38×10^{-5}
18%	6	30.57	46	6.43×10^{-5}
20%	6	31.75	42	2.48×10^{-5}

LRAP-JSTV algorithm. It can be observed that LRAP-JSTV converges well for the different CS datasets and yields good TCR values. In addition, the TCR increases and the number of iterations decreases with the number of measurements used for wall clutter mitigation and image reconstruction.

4) Effect of the Regularization Terms on Performance: This experiment aims to investigate the contributions of the LR, JS, and TV terms in the proposed model. The LR term is used to capture the wall clutter, the JS term is used to enforce joint sparsity among channel images, and the TV term is used to promote target continuity and remove artifacts. The performance of the model with the three regularization terms introduced simultaneously is compared to the performance when the regularization terms are used individually or in pairs. The experiment was conducted using a reduced dataset collected by 50% of the antennas and 50% of the frequencies from the wooden-wall dihedral scene.

Table V lists the TCR values of the images reconstructed by the proposed model using the different combinations of

TABLE V

PERFORMANCE OF THE PROPOSED MODEL IN TERMS OF TCR AVERAGED OVER ALL THE POLARIZATION IMAGES FORMED BY COMBINING DIFFERENT REGULARIZATION TERMS, USING 50% FREQUENCIES AND 50% ANTENNAS OF THE TOTAL MEASUREMENTS COLLECTED FROM THE WOODEN-WALL DIHEDRAL SCENE

Regularization terms	TCR (dB)
Low-rank + joint sparsity + total variation	**38.46**
Low-rank + joint sparsity	27.41
Low-rank + total variation	30.51
Joint sparsity + total variation	6.54
Low-rank	20.63
Joint sparsity	2.62
Total variation	5.27

the regularizers. It is evident that using LR, JS, and TV together yields the highest TCR of 38.46 dB. The pairs (LR, JS) and (LR, TV) produce TCR values of 27.41 dB and 30.51 dB, respectively. The model without the LR term has low TCR values: 6.54 dB for (JS, TV), 2.62 dB for JS, and 5.27 dB for TV. These results show that including the LR term in the model is necessary for separating the weak target signals from the dominant wall clutter. Furthermore, using the three regularizers together enhances the TCR of the formed image significantly.

5) Acceleration and Convergence Analysis of the Proposed Algorithm: The proposed LRAP-JSTV algorithm consists of forward gradient and backward proximity evaluations. It can be observed from Algorithm 1 that the proximal evaluations of the LRAP, LS, and TV are performed efficiently through the shrinkage technique with closed-form expressions. The computational efficiency is further achieved through the integration of the acceleration feature in the forward gradient evaluation at Step 2, see Algorithm 1. In this step, both current and previous estimates of wall clutter and target image $(\mathbf{Z}^w_t, \mathbf{S}_t)$ and $(\mathbf{Z}^w_{t-1}, \mathbf{S}_{t-1})$ are used for computing the auxiliary variables $(\mathbf{U}^w_t, \mathbf{U}^s_t)$. In the case of without acceleration, only current estimates are used for gradient evaluation, i.e., $\mathbf{U}^w_t \leftarrow \mathbf{Z}^w_t$ and $\mathbf{U}^s_t \leftarrow \mathbf{S}_t$. To investigate the effect of the acceleration scheme on the convergence rate, we use the 25% reduced dataset collected from half of the antennas and half of the frequencies, and evaluate the performance and convergence of the proposed model, with and without acceleration. Furthermore, we compare the running times of the proposed algorithm with those of several existing CS TWRI algorithms. All the algorithms were implemented in MATLAB and executed on a PC with Intel Core i7-3820 3.6GHz CPU and 32GB RAM.

Fig. 16 shows the convergence curves of $f(\mathbf{Z}^w, \mathbf{S})$ as a function of the number of iterations. The value of the objective function decreases during minimization. Furthermore, the objective function converges more rapidly with the acceleration scheme than with the standard gradient evaluation. The algorithm with acceleration reaches a steady-state (a value of 4.52) in 18 iterations, whereas the algorithm without acceleration takes 36 iterations to reach the same value.

Fig. 17 presents the rank values of $\mathbf{Z}^w$ recorded as a function of the number of iterations. It can be observed that for both cases (with and without acceleration), the algorithm

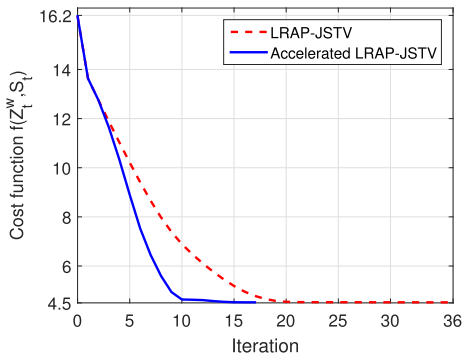

Fig. 16. Evolution of the objective function $f(\mathbf{Z}^w, \mathbf{S})$ during the minimization of LRAP-JSTV with acceleration (solid line) and without acceleration (dashed line), using 50% of frequencies and 50% of available antennas.

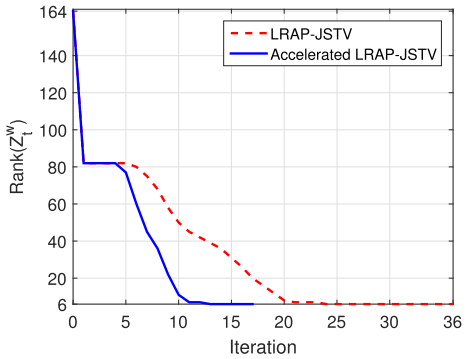

Fig. 17. The rank values of the estimated wall component $\mathbf{Z}^w$ as a function of the number of iterations by LRAP-JSTV with acceleration (solid line) and without acceleration (dashed line), using 50% of frequencies and 50% of available antennas.

TABLE VI

RUNNING TIMES OF SEVERAL CLUTTER MITIGATION AND IMAGE RECONSTRUCTION ALGORITHMS EVALUATED USING 50% FREQUENCIES AND 50% ANTENNAS OF THE TOTAL MEASUREMENTS COLLECTED FROM THE WOODEN-WALL DIHEDRAL SCENE

Clutter removal & imaging methods	**Iters.**	**Times (secs.)**
Accelerated LRAP-JSTV	**18**	**87.10**
LRAP-JSTV	36	180.30
ERPCA (LR-JS) [35]	40	152.34
RPCA (LR-S) [20, 21]	38	141.45
Two-stage SP & ℓ_1 minimization [42, 16]	–	253.58
Two-stage SF & ℓ_1 minimization [39, 16]	–	181.14
Two-stage SP & DS beamforming	–	83.13
Direct ℓ_1 minimization [10, 11]	69	144.12

achieves a low-rank representation of the wall signals, reaching a value of 6 at the steady-state, starting from a rank of 164, $\min(M, K \times L) = \min(801, 41 \times 4)$, at the first iteration. However, the rank of $\mathbf{Z}^w$ obtained with the acceleration scheme converges faster than the one without the acceleration scheme.

The running times of the different clutter mitigation and image reconstruction approaches were recorded and listed in Table VI. It is clear that although DS beamforming is the fastest technique, it requires signal recovery for wall clutter mitigation and image formation independently for 4 channels, which yields a run time of 83.13 s. By processing all the channel datasets simultaneously and integrating the acceleration scheme, the proposed algorithm takes 87.10 s or 18 iterations to produce the target images. The two-stage wall clutter mitigation and ℓ_1-minimization is the most time-consuming approach because its processing time includes the times for signal estimation, wall clutter mitigation, and ℓ_1-regularized image formation for all the channel images.

6) Discussion: It is worth noting that although the proposed model outperforms all the tested state-of-the-art methods in terms of TCR, the proposed approach requires the hyperparameters to be tuned appropriately. In this paper, Bayesian optimization is employed to search for the optimal hyperparameters, which increases the execution time. However, due to the capability of incorporating prior knowledge about the scores associated with the evaluated hyperparameters, Bayesian optimization with Gaussian process through its acquisition function optimizes the probability of searching for the hyperparameter with the best score, which makes Bayesian optimization much more efficient than heuristic, cross-validation grid, or random search strategies.

V. CONCLUSION

This paper presented an efficient approach that incorporates low-rank, jointly sparse and total variation for solving the crucial problems of wall clutter mitigation, target image reconstruction, and background artifact suppression in compressive TWRI. By including an analysis prior, the proposed model enhances wall clutter representation and segregation from the target signals. Furthermore, the introduction of TV improves the continuity of targets and alleviates background clutter. Here, TV regularization is extended as joint derivative sparsity and its computation is approximated with the wavelet decomposition and cycle spinning technique. This allows the development of an accelerating iterative algorithm to solve the LRAP, JS, and TV regularized TWRI efficiently. Through extensive experimental validation, we find that the proposed approach segregates wall clutter from target reflections well and yields high-quality images of indoor targets, even with reduced measurements. In compressive TWR sensing, the proposed method outperformed the ERPCA, RPCA, and existing CS-based wall clutter mitigation and image formation techniques in terms of TCR.

ACKNOWLEDGMENT

The authors would like to thank Drs. Moeness G. Amin and Fauzia Ahmad from the Center of Advanced Communications at Villanova University, Villanova, PA, USA, for providing part of the experimental data. The authors are also grateful for the constructive feedback from the anonymous reviewers.

REFERENCES

[1] M. G. Amin, Ed., *Through-The-Wall Radar Imaging*. Boca Raton, FL, USA: CRC Press, 2010.

[2] F. Ahmad, M. G. Amin, and G. Mandapati, "Autofocusing of through-the-wall radar imagery under unknown wall characteristics," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1785–1795, Jul. 2007.

[3] G. Wang and M. G. Amin, "Imaging through unknown walls using different standoff distances," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 4015–4025, Oct. 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[5] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[6] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[7] F. Ahmad, M. G. Amin, and S. A. Kassam, "Synthetic aperture beamformer for imaging through a dielectric wall," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 271–283, Jan. 2005.

[8] F. Ahmad and M. Amin, "Noncoherent approach to through-the-wall radar localization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 4, pp. 1405–1419, Oct. 2006.

[9] M. G. Amin and F. Ahmad, "Wideband synthetic aperture beamforming for through-the-wall imaging [Lecture Notes]," *IEEE Signal Process. Mag.*, vol. 25, no. 4, pp. 110–113, Jul. 2008.

[10] Y.-S. Yoon and M. G. Amin, "Compressed sensing technique for high-resolution radar imaging," in *Proc. 17th Signal Process., Sensor Fusion, Target Recognit.*, Orlando, FL, USA, Mar. 2008, pp. 69681A.1–69681A.10.

[11] Q. Huang, L. Qu, B. Wu, and G. Fang, "UWB through-wall imaging based on compressive sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1408–1415, Mar. 2010.

[12] M. Leigsnering, C. Debes, and A. M. Zoubir, "Compressive sensing in through-the-wall radar imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Prague, Czech Republic, May 2011, pp. 4008–4011.

[13] V. H. Tang, A. Bouzerdoum, and S. L. Phung, "Two-stage through-the-wall radar image formation using compressive sensing," *J. Electron. Imag.*, vol. 22, no. 2, Apr./Jun. 2013, Art. no. 021006.

[14] J. Yang, A. Bouzerdoum, F. H. C. Tivive, and M. G. Amin, "Multiple-measurement vector model and its application to through-the-wall radar imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Prague, Czech Republic, May 2011, pp. 2672–2675.

[15] A. Bouzerdoum, J. Yang, and F. H. C. Tivive, "Compressive sensing for multipolarization through-the-wall radar imaging," in *Compressive Sensing for Urban Radar*. M. G. Amin, Ed. Boca Raton, FL, USA: CRC Press, Aug. 2014, pp. 232–248.

[16] E. Lagunas, M. G. Amin, F. Ahmad, and M. Najar, "Joint wall mitigation and compressive sensing for indoor image reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 891–906, Feb. 2013.

[17] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Enhanced wall clutter mitigation for compressed through-the-wall radar imaging using joint Bayesian sparse signal recovery," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 7804–7808.

[18] A. Bouzerdoum, F. H. C. Tivive, and V. H. Tang, "Multi-polarization through-the-wall radar imaging using joint Bayesian compressed sensing," in *Proc. 19th Int. Conf. Digit. Signal Process.*, Hong Kong, Aug. 2014, pp. 783–788.

[19] F. Ahmad, J. Qian, and M. G. Amin, "Wall clutter mitigation using discrete prolate spheroidal sequences for sparse reconstruction of indoor stationary scenes," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1549–1557, Mar. 2015.

[20] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Radar imaging of stationary indoor targets using joint low-rank and sparsity constraints," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Shanghai China, Mar. 2016, pp. 1412–1416.

[21] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Indoor scene reconstruction for through-the-wall radar imaging using low-rank and sparsity constraints," in *Proc. IEEE Radar Conf. (RadarConf)*, Philadelphia PA, USA, May 2016, pp. 1–4.

[22] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[23] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *NumerischeMath.*, vol. 76, no. 2, pp. 167–188, Apr. 1997.

[24] J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira, "Total variation-based image deconvolution: A majorization-minimization approach," in *Proc. IEEE Int. Conf. Acoust. Speed Signal Process. Proc.*, Toulouse France, May 2006, pp. 861–864.

[25] I. W. Selesnick, H. L. Graber, D. S. Pfeil, and R. L. Barbour, "Simultaneous low-pass filtering and total variation denoising," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1109–1124, Mar. 2014.

[26] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[27] U. S. Kamilov, "A parallel proximal algorithm for anisotropic total variation minimization," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 539–548, Feb. 2017.

[28] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, Jan. 2005.

[29] P. L. Combettes and J. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. Luke, and H. Wolkowicz, Eds. New York, NY, USA: Springer, 2011, pp. 185–212.

[30] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 123–231, Nov. 2013.

[31] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci..*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[32] S. Becker, J. Bobin, and E. J. Candes, "NESTA: A fast and accurate first-order method for sparse recovery," *SIAM J. Imag. Sci.*, vol. 4, no. 1, pp. 1–39, Jan. 2011.

[33] Y. Nesterov, "Gradient methods for minimizing composite functions," *Math. Program.*, vol. 140, no. 1, pp. 125–161, Aug. 2013.

[34] M. Leigsnering, F. Ahmad, M. Amin, and A. Zoubir, "Multipath exploitation in through-the-wall radar imaging using sparse reconstruction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 920–939, Apr. 2014.

[35] V. H. Tang, A. Bouzerdoum, and S. L. Phung, "Multipolarization through-wall radar imaging using low-rank and jointly-sparse representations," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1763–1776, Apr. 2018.

[36] R. T. Rockafellar, *Convex Analysis* (Princeton Math). Princeton, NJ, USA: University Press, 1970.

[37] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[38] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, nos. 1–2, pp. 321–353, Jun. 2011.

[39] Y.-S. Yoon and M. G. Amin, "Spatial filtering for wall-clutter mitigation in through-the-wall radar imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 9, pp. 3192–3208, Sep. 2009.

[40] F. H. C. Tivive, A. Bouzerdoum, and M. G. Amin, "An SVD-based approach for mitigating wall reflections in through-the-wall radar imaging," in *Proc. IEEE RadarCon (RADAR)*, Kansas City, MO, USA, May 2011, pp. 519–524.

[41] F. H. C. Tivive and A. Bouzerdoum, "An improved SVD-based wall clutter mitigation method for through-the-wall radar imaging," in *Proc. IEEE 14th Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Darmstadt, Germany, Jun. 2013, pp. 430–434.

[42] F. H. C. Tivive, A. Bouzerdoum, and M. G. Amin, "A subspace projection approach for wall clutter mitigation in through-the-wall radar imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2108–2122, Apr. 2015.

[43] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.

[44] C. Vonesch and M. Unser, "A fast thresholded landweber algorithm for wavelet-regularized multidimensional deconvolution," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 539–549, Apr. 2008.

[45] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.

[46] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *Ann. Appl. Stat.*, vol. 1, no. 2, pp. 302–332, Dec. 2007.

[47] R. R. Coifman and D. L. Donoho, *Translation-Invariant De-Noising*. New York, NY, USA: Springer, 1995, pp. 125–150.

[48] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk, "On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs," *SIAM J. Numer. Anal.*, vol. 42, no. 2, pp. 686–713, Jan. 2004.
[49] U. Kamilov, E. Bostan, and M. Unser, "Wavelet shrinkage with consistent cycle spinning generalizes total variation denoising," *IEEE Signal Process. Lett.*, vol. 19, no. 4, pp. 187–190, Apr. 2012.
[50] J.-F. Cai, E. J. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.
[51] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, Eds. Curran Associates, Dec. 2011, pp. 2546–2554.
[52] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Dec. 2012, pp. 2951–2959.
[53] C. Yao, D. Cai, J. Bu, and G. Chen, "Pre-training the deep generative models with adaptive hyperparameter optimization," *Neurocomputing*, vol. 247, pp. 144–155, Jul. 2017.
[54] J. Snoek *et al.*, "Scalable Bayesian optimization using deep neural networks," in *Proc. Int. Conf. Mach. Learn.*, Lille, France, Jun. 2015, pp. 2171–2180.
[55] F. H. C. Tivive, A. Bouzerdoum, and C. Abeynayake, "GPR target detection by joint sparse and low-rank matrix decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2583–2595, May 2019.
[56] B. Matérn, *Spatial Variation* (Lecture Notes in Statistics), vol. 36. Berlin, Germany: Springer-Verlag, 1986.
[57] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," *J. Basic Eng.*, vol. 86, no. 1, pp. 97–106, Mar. 1964.
[58] J. Mockus, V. Tiesis, and A. Zilinskas, "The application of Bayesian methods for seeking the extremum," in *Towards Global Optimisation*, L. C. W. Dixon and G. P. Szego, Eds. Amsterdam, The Netherlands: North Holland, 1978, pp. 117–130.
[59] S. K. N. Srinivas, A. Krause, and M. Seeger, "Gaussian process optimization in the bandit setting: No regret and experimental design," in *Proc. Int. Conf. Mach. Learn.*, Haifa, Israel, Jun. 2010, pp. 1015–1022.

Van Ha Tang received the B.Eng. and M.Eng. degrees in computer engineering from Le Quy Don Technical University, Hanoi, Vietnam, in 2005 and 2008, respectively, and the Ph.D. degree in computer engineering from the University of Wollongong, Australia, in 2016. Since 2016, he has been working as a Postdoctoral Research Fellow at the School of Electrical Computer and Telecommunications Engineering, University of Wollongong. He is currently with the Faculty of Information Technology, Le Quy Don Technical University. His research interests include radar imaging and through-wall radar signal processing, statistical and probabilistic modeling, and inverse problems in image and signal processing.

Abdesselam Bouzerdoum (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, WA, USA.

From 1991 to 1997, he was with Adelaide University as a Research Associate, Assistant Professor, and Senior Assistant Professor. In 1998, he moved to Edith Cowan University, Western Australia, as an Associate Professor. In September 2004, he was appointed as a Professor of computer engineering and the Head of the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong. From 2007 to 2013, he was the Associate Dean (Research) of the Faculty of Informatics. In 2015, he was promoted to the position of Senior Professor of computer engineering at the University of Wollongong. He is currently the Head of the Information and Computing Technology Division, College of Science and Engineering, and also an Acting Associate Provost of Academic Affairs at Hamad Bin Khalifa University, Qatar. He held several Visiting Professor Appointments with international institutions in France, USA, and Hong Kong. From 2009 to 2011, he has served on the ARC College of Experts. He has published over 360 technical articles and graduated 50 Ph.D. and Research Masters students. His research interests include radar imaging and signal processing, image processing, vision, machine learning, and pattern recognition. He is a recipient of several awards, most notably the *Eureka Prize for Outstanding Science in Support of Defence or National Security* (2011) and the *Chercheur de Haut Niveau Award* from the French Ministry of Research and Higher Education (2001). He has served on the editorial boards of five international journals, including the IEEE TRANSACTIONS IMAGE ON PROCESSING and the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS (1999–2006).

Son Lam Phung (Senior Member, IEEE) received the B.Eng. (Hons.) and Ph.D. degrees from Edith Cowan University, Perth, Australia, in 1999 and 2003, respectively, both in computer engineering. He is currently an Associate Professor with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong. His general research interests are in the areas of image and signal processing, neural networks, pattern recognition, and machine learning. He received the University and Faculty Medals in 2000.