A VARIATIONAL BAYESIAN APPROACH FOR MULTICHANNEL THROUGH-WALL RADAR IMAGING WITH LOW-RANK AND SPARSE PRIORS

Van Ha Tang[‡] *Abdesselam Bouzerdoum*[†⋆] *Son Lam Phung*[⋆]

[‡]Faculty of Information Technology, Le Quy Don Technical University, Hanoi, Vietnam
[†]Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar
[⋆]School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia

ABSTRACT

This paper considers the problem of multichannel through-wall radar (TWR) imaging from a probabilistic Bayesian perspective. Given the observed radar signals, a joint distribution of the observed data and latent variables is formulated by incorporating two important beliefs: low-dimensional structure of wall reflections and joint sparsity among channel images. These priors are modeled through probabilistic distributions whose hyperparameters are treated with a full Bayesian formulation. Furthermore, the paper presents a variational Bayesian inference algorithm that captures wall clutter and provides channel images as full posterior distributions. Experimental results on real data show that the proposed model is very effective at removing wall clutter and enhancing target localization.

***Index Terms*—** Through-the-wall radar imaging, wall clutter mitigation, sparse Bayesian learning, variational inference.

1. INTRODUCTION

Over the past two decades, through-wall radar imaging (TWRI) and urban sensing technology have become a very active area of research and development due to its numerous desirable applications for law enforcement, emergency services, and military operations [1]. In such operations, the need for providing high-quality radar imagery of indoor targets is challenged by the presence of strong wall returns and weak target responses. This issue can be alleviated using multichannel imaging that combines useful electromagnetic (EM) data obtained from different polarizations.

Several imaging methods using multipolarization EM waves for enhancing TWR target localization have been proposed [2–5]. In [2], multichannel images were used for statistical-based motion indication. Adaptive target detection was performed using multipolarization imaging in [3]. The diversity of polarimetric signatures was exploited for target classification in [4]. In [5], multichannel images were fused to enhance target localization. These approaches, however, are applicable only for full multichannel sensing requiring complete data measurements.

For fast data collection and efficient storage, single and multichannel TWRI techniques were investigated using a compressive sensing (CS) framework [6, 7]. It has been shown in [8–12] that target image formation can be performed from reduced measurements by exploiting image sparsity through ℓ_1 constraint. The ℓ_1 constraint is extended to $\ell_{2,1}$ to account for the joint sparsity prior knowledge among channel images [13–16]. The target reconstruction is improved, but such algorithms provide a point estimate of the target image, thereby not quantifying the uncertainty associated with the recovery given the observed measurements. More importantly, such methods do not consider wall clutter in their imaging models and require a reference scene for wall clutter mitigation, which may be impractical in many operations. In TWRI, strong front-wall EM returns overwhelm target echoes, masking targets in the formed image [17–19]. To overcome this problem, multistage CS-based approaches have been considered that comprise signal estimation, wall clutter mitigation, and target image reconstruction [20–23]. However, such methods face the issue of multistage uncertainty because the signal estimation, wall clutter mitigation, and image formation tasks are performed separately. The issue of multistage processing has been addressed recently in [24, 25], but these techniques are still deterministic and require intensive hyperparameter tuning.

This paper introduces a probabilistic Bayesian learning approach for multichannel indoor radar imaging. We formulate a joint probability distribution of the observed measurements and the latent variables of the wall clutter matrix and target images that provides a complete summary of the uncertainty associated with those variables. Such variables and hyperparameters are considered as stochastic quantities, and treated with full Bayesian modeling. We propose an efficient variational inference that estimates wall clutter, channel images, and hyperparameters, and provides the uncertainty in the form of posterior distributions.

The remainder of the paper is organized as follows. Section 2 introduces the multichannel TWR signal model. Section 3 describes the proposed Bayesian approach for modeling the wall clutter and multichannel images. Section 4 presents the experimental evaluations. Finally, Section 5 gives concluding remarks.

2. MULTICHANNEL TWR SIGNAL MODEL

Consider a monostatic stepped-frequency TWR using L channels, N antennas, and M narrowband signals to image P targets behind-the-wall. Let $x_l(m,n)$ be the mth frequency radar signal received by the nth antenna for the lth channel. The signal $x_l(m,n)$ can be modeled as a superposition of the wall reflections $x_l^{\mathrm{w}}(m,n)$, target returns $x_l^{\mathrm{t}}(m,n)$, and noise $v_l(m,n)$:

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

The wall component $x_l^{\mathrm{w}}(m,n)$ can be expressed as [20]

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

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 $\tau^r_{n,\mathrm{w}}$ is the propagation delay of the rth wall reverberation. The target return is modeled as a superposition of echoes from P scatterers [22]:

$$x^{\mathrm{t}}_l(m,n) = \sum_{p=1}^{P} \sigma^l_p e^{-j2\pi f_m \tau_{n,p}}, \tag{3}$$

where σ^l_p is the reflectivity of the pth scatterer or target for the lth channel, and $\tau_{n,p}$ is the round-trip travel time of the signal from the nth antenna to the pth target. The signals received by the nth antenna for M frequencies are arranged into an $M \times 1$ vector $\mathbf{x}_{l,n}$,

$$\mathbf{x}_{l,n} = [x_l(1,n), \ldots, x_l(M,n)]^T = \mathbf{x}^{\mathrm{w}}_{l,n} + \mathbf{x}^{\mathrm{t}}_{l,n} + \boldsymbol{\upsilon}_{l,n}. \tag{4}$$

For image formation, the target space is partitioned into a rectangular grid consisting of 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:

$$s_l(q) = \begin{cases} \sigma^l_p, & \tau_{n,q} = \tau_{n,p}, \\ 0, & \tau_{n,q} \neq \tau_{n,p}. \end{cases} \tag{5}$$

Eq. (5) indicates that the value of the qth pixel is nonzero if it includes the pth target. Here, $\tau_{n,q}$ is the focusing delay between the nth antenna and the qth pixel. From (3), we can relate the target component $\mathbf{x}^{\mathrm{t}}_{l,n}$ to the lth channel image $\mathbf{s}_l = [s_l(1), \ldots, s_l(Q)]^T$,

$$\mathbf{x}^{\mathrm{t}}_{l,n} = \mathbf{D}_n \, \mathbf{s}_l, \tag{6}$$

where $\mathbf{D}_n$ is an $M \times Q$ matrix whose (m,q)th entry is given by $D_n(m,q) = \exp(-j2\pi f_m \tau_{n,q})$. From Eqs. (4) and (6), we have

$$\mathbf{x}_{l,n} = \mathbf{x}^{\mathrm{w}}_{l,n} + \mathbf{D}_n \, \mathbf{s}_l + \boldsymbol{\upsilon}_{l,n}. \tag{7}$$

Due to the presence of strong wall component $\mathbf{x}^{\mathrm{w}}_{l,n}$, applying imaging methods to the raw radar signal $\mathbf{x}_{l,n}$ yields an image $\mathbf{s}_l$ in which wall clutter dominates targets. Thus, before image formation, the wall interference needs to be removed. Wall clutter mitigation techniques, such as subspace projection [17], can be used, but they are less effective in CS TWR because only a subset of frequency samples is available. To address this issue, multistage CS-based imaging techniques were proposed, where the missing measurements are first recovered, followed by wall clutter mitigation applied to the estimated measurements. This approach, however, cannot cope with the multistage uncertainty.

3. BAYESIAN MULTICHANNEL TWRI MODEL

To cope with multistage uncertainty, a full Bayesian formulation is proposed for multipolarization TWRI. The proposed probabilistic modeling is described in the next subsection, followed by a variational Bayesian inference in Subsection 3.2.

3.1. Probabilistic modeling

Consider the signal model in (7). The wall clutter term $\mathbf{x}^{\mathrm{w}}_{l,n}$ in the signal space can be represented in the wall subspace as

$$\mathbf{x}^{\mathrm{w}}_{l,n} = \mathbf{W} \, \mathbf{z}_{l,n}, \tag{8}$$

where the matrix $\mathbf{W}$ of size $M \times J$ has columns representing the basis vectors of the wall subspace, and $\mathbf{z}_{l,n} \in \mathbb{C}^{J\times 1}$ is the representation of the wall signal $\mathbf{x}^{\mathrm{w}}_{l,n}$ in the wall subspace. As we will see later in Subsection 3.2, the latent variables $\mathbf{W}$ and $\mathbf{z}_{l,n}$ are inferred adaptively from the observed variable $\mathbf{x}_{l,n}$. It follows from Eqs. (7) and (8) that $\mathbf{x}_{l,n}$ can be written as

$$\mathbf{x}_{l,n} = \mathbf{W} \, \mathbf{z}_{l,n} + \mathbf{D}_n \, \mathbf{s}_l + \boldsymbol{\upsilon}_{l,n}. \tag{9}$$

Note that Eq. (9) represents the signal model in the full sensing case where all the measurements from the N antennas are available. It is also applicable to CS scenarios where only reduced sets of antennas and frequencies are used. However, in this case n denotes the index of a selected position belonging to the antenna subset $\Omega_n \subseteq \{1, \ldots, N_a\}$, where N_a is the total number of antennas and N is the number of selected antennas, $N = |\Omega_n|$. Likewise, if all M frequencies are used, then the vector $\mathbf{x}_{l,n}$ is complete; otherwise, it contains only a subset of frequency measurements. In other words, the mth entry $x_{l,n}(m)$ is observed if the mth frequency belongs to the selected frequency subset $\Omega_m \subseteq \{1, \ldots, M\}$.

The Bayesian treatment for (9) models the noise component $\boldsymbol{\upsilon}_{l,n}$ as a complex Gaussian distribution with zero-mean and covariance $\beta^{-1}\mathbf{I}$, $p(\boldsymbol{\upsilon}_{l,n}) = \mathcal{CN}(\boldsymbol{\upsilon}_{l,n}|\mathbf{0}, \beta^{-1}\mathbf{I})$. This implies that the likelihood function can be expressed as

$$\begin{aligned} p(\mathbf{x}_{l,n}|\mathbf{W}, \mathbf{z}_{l,n}, \mathbf{s}_l, \beta) &= \mathcal{CN}(\mathbf{x}_{l,n}|\mathbf{W}\mathbf{z}_{l,n} + \mathbf{D}_n\mathbf{s}_l, \beta^{-1}\mathbf{I}) \\ &= (\frac{\beta}{\pi})^M \exp\left\{-\beta\|\mathbf{x}_{l,n} - (\mathbf{W}\,\mathbf{z}_{l,n} + \mathbf{D}_n\mathbf{s}_l)\|_2^2\right\}. \end{aligned} \tag{10}$$

The conditional distribution of all the observed radar signals $\mathbf{X} = [\mathbf{x}_{l,n}] \in \mathbb{C}^{M\times LN}$ given the latent variables $\mathbf{Z} = [\mathbf{z}_{l,n}] \in \mathbb{C}^{J\times LN}$ and $\mathbf{S} = [\mathbf{s}_l] \in \mathbb{C}^{Q\times L}$ can be expressed as

$$p(\mathbf{X}|\mathbf{W}, \mathbf{Z}, \mathbf{S}, \beta) = \prod_{l=1}^{L}\prod_{n=1}^{N} p(\mathbf{x}_{l,n}|\mathbf{W}, \mathbf{z}_{l,n}, \mathbf{s}_l, \beta). \tag{11}$$

Now prior beliefs over the latent variables $\mathbf{W}$, $\mathbf{Z}$, and $\mathbf{S}$ are incorporated through their probabilistic distributions. The prior knowledge of $\mathbf{W}$ is that the effective dimension J of the wall-component subspace is much smaller than the dimension M of radar signal ($J \ll M$). This can be modeled by introducing sparse Gaussian priors over the columns of $\mathbf{W}$, $\mathbf{w}_j$, for $j = 1, \ldots, J$,

$$p(\mathbf{W}|\boldsymbol{\alpha}) = \prod_{j=1}^{J} \mathcal{CN}(\mathbf{w}_j|\mathbf{0}, \alpha_j^{-1}\mathbf{I}), \tag{12}$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_J]^T$. During inference, most entries of $\boldsymbol{\alpha}$ approach very large values, thereby driving the corresponding column vectors $\mathbf{w}_j$ to zero. The supporting (nonzero) columns of $\mathbf{W}$ are regarded as relevant basis vectors modeling the wall-clutter subspace and determining the rank of the wall clutter matrix. This automatic determination for the wall clutter subspace and its dimension is similar to the mechanism of relevant vector machine in the sparse Bayesian learning (SBL) framework [26–28]. Note that the effective columns of $\mathbf{W}$ are associated with the rows of the matrix $\mathbf{Z}$ due to the factorization, $\mathbf{X}^{\mathrm{w}} = [\mathbf{x}^{\mathrm{w}}_{l,n}] = \mathbf{W}\,\mathbf{Z}$, see (8). We therefore use $\boldsymbol{\alpha}$ to control the rows $\mathbf{z}_{j,:}$ of $\mathbf{Z}$:

$$p(\mathbf{Z}|\boldsymbol{\alpha}) = \prod_{j=1}^{J} \mathcal{CN}(\mathbf{z}_{j,:}|\mathbf{0}, \alpha_j^{-1}\mathbf{I}). \tag{13}$$

The prior of the target image $\mathbf{S}$ incorporates the knowledge that its columns represent sparse images of the same targets (joint sparsity support). We use a hyperparameter vector $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_Q]^T$ to control the sparsity of each image $\mathbf{s}_l$ modeled as a sparse Gaussian,

$$p(s_l(q)|\gamma_q) = \mathcal{CN}(s_l(q)|0, \gamma_q^{-1}). \tag{14}$$

During SBL, many values of $\boldsymbol{\gamma}$ tend to be large, promoting the sparsity of the target image. Because $\boldsymbol{\gamma}$ is used to model the sparsity for all the channel images, it promotes the joint sparsity among columns of $\mathbf{S}$. The overall prior over $\mathbf{S}$ is given by

$$p(\mathbf{S}|\boldsymbol{\gamma}) = \prod_{l=1}^{L}\prod_{q=1}^{Q} p(s_l(q)|\gamma_q). \tag{15}$$

For a full Bayesian model, the gamma conjugate distributions are introduced over the hyperparameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β as follows

$$p(\boldsymbol{\alpha}) = \prod_{j=1}^{J} \mathrm{Gam}(\alpha_j|a_\alpha, b_\alpha) \propto \prod_{j=1}^{J} \alpha_j^{a_\alpha - 1}\exp(-b_\alpha \alpha_j), \tag{16}$$

$$p(\boldsymbol{\gamma}) = \prod_{q=1}^{Q} \mathrm{Gam}(\gamma_q|a_\gamma, b_\gamma) \propto \prod_{q=1}^{Q} \gamma_q^{a_\gamma - 1}\exp(-b_\gamma \gamma_q), \tag{17}$$

$$p(\beta) = \mathrm{Gam}(\beta|a_\beta, b_\beta) \propto \beta^{a_\beta - 1}\exp(-b_\beta \beta). \tag{18}$$

Here, the parameters $a_{(.)}$, $b_{(.)}$ are treated as deterministic and set to very small values (e.g. 10^{-8}) to obtain corresponding broad hyperpriors. We now have the joint distribution of all the observed variable $\mathbf{X}$ and latent variables $\boldsymbol{\Theta} = \{\mathbf{W}, \mathbf{Z}, \mathbf{S}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\}$ given by

$$\begin{aligned} p(\mathbf{X}, \boldsymbol{\Theta}) = p(\mathbf{X}|\mathbf{W}, \mathbf{Z}, \mathbf{S}, \beta)\, p(\mathbf{W}|\boldsymbol{\alpha})\, p(\mathbf{Z}|\boldsymbol{\alpha}) \\ p(\mathbf{S}|\boldsymbol{\gamma})\, p(\boldsymbol{\alpha})\, p(\boldsymbol{\gamma})\, p(\beta). \end{aligned} \tag{19}$$

3.2. Variational Bayesian inference

Given the joint distribution in (19), an exact Bayesian inference requires the computation of the posterior distribution, $p(\boldsymbol{\Theta}|\mathbf{X})$, which is analytically intractable. Therefore, approximation methods must be used. Here, we seek a complete Bayesian treatment using the variational technique that approximates the posterior $p(\boldsymbol{\Theta}|\mathbf{X})$ by a distribution $q(\boldsymbol{\Theta})$ satisfying the following decomposition of the log marginal probability $p(\mathbf{X})$ [29, 30],

$$\ln p(\mathbf{X}) = \mathcal{L}(q) + \mathrm{KL}(q||p), \tag{20}$$

where $\mathrm{KL}(q||p)$ is the Kullback-Leibler (KL) divergence between $q(\boldsymbol{\Theta})$ and $p(\boldsymbol{\Theta}|\mathbf{X})$, and $\mathcal{L}(q)$ is the lower bound of the log of the model evidence $\ln p(\mathbf{X})$. The distribution $q(\boldsymbol{\Theta})$ is approximated by exploiting the factorization $q(\boldsymbol{\Theta}) = \prod q(\theta_k)$; the posterior $q(\theta_k)$ of each latent variable $\theta_k \in \boldsymbol{\Theta}$ is computed as

$$\ln q(\theta_k) = \mathbb{E}_{\boldsymbol{\Theta}\setminus\theta_k}[\ln p(\mathbf{X}, \boldsymbol{\Theta})] + \text{const}, \tag{21}$$

where $\boldsymbol{\Theta} \setminus \theta_k$ denotes the set $\boldsymbol{\Theta}$ removing θ_k. Eq. (21) means that the posterior distribution of each hidden variable θ_k is obtained by considering the log of the joint distribution in (19) over all observed and latent variables and then taking the expectation $\mathbb{E}[\cdot]$ with respect to all other variables except for θ_k. The update rules for the latent variables are presented in the following.

A. Estimation of wall subspace $\mathbf{W}$

Using (21), the posterior distribution of $\mathbf{W}$ is given by

$$\ln q(\mathbf{W}) = \mathbb{E}_{\mathbf{Z},\mathbf{S},\beta}[\ln p(\mathbf{X}|\mathbf{W}, \mathbf{Z}, \mathbf{S}, \beta)] + \mathbb{E}_{\boldsymbol{\alpha}}[\ln p(\mathbf{W}|\boldsymbol{\alpha})]. \tag{22}$$

Substituting (11) and (12) into (22) and keeping only terms related to $\mathbf{W}$ yields

$$q(\mathbf{W}) = \prod_{m=1}^{M} \mathcal{CN}(\mathbf{w}_{m,:}|\mathbf{m}_w, \mathbf{C}_w), \tag{23}$$

where $\mathbf{A} = \mathrm{diag}(\mathbb{E}[\boldsymbol{\alpha}])$, $\mathbf{w}_{m,:}$ is the mth row of matrix $\mathbf{W}$, and

$$\mathbf{m}_w = \mathbb{E}[\beta]\mathbf{C}_w \sum_{l=1}^{L}\sum_{n=1}^{N} \mathbb{E}[\mathbf{z}_{l,n}](x_{l,n}(m) - \mathbf{D}_n^{m,:}\mathbb{E}[\mathbf{s}_l]), \tag{24}$$

$$\mathbf{C}_w = (\mathbf{A} + \mathbb{E}[\beta]\sum_{l=1}^{L}\sum_{n=1}^{N} \mathbb{E}[\mathbf{z}_{l,n}\mathbf{z}_{l,n}^H])^{-1}. \tag{25}$$

B. Estimation of matrix $\mathbf{Z}$

The posterior distribution of $\mathbf{Z}$ is estimated using (21):

$$\ln q(\mathbf{Z}) = \mathbb{E}_{\mathbf{W},\mathbf{S},\beta}[\ln p(\mathbf{X}|\mathbf{W}, \mathbf{Z}, \mathbf{S}, \beta)] + \mathbb{E}_{\boldsymbol{\alpha}}[\ln p(\mathbf{Z}|\boldsymbol{\alpha})]. \tag{26}$$

Using (11), (13), and (26) and keeping the terms related to $\mathbf{Z}$ yields

$$q(\mathbf{Z}) = \prod_{l=1}^{L}\prod_{n=1}^{N} \mathcal{CN}(\mathbf{z}_{l,n}|\mathbf{m}_z, \mathbf{C}_z), \tag{27}$$

$$\mathbf{m}_z = \mathbb{E}[\beta]\mathbf{C}_z\mathbb{E}[\mathbf{W}]^H(\mathbf{x}_{l,n} - \mathbf{D}_n\mathbb{E}[\mathbf{s}_l]), \tag{28}$$

$$\mathbf{C}_z = (\mathbf{A} + \mathbb{E}[\beta]\mathbb{E}[\mathbf{W}^H\mathbf{W}])^{-1}. \tag{29}$$

C. Estimation of channel image $\mathbf{S}$

Using (21), the posterior distribution of $\mathbf{S}$ is expressed as

$$\ln q(\mathbf{S}) = \mathbb{E}_{\mathbf{W},\mathbf{Z},\beta}[\ln p(\mathbf{X}|\mathbf{W}, \mathbf{Z}, \mathbf{S}, \beta)] + \mathbb{E}_{\boldsymbol{\gamma}}[\ln p(\mathbf{S}|\boldsymbol{\gamma})]. \tag{30}$$

Substituting (11) and (15) into (30), keeping the terms associated with $\mathbf{S}$, and defining $\boldsymbol{\Gamma} = \mathrm{diag}(\mathbb{E}[\boldsymbol{\gamma}])$, we obtain

$$q(\mathbf{S}) = \prod_{l=1}^{L} \mathcal{CN}(\mathbf{s}_l|\mathbf{m}_s, \mathbf{C}_s), \tag{31}$$

$$\mathbf{m}_s = \mathbf{C}_s[\mathbb{E}[\beta]\sum_{n=1}^{N} \mathbf{D}_n^H(\mathbf{x}_{l,n} - \mathbb{E}[\mathbf{W}]\mathbb{E}[\mathbf{z}_{l,n}])], \tag{32}$$

$$\mathbf{C}_s = (\boldsymbol{\Gamma} + \mathbb{E}[\beta]\sum_{n=1}^{N} \mathbf{D}_n^H\mathbf{D}_n)^{-1}. \tag{33}$$

D. Estimation of hyperparameters $\boldsymbol{\alpha}$*,* $\boldsymbol{\gamma}$*, and* β

Similarly, by (21), the posterior distributions of the hyperparameters are approximated as $q(\boldsymbol{\alpha}) = \prod_{j=1}^{J} \mathrm{Gam}(\alpha_j|\hat{a}_\alpha, \hat{b}_\alpha)$, $q(\boldsymbol{\gamma}) = \prod_{q=1}^{Q} \mathrm{Gam}(\gamma_q|\hat{a}_\gamma, \hat{b}_\gamma)$, and $q(\beta) = \mathrm{Gam}(\beta|\hat{a}_\beta, \hat{b}_\beta)$. Since such posteriors follow gamma distributions, the expectations are given by

$$\mathbb{E}[\alpha_j] = \frac{\hat{a}_\alpha}{\hat{b}_\alpha} = \frac{a_\alpha + M + LN}{b_\alpha + \mathbb{E}[\mathbf{w}_j^H\mathbf{w}_j] + \mathbb{E}[\mathbf{z}_{j,:}\mathbf{z}_{j,:}^H]}, \tag{34}$$

$$\mathbb{E}[\gamma_q] = \frac{\hat{a}_\gamma}{\hat{b}_\gamma} = \frac{a_\gamma + 1}{b_\gamma + \sum_{l=1}^{L}\mathbb{E}[s_l(q)]^2}, \tag{35}$$

$$\mathbb{E}[\beta] = \frac{a_\beta + MLN}{b_\beta + \sum_{l=1}^{L}\sum_{n=1}^{N}\|\mathbf{x}_{l,n} - \mathbb{E}[\mathbf{W}]\mathbb{E}[\mathbf{z}_{l,n}] - \mathbf{D}_n\mathbb{E}[\mathbf{s}_l]\|_2^2}. \tag{36}$$

In summary, the proposed inference algorithm starts by estimating the wall subspace $\mathbf{W}$ and the wall signal coefficient $\mathbf{Z}$ using their expectations (24) and (28), followed by learning the multichannel image $\mathbf{S}$ using (32), and finally updating the hyperparameters using (34) for $\boldsymbol{\alpha}$, (35) for $\boldsymbol{\gamma}$, and (36) for β. Note that these update rules depend on the moments of the latent variables only and therefore we just need to store such variables. As the inference scheme maximizes the variational lower bound $\mathcal{L}(q)$, its convergence can be ascertained by tracking $\mathcal{L}(q)$.

4. EXPERIMENTAL RESULTS

A stepped-frequency TWR system was used for data collection. Four polarization datasets ($L = 4$) were collected by placing the radar system in front of a 0.16 m thick wooden wall, at a standoff distance of 1 m. A transceiver is moved along the wall to synthesize a 61-element antenna array with an inter-element spacing of 0.01 m. At each antenna, a set of 401 monochromatic signals, covering a frequency range of [2–4] GHz with a step of 5 MHz, is transmitted to illuminate a scene containing one triangular plate trihedral centered at (-0.5 m, 1.0 m), and a 0.2 m square plate dihedral centered at (1.0 m, 1.0 m), as shown in Fig. 1.

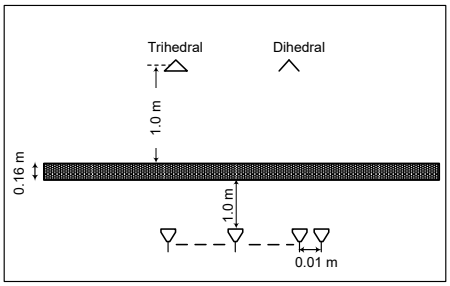

Fig. 1. Layout of the TWRI scene used for data collection.

Wall clutter mitigation and image reconstruction are performed using reduced datasets generated by randomly selecting half of the total frequencies at all the antennas. Using input data matrix $\mathbf{X}$, the Bayesian learning starts by initializing the latent variables, which can be set randomly and yield satisfactory results. However, we can achieve fast convergence with more suitable initial values. Here, the wall subspace factors $\mathbf{W}$ and $\mathbf{Z}$ are initialized using dominant singular components (SC) obtained from the singular value decomposition (SVD) of the input radar signal matrix, $\mathbf{X} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{V}^H$. We select $J = \lfloor \min(M, LN)/2 \rfloor$ dominant SCs and set $\mathbf{W} = \mathbf{U}\,\mathbf{\Lambda}^{1/2}$ and $\mathbf{Z} = \mathbf{\Lambda}^{1/2}\,\mathbf{V}^H$.

Figs. 2(a) and (b) show the co-polarization (HH) and cross-polarization (HV) images, respectively, formed after convergence by the proposed inference algorithm. The targets are well localized and clutter is significantly suppressed. For comparison, Figs. 2(c) and (d) present the multichannel images reconstructed by the multistage CS-based method [20–23] using the same reduced datasets. Here, the signal estimate was performed first, followed by subspace projection [17] for wall clutter mitigation, and ℓ_1 minimization for image formation. Strong wall clutter is removed, but target pixels are not recovered fully, and the formed images are affected by other artifacts. To quantify the improvement in target recovery and clutter mitigation, the target-to-clutter ratio (TCR) is used as the performance measure. Let A_t and A_c be, respectively, the target and clutter regions of the formed image I, and let N_t and N_c denote, respectively, the number of target and clutter pixels. The TCR (in dB) is defined as

$$\mathrm{TCR} = 10\log_{10}\Big(\frac{\frac{1}{N_t}\sum_{q\in A_t}|I_q|^2}{\frac{1}{N_c}\sum_{q\in A_c}|I_q|^2}\Big). \tag{37}$$

The average TCRs of the images formed by the proposed Bayesian and multistage CS-based methods are both computed using (37). The Bayesian method has a TCR of 35.59 dB, much higher than the 24.81 dB obtained by the multistage CS-based method. Fig. 3(a) shows the automatic relevant determination of the rank of the wall clutter subspace $\mathbf{W}$ by the Bayesian inference. The rank starts at 122 ($J = \lfloor \min(401, 4\times 61)/2 \rfloor$), decreases during inference, and reaches 2 at convergence. Fig. 3(b) illustrates the learning convergence, which maximizes the lower bound $\mathcal{L}(q)$; $\mathcal{L}(q)$ increases during the inference, and reaches a steady state after 10 iterations.

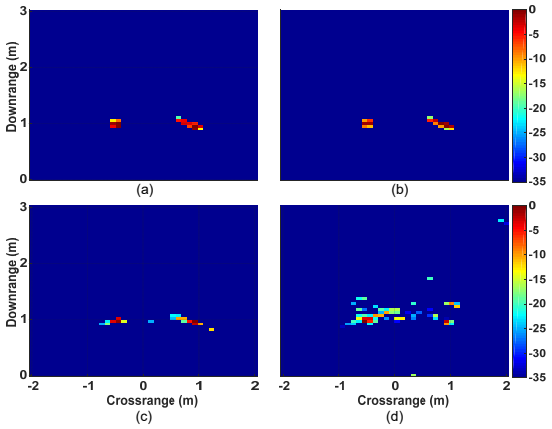

Fig. 2. Multichannel images reconstructed by different methods with 50% data measurements: (a) co-polarization (HH) and (b) cross-polarization (HV) obtained by the proposed Bayesian method, (c) co-polarization (HH) and (d) cross-polarization (HV) obtained by the multistage CS-based method.

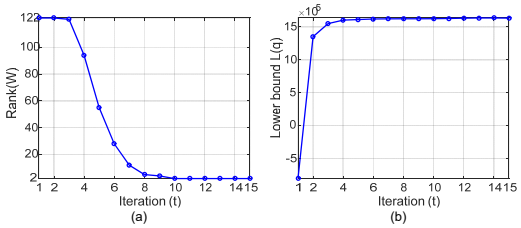

Fig. 3. Relevant determination of the effective rank of wall subspace and convergence of the inference algorithm: (a) the rank value of the wall clutter subspace $\mathbf{W}$ as the function of iterations, (b) the lower bound $\mathcal{L}(q)$ of the log of the marginal probability $\ln p(\mathbf{X})$ as the function of iterations.

5. CONCLUSION

This paper presented a variational Bayesian model for multichannel TWRI. Through experimental validation, it was demonstrated that the proposed model captures wall clutter well and enhances multichannel imaging, even with reduced measurements. Furthermore, by providing latent variables as posterior distributions and learning the wall subspace and target images adaptively from the observed radar measurements, the proposed model is capable of capturing uncertainty and alleviating the issue of hyperparameter tuning.

Acknowledgments

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant 102.01-2017.307. The work of A. Bouzerdoum and S. L. Phung was supported by a grant from the Australian Research Council (ARC).

6. REFERENCES

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

[2] K. M. Yemelyanov, N. Engheta, A. Hoorfar, and J. A. McVay, "Adaptive polarization contrast techniques for through-wall microwave imaging applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1362–1374, May 2009.

[3] C. Debes, A. M. Zoubir, and M. G. Amin, "Enhanced detection using target polarization signatures in through-the-wall radar imaging," *IEEE Trans. Geosci. Remote Sens.,*, vol. 50, no. 5, pp. 1968–1979, May 2012.

[4] A. A. Mostafa, C. Debes, and A. M. Zoubir, "Segmentation by classification for through-the-wall radar imaging using polarization signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3425–3439, Sep. 2012.

[5] C. H. Seng, A. Bouzerdoum, M. G. Amin, and S. L. Phung, "Probabilistic fuzzy image fusion approach for radar through wall sensing," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4938–4951, Dec. 2013.

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

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

[8] Y.-S. Yoon and M. G. Amin, "Compressed sensing technique for high-resolution radar imaging," in *Proc. SPIE: Signal Process., Sensor Fusion, and Target Recognition XVII*, Orlando, FL, Mar. 2008, pp. 69 681A.1– 69 681A.10.

[9] 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.

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

[11] 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," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process.*, pp. 2672–2675, Prague, Czech Republic, May 2011.

[12] V. H. Tang, A. Bouzerdoum, and S. L. Phung, "Two-stage through-the-wall radar image formation using compressive sensing," *J. Electronic Imaging*, vol. 22, no. 2, pp. 021 006.1–021 006.10, Apr.–Jun. 2013.

[13] 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: CRC Press, Aug. 2014, pp. 232–248.

[14] C. Peng, G. Li, and R. J. Burkholder, "Fusion of polarimetric radar images using hybrid matching pursuit," in *IEEE Int. Conf. Acoust., Speech and Signal Process.*, Brisbane QLD, Apr. 2015, pp. 2744–2748.

[15] G. Li and R. J. Burkholder, "Hybrid matching pursuit for distributed through-wall radar imaging," *IEEE Trans. Antennas Propag.*, vol. 63, no. 4, pp. 1701–1711, Apr. 2015.

[16] X. Wang, G. Li, Q. Wan, and R. J. Burkholder, "Look-ahead hybrid matching pursuit for multipolarization through-wall radar imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 4072–4081, Jul. 2017.

[17] 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.

[18] F. Ahmad, "Wall clutter mitigations for compressive imaging of building interiors," in *Compressive Sensing for Urban Radar*. M. G. Amin (Ed.), Boca Raton, FL: CRC Press, Aug. 2015, pp. 123–150.

[19] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "A sparse Bayesian learning approach for through-wall radar imaging of stationary targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 5, pp. 2485–2501, Oct. 2017.

[20] 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.

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

[22] 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.

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

[24] 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 *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, Shanghai China, Mar. 2016, pp. 1412–1416.

[25] 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.

[26] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[27] A. C. Faul and M. E. Tipping, "Analysis of sparse Bayesian learning," in *Proc. Advances in Neural Inf. Process. Syst.* MIT Press, Dec. 2001, pp. 383–389.

[28] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[29] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[30] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Magazine*, vol. 25, no. 6, pp. 131–146, 2008.