Through-The-Wall Radar Imaging Using a Distributed Quasi-Newton Method

Haroon Raja*, Waheed U. Bajwa*, and Fauzia Ahmad†
*Department of Electrical and Computer Engineering, Rutgers University–New Brunswick, Piscataway, NJ
†Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA
E-mail:*{haroon.raja,waheed.bajwa}@rutgers.edu,†fauzia.ahmad@temple.edu

***Abstract*—This paper considers a distributed network of through-the-wall imaging radars and provides a solution for accurate indoor scene reconstruction in the presence of multipath propagation. A sparsity-based method is proposed for eliminating ghost targets under imperfect knowledge of interior wall locations. Instead of aggregating and processing the observations at a central fusion station, joint scene reconstruction and estimation of interior wall locations is carried out in a distributed manner across the network. Using alternating minimization approach, the sparse scene is reconstructed using the recently proposed MDOMP algorithm, while the wall location estimates are obtained with a distributed quasi-Newton method (D-QN) proposed in this paper. The efficacy of the proposed approach is demonstrated using numerical simulation.**

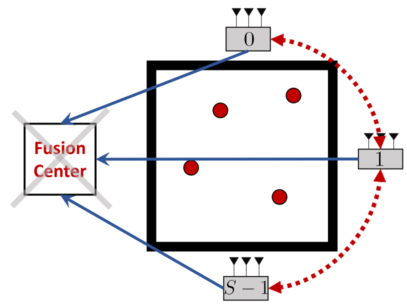

Fig. 1. Distributed TWRI setup where data communication only happens between radar units, shown by dotted red lines. In contrast, [11] requires accumulation of data at a fusion center using links shown by solid blue lines.

I. Introduction

Through-the-wall radar imaging (TWRI) technology has improved significantly over the last decade. However, effectively dealing with the uncertainty caused by high amount of multipath propagation remains a challenge [1], [2]. A number of approaches, both under conventional and sparse reconstruction frameworks, have been recently proposed in the literature to deal with this challenge [3]–[10]. However, these methods require prior knowledge of the exact interior layout of the building being imaged to eliminate ghost targets (accumulation of unwanted energy at incorrect target locations) and provide enhanced image quality. In practice, such information may not be perfectly available in advance, resulting in ghost targets and poor image quality.

The problem of TWRI with uncertainties in the room layout information has been addressed by Leigsnering et al. [11]. In [11], this problem is posed as a parametric dictionary learning problem, where the dictionary to be learned is parametrized by unknown wall locations $\mathbf{w}$. Similar to standard dictionary learning [12], [13], the authors in [11] pose parametric dictionary learning as a nonconvex optimization problem, which is solved using alternating minimization approach involving a dictionary update step and a sparse recovery step (coefficient update). Although the overall objective function considered in traditional dictionary learning is nonconvex, it is convex for each individual step, i.e., when one considers dictionary and coefficient variables separately. Because of this, one can use tools from convex optimization to solve these individual steps. In contrast, parametric dictionary learning in TWRI results in a dictionary update step that is nonconvex and hence needs special attention. One main contribution of [11] is to show that using particle swarm optimization (PSO) and a quasi-Newton method one can recover the dictionary parameter $\mathbf{w}$.

The parametric dictionary learning method proposed in [11] requires data measurements from each individual radar unit to be collected at a centralized location for processing (see Fig. 1). In practical settings, where we have a large-scale network of radar units interrogating a scene, such as a large building, it is not feasible to accumulate data at one centralized location. Furthermore, distributed solutions are more robust to a variety of issues, such as bad communication links, node failures, etc., which can occur in adverse settings where we need to deploy radar units, e.g., military or rescue missions. As such a distributed approach may be preferred over a centralized solution [11] in practical settings. Note that both PSO and quasi-Newton method have been shown to have good performance for solving the parametric dictionary learning problem in TWRI [11]. We proposed a distributed variant of PSO method in our previous work [14]. In this paper, we propose a quasi-Newton based distributed solution for parametric dictionary learning for TWRI. Our motivation for using quasi-Newton method is that heuristic algorithms like PSO lack convergence guarantees in comparison to gradient-based methods, such as gradient descent, quasi-Newton, etc.

In terms of prior work, Eisen et al. [15] have proposed a distributed version of quasi-Newton method for convex optimization problems. That solution is not directly applicable to the parametric dictionary problem here because our objective function is nonconvex and can have gradient with very large value of the Lipschitz continuity constant. Because of these

*This work is supported in part by the ARO under grant W911NF-17-1-0546 and NSF under grant CCF-1453073.

reasons, we need to design a distributed variant of quasi-Newton method that takes into account the specific properties of TWRI objective function. Specifically, since the function gradient can have large Lipschitz continuity constant, we cannot use a gradient-based stopping rule. Instead we propose a reconstruction error-based stopping rule that is provided in Section IV-C. Secondly, we cannot use a constant or decreasing step size, which are common choices in distributed optimization literature [16], [17]; rather, we need to perform exact line search in order for our method to converge.

In the following, we first formally state the problem in Section II. We provide an overview of methods that will be used for solving distributed TWRI problem in Section III. In Section IV, we describe in detail the proposed approach, while supporting numerical results are provided in Section V. Finally, conclusions are drawn in Section VI.

II. Problem Formulation

A. System Model

Consider S radar units, distributed at known positions either along the front wall or surrounding the building being imaged. Each radar unit is equipped with M transmitters and N receivers, where both M and N are assumed to be small. An 'across-units' mode of operation is considered, wherein transmission-reception occurs across multiple radar units. While operating in this mode, each transmitted pulse is received simultaneously by all receivers from all units. It is assumed that the individual radar units can transmit and receive without interference from others and each radar can associate the received signal with a specific transmitter.

Let s_1 and s_2 be the indices of the transmitting and receiving radar units, respectively, where $s_1 = 0, 1, \ldots, S-1$, and $s_2 = 0, 1, \ldots, S-1$. The scene of interest is divided into P grid points, which defines the target space. Let $\sigma_p^{s_1 s_2}$ be the complex reflectivity associated with grid point p corresponding to the transmitting unit s_1 and receiving unit s_2, with $\sigma_p^{s_1 s_2} = 0$ representing the absence of a target. Neglecting multipath contributions, the baseband signal recorded at receiver $n = 0, 1, \ldots, N-1$ of the s_2-th radar unit, with transmitter $m = 0, 1, \ldots, M-1$ of the s_1-th radar unit active, can be expressed as,

$$z_{mn}^{s_1 s_2}(t) = \sum_{p=0}^{P-1} \sigma_p^{s_1 s_2} s(t - mT_r - s_1 M T_r - \tau_{pmn}^{s_1 s_2}) \times \exp\left(-j2\pi f_c (mT_r + s_1 M T_r + \tau_{pmn}^{s_1 s_2})\right). \quad (1)$$

Here, $s(t)$ is the transmitted wideband pulse in complex baseband, f_c is the carrier frequency, T_r is the pulse repetition interval, and $\tau_{pmn}^{s_1 s_2}$ is the propagation delay from transmitter m of unit s_1 to the grid point p and back to the receiver n of unit s_2. We sample $z_{mn}^{s_1 s_2}(t)$ at or above the Nyquist rate to obtain a signal vector $\mathbf{z}_{mn}^{s_1 s_2}$ of length N_T. Stacking signal vectors corresponding to the M transmitters and N receivers, we obtain an $MNN_T \times 1$ measurement vector $\bar{\mathbf{z}}_{s_1 s_2}$, which, using (1), can be expressed as, $\bar{\mathbf{z}}_{s_1 s_2} = \mathbf{\Psi}_{s_1 s_2}^{(0)} \boldsymbol{\sigma}_{s_1 s_2}^{(0)}$, where $\boldsymbol{\sigma}_{s_1 s_2}^{(0)} = [\sigma_0^{s_1 s_2}, \sigma_1^{s_1 s_2}, \ldots, \sigma_{P-1}^{s_1 s_2}]^{\mathrm{T}}$ with the superscript '(0)' indicating direct path propagation, '$[.]^{\mathrm{T}}$' denotes matrix transpose, and, for $i = 0, \ldots, N_T - 1$, the elements of the dictionary matrix $\mathbf{\Psi}_{s_1 s_2}^{(0)} \in \mathbb{C}^{MNN_T \times P}$ are given by

$$\left[\mathbf{\Psi}_{s_1 s_2}^{(0)}\right]_{i+nN_T+mN_TN,p} = s(t_i - mT_r - s_1 M T_r - \tau_{pmn}^{s_1 s_2}) \times \exp\left(-j2\pi f_c (t_i - (mT_r + s_1 M T_r + \tau_{pmn}^{s_1 s_2}))\right). \quad (2)$$

Next, we assume that multipath for each target is generated due to secondary reflections at one or more interior walls. Parameterizing the interior wall locations as $\mathbf{w} \in \mathbb{R}^3$ and employing geometric optics to model $R-1$ additive multipath contributions in the received signal, we obtain the signal model under multipath propagation for the s_1-th transmitting unit and s_2-th receiving unit as

$$\bar{\mathbf{z}}_{s_1 s_2} = \mathbf{\Psi}_{s_1 s_2}^{(0)} \boldsymbol{\sigma}_{s_1 s_2}^{(0)} + \sum_{r=1}^{R-1} \mathbf{\Psi}_{s_1 s_2}^{(r)}(\mathbf{w}) \boldsymbol{\sigma}_{s_1 s_2}^{(r)}, \quad (3)$$

where $\mathbf{\Psi}_{s_1 s_2}^{(r)}$ is defined according to (2) with $\tau_{pmn}^{s_1 s_2}$ replaced by the propagation delay $\tau_{pmn}^{s_1 s_2,(r)}$ between transmitter m, grid point p, and receiver n along the r-th multipath [7]. Note that the multipath time delays $\tau_{mnp}^{s_1 s_2,(r)}, r = 1, \ldots, R-1$, depend on the wall locations and, therefore, the dictionary matrices $\{\mathbf{\Psi}_{s_1 s_2}^{(r)}\}_{r=1}^{R-1}$ are all functions of $\mathbf{w}$. Defining $\widetilde{\mathbf{\Psi}}_{s_1 s_2}(\mathbf{w}) = \left[\mathbf{\Psi}_{s_1 s_2}^{(0)} \cdots \mathbf{\Psi}_{s_1 s_2}^{(R-1)}(\mathbf{w})\right]$ and $\widetilde{\boldsymbol{\sigma}}_{s_1 s_2} = [\boldsymbol{\sigma}_{s_1 s_2}^{(0)\mathrm{T}} \cdots \boldsymbol{\sigma}_{s_1 s_2}^{(R-1)\mathrm{T}}]^{\mathrm{T}}$, and assuming additive noise $\bar{\mathbf{n}}$, we can rewrite

$$\bar{\mathbf{z}}_{s_1 s_2} = \widetilde{\mathbf{\Psi}}_{s_1 s_2}(\mathbf{w}) \widetilde{\boldsymbol{\sigma}}_{s_1 s_2} + \bar{\mathbf{n}}_{s_1 s_2}. \quad (4)$$

B. Centralized Problem Formulation

In the case of centralized processing, the S^2 measurement vectors, $\{\bar{\mathbf{z}}_{s_1 s_2}, s_1 = 0, \ldots, S-1, s_2 = 0, \ldots, S-1\}$, corresponding to the 'across-units' operation of the S radar units, are communicated to a fusion center where the scene reconstruction is performed [11]. Specifically, the S^2 measurements can be collectively represented as

$$\breve{\mathbf{z}} = \breve{\mathbf{A}}(\mathbf{w}) \breve{\boldsymbol{\sigma}} + \breve{\mathbf{n}}, \quad (5)$$

where

$$\breve{\mathbf{z}} = \left[\bar{\mathbf{z}}_{0\,0}^{\mathrm{T}}, \ldots, \bar{\mathbf{z}}_{S-1\,S-1}^{\mathrm{T}}\right]^{\mathrm{T}}, \ \breve{\boldsymbol{\sigma}} = \left[\widetilde{\boldsymbol{\sigma}}_{0\,0}^{\mathrm{T}}, \ldots, \widetilde{\boldsymbol{\sigma}}_{S-1\,S-1}^{\mathrm{T}}\right]^{\mathrm{T}},$$
$$\breve{\mathbf{n}} = \left[\bar{\mathbf{n}}_{0\,0}^{\mathrm{T}}, \ldots, \bar{\mathbf{n}}_{S-1\,S-1}^{\mathrm{T}}\right]^{\mathrm{T}}, \text{ and}$$
$$\breve{\mathbf{A}}(\mathbf{w}) = \text{blkdiag}\{\widetilde{\mathbf{\Psi}}_{0\,0}(\mathbf{w}), \ldots, \widetilde{\mathbf{\Psi}}_{S-1\,S-1}(\mathbf{w})\}, \quad (6)$$

and blkdiag$\{.\}$ denotes a block-diagonal matrix.

Given the measurements $\breve{\mathbf{z}}$ in (5), the aim is to determine the wall locations $\mathbf{w}$ as well as reconstruct the scene reflectivity vector $\breve{\boldsymbol{\sigma}}$. Since the same physical scene is observed via all paths by the various radar units, the scene reflectivity vector exhibits a group sparse structure [11]. As such, for a regularization parameter λ, the scene recovery and wall location estimation can be posed as the following optimization problem:

$$\min_{\breve{\boldsymbol{\sigma}}, \mathbf{w}} \|\breve{\mathbf{z}} - \breve{\mathbf{A}}(\mathbf{w}) \breve{\boldsymbol{\sigma}}\|_2^2 + \lambda \|\breve{\boldsymbol{\sigma}}\|_{1,2}, \quad (7)$$

where $\|\breve{\boldsymbol{\sigma}}\|_{1,2} = \sum_{p=0}^{P-1} \left\|[\boldsymbol{\sigma}_{0\,0_p}^{(0)}, \ldots, \boldsymbol{\sigma}_{0\,0_p}^{(R-1)}, \ldots, \boldsymbol{\sigma}_{S-1\,S-1_p}^{(0)}, \ldots, \boldsymbol{\sigma}_{S-1\,S-1_p}^{(R-1)}]^{\mathrm{T}}\right\|_2$ and $\boldsymbol{\sigma}_{s_1\,s_{2_p}}^{(0)}$ is the p-th element of vector

$\boldsymbol{\sigma}^{(0)}_{s_1 s_2}$. The optimization problem in (7) is nonconvex as the matrix $\breve{\mathbf{A}}(\mathbf{w})$ has a nonlinear dependence on the wall locations. An iterative approach proposed in [11] solves (7) by alternating between optimization over $\breve{\boldsymbol{\sigma}}$ and $\mathbf{w}$.

C. Distributed Problem Formulation

The focus of this work is on wall location estimation and scene reconstruction, i.e., solving (7), in a distributed manner across the S radar units, with each radar unit having access to only a subset of the measurements $\breve{\mathbf{z}}$. Substituting $\breve{\mathbf{A}}(\mathbf{w})$, $\breve{\mathbf{z}}$, and $\breve{\boldsymbol{\sigma}}$ from (6) in (7), we can write the problem as

$$\min_{\breve{\boldsymbol{\sigma}},\mathbf{w}} \sum_{s_2=0}^{S-1}\sum_{s_1=0}^{S-1} \|\bar{\mathbf{z}}_{s_1 s_2} - \tilde{\boldsymbol{\Psi}}_{s_1 s_2}(\mathbf{w})\tilde{\boldsymbol{\sigma}}_{s_1 s_2}\|_2^2 + \lambda\|\breve{\boldsymbol{\sigma}}\|_{1,2}. \quad (8)$$

Using alternating minimization framework, we need to solve this optimization problem in a distributed manner. For distributed optimization over $\breve{\boldsymbol{\sigma}}$, we use the Modified Distributed OMP (MDOMP) method proposed in [18]. Then for a fixed $\breve{\boldsymbol{\sigma}}$, the objective function is just the first term in (8). That is,

$$\min_{\mathbf{w}} f(\mathbf{w}) := \min_{\mathbf{w}} \sum_{s_2=0}^{S-1} f_{s_2}(\mathbf{w}), \quad (9)$$

where,

$$f_{s_2}(\mathbf{w}) := \sum_{s_1=0}^{S-1} \|\bar{\mathbf{z}}_{s_1 s_2} - \tilde{\boldsymbol{\Psi}}_{s_1 s_2}(\mathbf{w})\tilde{\boldsymbol{\sigma}}_{s_1 s_2}\|_2^2$$

is the objective function at radar unit s_2. In this paper, we develop a distributed quasi-Newton method to solve this optimization problem.

III. Technical background

A. Quasi-Newton Method for Optimization

Gradient-based descent direction methods are a popular choice for solving optimization problems [19]. First-order methods like gradient descent are computationally efficient but can result in slow convergence for some problems. To overcome this issue we can use second-order methods such as Newton's method, which for step size γ_t, is given as follows:

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \gamma_t \nabla^2 f(\mathbf{w}_t)^{-1} \nabla f(\mathbf{w}_t). \quad (10)$$

However, computing the Hessian matrix and its inverse for Newton's method can be computationally prohibitive in practice for high-dimensional problems. Quasi-Newton method resolves this issue by using only gradient information to approximate the Hessian matrix. For an approximation V_t of inverse of the Hessian matrix, i.e., $V_t \approx \nabla^2 f(\mathbf{w}_t)^{-1}$, we can rewrite (10) as:

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \gamma_t V_t \nabla f(\mathbf{w}_t). \quad (11)$$

A number of methods exist in the literature to compute matrix V_t, with BFGS [19, Chap. 8] being one of the most widely used. Using vectors

$$\mathbf{p}_t := \mathbf{w}_{t+1} - \mathbf{w}_t \quad \text{and} \quad \mathbf{q}_t := \nabla f(\mathbf{w}_{t+1}) - \nabla f(\mathbf{w}_t),$$

BFGS method computes $\mathbf{V}_t$ as follows:

$$\mathbf{V}_{t+1} = \mathbf{V}_t + \left(1 + \frac{\mathbf{q}_t^{\mathrm{T}}\mathbf{V}_t\mathbf{q}_t}{\mathbf{p}_t^{\mathrm{T}}\mathbf{q}_t}\right)\frac{\mathbf{p}_t\mathbf{p}_t^{\mathrm{T}}}{\mathbf{p}_t^{\mathrm{T}}\mathbf{q}_t} - \frac{\mathbf{p}_t\mathbf{q}_t^{\mathrm{T}}\mathbf{V}_t + \mathbf{V}_t\mathbf{q}_t\mathbf{p}_t^{\mathrm{T}}}{\mathbf{p}_t^{\mathrm{T}}\mathbf{q}_t}.$$

For the distributed TWRI problem, the objective function is distributed across radar units; hence, we cannot use centralized quasi-Newton method. Instead, we will employ consensus averaging [20], [21] to develop a distributed variant of the quasi-Newton method. In the following, we overview consensus averaging before presenting the proposed distributed quasi-Newton method.

B. Consensus Averaging

For some scalar values $\{x_i\}_{i=0}^{S-1}$ that are distributed across S radar units, consensus averaging provides an iterative method for computing the average $(1/S)\sum_{i=0}^{S-1} x_i$. Let us first represent the connectivity among the distributed radar units by a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N} = \{0, 1, \ldots, S-1\}$ denotes the set of nodes (radar units in the underlying application) in a network and $\mathcal{E}$ are the edges defining the interconnection among the nodes, i.e., $(i,i) \in \mathcal{E}$ and $(i,j) \in \mathcal{E}$ when node i can communicate with node j. From graph $\mathcal{G}$, we generate a doubly stochastic matrix $\mathbf{A}$ such that its (i,j)-th entry, $\mathbf{A}_{i,j}$, satisfies the condition $\mathbf{A}_{i,j} = 0,\ \forall (i,j) \notin \mathcal{E}$. Then, starting from $\mathbf{x}^{(0)} = [x_0, \ldots, x_{S-1}]^{\mathrm{T}}$, the update at iteration t_c of consensus averaging is given by

$$\mathbf{x}^{(t_c)} = \mathbf{A}\mathbf{x}^{(t_c-1)}. \quad (12)$$

Previous work on consensus averaging [20], [21] shows that if $\mathbf{A}$ is doubly stochastic then as $t_c \to \infty$, each element of $\mathbf{x}^{(t_c)}$ approaches the mean of the values in $\mathbf{x}^{(0)}$.

C. MDOMP for Distributed Sparse Scene Recovery

MDOMP has been proposed in [18] for sparse scene recovery in the case of a distributed network of TWRI units. MDOMP is a distributed version of the OMP algorithm [22]. At each radar unit, a communication step is performed in each iteration of MDOMP, wherein each radar unit computes a correlation vector using the local measurements only and shares it with all other radar units. Each unit then adds all correlation vectors, selects the index corresponding to the largest element in the correlation vector sum, and updates its set of active indices. The remaining part of the algorithm is similar to OMP and is omitted here for the sake of brevity.

IV. Proposed approach

A. Distributed Quasi-Newton Method for TWRI

Lack of access to the complete objective function in distributed settings does not permit quasi-Newton update as given in (11). Instead, we prospose that starting from same initial value of the minimizer $\mathbf{w}_{\text{init}}$ and Hessian approximate $\mathbf{V}_{\text{init}}$ at each radar unit, we compute the update in (11) locally at each radar unit as follows:

$$\mathbf{w}_{s_2,t+1} = \mathbf{w}_{s_2,t} - \tau_{s_2,t}\mathbf{V}_{s_2,t}\nabla f_{s_2}(\mathbf{w}_{s_2,t}). \quad (13)$$

Algorithm 1: Distributed Quasi-Newton Method (D-QN).

Input: Local data $\{\bar{\mathbf{z}}_{0\,0}, \ldots, \bar{\mathbf{z}}_{S-1\,S-1}\}$, $\breve{\boldsymbol{\sigma}}_i$ computed using MDOMP, constraint set $\mathcal{W}$, and a doubly stochastic matrix $\mathbf{A}$.

Initialize: Randomly pick a starting wall position $\widehat{\mathbf{w}}_{s_2,0} \leftarrow \mathbf{w}_{\text{init}}$ and a positive-definite matrix $\mathbf{V}_{s_2,0} \leftarrow \mathbf{V}_{\text{init}}$ at each radar unit s_2, and $t \leftarrow 0$.

1: **while** $stopping\ rule$ **do**
2: $\widehat{\mathbf{d}}_{s_2,t} \leftarrow -\widehat{\mathbf{V}}_{s_2,t} \nabla f(\widehat{\mathbf{w}}_{s_2,t})$
3: $\gamma_{s_2,t} \leftarrow \arg\min_{\gamma} \sum_{s_2} f_{s_2}\big(\widehat{\mathbf{w}}_{s_2,t} + \gamma \widehat{\mathbf{d}}_{s_2,t}\big)$
4: $\bar{\mathbf{w}}_{s_2,t+1} \leftarrow \widehat{\mathbf{w}}_{s_2,t} + \gamma_{s_2,t} \widehat{\mathbf{d}}_{s_2,t}$
5: **if** $\bar{\mathbf{w}}_{s_2,t+1} \in \mathcal{W}$ **then**
6: $\widehat{\mathbf{w}}_{s_2,t+1} \leftarrow \bar{\mathbf{w}}_{s_2,t+1}$
7: **else**
8: ***(Projection onto constraint set)***
9: $\mathbf{P}(\bar{\mathbf{w}}_{s_2,t+1}) = \arg\min_{\mathbf{y} \in \mathcal{W}} \|\bar{\mathbf{w}}_{s_2,t+1} - \mathbf{y}\|_2$
10: $\widehat{\mathbf{d}}_{s_2,t} \leftarrow \mathbf{P}(\bar{\mathbf{w}}_{s_2,t+1}) - \widehat{\mathbf{w}}_{s_2,t}$
11: $\gamma_{s_2,t} \leftarrow \arg\min_{\gamma<1} \sum_{s_2} f_{s_2}\big(\widehat{\mathbf{w}}_{s_2,t} + \gamma \widehat{\mathbf{d}}_{s_2,t}\big)$
12: $\widehat{\mathbf{w}}_{s_2,t+1} \leftarrow \widehat{\mathbf{w}}_{s_2,t} + \gamma_{s_2,t} \widehat{\mathbf{d}}_{s_2,t}$
13: **end if**
14: ***(Consensus Averaging)***
15: Initialize $t_c \leftarrow 0$ and $\bar{\mathbf{w}}_{s_2,0} \leftarrow \widehat{\mathbf{w}}_{s_2,t+1}$
16: **while** $stopping\ rule$ **do**
17: $\bar{\mathbf{w}}_{s_2,t_c+1} \leftarrow \sum_{j \in \mathcal{N}_{s_2}} \mathbf{A}_{s_2,j} \bar{\mathbf{w}}_{s_2,t_c}$
18: $t_c \leftarrow t_c + 1$
19: **end while**
20: ***(Update*** $\widehat{\mathbf{p}}_{s_2,t}$***,*** $\widehat{\mathbf{q}}_{s_2,t}$***, and*** $\widehat{\mathbf{V}}_{s_2,t}$***):***
21: $\widehat{\mathbf{p}}_{s_2,t} \leftarrow \widehat{\mathbf{w}}_{s_2,t+1} - \widehat{\mathbf{w}}_{s_2,t}$
22: $\widehat{\mathbf{q}}_{s_2,t} \leftarrow \nabla \widehat{f}_{s_2}(\widehat{\mathbf{w}}_{s_2,t+1}) - \nabla \widehat{f}_{s_2}(\widehat{\mathbf{w}}_{s_2,t})$
23: $\widehat{\mathbf{V}}_{s_2,t+1} \leftarrow \widehat{\mathbf{V}}_{s_2,t} + \left(1 + \frac{\widehat{\mathbf{q}}_{s_2,t}^T \widehat{\mathbf{V}}_{s_2,t} \widehat{\mathbf{q}}_{s_2,t}}{\widehat{\mathbf{p}}_{s_2,t}^T \widehat{\mathbf{q}}_{s_2,t}}\right) \frac{\widehat{\mathbf{p}}_{s_2,t} \widehat{\mathbf{p}}_{s_2,t}^T}{\widehat{\mathbf{p}}_{s_2,t}^T \widehat{\mathbf{q}}_{s_2,t}} - \frac{\widehat{\mathbf{p}}_{s_2,t} \widehat{\mathbf{q}}_{s_2,t}^T \widehat{\mathbf{V}}_{s_2,t} + \widehat{\mathbf{V}}_{s_2,t} \widehat{\mathbf{q}}_{s_2,t} \widehat{\mathbf{p}}_{s_2,t}^T}{\widehat{\mathbf{p}}_{s_2,t}^T \widehat{\mathbf{q}}_{s_2,t}}$
24: $t \leftarrow t + 1$
25: **end while**

Return: $\widehat{\mathbf{w}}_{s_2,t}$

Note that since we started from same initial values at each radar unit, we can express the global update as:

$$\begin{aligned} \mathbf{w}_{t+1} &= \sum_{s_2=0}^{S-1} \mathbf{w}_{s_2,t+1} \\ &= \sum_{s_2=0}^{S-1} \Big(\frac{1}{S} \mathbf{w}_{s_2,t} - \tau_{s_2,t} \mathbf{V}_{s_2,t} \nabla f_{s_2}(\mathbf{w}_{s_2,t})\Big). \end{aligned} \tag{14}$$

Now we can use consensus averaging to compute this summation. Note that we will have numerical errors in the summation due to finite number of consensus iterations. This leads to biased estimates of $\mathbf{q}_t$ and $\mathbf{p}_t$, as shown in Steps 21–22 of Algorithm 1. These biases in $\mathbf{q}_t$ and $\mathbf{p}_t$ will result in an error in the estimate of $\mathbf{V}_t$ (Step 23 of Algorithm 1). Defining $\boldsymbol{\epsilon}_{c,s_2}$ as the error in the estimate of $\mathbf{V}_t$ then we can re-write (14) as follows:

$$\begin{aligned} \mathbf{w}_{t+1} &= \sum_{s_2=0}^{S-1} \Big(\frac{1}{S} \mathbf{w}_{s_2,t} - \tau_{s_2,t} \mathbf{V}_{s_2,t} \nabla f_{s_2}(\mathbf{w}_{s_2,t})\Big) \\ &= \sum_{s_2=0}^{S-1} \Big(\frac{1}{S} \mathbf{w}_{s_2,t} - \tau_{s_2,t} (\mathbf{V}_t + \epsilon_{c,s_2}) \nabla f_{s_2}(\mathbf{w}_{s_2,t})\Big) \\ &= \sum_{s_2=0}^{S-1} \Big(\frac{1}{S} \mathbf{w}_{s_2,t} - \tau_{s_2,t} \mathbf{V}_t \nabla f_{s_2}(\mathbf{w}_{s_2,t})\Big) + \boldsymbol{\epsilon}_t. \end{aligned} \tag{15}$$

Thus, if we perform enough consensus iterations such that $\boldsymbol{\epsilon}_t$ stays sufficiently small then we can acheive similar results as the centralized quasi-Newton method.

B. Projection onto Constraint Set

After consensus averaging, we have estimate of wall locations $\bar{\mathbf{w}}_{s_2,t+1}$ at each radar unit s_2. Assuming we know initially the wall locations with an accuracy of $\pm 0.5m$, if $\mathbf{w}^*$ is the initial given wall location then we constrain our estimates to be in an interval $[\mathbf{w}^* - 0.5, \mathbf{w}^* + 0.5]$. We formally define constraint set as follows:

$$\mathcal{W} := \big\{\mathbf{w} \in \mathbb{R}^3 : \mathbf{w}^* - 0.5 \leq \mathbf{w} \leq \mathbf{w}^* + 0.5\big\}. \tag{16}$$

In order to project our estimates onto these box constraints, we use the method proposed in [23]. In Algorithm 1, each radar unit has new estimate of wall locations after Step 4. Next, we test whether the new estimate is within the costraint set defined in (16). If wall location estimates are outside the constraint set $\mathcal{W}$ then we use the method in [23]. The first step involves projection of $\bar{\mathbf{w}}_{s_2,t}$ onto set $\mathcal{W}$ as shown in Step 9 of Algorithm 1, while the second step finds a descent direction within the constraint set that is the difference between the projection we obtained and the previous iterate as shown in Step 10 of Algorithm 1. Finally, we compute the step size in the descent direction given as Step 11 of Algorithm 1.

C. Stopping Criterion

Empirical evidence suggests that the function gradient in our problem changes very rapidly, especially in the neighborhood of stationary points. Due to this, using gradient information as a stopping criterion is not possible. On the other hand, from Fig. 2, we can see that the TWRI objective function has very small value within a small neighborhood of the global minimum as compared to anywhere else within the constraint set. Using this insight, we use reconstruction error as a stopping criterion.

V. Simulation Results

We consider a square room with four walls, each of length 2 m. We deploy $S = 5$ radar units, uniformly distributed over an extent of 2 m in crossrange, at a standoff distance of 1.5 m from the front wall. Each radar unit is equipped with $M = 1$ transmitter and $N = 3$ receivers. We assume that at any given time instant, only one radar unit transmits and all units receive the reflections. For each transmission, we use a Gaussian pulse with 50% relative bandwidth, modulating a sinusoid of carrier frequency $f_c = 2$ GHz. The received signal at each radar unit is sampled at the Nyquist rate and $N_T = 150$ samples are collected over the interval of interest. In addition to the direct signal, we assume two multipath contributions arising from the side walls, i.e., $R = 3$. Multipath returns are assumed to

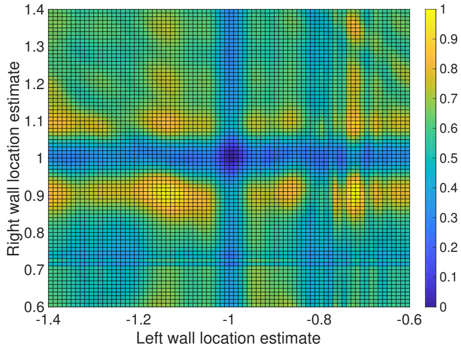

Fig. 2. Normalized objective function (9) when varying wall locations over an interval of ± 0.4 in each dimension around true wall locations.

TABLE I
PERFORMANCE COMPARISON OF DISTRIBUTED GRADIENT DESCENT (D-GD) AND DISTRIBUTED QUASI-NEWTON METHOD (D-QN).

Optimization Method	Estimates within ± 0.1 accuracy	Estimates within ± 0.01 accuracy	Average iteration count
D-GD	94.8%	86.8%	136
D-QN	96.8%	92.6%	110.3

be attenuated by 6 dB as compared to the direct path signal. Further, the received signals are assumed to be corrupted by complex circular Gaussian noise, resulting in a signal-to-noise ratio (SNR) of 20 dB.

The region of interest covers the room interior and is divided into 32×32 pixels in crossrange and downrange. Four point targets are assumed to be located within the room at distinct locations. In the simulation, starting from a random point in interval $\mathbf{w}^* \pm 0.5$, our goal is to estimate the correct wall positions and reconstruct the scene. For comparison with other gradient methods, we also implemented a distributed variant of gradient descent method (D-GD) tailored for the TWRI problem. Instead of using the descent direction given in Step 2 of Algorithm 1, we employ gradient as a descent direction for D-GD. Rest of the algorithm works same as Algorithm 1. The results comparing both these methods are provided in Table I. In the simulation, we are only assuming unkonwn side wall locations and we assume that the true wall location is $[-1 \quad 1]^{\mathrm{T}}$. First and second columns of Table I provide the percentage of times a method estimates the wall location within ± 0.1 and ± 0.01 of the true value, respectively. We can see that D-QN method estimates wall locations with higher accuracy as compared to the D-GD. Our results also show that on average D-QN method requires less number of iterations as compared to the D-GD as well.

VI. CONCLUSION

In this paper, we have proposed a distributed variant of the quasi-Newton method for wall position estimation in TWRI problem. Due to the properties of the objective function, we propose using exact line search for step-size computation and we provide a stopping criterion specifically for the TWRI problem. Finally, simulation results are provided to show the effectiveness of the proposed solution.

REFERENCES

[1] M. G. Amin, Ed., *Through-the-Wall Radar Imaging*. Boca Raton, FL: CRC press, 2011.

[2] M. Amin and F. Ahmad, "Compressive sensing for through-the-wall radar imaging," *J. Electron. Imag.*, vol. 22, no. 3, pp. 030 901–030 922, 2013.

[3] S. Kidera, T. Sakamoto, and T. Sato, "Extended imaging algorithm based on aperture synthesis with double-scattered waves for UWB radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5128–5139, Dec. 2011.

[4] P. Setlur, M. Amin, and F. Ahmad, "Multipath model and exploitation in through-the-wall and urban radar sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 4021–4034, Oct. 2011.

[5] M. Amin, Ed., *Compressive Sensing for Urban Radar*. Boca Raton, FL: CRC Press, 2015.

[6] M. Leigsnering, F. Ahmad, M. G. Amin, and A. M. Zoubir, "Compressive sensing based specular multipath exploitation for through-the-wall radar imaging," in *Proc. IEEE Int. Conference on Acoustics, Speech and Signal Process. (ICASSP)*, 2013, pp. 6004–6008.

[7] 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, 2014.

[8] G. Gennarelli, I. Catapano, and F. Soldovieri, "RF/microwave imaging of sparse targets in urban areas," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 643–646, 2013.

[9] M. Leigsnering, F. Ahmad, M. Amin, and A. Zoubir, "Compressive sensing-based multipath exploitation for stationary and moving indoor target localization," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1469–1483, 2015.

[10] A. AlBeladi and A. Muqaibel, "Compressive sensing based joint wall position detection and multipath exploitation in through-the-wall radar imaging," in *Proc. 9th Int. Symp. Image Signal Process. and Analysis*, 2015, pp. 260–264.

[11] M. Leigsnering, F. Ahmad, M. Amin, and A. Zoubir, "Parametric dictionary learning for sparsity-based TWRI in multipath environments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 2, pp. 532–547, 2016.

[12] M. Aharon, M. Elad, and A. Bruckstein, "k-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[13] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learning Res.*, vol. 11, no. Jan, pp. 19–60, 2010.

[14] H. Raja, W. U. Bajwa, F. Ahmad, and M. G. Amin, "Parametric dictionary learning for TWRI using distributed particle swarm optimization," in *IEEE Radar Conf. (RadarConf)*. IEEE, 2016, pp. 1–5.

[15] M. Eisen, A. Mokhtari, and A. Ribeiro, "Decentralized quasi-Newton methods," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2613–2628, 2017.

[16] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, 2009.

[17] K. Yuan, Q. Ling, and W. Yin, "On the convergence of decentralized gradient descent," *SIAM J. Opt.*, vol. 26, no. 3, pp. 1835–1854, 2016.

[18] M. Stiefel, M. Leigsnering, A. Zoubir, F. Ahmad, and M. Amin, "Distributed greedy signal recovery for through-the-wall radar imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 10, pp. 1477–1481, 2016.

[19] J. Nocedal and S. J. Wright, *Sequential quadratic programming*. Springer, 2006.

[20] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[21] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM J. Control and Opt.*, vol. 48, no. 1, pp. 33–55, 2009.

[22] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[23] W. W. Hager and H. Zhang, "A new active set algorithm for box constrained optimization," *SIAM J. Opt.*, vol. 17, no. 2, pp. 526–557, 2006.