Compressed Sensing-based Multiband Sensing in FR3

Dexin Wang*† and Isha Bhadreshkumar Jariwala†

*Engineering Division, New York University (NYU) Abu Dhabi, Abu Dhabi, UAE.
†NYU WIRELESS, NYU Tandon School of Engineering, New York, USA

Abstract—The following project report describes the problem formulation and algorithm for applying compressed sensing (CS) to perform multiband sensing using QAM-modulated communication symbols in the uplink. The goal includes joint detection and localization of the Tx user and the scatterers, as well as gains estimation which contains the radar cross section (RCS) information for the scatterers. The simulations are performed in FR3. Compared to beamforming, the results show that...

Index Terms—compressed sensing, multiband sensing, joint angle-delay estimation, localization, ISAC, FR3, 6G, ADMM

I. INTRODUCTION

Compressed sensing (CS) is a powerful technique. It is known to be capable of reducing noise [1] [2] [3], hence false detection of targets. It also enables sub-Nyquist sampling given that the signal is sparse in some basis functions space. Generally, given a signal with N_s number of samples but with only N_b elements in the basis functions space, the minimum number of measurements required follows $N_{\min} = O(N_b \log N_s)$ [4].

In this project, we wish to perform multiband sensing using CS. As we normally expect location parameters to not change between subbands, works have shown that in this case the measurement burden can be distributed [5] [6], further reducing the resource needed such as hardware footprint and processing power. We consider an integrated sensing and communication (ISAC) scenario [7], which focuses on correctly estimating the locations and path gains while using QAM-modulated communication pilot symbols (specifically in FR3) in the uplink. From a communication-centric ISAC point of view, using CS can mean that to maintain the sensing performance, less sacrifice in grid resources is needed from the communication side.

Notations: $\otimes$ is the Kronecker product and $\odot$ is the element-wise product. Moreover, $\mathrm{tr}(\cdot)$ is the trace of matrix, and $\mathrm{diag}(\cdot)$ is the diagonal matrix formed by the elements. Additionally, $\mathrm{vec}(\cdot)$ is reshaping a matrix into a column vector, and $\mathrm{mat}(\cdot)$ is reshaping a column vector into a matrix. Also, $\{a_n\}_{n=1}^N$ is the set $\{a_1, a_2, \ldots, a_N\}$, $[\boldsymbol{a}_n]_{n=1}^N$ is the vertically stacked vector $[\boldsymbol{a}_1^T, \boldsymbol{a}_2^T, \ldots, \boldsymbol{a}_N^T]^T$, and $\mathrm{cat}_n\{\cdot\}$ is concatenating tensors in the nth way.

II. SYSTEM MODEL

A. Multiband Channel and Signal Model

We consider a bi-static radar scenario on K subbands with one uplink Tx communication user, one Rx dual-function radar communication (DFRC) base station, and $N-1$ scatterers.
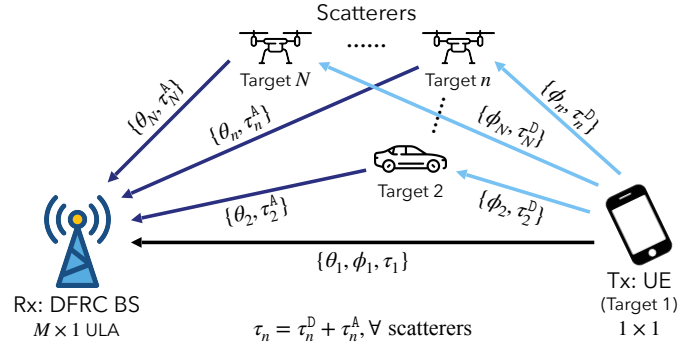

Fig. 1. Bi-static radar scenario with N targets, including one uplink Tx communication user and $N-1$ scatterers.

This is equivalent to a system with N targets, where *target* encompasses both the user and the scatterers. There are M uniform linear array (ULA) Rx antennas and Q subcarriers per subband. We need to know the angles of arrival (AoAs) and times of arrival (ToAs) to localize the targets: since the Tx user is associated with the line-of-sight (LoS) path, its location can be easily determined from these parameters of the path with the lowest ToA; and once the Tx location is known, the rest of the AoAs and the two-way ToAs are sufficient for knowing the scatterer locations [8]. We also want to know the path gains, which contains the radar cross section (RCS) information for the scatterers. The scenario is illustrated in Fig. 1.

For each subband, we consider the single-input, multiple-output (SIMO) channel described below

$$\boldsymbol{H}_k = \sum_{n=1}^{N} g_{n,k} \left(\boldsymbol{a}_k^{\mathrm{R}}(\theta_n) \right) \left(\boldsymbol{a}_k^{\mathrm{F}}(\tau_n) \right)^T \tag{1}$$

where k is the subband index, n is the target index, θ_n is the AoA, and τ_n is the ToA. The steering vectors are given by

$$\boldsymbol{a}_k^{\mathrm{R}}(\theta) = [\exp\left(-j2\pi f_k (d/c) m \sin\theta\right)]_{m=0}^{M-1} \tag{2}$$

$$\boldsymbol{a}_k^{\mathrm{F}}(\tau) = [\exp\left(-j2\pi (f_k + q\Delta f_k)\tau\right)]_{q=0}^{Q-1} \tag{3}$$

where m is the Rx antenna index, q is the subcarrier index, f_k is the subband frequency and Δf_k is the subcarrier spacing. The gains are modeled by the free-space pathloss equation

$$g_{n,k} = \xi_{n,k} \sqrt{\frac{P^{\mathrm{T}} \lambda_k^2 u^{\mathrm{T}}(\phi_n) u^{\mathrm{R}}(\theta_n)}{(4\pi)^2 (c\tau_{n,k})^{\beta_k}}} \tag{4}$$

where λ_k is the wavelength for f_k, $\xi_{n,k}$ is the frequency dependent complex path coefficient, ϕ_n is the target angle of departure (AoD), P^{T} is the average transmit power, $u^{\text{T}}(\phi)$ and $u^{\text{R}}(\theta)$ are the antenna directivities, and β_k is the distance pathloss exponent for subband k. Here, effects such as shadow fading can be absorbed into $\xi_{n,k}$. Note that the gains for the scatterers should in fact be modeled more appropriately with the two-way radar equation, and although $g_{n,k}$ is generated via this model, we ignore the dependence of $g_{n,k}$ on the location parameters and estimate them as they are.

Since the same stream of pilot symbols is received across antennas, we can write the received signal at each subband as

$$\boldsymbol{Y}_k = \boldsymbol{H}_k \operatorname{diag}(\boldsymbol{s}_k) + \boldsymbol{W}_k \qquad (5)$$

where $\boldsymbol{s}_k$ are the transmitted orthogonal frequency-division multiplexing (OFDM) symbols across the subcarriers and the noise follows $\operatorname{vec}(\boldsymbol{W}_k) \sim \mathcal{CN}(\boldsymbol{0}, \sigma_k^2 \boldsymbol{I})$, where $\sigma_k^2 = \sigma^2 B_k / B$, $B_k = Q \Delta f_k$ is the bandwidth of subband k, B is the total bandwidth, and σ^2 is the total noise power. Note that we do not consider the effect of dense multipath components (DMCs).

B. Problem Formulation

Note that $\boldsymbol{Y}_k$ can be rewritten in vector form as

$$\operatorname{vec}(\boldsymbol{Y}_k) = \sum_{n=1}^{N} g_{n,k} \boldsymbol{s}_k \odot \boldsymbol{a}_k^{\text{F}}(\theta_n) \otimes \boldsymbol{a}_k^{\text{R}}(\tau_n) + \operatorname{vec}(\boldsymbol{W}_k) \quad (6)$$

We can think of the summation term due to the paths as a linear combination of the basis vectors

$$\mathcal{B} = \left\{ \boldsymbol{s}_k \odot \boldsymbol{a}_k^{\text{F}}(\theta_n) \otimes \boldsymbol{a}_k^{\text{R}}(\tau_n) \right\}_{n=1}^{N} \qquad (7)$$

each corresponding to a ToA-AoA pair. In order to resolve the paths via CS, sparsity should be established in $\mathcal{B}$. We can do so by creating a "fat" dictionary matrix $\boldsymbol{A}_k = \boldsymbol{A}_k^{\text{F}} \otimes \boldsymbol{A}_k^{\text{R}}$ using very fine delay and angle grids to represent the basis vectors, which can be expressed by

$$\boldsymbol{A}_k^{\text{F}} = \operatorname{cat}_2 \left\{ \boldsymbol{s}_k \odot \boldsymbol{a}_k^{\text{F}}(\tau) \right\}_{\tau \in \mathcal{T}} \in \mathbb{C}^{Q \times L^{\text{F}}} \qquad (8)$$

$$\boldsymbol{A}_k^{\text{R}} = \operatorname{cat}_2 \left\{ \boldsymbol{a}_k^{\text{R}}(\theta) \right\}_{\theta \in \Theta} \in \mathbb{C}^{M \times L^{\text{R}}} \qquad (9)$$

and the grids are given by

$$\mathcal{T} = \left\{ \tau \,|\, 0 \leq i \cdot \tau_{\max} / (L^{\text{F}} - 1) \leq \tau_{\max}, i \in \mathbb{Z} \right\} \qquad (10)$$

$$\Theta = \left\{ \theta \,|\, -\pi/2 \leq i \cdot \pi / (L^{\text{R}} - 1) \leq \pi/2, i \in \mathbb{Z} \right\} \qquad (11)$$

Now, we can express $\boldsymbol{Y}_k$ in the following form:

$$\operatorname{vec}(\boldsymbol{Y}_k) = \boldsymbol{A}_k \boldsymbol{x}_k + \operatorname{vec}(\boldsymbol{W}_k) \qquad (12)$$

where $\boldsymbol{x}_k$ is called the *sparse sensing vector*. We can think of $\boldsymbol{x}_k$ as "activating" the corresponding *sparse* columns in the fine matrix $\boldsymbol{A}_k$, each corresponding to a delay-angle pair. We can further exploit the Kronecker structure to speed up computations by rewriting (12) in the following form

$$\boldsymbol{Y}_k = \boldsymbol{A}_k^{\text{R}} \operatorname{mat}(\boldsymbol{x}_k) \boldsymbol{A}_k^{\text{F}^T} + \boldsymbol{W}_k \qquad (13)$$

In order to localize the targets, we can notice the following. Since we expect the ToAs and AoAs to be the same across

subbands, each $\boldsymbol{x}_k$ should "activate" the same columns of the dictionary and hence have the same positions with non-zero elements. To get the sparsest solution for each band, an ℓ_1-norm objective is typically proposed [9]. However, since the channel behavior varies across subbands, the exact values of those elements will be different. To jointly optimize for sparsity within each subband while preserving the same support locations across subbands, we propose the following optimization problem for localizing targets.

$$(\mathcal{P}_{\text{CS}}) : \begin{cases} \min_{\boldsymbol{X}} & \|\boldsymbol{X}\|_{2,1} \\ \text{s.t.} & \left\| \boldsymbol{Y}_k - \boldsymbol{A}_k^{\text{R}} \operatorname{mat}(\boldsymbol{x}_k) \boldsymbol{A}_k^{\text{F}^T} \right\|_F \leq \varepsilon_k, \\ & \forall k = 1, 2, \dots, K \end{cases} \qquad (14)$$

where $\boldsymbol{X} = \operatorname{cat}_2 \{\boldsymbol{x}_k\}_{k=1}^{K}$ is the multiband sparse sensing matrix, $\varepsilon_k = \varepsilon \sqrt{B_k / B}$ are the denoise thresholds, ε can be related to the actual noise via the quantity $\sigma \sqrt{MQ}$, and

$$\|\boldsymbol{X}\|_{2,1} = \sum_{\ell=1}^{L^{\text{F}} L^{\text{R}}} \sqrt{\sum_{k=1}^{K} |\boldsymbol{X}_{\ell,k}|^2} \qquad (15)$$

is called the $\ell_{2,1}$-mixed norm of matrix $\boldsymbol{X}$.

III. ESTIMATION ALGORITHM

A. Target Localization

To solve the target localization problem posed in (14), we can use alternating direction method of multipliers (ADMM). We describe the methodology used in each slot t. Following the structure in [10], we can introduce auxiliary variables $\boldsymbol{Z}_{1\dots K}$ and rewrite the problem in the form

$$(\mathcal{P}_{\text{ADMM}}) : \begin{cases} \min_{\boldsymbol{X}, \boldsymbol{Z}_{1\dots K}} & \|\boldsymbol{X}\|_{2,1} \\ \text{s.t.} & \boldsymbol{Z}_k \in \mathcal{C}_k, \\ & \boldsymbol{Z}_k = \boldsymbol{A}_k^{\text{R}} \operatorname{mat}(\boldsymbol{x}_k) \boldsymbol{A}_k^{\text{F}^T}, \\ & \forall k = 1, 2, \dots, K \end{cases} \qquad (16)$$

where, as a common approach, we treat each inequality constraint $\|\boldsymbol{Y}_k - \boldsymbol{Z}_k\|_F \leq \varepsilon_k$ as defining the convex feasible set $\mathcal{C}_k$ of $\boldsymbol{Z}_k$. Hence, the augmented Lagrangian is given by

$$\mathcal{L}_\rho(\boldsymbol{X}, \boldsymbol{Z}_{1\dots K}, \boldsymbol{V}_{1\dots K}) = \|\boldsymbol{X}\|_{2,1} + \sum_{k=1}^{K} \mathcal{I}_k(\boldsymbol{Z}_k)$$

$$+ \sum_{k=1}^{K} \operatorname{tr}(\boldsymbol{U}_k^T (\boldsymbol{A}_k^{\text{R}} \operatorname{mat}(\boldsymbol{x}_k) \boldsymbol{A}_k^{\text{F}^T} - \boldsymbol{Z}_k))$$

$$+ \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \boldsymbol{A}_k^{\text{R}} \operatorname{mat}(\boldsymbol{x}_k) \boldsymbol{A}_k^{\text{F}^T} - \boldsymbol{Z}_k \right\|_F^2 \qquad (17)$$

where ρ_k is the per-subband ADMM penalty factor, $\boldsymbol{V}_{1\dots K}$ contains the Lagrangian dual variables, and each feasible set $\mathcal{C}_k$ can be enforced by this indicator function [10] [11]

$$\mathcal{I}_k(\boldsymbol{Z}_k) = \begin{cases} 0 & \text{if } \|\boldsymbol{Y}_k - \boldsymbol{Z}_k\|_F \leq \varepsilon_k \\ \infty & \text{otherwise} \end{cases} \qquad (18)$$

We then explain our variable update policies. First we update $\boldsymbol{X}$ while freezing $\boldsymbol{Z}_{1\ldots K}$ and $\boldsymbol{V}_{1\ldots K}$. From [10], in this step, we want to find $\boldsymbol{X}$ that minimizes the augmented Lagrangian. The part of the augmented Largrangian that we care about is only the first, third and fourth terms. Since the $\ell_{2,1}$-mixed norm term is not smooth while the others are, we can use the proximal gradient descent method [11]. We start by computing the gradient for the smooth terms.

$$\nabla_{\boldsymbol{x}_k}\mathcal{L}_{\rho}^{\text{smth}} = \boldsymbol{A}_k^{\text{R}\,H}\boldsymbol{V}_k\overline{\boldsymbol{A}_k^{\text{F}}} + \rho_k\boldsymbol{A}_k^{\text{R}\,H}(\boldsymbol{A}_k^{\text{R}}\operatorname{mat}(\boldsymbol{x}_k)\boldsymbol{A}_k^{\text{F}\,T} - \boldsymbol{Z}_k)\overline{\boldsymbol{A}_k^{\text{F}}} \tag{19}$$

To speed up computations, we can scale the dual variables. Let $\boldsymbol{U}_k = \boldsymbol{V}_k/\rho_k$, we can rewrite the gradient as

$$\nabla_{\boldsymbol{x}_k}\mathcal{L}_{\rho}^{\text{smth}} = \rho_k\boldsymbol{A}_k^{\text{R}\,H}(\boldsymbol{A}_k^{\text{R}}\operatorname{mat}(\boldsymbol{x}_k)\boldsymbol{A}_k^{\text{F}\,T} - \boldsymbol{Z}_k + \boldsymbol{U}_k)\overline{\boldsymbol{A}_k^{\text{F}}} \tag{20}$$

Denoting $\nabla_{\boldsymbol{X}}\mathcal{L}_{\rho}^{\text{smth}} = \operatorname{cat}_2\left\{\nabla_{\boldsymbol{x}_k}\mathcal{L}_{\rho}^{\text{smth}}\right\}_{k=1}^{K}$, we update the target variable by applying a proximal operator of the objective function onto the smooth gradient descent result

$$\boldsymbol{X}^{(i+1)} = \operatorname{prox}_{\gamma\|\cdot\|_{2,1}}\left(\boldsymbol{X}^{(i)} - \gamma\nabla_{\boldsymbol{X}}\mathcal{L}_{\rho}^{\text{smth}}\right) \tag{21}$$

In our case, we use group soft thresholding, which is a common proximal operator for the $\ell_{2,1}$-mixed norm

$$\operatorname{prox}_{\gamma\|\cdot\|_{2,1}}(\boldsymbol{C}) = \left[\max\left(0, 1 - \frac{\gamma}{\|\boldsymbol{C}_{\ell}\|_2}\right)\boldsymbol{C}_{\ell}\right]_{\ell=1}^{N_C} \tag{22}$$

where $\boldsymbol{C}_{\ell}$ is the ℓth row of some matrix $\boldsymbol{C}$ with N_C rows. Since the $\nabla_{\boldsymbol{x}_k}\mathcal{L}_{\rho}'$ has the form $\rho_k\left(\boldsymbol{A}_k^H\boldsymbol{A}_k\right)\boldsymbol{x}_k + \boldsymbol{b}$, the step size is conservatively determined from the Lipschitz's constraint

$$\gamma = \left(\max_k\left(\rho_k\left\|\boldsymbol{A}_k^{\text{F}}\right\|_2^2\left\|\boldsymbol{A}_k^{\text{R}}\right\|_2^2\right)\right)^{-1} \tag{23}$$

Next we update $\boldsymbol{Z}_{1\ldots K}$. From [10], in this step, we apply the proximal operator as follows.

$$\boldsymbol{Z}_k^{(i+1)} = \operatorname{prox}_{\mathcal{I}_k(\cdot)}\left(\boldsymbol{A}_k^{\text{R}}\operatorname{mat}\left(\boldsymbol{x}_k^{(i+1)}\right)\boldsymbol{A}_k^{\text{F}\,T} + \boldsymbol{U}_k^{(i)}\right), \forall k \tag{24}$$

From [11], since we converted the inequality constraints using an indicator function, the proximal operator reduces to a Euclidean norm ball projection. In our case, this is a Frobenius norm ball projection centered at $\boldsymbol{Y}_k$ with radius of ε_k.

$$\operatorname{prox}_{\mathcal{I}_k(\cdot)}(\boldsymbol{C}) = \begin{cases} \boldsymbol{Y}_k + \varepsilon_k \cdot \frac{\boldsymbol{C} - \boldsymbol{Y}_k}{\|\boldsymbol{C} - \boldsymbol{Y}_k\|_F}, & \text{if } \|\boldsymbol{C} - \boldsymbol{Y}_k\|_F \leq \varepsilon_k \\ \boldsymbol{C}, & \text{otherwise} \end{cases} \tag{25}$$

After that, we update $\boldsymbol{U}_{1\ldots K}$ through the step described in [10]. Note that ρ_k-scaling is omitted due to the scaled dual variables.

$$\boldsymbol{U}_k^{(i+1)} = \boldsymbol{U}_k^{(i)} + \boldsymbol{A}_k^{\text{R}}\operatorname{mat}\left(\boldsymbol{x}_k^{(i+1)}\right)\boldsymbol{A}_k^{\text{F}\,T} - \boldsymbol{Z}_k^{(i+1)}, \forall k \tag{26}$$

Finally, we combine the contributions from all subbands through arithmetic averaging. The output profile is given by

$$X(\theta, \tau) = \left[\frac{1}{K}\sum_{k=1}^{K}\operatorname{mat}(|\boldsymbol{x}_k^*|)\right]_{i,j}, \forall\{\theta_i, \tau_j\} \in \Theta \times \mathcal{T}. \tag{27}$$

The support locations and the number of them are jointly estimated through performing peakfinding with threshold on the normalized output profile.

$$\{\hat{\theta}_n, \hat{\tau}_n\} \in \arg\max_{\theta, \tau}\left(\max\left(r_x, \frac{X(\theta, \tau)}{\max X(\theta, \tau)}\right)\right) \tag{28}$$

where $\boldsymbol{x}_k^*$ is the optimal objective $\boldsymbol{x}_k$ and r_x is the peakfinding threshold. We denote the number of estimated peaks by $\hat{N}$.

B. Gains Estimation

Since our objective function only cares about regularizing the support in $\boldsymbol{X}$ instead of optimizing for similarity with the received signal, the magnitudes of the non-zero elements in $\boldsymbol{X}$ may not be accurately representing the gains of the paths. However, after we know the support locations, we can recover the gains easily using the least squares solution of channel estimation. Also, this separated approach of estimating gains (instead of jointly with angles / delays) enables our adaptive ε scheme discussed in the next subsection.

Specifically speaking, for each subband we first construct the reduced dictionary

$$\boldsymbol{D}_k = \operatorname{cat}_2\left\{\boldsymbol{s}_k \odot \boldsymbol{a}_k^{\text{F}}(\hat{\theta}_n) \otimes \boldsymbol{a}_k^{\text{R}}(\hat{\tau}_n)\right\}_{n=1}^{\hat{N}} \tag{29}$$

Then, the gains for that subband can be recovered through the least squares solution by

$$[\hat{g}_{n,k}]_{n=1}^{\hat{N}} = \left(\boldsymbol{D}_k^H\boldsymbol{D}_k\right)^{-1}\boldsymbol{D}_k^H\operatorname{vec}(\boldsymbol{Y}_k) \tag{30}$$

We can also reconstruct the channel and estimate the noise variance by

$$\hat{\boldsymbol{H}}_k = \sum_{n=1}^{\hat{N}}\hat{g}_{n,k}\left(\boldsymbol{a}_k^{\text{R}}(\hat{\theta}_n)\right)\left(\boldsymbol{a}_k^{\text{F}}(\hat{\tau}_n)\right)^T \tag{31}$$

$$\hat{\sigma}^2 = \frac{1}{MQ}\sum_{k=1}^{K}\left\|\boldsymbol{Y}_k - \hat{\boldsymbol{H}}_k\operatorname{diag}(\boldsymbol{s}_k)\right\|_F^2 \tag{32}$$

C. Adaptive Learning Parameters

In reality, the pilot symbols typically occupy multiple time slots t. In our scenario, we assume the channel to be flat during these T slots. In order to leverage the efficiency of CS and ensure real-time responses, we would prefer that for each slot, the receiver can provide an estimated channel using the methodology provided in the previous two subsections while gradually adapt to the real channel parameters over time by choosing an appropriate ε. Also, for each slot, there are many iterations for the ADMM update, so we wish to adaptively adjust the learning rate parameters ρ_k and automatically break the loop if we know the algorithm has converged.

For evert slot t, if ε is too low compared to the actual $\sigma\sqrt{MQ}$, we would expect the algorithm to overfit to the noise in the received signal, producing many false peaks, leading to error induced by model variance. In this case, at the start of each slot t we can calculate

$$\hat{\sigma}_{\text{cr}}^2(t) = \frac{1}{MQ}\sum_{k=1}^{K}\left\|\boldsymbol{Y}_k^{(t)} - \hat{\boldsymbol{H}}_k^{(t-1)}\operatorname{diag}\left(\boldsymbol{s}_k^{(t)}\right)\right\|_F^2 \tag{33}$$

If this value is higher than the estimated noise variance $\hat{\sigma}^2$ from the last slot $t-1$, we should increase ε. On the other hand, if ε is too high, we would expect that few targets would be detected, leading to under-modeling during the gains estimation phase and error induced by bias. This usually means too little variance. Hence, in summary

$$\varepsilon^{(t+1)} = \begin{cases} 1.5\hat{\sigma}^{(t)}\sqrt{MQ}, & \text{if } \hat{\sigma}_{\text{cr}}^2(t+1) - \hat{\sigma}^{2^{(t)}} \geq r_\varepsilon^{\text{v}}\hat{\sigma}^{2^{(t)}} \\ 0.9\hat{\sigma}^{(t)}\sqrt{MQ}, & \text{if } \left|\hat{\sigma}_{\text{cr}}^2(t+1) - \hat{\sigma}^{2^{(t)}}\right| \leq r_\varepsilon^{\text{b}}\hat{\sigma}^{2^{(t)}} \\ 1.2\hat{\sigma}^{(t)}\sqrt{MQ}, & \text{otherwise} \end{cases}$$

(34)

Auto-Convergence / Residuals / Tolerance

The adaptive ρ_k scheme is also from [10], it balances primal and dual residuals.

D. Summary

A summary of the entire algorithm is given in Algorithm 1.

Algorithm 1: Multiband CS Adaptive ADMM

Input: $\{Y_k^{(t)}\}_{k=1}^K, \forall t$
Output: $\{\hat{\theta}_n, \hat{\tau}_n\}_{n=1}^{\hat{N}}, \{\hat{g}_{n,k}\}_{n,k=1}^{\hat{N},K}$
Compute $A_k^{\text{R}}, \forall k$ by (9)
for $t = 1$ **to** T **do**
    Compute $A_k^{\text{F}}, \forall k$ by (8)
    Initialize $X \leftarrow 0, Z_k \leftarrow 0, U_k \leftarrow 0, \forall k$
    Initialize γ by (23)
    if $t = 1$ **then**
        Initialize $\hat{\sigma}^2 \leftarrow 0, \varepsilon \leftarrow \varepsilon_0, \rho_k \leftarrow \rho_0, \forall k$
    else
        Compute $\hat{\sigma}_{\text{cr}}^2$ by (33)
        Update ε by (34)
    for $i = 1$ **to** max_iter **do**
        Compute $\nabla_{x_k}\mathcal{L}_\rho^{\text{smth}}, \forall k$ by (20)
        $\nabla_X\mathcal{L}_\rho^{\text{smth}} \leftarrow \text{cat}_2\left\{\nabla_{x_k}\mathcal{L}_\rho^{\text{smth}}\right\}_{k=1}^K$
        Update X by (21) (22)
        for $k = 1$ **to** K **do**
            Update Z_k by (24) (25)
            Update U_k by (26)
            Compute p_res by
            Compute d_res by
        Update ρ_k by every 10 iterations
        Update γ by (23) every 10 iterations
        Compute ϵ^{prim} by
        Compute ϵ^{dual} by
        if p_res $< \epsilon^{\text{prim}}$ **and** d_res $< \epsilon^{\text{dual}}$ **then**
            break
    Estimate $\{\hat{\theta}_n, \hat{\tau}_n\}, \forall n$ and $\hat{N}$ by (27) (28)
    for $k = 1$ **to** K **do**
        Estimate $[\hat{g}_{n,k}]_{n=1}^{\hat{N}}$ by (29) (30)
        Update $\hat{\sigma}^2$ by (31) (32)
Return $\{\hat{\theta}_n, \hat{\tau}_n\}_{n=1}^{\hat{N}}, \{\hat{g}_{n,k}\}_{n,k=1}^{\hat{N},K}$

IV. SIMULATION RESULTS

A. Comparison Benchmark

A beamforming (steering vector)-based approach is used in [12]. For fair comparison, we corrected a typo in the paper (not including s in the frequency steering vector), added accomodation of multiband by averaging the power profile across bands for fair comparison, and changed the goal from localizing the user to detecting targets by adding a peakfinding similar to (28) for the output power profile $P(\theta, \tau)$.

B. Metrics

One of the metrics used is the Hungarian-matched root mean squared error (H-RMSE), where we only consider errors of the true and estimated points pairs that are optimally matched without replacement using the Hungarian algorithm. We evaluate this metric separately for all parameters, and we consider the percentage H-RMSE for $\Re g_{n,k}$ and $\Im g_{n,k}, \forall n, k$.

Another metric that we consider is the successful recovery probability (SRP), which focuses only on localization performance. It is defined as the fraction of all Monte Carlo simulations that has $\hat{N} = N$, $|\hat{\theta}_n - \theta_n| \leq 1°$, and $|\hat{\tau}_n - \tau_n| \leq 1$ ns satisfied simultaneously.

C. Simulation Implementation

When generating the parameters, we assume $u^{\text{T}}(\phi) = u^{\text{R}}(\theta) = 1$ and $\beta_k = 2, \forall k$ for simplicity. A table of all parameters is provided in Table I, where N_0 is the thermal noise power spectral density and NF is the noise figure. Currently for simplicity, we assume $T = 1$ and the estimator knows $\varepsilon_0 = 1.2\sigma\sqrt{MQ}$. The test variables are P^{T}, N, and the separation between two subbands, known as band aperture.

D. Results and Discussion

From Fig. 2 and Fig. 3, we can easily see that under the same system parameters, our approach produces much sharper and less noisy peaks then the benchmark, which implies improved detection and resolution. However, keep in mind that the system has theoretical bounds in delay and angular resolutions no matter which method is used.

V. CONCLUSION AND FUTURE WORK

A. Conclusion

B. Future Work

Due to the angular ambiguity brought by the grating lobes effect, we have not considered the realistic case where we have the same antenna spacing for all subbands. A primary thought would be to first get a coarse localization by scanning the full angular range using the lowest subband, and then use all subbands to finely cover all those suspected with the minimal number of steerable ranges.

Another possible extension is to consider a more realistic wideband channel model with background [13] or even per-target [14] DMCs which can be subband frequency dependent. DMCs, especially per-target ones, can give information about surface geometry or even roughness. A common way to model

TABLE I
SIMULATION PARAMETERS

System Parameters	
N	1 to 20
M	60
Q	1000
T	1
f_k	7 to 24 GHz (FR3)
Δf_k	200 kHz
N_0	-174 dBm/Hz
NF	7 dB
P^{T}	-10 to 30 dBm
τ_n	$\sim \mathcal{U}[40 \text{ ns}, 200 \text{ ns}]$, rounded to grid $\mathcal{T}$
$\tau_{\max}$	200 ns
θ_n	$\sim \mathcal{U}[-90°, 90°]$, rounded to grid Θ
$\xi_{n,k}$	$\sim \mathcal{U}[0,1] + j \cdot \mathcal{U}[0,1]$
modulation	4-QAM
Algorithm Parameters	
max_iter	700
ρ_0	1
μ^+	1.11
μ^-	0.9
r_ρ^+	1000
r_ρ^-	0.001
r_x	0.05 for CS ADMM, 0.15 for benchmark from [12]
L^{F}	201 (1 ns steps)
L^{R}	181 (1° steps)

Fig. 3. Output magnitude profile of our multiband CS ADMM. The black lines are the true target AoAs and ToAs.

such as maximal ratio combining (MRC). Also, instead of grid-based approaches, we might also consider gridless methods such as atomic norm minimization to solve the CS problem and achieve better resolution and accuracy.

For performance comparison, we have not compared with fundamental limits such as Cramér-Rao bound (CRB), which can be easily derived from the system model and can give us more theoretical insights on whether we have taken the full advantage of our multiband system. Also, in the future, we want to compare with single-band CS on the constituent subband frequencies with the same Q and B to see the benefits of multiband.
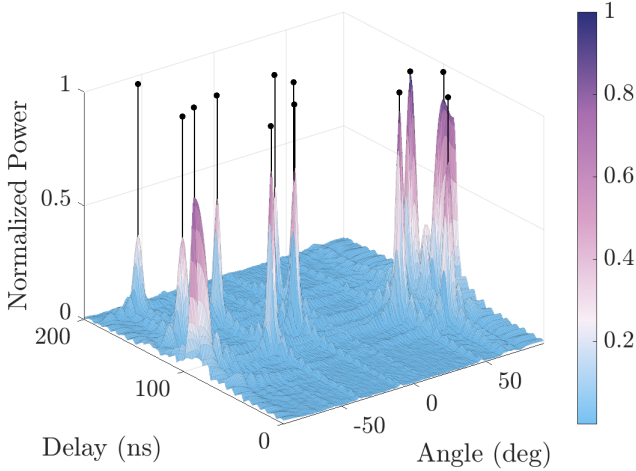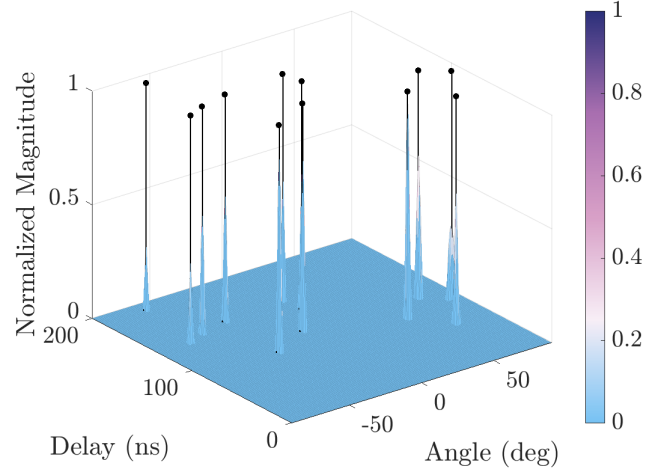

Fig. 2. Output power profile of the benchmark in [12] adapted to multiband. The black lines are the true target AoAs and ToAs.

DMCs is through non-white Gaussian noise [15]. Since our CS estimator does not assume white Gaussian noise, this can be an easy but informative extension. In this case, we estimate the covariance matrix from the residue.

Furthermore, we may have not fully leveraged the multiband benefits as we only take the contributions from each and every subband equally. In the future, we might consider techniques

REFERENCES

[1] D. Wu, W.-P. Zhu, and M. Swamy, "A Compressive Sensing Method for Noise Reduction of Speech and Audio Signals," in *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2011, pp. 1–4.

[2] A. E. Mahdaoui, A. Ouahabi, and M. S. Moulay, "Image Denoising Using a Compressive Sensing Approach Based on Regularization Constraints," *Sensors*, vol. 22, no. 6, 2022.

[3] G. Caire, T. Y. Al-Naffouri, and A. K. Narayanan, "Impulse noise cancellation in OFDM: an application of compressed sensing," in *2008 IEEE International Symposium on Information Theory*, 2008, pp. 1293–1297.

[4] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] M. Mishali and Y. C. Eldar, "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, 2010.

[6] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," *arXiv preprint arXiv:0901.3403*, 2009.

[7] M. Chafii, L. Bariah, S. Muhaidat, and M. Debbah, "Twelve Scientific Challenges for 6G: Rethinking the Foundations of Communications Theory," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2023.

[8] D. Wang, A. Bazzi, and M. Chafii, "RIS-Enabled Integrated Sensing and Communication for 6G Systems," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, 2024, pp. 1–6.

[9] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics: A Journal Issued*

by the Courant Institute of Mathematical Sciences, vol. 59, no. 6, pp. 797–829, 2006.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[11] N. Parikh, S. Boyd *et al.*, "Proximal algorithms," *Foundations and trends® in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[12] T. Raviv, S. Kang, M. Mezzavilla, S. Rangan, and N. Shlezinger, "Multi-Frequency Upper Mid-Band Localization," in *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 736–740.

[13] A. Richter, "Estimation of Radio Channel Parameters: Models and Algorithms," Ph.D. dissertation, Dept. Elektrotechnik und Informationstechnik, Technische Universität Ilmenau, Ilmenau, Germany, 2005.

[14] Z. Cheng and A. F. Molisch, "Multi-Cluster Frequency-Angle Diffuse Scattering Model and Its Estimation," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 7176–7196, 2024.

[15] S. Jiang, W. Wang, Y. Miao, W. Fan, and A. F. Molisch, "A Survey of Dense Multipath and Its Impact on Wireless Systems," *IEEE Open Journal of Antennas and Propagation*, vol. 3, pp. 435–460, 2022.