Knowledge-Aided and Adaptive Beam-Squint Aware MIMO-OFDM Radar Detectors for ISAC

Uğur Berkay Sarac¹ and Gökhan Muzaffer Güvenşen¹

¹Middle East Technical University

June 24, 2024
Abstract—Integrating radar and communication systems for economical use of hardware and spectrum resources is projected to be a crucial aspect of sixth-generation (6G) systems, leading to extensive research in the integrated sensing and communication (ISAC) area. In this article, we propose a radar detector structure for a multiple-input multiple-output (MIMO) ISAC system using orthogonal frequency-division multiplexing (OFDM) modulated waveforms. These waveforms are utilized to communicate with downlink (DL) users while receiving echoes from targets and clutter, in addition to separate OFDM-modulated uplink (UL) communication waveforms. The transmitter (Tx) and receiver (Rx) employ hybrid beamformers, with analog beamformers precisely designed to cover flat angular sectors. Tx and Rx beams are directed towards DL and UL users, respectively, with the possibility of overlapping or separate angular sectors. We introduce a Doppler-aware Code Bank (DACB) as the initial processing stage, thoroughly investigating the effects of Doppler mismatch. Following DACB, a sub-optimal sequential angle-range processing (SARP) method is proposed to maximize the output signal-to-interference-plus-noise ratio (SINR) while maintaining feasible processing. A two-variant detector scheme is proposed to address this processing’s suboptimality. Four different detectors with varying complexities, including a fully adaptive detector, are introduced. The potential beam-squint effect due to increased bandwidth is also considered, and a subband approach is proposed to mitigate these effects. Simulation results for all four detectors, as well as a conventional 3-dimensional periodogram detector commonly used as a benchmark in the literature, are provided. The results demonstrate that the proposed detectors can significantly enhance SINR and the probability of detection, particularly when accounting for the coupling between the clutter Doppler and Tx OFDM symbols.

Index Terms—Radar and communication coexistence, MIMO OFDM radar, detector structures, clutter channel model.

I. INTRODUCTION

There has been a significant interest in integrating radar and communication systems in the last few years, especially due to the increasing demand for sensing capability in communication-oriented systems. [1]–[4] are a few examples of valuable surveys about this research area. Among the numerous studies, there has yet to be a consensus on the name of the systems; integrated sensing and communication (ISAC), joint sensing and communication (JSC) and RadCom are commonly used in the literature. We prefer to use ISAC to depict the integrated radar and communication system. Orthogonal frequency division multiplexing (OFDM) modulation is one of the most commonly used modulation schemes used in modern communication systems because it enables frequency-domain equalization which easily solves multipath channel problems, and its implementation is relatively easy. On the other hand, there are various studies which show that OFDM waveforms can also be used in sensing operation [4]–[15].

In the literature for fifth-generation (5G) and sixth-generation (6G) systems, mm-wave multiple-input multiple-output (MIMO) communications are commonly focused due to the large communication band and high processing gain it provides. As expected, there are also several MIMO ISAC studies investigating sensing operation integrated into MIMO communication systems. In [16], a portion of the transmit (Tx) power is allocated for beam-scanning for radar detection. In [17], both transmit and receive (Rx) beamformers are optimized for sensing performance at a known target inside the cell-under-test (CUT). In [11], separate two antennas are utilized for radar signal transmission while a large antenna array continues MIMO communication and sensing. In [18], hybrid beamformers and compressive sensing are utilized for ISAC signal processing. In [19], radar detection is performed using the angle-sidelobes of the communication beams. These are only a few examples showing how different the MIMO ISAC approaches can be, mainly due to the large number of parameters, e.g. antenna and waveform structures, Tx and Rx beamformers, signal processing methods, etc. [20]–[26] are valuable works about MIMO ISAC using OFDM waveforms.

An aspect of the MIMO-OFDM ISAC systems is the effective bandwidth (W) of the system. In general, using as large W as possible is beneficial for both communication and sensing purposes; however, significantly increasing the W of the system can reduce the range resolution of the system so much that different antennas in the array can see different range responses coming from the same scatterer. In another perspective, a single scatterer can be seen at different angles in different subcarriers in a MIMO-OFDM detector. This beam-squint effect is investigated in [27] and has been shown to be effective for high carrier frequency over W ratios. In [5], this effect is mentioned and ignored as the W is selected to be relatively small. In [28], the beam-squint effect is intentionally used for searching purposes.

In this study, we focus on the radar detection performance of a MIMO-OFDM ISAC system under a realistic clutter model and the disruptive effects of beam-squint and uplink (UL) communication users. In our scenario, there are multiple scatterers in the sector of interest. The receiver gets the echoes of the transmitted signal from all of these scatterers but is interested in detecting only one of them. The scatterer that the receiver is trying to detect, mostly because it is a newcomer to the sector of interest, is called the intended target (IT). All other scatterers are also targets for the radar.
system but they are unwanted, this is why they are called unintended targets (UITs). UITs act like clutter, and they may be more powerful than the IT itself. The UITs also consist of multiple scattering centers, making sure that some part of them lies on fractional range and angle bins, making it harder to suppress them. The same interference model is used in [29] for a single-input, single-output (SISO) OFDM radar scenario, for reference. On the other hand, there are UL communication users and their signals are also disruptive to the radar operation. Besides, when the communication bandwidth becomes comparable to the center frequency, the beam-squint effect introduces a range-angle coupling, negatively affecting the MIMO radar performance. Last but not least, a Doppler mismatch in the filters or Doppler spread of UITs results in a symbol-dependent interference and undermines both angle and range processing if not handled correctly. In [30], this effect is explained and mitigated by imposing a constraint on the linear dependency of modulation symbols. In [31], the same effect is mitigated by designing a special beamformer. Under all these disruptive effects, we propose a detector scheme that successfully suppresses the UITs and detects the IT.

The contributions of this study can be listed as:

- A mathematical model for the MIMO-OFDM channel under the disruptive effects of clutter, UL users, beam-squint effect and Doppler mismatch is constructed. The effect of Doppler mismatch on angle and range covariances are explained and exploited for performance increase.
- A suboptimal knowledge-aided (KA) sequential angle-range processing (SARP) is proposed. Within SARP, theoretical signal-to-interference-plus-noise ratio (SINR) expressions and SINR maximizing filters are provided.
- A novel detector scheme is proposed to overcome the performance losses arising due to suboptimality. Its performance is evaluated in terms of SINR and probability of detection ($P_d$) metrics, including the comparison with benchmark detectors in the literature.
- Along with benchmark detectors, computationally softer detectors are proposed. Their mathematical backgrounds are provided. A novel, adaptive matched filter (AMF)-like [32] fully adaptive detector is also provided, which uses symbol-independent filters that do not require recalculation for multiple frames.

The rest of this paper is organized as follows. The system model including the MIMO channel and well-defined beamformers are explained in II. The proposed Doppler-Aware Code Bank (DACB) and SARP methods are explained in detail in III. An explicit interpretation of the Doppler mismatch effect is also included in this section. In IV, the novel 2-variants detector scheme and alternative detectors used in this study are investigated. The simulation results and discussions are provided in V, followed by the concluding remarks in VI. Some mathematical derivations are provided in the Appendix.

**Notations:** Regular letters, bold lowercase and bold uppercase letters denote scalars, vectors and matrices, respectively. $[\cdot]^T$, $[\cdot]^H$, $(\cdot)^*$, $[\cdot]^\#$, $\text{tr}\{\cdot\}$, $\odot$, $\otimes$ and $\oplus$ denote transpose, Hermitian transpose, conjugate, pseudo-inverse, trace, element-wise multiplication, Kronecker product and Khatri-Rao (column-wise Kronecker) product operations, respectively. $\text{diag}\{\cdot\}$ represents a vector including diagonal elements of $\cdot$. $\text{diag}\{\cdot\}$ represents a diagonal matrix $\cdot$ whose diagonal elements are $\cdot$. $[\cdot]_{(n)}$, $[\cdot]_{(n,m)}$ and $[\cdot]_{(n,m)}$ represent $n^{th}$ element of vector $\cdot$, $(n, m)^{th}$ element of matrix $\cdot$ and $(n, m)^{th}$ block matrix inside matrix $\cdot$, respectively.

## II. System Model

In this section, the mathematical model used in this article will be explained. In the first part, the MIMO channel is constructed and the necessary assumptions are made. In the second part, the beamformers are constructed and the different scenarios investigated in the article are explained.

### A. MIMO Channel

The system in this study uses OFDM-modulated signals for both uplink (UL) and downlink (DL) communication purposes. For OFDM modulation, required cyclic prefix operations are assumed to be conducted successfully throughout the article. There are $M$ OFDM symbols with duration $T_s$ in a coherent processing interval (CPI), among which the reflected complex gain from any scatterer is highly correlated, enabling Doppler processing operations. On the other hand, the reflected complex gains from the scatterers are independent and identically distributed (i.i.d) from CPI to CPI. The slow-time parameters of the scatterers (range, angle, velocity) are assumed to be the same through $T_s$ CPIs, which are called frames, with duration $T_f$. In each OFDM symbol, there are $N$ subcarriers, the distribution of which among communication users is not in the scope of this work. Among these $N$ subcarriers, only $N$ of them are used for radar purposes. For simplicity, these $N$ subcarriers are assumed to be evenly distributed among all $N$ subcarriers. Since this article focuses on radar signal processing, the word ‘subcarriers’ will always mean the $N$ subcarriers used for radar purposes, unless otherwise is clearly stated. The subcarrier spacing is $\Delta f$ and the carrier frequency of $n^{th}$ subcarrier is $f_n = f_1 + (n-1)\Delta f$, for $n = 1, 2, \ldots, N$. This means that total RF bandwidth is $W = N\Delta f$ and the center carrier frequency is $f_c = (f_1 + f_N)/2$. Corresponding to these subcarrier frequencies, the wavelength of $n^{th}$ subcarrier is $\lambda_n = c/f_n$, where $c$ is the speed of light in free space.

In our work, the UITs are considered to be large objects, possibly covering multiple range and angle resolution cells. For the sake of simplicity, the targets are approximated as collections of several point scattering centers. There are multiple UITs in the section of interest, each having a specific number of scatterers, which depends on the size of the corresponding UIT. A total of $K$ scatterers exist in our scenario and in the rest of the paper, $k$ is the index used for counting the scatterers. The $k^{th}$ scatterer is assumed to be at range $r_k$, corresponding to a round-trip time of $\tau_k = 2r_k/c$, and has a radial velocity of $v_k$, corresponding to a normalized Doppler...
frequency of $\rho^2_{kn} = (2\nu_k/\lambda_c)T_d$, where $\lambda_c = c/f_c$ is the wavelength corresponding to the center carrier frequency.

We assume that there are $N_t$ Tx antennas, $N_r$ Rx antennas, $D_t$ Tx RF chains and $D_r$ Rx RF chains in the system. The Tx and Rx analog beamformers are denoted as $W_t$ and $W_r$, respectively. Using these variables, the frequency-domain MIMO channel matrix of the $k$th scatterer for $n$th subcarrier and $m$th OFDM symbol can be written as:

$$H_{kmn} = \alpha_{km}e^{j2\pi\rho_{kn}^2m}H_{kn} = \alpha_{km}e^{j2\pi\rho_{kn}^2m}[W_r^H\alpha_{kn}b_k^T W_r^* e^{-j2\pi\tau_{kn}\Delta fn}]$$

(1)

where $\alpha_{kn}$ and $b_k$ are the receive and transmit steering vectors towards $k$th scatterer at $n$th subcarrier. $H_{kn}$ is the effective MIMO channel, including range and angle information of the scatterer after dimensions are reduced via analog beamforming, which is assumed to be stationary during $M$ symbols; thus, it has no $m$ index inside it. The random variable $\alpha_{kn}$ is the reflected complex gain, depicting the complex amplitude and Doppler spread information of the scatterer, and $e^{j2\pi\rho_{kn}^2m}$ is for the mean Doppler shift of the scatterer. Note that $\alpha_{kn}$ is a realization of a stationary slow-time random process with a given PSD among the Doppler axis. Let’s define $\alpha_k \triangleq [\alpha_{k1} \alpha_{k2} \cdots \alpha_{kM}]^T, R_k^\alpha \triangleq \mathbb{E}\{\alpha_k \alpha_k^H\}$,

(2)

where all of the diagonal elements of $R_k^\alpha$ are assumed to be equal to $\sigma_k^2 \triangleq \mathbb{E}\{|\alpha_{km}|^2\}$ because the average power of the returns from the scatterer is assumed to be the same for all $M$ symbols. The off-diagonal elements of $R_k^\alpha$ determines the slow-time correlation properties of the $k$th UIT.

The receive and transmit steering vectors towards angle $\theta$ at the $n$th subcarrier are denoted as $a_n(\theta)$ and $b_n(\theta)$, which are $N_r \times 1$ and $N_t \times 1$ vectors, respectively. When the steering vectors are for $k$th UIT or $k$th user, we drop the angle and use simply $a_n$ or $b_n$, for notational simplicity. The reader should understand that $a_n$ is the steering vector for $k$th UIT, which is at angle $\theta_n$. For example, if the antenna spacing is denoted as $d$, the receive steering vector can be written as:

$$a_n(\theta) = [\exp{(j2\pi\frac{d \sin \theta}{\lambda_c})} \exp{(j2\pi\frac{d \sin \theta}{\lambda_c})} \cdots \exp{(j2\pi\frac{d \sin \theta}{\lambda_c})}]^T$$

(3)

which depends on $n$ due to the beam-squint effect. The $(f_n/f_c)$ ratio determines how much the wave number shifts from its center value $(2\pi/\lambda_c)$, and its value is assumed to be unity in most scenarios. However, when $W$ is comparable with $f_c$, it can be seen that $1 - (W/2f_c) \leq (f_n/f_c) \leq 1 + (W/2f_c)$.

In other words, the beam-squint effect becomes more effective as the $(f_n/f_c)$ ratio decreases.

The Doppler shift is clearly dependent on the carrier frequency, and this results in different Doppler shifts occur in different subcarriers. However, when the center frequency $f_c$ and bandwidth $W$ of the system satisfies $f_c/W > M/2$, the maximum amount of Doppler shift across subcarriers becomes insignificant with respect to the Doppler resolution of the system. In this work, this inequality is assumed to be satisfied so that the normalized Doppler frequency is not dependent on subcarrier index.

The random variable $\alpha_{kn}$ is chosen to be a realization of a zero-mean complex Gaussian process, whose PSD function is Gaussian shaped with a variance of $(\sigma_k^2)^2$ and mean of zero. Such processes can be easily shown to have their autocorrelation matrix entries as $[R_k^\alpha]_{(i,j)} = \gamma_k^2 \exp(-2\nu^2(\sigma_k^2)^2 (i-j)^2T_d^2)$.

Along with the clutter, a total of $U$ single antenna communication users are in the section of interest. The $u$th user sends the frequency domain communication symbol $f_{umn}$ at $n$th subcarrier in $m$th symbol duration. Similar to the definition in (1), the effective channel vector for the $u$th user at $n$th subcarrier for $m$th symbol is defined as:

$$g_{umn} = \beta_{um}e^{j2\pi\rho_{umn}^2m}g_{um} = \beta_{um}e^{j2\pi\rho_{umn}^2m}[W_r^H a_{um}]$$

(4)

where $g_{um}$ is the effective SIMO channel including complex channel and angle information of the user after dimensions are reduced via analog beamforming. It should be noted that this channel is also assumed to be stationary during $M$ symbols, therefore it has no $m$ index inside it. The random variable $\beta_{um}$ is the received complex gain including the received power and Doppler spread information of the user channel and $e^{j2\pi\rho_{umn}^2m}$ is for the mean Doppler shift of the user. Similar to (2),

$$\beta_u \triangleq [\beta_{u1} \beta_{u2} \cdots \beta_{uM}]^T, R_u^\beta \triangleq \mathbb{E}\{\beta_u \beta_u^H\},$$

(5)

where all of the diagonal elements of $R_u^\beta$ are assumed to be equal to $\gamma_u^2 \triangleq \mathbb{E}\{|\beta_{um}|^2\}$.

The received signal vector $y_{nm}$ can be written as:

$$y_{nm} = H_{0nm}x_{nm} + \sum_{k=1}^{K} H_{knm}x_{nm} + \sum_{u=1}^{U} g_{umn}f_{umn} + n_{nm}$$

(6)

$$\begin{bmatrix}
t_{nm} \\
c_{nm} \\
\psi_{nm}
\end{bmatrix} =
\begin{bmatrix}
t_{nm} \\
c_{nm} \\
\psi_{nm}
\end{bmatrix}$$

where $x_{nm}$ is the $D_t \times 1$ vector of Tx streams to DL users, and $f_{umn}$ is the Rx symbol received from $u$th UL user, at $n$th subcarrier in $m$th OFDM symbol. $t_{nm}, c_{nm}, \psi_{nm}$ and $n_{nm}$, are $D_r \times 1$ vectors who represent the received signal vector from intended target, total received signal vector from unintended targets (clutter), total received signal vectors from UL user signals, and noise, respectively. $H_{0nm}$ is the effective MIMO channel for the intended target, whose expression is the same as the channels for UITs, given in (1). $\psi_{nm}$ is the total additive interference on the intended target’s response, namely the clutter, noise and the communication signal of interest (SoI). The noise vector $n_{nm}$ is also defined after beamforming operation, namely it consists of $D_r$ circularly symmetric zero-mean Gaussian random variables with covariance $\sigma^2_W W_r^H W_r$. For notational simplicity, let’s define:

$$X_n \triangleq [x_{n1} x_{n2} \cdots x_{nM}], T_n \triangleq [t_{n1} t_{n2} \cdots t_{nM}],$$

$$C_n \triangleq [c_{n1} c_{n2} \cdots c_{nM}], S_n \triangleq [s_{n1} s_{n2} \cdots s_{nM}],$$

$$N_n \triangleq [n_{n1} n_{n2} \cdots n_{nM}], \psi_n \triangleq [\psi_{n1} \psi_{n2} \cdots \psi_{nM}],$$

$$Y_n \triangleq [y_{n1} y_{n2} \cdots y_{nM}], J_n \triangleq [j_{n1} j_{n2} \cdots j_{nM}]^T, \psi_{nM} \triangleq [\psi_{n1} \psi_{n2} \cdots \psi_{nM}]$$

(7)

Since it is assumed that the synchronization is done for the SoI channel, the complex phase coming from propagation delay is not included in $g_{umn}$. On the other hand, angular coherence time is assumed to be much larger than the CPI so that $g_{umn}$ and $H_{kn}$ stay the same for $M$ symbol durations.

The random variable $\beta_{um}$ is chosen to be a realization of a zero-mean complex Gaussian process, whose PSD function is Gaussian shaped with a variance of $(\sigma_u^2)^2$ and mean of zero, whose autocorrelation matrix becomes $[R_u^\beta]_{(i,j)} = \gamma_u^2 \exp(-2\nu^2(\sigma_u^2)^2 (i-j)^2T_d^2)$.
In this work, the slow-time processing gain is included in the transmit beamformer and not visible in the transmitted symbols. Therefore, $X_n$ and $J_n$ are normalized accordingly. The power per stream per subcarrier is selected to be unity, therefore $\text{Tr}\{E\{X_nX_n^H\}\} = D_t$ and $\text{Tr}\{E\{J_nJ_n^H\}\} = U$. To simplify the notation, we define four diagonal matrices:

$$D_k^C \triangleq \text{diag}\{1, e^{j2\pi \rho_{k1}}, \ldots , e^{j2\pi(M-1)\rho_{k1}}\}, \Lambda_k^C \triangleq \text{diag}\{\alpha_k\}$$

(8)

are the matrices that represent the mean Doppler shift and the Doppler spread of the $k$th scatterer. $\Lambda_k^C$ also includes the amplitude information of the $k$th scatterer as $\alpha_k$ is the reflected complex gain vector as described before. $D_u^S$ and $\Lambda_u^S$ have the same forms with $D_k^C$ and $\Lambda_k^C$ but they are for $u$th communication user. Using (6), (7) and (8), $Y_n$ is written and its elements can be defined as below:

$$Y_n = T_n + \sum_{k=1}^{K} C_{kn} + \sum_{u=1}^{U} S_{u,n} + N_n$$

(9)

$$T_n \triangleq H_{0n}X_nD_k^C\Lambda_k^C, \quad C_{kn} \triangleq H_{kn}X_nD_k^C\Lambda_k^C, \quad S_{u,n} \triangleq g_{un}J_{u,n}^T D_u^S\Lambda_u^S$$

(10)

It should be noted that (9) can be written because the MIMO channel matrix $H_{kn}$ defined in (1) and SoI channel vector $g_{un}$ are assumed to stay the same for the duration of $M$ symbols.

### B. Beamformers

In this article, transmit and receive beamformers, $W_t$ and $W_r$, are adjusted separately. $W_t(W_r)$ is formed to cover the angular sector where DL(UL) users are located, and the UL and DL users may or may not be inside the same sector in a real-life application. Therefore, different scenarios in which the sectors fully overlap or do not overlap are investigated in the simulations section. Fig. 1 shows how the Tx and Rx sectors, DL and UL users, UITs, and the IT can be located in an example scenario of partially overlapping Tx-Rx sectors.

To reduce the power fluctuations inside the angular sectors, the columns of $W_t(W_r)$ are selected to be the eigenvectors of the intended angular sector corresponding to $D_t(D_r)$ largest eigenvalues. The transmit power constraint is adjusted so that $\text{Tr}\{W_tW_t^H\} = M$ is satisfied. This means that the system’s slow-time processing gain is included through transmit beamforming, and the total transmit power is constrained. In the receiver part, however, to preserve the second-order stationarity of the noise after receive beamforming, $W_r$ is scaled so that $W_r^H W_r = I_{D_r}$ is satisfied.

### III. Proposed Processing Method

In this article, an IT is to be detected under the disruptive effects of UITs and UL users, by implementing knowledge-aided or adaptive detection strategies. Within the given system model, there are three domains in which detection will be conducted, namely range, angle and Doppler domains. To have a maximum-SINR filter for all domains, a $D_tNM \times D_tNM$ matrix must be known using the tracker or learned from the environment and its inverse must be taken in the filter implementation. In practical systems, this dimension is so large that neither effectively learning the matrix is possible before the channel decorrelates, nor is there enough processing power to implement this matrix inversion in real-time.

Therefore, this article pursues a sub-optimal processing that separates range, angle and Doppler domains. Due to possible beam-squint effects, MIMO processing should be done separately for each subcarrier, forcing angle processing to come before range processing. However, $D_tM \times D_tM$ matrix processing for each of the $N$ subcarriers is still hard to work with. Besides, eliminating the Tx symbols at the first stage is crucial for OFDM radars, as the residual symbols would deteriorate the performance for the subsequent processing, as will be shown. For these reasons, a Doppler-Aware Code Bank (DACB) is proposed as the first stage of the processing, as in [5]. With this processing, the $M$ symbols of transmitted data are combined separately for each Doppler cell under test, and their dimension is reduced to $D_t < M$. It should be noted
that for classical phased-array systems with $D_t = 1$, this processing would be equivalent to slow-time match filtering. In the MIMO case, the residual $D_t$ dimension allows transmit beamforming on the receiver side to utilize full MIMO gain while easing the computations on the subsequent processing.

In summary, the proposed processing is a post-Doppler processing, focusing on a single Doppler bin before any angle or range operations. In the next step, MIMO processing is done for each subcarrier to eliminate the negative effects of beam-squint. Lastly, the range processing is conducted to reach the detection metric. The metric is then compared with the threshold and a decision is made on whether there is no detection ($H_0$ hypothesis) or detection ($H_1$ hypothesis). A block diagram showing the frame, symbol and subcarrier structures, along with the proposed signal processing until the sequential range-angle processing (SARP) part, is given in Fig. 2. Another diagram will be provided in later sections to explain the SARP part of the proposed signal processing scheme.

\[ V_n(\tilde{m}) \triangleq D_{\tilde{m}}^n(X_n^H)^* = D_{\tilde{m}}^n X_n^T (X_n^T)^{-1} \]  

where $X_n^\# = X_n^H (X_n X_n^H)^{-1}$ is pseudo-inverse of $X_n$ and $D_{\tilde{m}} = \text{diag}(1 e^{j2\pi \tilde{m}/M} e^{j2\pi 2\tilde{m}/M} \ldots e^{j2\pi (M-1)\tilde{m}/M})$ (12) is the Doppler preprocessing matrix for $\tilde{m}^{th}$ Doppler bin. After the preprocessing operation, the MIMO channel estimate at $n^{th}$ subcarrier for the CUT becomes:

\[ \hat{H}_n(\tilde{m}) \triangleq Z_n(\tilde{m}) \triangleq Y_n^H X_n^\# = Y_n^H V_n \]  

(13)

\[ = T_n V_n^* + C_n V_n^* + S_n V_n^* + N_n V_n^* = T_n V_n^* + \Psi_n V_n^* \]

Here are some remarks about our preprocessing, DACB:

- The dimension of the processed data in each subcarrier is reduced from $D_r \times M$ to $D_r \times D_t$ with this preprocessing, meaning that OFDM symbols are coherently combined but there are still enough dimensions to achieve additional Doppler suppression with angle or range processing.
- The Doppler processing and the elimination of Tx symbols are coupled. More specifically, $M \geq D_t$ must be satisfied to effectively find $X_n^\#$ to eliminate the Tx symbols, but when $M$ symbols are processed together, Doppler effects changes the slow-time codes of Tx symbols. This is a unique property of MIMO systems.
- If the Doppler processing for the CUT, $D_{\tilde{m}}$, matches the Doppler of the IT, $D_{\hat{m}}$, and there is no Doppler spread of the target, then $T_n V_n^*$ becomes equal to $H_{\hat{m}}$, which means that the channel estimation is successful.
- However, it is almost always the case that there are other scatterers in the environment which have different Doppler shifts and Doppler spreads, meaning that there will be a Doppler mismatch between the preprocessing and some of the scatterers in the environment. When this is the case, $X_n$ could not be eliminated effectively, causing the interference become dependent on $X_n$.
- As will be shown in (27), Doppler mismatch results in a spread in the range response of the scatterers. If this effect is taken into consideration, it acts as a diversity in Doppler domain and results in an increased output SINR (due to clutter suppression in Doppler domain). If this effect is overseen, it makes the clutter spread over the whole range spectrum, disables the range processing to suppress the clutter and results in a reduced output SINR.
- Doppler mismatch also affects the angular response of the scatterers, which can be seen when (19) is investigated, but the details are not written in this paper for the sake of neatness. The angular responses of the scatterers are linearly transformed due to the Doppler mismatch, which is also dependent on $X_n$. Similar to the range processing part, detectors that take this effect into consideration achieve greater clutter suppression with MIMO filters.

### A. Doppler-Aware Code Bank (DACB)

When (9) and (10) are investigated, it is seen that to reach the channel $H_{\hat{m}}$, of the IT for detection, the transmitted symbols and Doppler effects must be removed from the received signal. For SISO channels, where $H_{\hat{m}}$ is a scalar for each subcarrier and OFDM symbol, the suppression of $X_n$ can be done via Hadamard division, as in [12], [29]. However, in MIMO channels, as seen in (10), Hadamard division cannot give a good estimate of $H_{\hat{m}}$. In this work, we estimate the channel using least-squares (LS) approach, as in [5], but we also conduct the Doppler processing at the same time.

For our preprocessing, and for the rest of the paper, we assume that the cell under test (CUT) is at $\tilde{m}^{th}$ range bin, $\tilde{m}^{th}$ Doppler bin and $\theta^{th}$ angle bin. In general, $\tilde{m}, \hat{m}$ and $\theta$ can be fractional bins. Parentheses notation, e.g. $(\tilde{m})$, will be explicitly used when it is needed to clarify the dependence of a variable to any of the CUT parameters, like $\tilde{m}$ in this example. The preprocessing matrix is defined as:

\[ V_n(\tilde{m}) \triangleq D_{\tilde{m}}^n(X_n^H)^* = D_{\tilde{m}}^n X_n^T (X_n^T)^{-1} \]  

B. Knowledge-Aided Sequential Angle-Range Processing

As the next step of processing, we apply spatial processing for each subcarrier to resolve any beam-squint effect before combining the symbols in range domain. At each subcarrier, the MIMO channel is a $D_r \times D_t$ matrix and to jointly use all
Doppler preprocessing and given, before the DACB. When the transmitted symbols
the rest of the paper. Let’s define the result of the processing
in
\[ R_n^\psi = \sum_{k=1}^{K} (I_M \otimes H_{kn})(I_M \otimes (X_n D_k^r)) R_k^\psi (I_M \otimes (X_n D_k^r))^H (I_M \otimes H_{kn})^H + \sum_{u=1}^{\frac{S}{M}} \frac{1}{M} (I_M \otimes g_{un} g_{mn}^H) + \sigma_n^2 I_M D_r \]  
\[ R_{n_1 n_2}^\psi = \begin{cases} 
R_{n_1}^\psi, & \text{if } n_1 = n_2 \\
\sum_{k=1}^{K} (I_M \otimes H_{kn_1})(I_M \otimes (X_{n_1} D_k^r)) R_k^\psi (I_M \otimes (X_{n_2} D_k^r))^H (I_M \otimes H_{kn_2})^H, & \text{if } n_1 \neq n_2 
\end{cases} \]  
\[ \text{As the last step of processing, we apply a temporal filter } u \text{ on } r \text{ to get our decision metric } \xi \text{ as:} \]
\[ \xi(n\tilde{m}, \tilde{\theta}) = |u^H r|^2. \]  
The temporal filter \( u \) can be written as:
\[ u(n\tilde{m}, \tilde{\theta}) = \frac{\Sigma^{-1}(\tilde{m}, \tilde{\theta}) q(n\tilde{m})}{\sqrt{q^H(n\tilde{m}) \Sigma^{-1}(\tilde{m}, \tilde{\theta}) q(n\tilde{m})}} \]  
where \( q(n\tilde{m}) \) is the range steering vector towards \( n \tilde{m} \) and \( \Sigma(n\tilde{m}, \tilde{\theta}) \) is defined as the range covariance matrix of the preprocessed and spatial filtered received frame, again without the IT contamination. It should be noted that the normalization in \( u \) makes sure that \( E \{ |r| H_0 \} = 1 \) if \( \Sigma = E \{ r r^H | H_0 \} \) is indeed satisfied. Assuming that the IT is at \( n\tilde{m} \) subcarrier, \( n \)th element of the temporal steering vector towards the IT is:
\[ [q(n\tilde{m})]_n = e^{-j(2\pi/N)n\tilde{m}} \omega_n^H p_n = e^{-j(2\pi/N)n\tilde{m}}, \]  
which is the temporal response coming from a hypothetical point target at \( n \)th range cell. The range covariance matrix of \( u \) under \( H_0 \) can be defined as:
\[ \Sigma^u(n\tilde{m}, \tilde{\theta}) = E \{ r(n\tilde{m}, \tilde{\theta}) r^H(n\tilde{m}, \tilde{\theta}) | H_0 \} \]  
When the transmitted symbols for all subcarriers, namely \( X_1, X_2, \ldots, X_N \), are given, \( (n_1, n_2) \)th element of \( \Sigma^u \) can be written as:
\[ [\Sigma^u]^u_{n_1 n_2} = E \{ \omega_{n_1}^{n_2} H_{n_1} \omega_{n_2}^{n_2} H_{n_2} | H_0, X_{n_1}, X_{n_2} \} \]
\[ = E \{ \psi_{n_1} | X_{n_1} \} E \{ \psi_{n_2} | X_{n_2} \} \]  
where \( \Sigma_{n_1 n_2}^u \) is the cross-correlation matrix of preprocessed interference terms between \( n_1 \)st and \( n_2 \)nd subcarriers, which can also be written as:
\[ \Sigma_{n_1 n_2}^u = \nabla_{n_1}^H \psi_{n_1} \nabla_{n_2} \psi_{n_2} \]  
where \( \Sigma_{n_1 n_2}^u \) is the cross-correlation matrix of the raw interference terms between \( n_1 \)st and \( n_2 \)nd subcarriers, whose closed form expression is given in (20). The derivation of (20) and the expectation of \( \Sigma_{n_1 n_2}^u \) over \( X_{n_1} \) and \( X_{n_2} \) are given in the Appendix. When (1), (26) and (20) are investigated, it can be seen that the range information of the scatterers are carried in the off-diagonal elements of the \( \Sigma^u \) matrix. More explicitly, the contribution from the \( k \)th subcarrier to the \( (n_1, n_2) \)th element of \( \Sigma^u \) is \( e^{-j2\pi n_1 \Delta f(n_1-n_2)} \) times another scalar which does not depend on the range of the corresponding scatterer.
If we define $q_k^C$ to be the range response of the $k^{th}$ scatterer where $[q_k^C]_n \triangleq e^{-j2\pi t_n \Delta f_n}$ for $n = 0, 1, \cdots, N - 1$, then

$$\Sigma^\eta = \sum_{k=1}^{K} q_k^C (q_k^C)^H \otimes A_k (\tilde{m}, \tilde{\theta}, X_n)$$

where $A_k$, which depends on $\tilde{m}, \tilde{\theta}, X_n$ for all subcarriers, as well as the angle and Doppler properties of $k^{th}$ scatterer but not the range of it, can be found easily but not written here for the sake of neatness. It should be noted that if there were only a single scatterer with no Doppler spread and the Doppler processing perfectly matched the Doppler of the scatterer, $A_k$ would become an all-ones matrix and $\Sigma^\eta$ would become a rank-1 matrix which includes the range information of the scatterer. However in a more realistic scenario, even a slight Doppler spread or a Doppler mismatch between the DACB and any of the scatterers result in an increase in the range of the scatterer. However in a more realistic scenario, even a slight Doppler spread or a Doppler mismatch between the DACB and any of the scatterers result in an increase in the rank of $A_k$ and $\Sigma^\eta$. In other words, the Doppler mismatch causes the range response of the scatterers, which would be impulsive otherwise, to spread over the whole range spectrum. In addition to this, this spread depends on the transmitted symbols $X_n$. As a consequence, if $A_k$ matrix information is not used at the receiver side, the range information of the scatterers cannot be learnt and the output SINR is reduced when there is Doppler mismatch. On the other hand, if the $A_k$ matrix is exploited at the receiver side, which requires tracker information and a relatively high computation power, the effects of the Doppler mismatch can be resolved and due to the increased rank of $\Sigma^\eta$, output SINR can be increased even more with respect to the case with no Doppler mismatch.

After the decision metric $\xi$ is determined, it is compared to the pre-determined CFAR threshold $\gamma_{th}$ to make the decision:

$$\xi(\tilde{m}, \tilde{\theta}) \gtrless H_1 \quad \text{or} \quad \xi(\tilde{m}, \tilde{\theta}) \gtrless H_0,$$

where $\gamma_{th}$ can be adaptively found using previous uncontaminated frames of data. A similar threshold map calculation is explained in detail in [29] for SISO OFDM ISAC systems. In this study, Tx symbols cannot be completely eliminated as explained before, therefore instantaneous thresholds depend on Tx symbols. However, the system can calculate average thresholds over Tx symbols and use them in the long term as the clutter is assumed to be stationary for multiple frames.

As a metric to compare the processing methods fairly, the output SINR and the range covariance matrix of the IT after MIMO processing are defined as:

$$\Sigma^\phi(\tilde{m}, \tilde{\theta}) \triangleq \mathbb{E}\{\mathbf{r} \mathbf{r}^H | H_1\} - \mathbb{E}\{\mathbf{r} \mathbf{r}^H | H_0\},$$

$$\text{SINR} = \frac{\mathbf{u}^H \Sigma^\phi \mathbf{u}}{\mathbf{u}^H \Sigma^\eta \mathbf{u}}.$$

It should be noted that in this section, the filters in (16) and (22) are written for the benchmark knowledge-aided detector. In the later sections, the covariances that these filters use will change depending on the detector structure, but the filter calculations will remain the same.

IV. PROCESSING ALTERNATIVES

In this section, alternative processing methods that are investigated in this article will be explained.

A. Two-variants detector scheme

The proposed processing method is thoroughly explained in the previous section. In practice, calculating the covariances $\mathbf{R}, \Sigma$ and thus the filters $\omega, \mathbf{u}$ according to the explained max-SINR forms can result in a performance degradation in some specific scenarios, specifically when the clutter is close to the IT in angle-domain but separated from the IT in the range domain. This is because the first filter, which is the proposed MIMO filter, tries to suppress the clutter without considering that there will be a range processing which can also suppress the clutter. Since the IT is near the clutter, it would also be suppressed to a level under the noise level, and range processing cannot do anything to increase the SINR. In other words, first stage can lose the useful information that second stage can exploit in order to maximize its own output SINR. Using jointly-optimal range-angle filter can solve this problem easily, but this is not realizable in real-time as explained before. Therefore, we propose another suboptimal detector in order to get rid of this specific problematic case.

We propose a detection scheme in which two different detectors are utilized and their detection results are combined (with or operation). The first set of detectors, identified with $\mathcal{V}_1$, use an interference-aware processing in MIMO domain and conventional IDFT processing in range domain. The other set of detectors, identified with $\mathcal{V}_2$, use conventional DFT processing in MIMO domain and an interference-aware processing in range domain. This way, in both angle and range domains, the max-SINR filters can use all useful information about the interference, not affecting each other. The working principle of this detector scheme is shown in Fig. 3. Details of the interference-aware and conventional processings are given in the following subsections.

B. $\mathcal{V}_1$ detectors

As explained before, we propose to use a two-detector scheme in order to compensate for the possible SINR losses due to non-optimality of the signal processing. The first set of detectors, which we call as $\mathcal{V}_1$ detectors, uses conventional IDFT for range processing. In other words, they select $\Sigma^\eta = I_N$ when calculating the range processing filter $\mathbf{u}$ in (22). For MIMO processing part, three detectors are investigated in this paper. It should be noted that if the beam-squint effect is neglected, the order of MIMO and range processings of $\mathcal{V}_1$ detectors can be changed. In other words, without the beam-squint effects, conventional IDFT in range domain can be applied directly to $\{Z_n\}$ and MIMO processing can be applied to each subcarrier afterwards, without any change in the resultant detection metrics.
1) **KA-SARP Given** $X_n$, $V_1$ detector: This detector utilizes the proposed MIMO filter structure in (16) using the true knowledge of interference angular covariance matrix $R^n_\omega$, which can be supplied to the detector as target maps from trackers. The MIMO filter $\omega_n$ is constructed for each subcarrier and each Tx symbol $X_n$, which requires $D_r D_t \times D_t D_t$ matrix inversions $N$ times per frame. In each subcarrier, the angular filter depends on the transmitted symbols. Therefore, this detector is denoted as “Given $X_n$”. This detector is the only one in this paper that uses different MIMO filters for each subcarrier due to changing Tx symbols $X_n$. This unique property results in an improved angular interference suppression with a high computational cost.

2) **KA-SARP with** Expected $X_n$, $V_1$ detector: As mentioned previously, the true covariances $R$, $\Sigma$ and thus the max-SINR filters $\omega$, $u$ depend on the symbol sequence $X_n$ at each subcarrier. In order to ease the calculation of the filters, expected filters over symbols $X_n$ at each subcarrier can be found. The calculation of the expectations are given in the Appendix. The expectations of $R^n_\omega$ and $\Sigma^n$ over $X_n$ are denoted as $\bar{R}^n_\omega$ and $\bar{\Sigma}^n$, respectively. Since the expectations are taken over symbols, this detector is denoted as “Expected $X_n$”.

Expected $X_n$, $V_1$ detector forms the max-SINR MIMO filter $\omega_n$ in (16) by using $R^n_\omega$ instead of $R^n_\omega$. The angular correlation time is assumed to be much larger than the OFDM frame duration. Therefore, the filters in Expected $X_n$, $V_1$ detector can be calculated only once in several frames.

3) **Fully Adaptive AMF-like SARP with SA, $V_1$ detector:** The previously explained detectors are knowledge-aided, meaning that their performances significantly depend on the true knowledge of the interference covariance matrices. As an alternative processing method, we propose a fully adaptive filter learning the interference covariance using only $T$ frames of data via sample averaging (SA) method. This detector can be considered the application of the well-known AMF detector in MIMO scenarios where clutter consists of multiple and extended scatterers and beam-squint effects are visible. This is why this kind of detector is denoted as “SA-AMF”. In (16), instead of $R^n_\omega$, SA, $V_1$ detector uses the estimated interference angle covariance matrix $\bar{R}^n_\omega$ which was learnt via SA method. Similar to Given $X_n$, $V_1$ detector, the $\omega_n$ filters must be constructed $N$ times per frame. However, if the beam-squint and Doppler mismatch are ignored, $R^n_\omega$ can be assumed to be the same for all subcarriers. Therefore, it might be sufficient to use all subcarriers to learn a single covariance matrix $\bar{R}^n_\omega$ and use it to construct a single MIMO filter $\bar{\omega}$. On the other hand, if the beam-squint is effective, the angular channel of interference would depend on the subcarrier frequency, and a subband approach could be used to ease the calculations.

The working principle of SA-AMF detectors with a subband approach is explained in more detail in Fig. 4. As seen in Fig. 4, SA-AMF detectors divide the spectrum into subbands to mitigate the effects of beam-squint. There are $N/N' = N'$ subcarriers in each of the $N'$ subbands, where $N'$ can be selected so that the beam-squint effect is negligible inside the subbands [27]. $T$ frames of observation are assumed to be taken before the IT is present. The SA-AMF, $V_1$ detector uses these $T$ frames and $N'$ subcarriers to estimate the angular interference covariance matrix $\bar{R}^n_\omega$ of $(n')^{th}$ subband via SA. Then, the same $\omega_{n'}$ is used for all $T$ frames and $N'$ subcarriers in each subband. As the result of the angular processing, there are $T$ vectors to be processed in range domain, $t^{th}$ of them being denoted as $r(t)$. $D^n$ in Fig. 4 represents a diagonal load matrix, which guarantees that the matrix inversion converges. In this study, $D^n$ is proportional to expected noise angle covariance matrix, which is $D^n = \gamma_{load} \bar{R}^n_\omega \text{diag}(I_{D_r} \otimes (W_t^H W_r))$, where $\gamma_{load}$ is the diagonal loading factor.

### C. $V_2$ detectors

The second set of detectors in our two-variants detector scheme, $V_2$ detectors, use conventional DFT for MIMO processing. Namely, they select $R^n_\omega = I_{D_r} D_t$ in (16). For range processing part, three detectors are investigated in this paper.

1) **KA-SARP Given** $X_n$, $V_2$ detector: In this detector, the range filter $u$ in (22) is constructed using the true knowledge of interference range covariance matrix $\Sigma^n$, provided by the tracker. As shown in (25), this range filter is dependent on Tx symbols $X_n$, but there is only one matrix inversion of size $N \times N$ required per frame. Therefore, this detector is computationally less burdensome than Given $X_n$, $V_1$ detector.

2) **KA-SARP with** Expected $X_n$, $V_2$ detector: This detector forms the max-SINR range filter $\hat{u}$ in (22) by using $\bar{\Sigma}^n$ instead of $\Sigma^n$. Similar to Expected $X_n$, $V_1$ detector, the filters can be calculated only once in several frames because the range response of the scatterers are assumed to stay correlated for several OFDM frame durations.

3) **Fully Adaptive AMF-like SARP with SA, $V_2$ detector:** In (22), instead of $\Sigma^n$, SA-AMF, $V_2$ detector uses the estimated interference range covariance matrix $\bar{\Sigma}^n$ which was learnt via SA method. As shown in Fig. 4, $\bar{\Sigma}^n$ is learned in each frame, and then the average of these matrices is used to create the range filter $u$ to be used in all $T$ frames. $D^n$ in Fig. 4 represents a diagonal load matrix, which guarantees that the matrix inversion converges. In this study, $D^n$ is proportional to the expected noise range covariance matrix, whose $(n,n)^{th}$ element is $[D^n]_{(n,n)} = \gamma_{load} \bar{R}_{\omega} \|W_t^H b_n(\theta)\|^2 \|W_r^H a_n(\theta)\|^2$, where $\gamma_{load}$ is the diagonal loading factor.

As explained in (27), Doppler mismatch results in $\Sigma^n$ being dependent on $X_n$. This dependence can be described as element-wise multiplication with a random matrix $A_n$, which is zero-mean when the expectation is taken over $X_n$. To estimate the deterministic part in $\Sigma^n$, which is $Q_{\Sigma}^{n}(q_{\Sigma}^{n})^H$ in (27), the same $X_n$ matrix must be repeated for multiple frames to prevent the range information from being nullified. Therefore, $X_n$ symbols are repeated frame-to-frame in this paper, which has an adverse effect on the communication rate. However, it should be noted that the repetition occurs only in $N'$ subcarriers among all $N'$ subcarriers; and after the training period is finished, the range filters of SA-AMF, $V_2$ detector can be fixed and the communication rate can be recovered.5

5The training of SA-AMF filters is done under $H_0$ hypothesis in this study. However, it can also be done under $H_1$ hypothesis without losing SINR by using an increased number of training symbols.
SA-AMF, $V_1$ detector approaches. Given $X_n$, $V_2$ detector. This is because in SA-AMF, $V_1$, different $X_n$ sequences are used in training, effectively averaging the MIMO filters over $X_n$. On the other hand, in SA-AMF, $V_2$, the same $X_n$ is repeated so that the range covariance is learned as if $X_n$ is given.

D. 3-D Periodogram detector

As a benchmark detector, we also utilize 3-D periodogram, which is commonly used in the literature, as in [11]. This detector uses no knowledge of interference and performs conventional DFT or IDFT in range, angle and Doppler domains.

The calculation of filters for different processing methods, along with the required knowledge and computational complexities of them, are summarized in Table I. For the computational complexities, the most complex operation that is done for each frame is considered for each method. Given $X_n$ methods must calculate the true covariances for each frame. Therefore, they are by far the most complex methods.

V. SIMULATION RESULTS

In this section, the simulation results for the proposed detector and benchmark detectors will be provided.

A. Simulation Parameters

The simulation parameters are provided in Table II. In Table II, (BS) represents the cases where beam-squint effect is severe, which is satisfied by increasing the actual number of subcarriers, $\bar{N}$, 16 times while keeping the number of subcarriers for radar signal processing, $N$, the same. Therefore, effectively, $\Delta f$ is increased and range resolution is decreased 16 times in the BS simulations. There are $K = 10$ UIT scatterers and $U = 2$ UL users. The first UIT is assumed to be a single scatterer with $r_1 = 0$ m, $v_1 = 0$ m/s, $\sigma_c = 0$ Hz and $\gamma_c = 30$ dB, depicting a strong self-interference. The other UIT scatterers are in a cluster, which depicts an object of finite and nonzero length and width. The center range is 18.11 m (10th range bin) and the center angle is $-12$ degrees for this object. The length and width of the object are half of the range and angle resolutions of the system, respectively. We consider that the scattering points are on the edge centers and corners of the object in both range and angle dimensions, as well as the center of it, resulting in 9 scattering points. The velocity, Doppler spread and total reflected power from the object are 20 m/s (63.3 Hz Doppler bin), 200 Hz and 20 dB, respectively. The UL users are at 7.5 and 12.5 degrees, and their reflected powers are 10 dB each. The UL users are assumed to have no velocity or Doppler spread. To ease the understanding of the graphs, red dashed and blue dash-dotted vertical lines are drawn wherever the UITs and SoIs are located, respectively.

B. Beamforming Scenarios

Two different beamforming schemes are simulated in this study. In the first scenario, both Tx and Rx beams are formed to cover the angle interval of $(-7.5, 7.5)$ degrees, which is called as fully overlapping (FO) scheme. In the second scenario, Tx beamformer covers the angle interval of $(-20, -5)$ degrees and Rx beamformer covers the angle interval of $(5, 20)$ degrees, which is called as no overlapping (NO) scheme. The Tx and Rx beamforming gain patterns, and their multiplication depicting the total beamforming gain pattern, are provided in Fig. 5. Tx and Rx beamforming gain pattern values at the $n^{th}$ subcarrier for angle $\theta$ is equal to $\text{tr}\{a_n(\theta)^H W_t W_r^H a_n(\theta)\}$ and $\text{tr}\{b_n(\theta)^H W_t W_r^H b_n(\theta)\}$, respectively. It should be noted that the gain is $N_t/D_t$ inside the flat Tx sector because a total Tx power constraint is assumed. On the other hand, the whole $N_r$ beamforming gain is achieved on the receiver side.
**TABLE I**

<table>
<thead>
<tr>
<th>Name of processing method</th>
<th>MIMO Filter</th>
<th>Range Filter</th>
<th>Knowledge Required</th>
<th>Computational Complexity per CUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given $X_n$, $\mathcal{V}_1$</td>
<td>$\omega_n(R_n^2)$</td>
<td>$u(I_N)$</td>
<td>$X_n, D_k^2, \theta_k, R_k^2$</td>
<td>$O(NM^2D_k^2D_f^2T)$</td>
</tr>
<tr>
<td>Given $X_n$, $\mathcal{V}_2$</td>
<td>$\omega_n(I_{D_1,D_2})$</td>
<td>$u(\Sigma_i^2)$</td>
<td>$X_n, D_k^2, \theta_k, R_k^2, \tau_k$</td>
<td>$O(N^2M^2D_k^2D_f^2T)$</td>
</tr>
<tr>
<td>Expected, $\mathcal{V}_1$</td>
<td>$\omega_n(R_n^2)$</td>
<td>$u(I_N)$</td>
<td>$D_k^2, \theta_k, R_k^2$</td>
<td>$O(N^2D_k^2D_f^2T)$</td>
</tr>
<tr>
<td>Expected, $\mathcal{V}_2$</td>
<td>$\omega_n(I_{D_1,D_2})$</td>
<td>$u(\Sigma_i^2)$</td>
<td>$D_k^2, \theta_k, R_k^2, \tau_k$</td>
<td>$O(N^2D_k^2D_f^2T)$</td>
</tr>
<tr>
<td>SA-AMF, $\mathcal{V}_1$</td>
<td>$\omega_n(R_n^2)$</td>
<td>$u(I_N)$</td>
<td>$-$</td>
<td>$O(N^2D_k^2D_f^2T)$</td>
</tr>
<tr>
<td>SA-AMF, $\mathcal{V}_2$</td>
<td>$\omega_n(I_{D_1,D_2})$</td>
<td>$u(\Sigma_i^2)$</td>
<td>$-$</td>
<td>$O(N^2D_k^2D_f^2T)$</td>
</tr>
<tr>
<td>3D Perio</td>
<td>$\omega_n(I_{D_1,D_2})$</td>
<td>$u(I_N)$</td>
<td>$-$</td>
<td>$O(N^2D_fD_T)$</td>
</tr>
</tbody>
</table>

**TABLE II**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N, N'$</td>
<td>1024, 32</td>
<td>$M, T$</td>
<td>16, 16</td>
</tr>
<tr>
<td>$N, N'$</td>
<td>32</td>
<td>$D_f, D_T$</td>
<td>10</td>
</tr>
<tr>
<td>$T_s$</td>
<td>12.38 $\mu$s</td>
<td>$T_f$</td>
<td>198 $\mu$s</td>
</tr>
<tr>
<td>$\Delta f$</td>
<td>2.59 MHz</td>
<td>$W$</td>
<td>82.75 MHz</td>
</tr>
<tr>
<td>$\Delta f$ (BS)</td>
<td>41.37 MHz</td>
<td>$W$ (BS)</td>
<td>1.32 GHz</td>
</tr>
<tr>
<td>$f_c$</td>
<td>24 GHz</td>
<td>$\lambda_c$</td>
<td>12.5 mm</td>
</tr>
<tr>
<td>$\sigma_n, \sigma_t$</td>
<td>1</td>
<td>$\gamma_n$</td>
<td>10 dB</td>
</tr>
<tr>
<td>$\gamma_L^2$</td>
<td>30, 20 dB</td>
<td>$\gamma_{load}$</td>
<td>3</td>
</tr>
</tbody>
</table>

**C. Simulation Results**

In this section, output SINR values for five different detectors will be given for different values of CUT angle, range and Doppler bins. In addition, $P_d$ vs input SNR curves for two different CUT locations are also provided for these detectors. The first detector uses ‘Given $X_n$’ processing for both $\mathcal{V}_1$ and $\mathcal{V}_2$, and it is labeled as ‘$\mathcal{V}_1\&\mathcal{V}_2$ Given $X_n$’ in the legends. Since this detector uses $X_n$ data for both variants, its SINR is expected to be the upper limit for all other detectors. The second detector uses ‘Expected $X_n$’ processing for both $\mathcal{V}_1$ and $\mathcal{V}_2$, and it is labeled as ‘$\mathcal{V}_1\&\mathcal{V}_2$ Expected $X_n$’ in the legends. Since this detector does not use $X_n$ data for any variants, it is expected to suffer from the Doppler mismatch effect mentioned before, but its filters must be calculated only once per several frames. The third detector uses ‘Given $X_n$’ processing for $\mathcal{V}_1$ but ‘Given $X_n$’ processing for $\mathcal{V}_2$, and it is labeled as ‘$\mathcal{V}_1$ Expected $X_n\&\mathcal{V}_2$ Given $X_n$’ in the legends. This detector does require $X_n$ data for range processing but uses expected filters for angle processing. Since the Doppler mismatch effect is more severe in range rather than angle processing, this detector is expected to overcome this difficulty while its complexity is reduced with respect to the ‘$\mathcal{V}_1\&\mathcal{V}_2$ Given $X_n$’ detector. The fourth detector uses ‘SA-AMF’ processing for both $\mathcal{V}_1$ and $\mathcal{V}_2$, and it is labeled as ‘$\mathcal{V}_1\&\mathcal{V}_2$ SA-AMF’ in the legends. This detector adaptively learns both angle and range covariance matrices. Since its learning inherently includes the effects of $X_n$, this detector is expected to overcome the Doppler mismatch problem, and its MIMO filters can be calculated only once after the training period of several frames. The last detector uses ‘conventional DFT,’ or ‘3-dimensional periodogram’ processing for both $\mathcal{V}_1$ and $\mathcal{V}_2$, and it is labeled as ‘3-D Perio’ in the legends. This detector does not use any information about the interference or $X_n$ and is expected to perform the worst among others in general. For conventional DFT and IDFT processings in angle and range domains, Chebyshev windows are used with sidelobe suppression values of 50 dB and 100 dB, respectively. Since all detectors use conventional DFT or IDFT processing at one of their variants, they suffer from a windowing loss of around 3.2 dB. The detectors other than 3-D Perio can bypass windowing without losing sidelobe performances, but we kept the windows in the simulations for the sake of fairness.

SINR vs CUT angle graphs for two different scenarios are provided in Fig. 6. The top graph in Fig. 6 represents a FO case where both a UIT and two Sols are inside the
overlapping Tx-Rx sectors. It is seen that all detectors except 3-D Perio perform similarly when the CUT angle is away from UITs or SoIs. However, when CUT is on top of the UIT in both angle and range, $V_1 \& V_2$ Expected $X_n$ and 3-D Perio detectors fail to suppress the interference while other detectors can. This region is where the only difference between the IT and the UIT is in Doppler domain. However, as explained before, $X_n$ knowledge is required to suppress the interference using the Doppler diversity. Therefore, the detectors that know (or learn) $X_n$ information can suppress the UIT while other detectors fail to reach a high SINR. An important aspect of the clutter suppression here is that all detectors use the same Doppler preprocessing. However, the Doppler mismatch shows itself in both angle and range covariance matrices, and the clutter can be suppressed in Doppler domain with MIMO and range processing. The only difference between $V_1 \& V_2$ Expected $X_n$ and $V_1$ Expected $X_n$& $V_2$ Given $X_n$ detectors is the range processing part, but the SINR difference is more than 60 dB between them at CUT angle 0 degree. Similarly, the only difference between $V_1$ Expected $X_n$& $V_2$ Given $X_n$ and $V_1$& $V_2$ Given $X_n$ detectors is the MIMO processing part, and the SINR difference is around 10 dB between them at CUT angle 0 degree. This shows that Doppler mismatch effect is more severe in the range rather than the angle processing. The bottom graph in Fig. 6 represents a NO case, where Tx and Rx sectors are separated. It can be seen that the importance of good MIMO processing is more visible when the sectors are separated. When the CUT angle is near (but not on top of) the UIT, $V_1 \& V_2$ Given $X_n$ detector can outperform the closest detector by around 7 dB. The effects of Doppler mismatch on both range and angle processings are still visible in this case.

$X_n$ vs CUT range bin graphs for FO and NO scenarios are provided in Fig. 7 and Fig. 8, respectively. Similar to the case in Fig. 6, when the CUT angle is away from the interference, all detectors except 3-D Perio perform similarly, and when the CUT is on top of the interference, using $X_n$ knowledge is important. In Fig. 8, it is seen that $V_1 \& V_2$ Given $X_n$ and $V_1$ Expected $X_n$& $V_2$ Given $X_n$ detectors perform the same for most CUT ranges, only except on top of the UIT range. This shows that $V_1$ variant performs better than $V_2$ variant only when the CUT range is so close to UIT range. This is expected because $V_1$ uses max-SINR MIMO processing and $V_2$ uses max-SINR range processing. Fig. 8 shows that this 2-variant detector scheme benefits the scenarios where the interference can only be suppressed in one dimension.

$X_n$ vs CUT Doppler bin graphs for two different CUT angles for a FO scenario are provided in Fig. 9. In the top graph of Fig. 9, all detectors except 3-D Perio perform similarly as in the other graphs before. In the bottom graph of Fig. 9, it is seen that when the suppression in angle domain is impossible, the suppression in the range domain depends on the CUT Doppler. When the CUT Doppler is the same as the UIT Doppler, range processing can successfully suppress the clutter for all detectors. However, when there is Doppler mismatch, the detectors that do not use $X_n$ information fail to suppress the UIT even if it is separated in range domain. This clearly shows that Doppler mismatch can harm range
processing if $X_n$ information is not exploited.

SINR vs CUT Doppler bin graphs for the scenarios in which CUT angle and range bins are on top of UITs are provided in Fig. 10. When the CUT is on top of any UUT in both angle and range domains, the UIT can be suppressed only in Doppler domain. Therefore, the detectors using $X_n$ information outperform the others when the CUT Doppler is separated from the UIT Doppler. It can also be seen that Doppler spread of the UIT is also important and determines how much separation in Doppler is needed to create diversity and suppress the UIT.

The beam-squint effect when $f_c/W$ ratio is relatively small is investigated in Fig. 11. In the top graph of Fig. 11, the detectors do not care about the beam-squint effect and use only a single subband, applying the same angular processing for all subcarriers. It can be seen that the performance of the detectors decreases due to the misinformation on MIMO processing caused by the beam-squint effect. On the other hand, $V_1 \& V_2$ given $X_n$ detector tries to put a sharp null on a wrong angle and therefore its SINR cannot be interpreted as an upper limit to other detectors anymore. Besides this, SA-AMF detector learns the angle information from all subcarriers, which all have different angle information due to the beam-squint effect, and therefore tries to nullify a wider set of angles than it should be. When the bottom graph of Fig. 11 is investigated, it can be seen that subband approach can be used to mitigate the negative effects of beam-squint. When $N' = 8$, the MIMO processing filters consider different amounts of suppression for 8 different angles, therefore the UIT can be suppressed better for both detectors.

$P_d$ vs input SNR curves for two different CUT locations are provided in Fig. 12. In this figure, input SNR means the SNR value per subcarrier per antenna per symbol. For each detector, the CFAR threshold is determined by the formula $\gamma_{th} = -\ln(P_{fa}) \beta$ where $\beta$ is the detection metric at the detector input when $H_0$ hypothesis is true. Theoretically, this metric is equal to the denominator of the SINR expression in (30) for each detector. Since the covariance matrix and the filters in the expression depend on $X_n$, the theoretic threshold also depends on $X_n$. However, for successful detectors, interference is suppressed so that the $X_n$ dependence of the thresholds is weak. On the other hand, adaptive CFAR thresholding using multiple training frames with different $X_n$ data can also be used to average out the effects of $X_n$ on the thresholds. A similar adaptive thresholding method is explained in [29]. In Fig. 12, average thresholds are found for $P_{fa} = 10^{-3}$ and 300 Monte-Carlo trials are conducted to find $P_{fa}$ values. It can be seen in Fig. 12 that $P_d$ vs input SNR curves give similar results to SINR curves, as expected.

VI. CONCLUSION

This paper proposed a radar detector structure for a MIMO OFDM ISAC system under the disruptive effects of Doppler mismatch, beam-squint, UL users and multiple scatterers in
fractional Doppler-angle-range bins. The introduction of the DACB and coupling between Doppler mismatch and Tx symbols was crucial in understanding the nature of the MIMO OFDM ISAC systems. The sub-optimal SARP method proved effective in maximizing the output SINR, with the help of a novel two-variant detector scheme. On the other hand, the possible beam-squint effect dictated that MIMO processing should be done before range processing in SARP. The proposed subband approach effectively addressed the beam-squint effect. Simulation results confirmed the superior performance of our proposed detectors compared to conventional methods, highlighting their potential in enhancing MIMO ISAC system performance for 6G systems.

REFERENCES


APPENDIX

CALCULATION OF AUTOCORRELATION MATRICES

A. Transmitted symbols are given

When the transmitted symbols \( X_n \) are known, \( R_n^\psi (\hat{m}) \) in (19) can be written as:

\[
R_n^\psi = \mathbb{E} \left\{ \text{vec}(\Psi_n) \ \text{vec}(\Psi_n)^H \right\} \\
= \sum_{k=1}^{K} \mathbb{E} \left\{ \text{vec}(C_{kn}) \ \text{vec}(C_{kn})^H \right\} \\
+ \sum_{u=1}^{U} \mathbb{E} \left\{ \text{vec}(S_{un}) \ \text{vec}(S_{un})^H \right\} \\
+ \mathbb{E} \left\{ \text{vec}(N_n) \ \text{vec}(N_n)^H \right\} , \quad (31)
\]

We can calculate clutter, Sol and noise parts separately. Let’s define:

\[
R_{kn}^c \triangleq \mathbb{E} \left\{ \text{vec}(C_{kn}) \ \text{vec}(C_{kn})^H \right\} , \\
R_{un}^s \triangleq \mathbb{E} \left\{ \text{vec}(S_{un}) \ \text{vec}(S_{un})^H \right\} , \\
R_{nn}^n \triangleq \mathbb{E} \left\{ \text{vec}(N_n) \ \text{vec}(N_n)^H \right\} . \quad (32)
\]

Following the definition in (10),

\[
\text{vec}(C_{kn}) = \text{vec}(H_{kn} X_n D_k^C \Lambda_k^C) \\
= (I_M \otimes H_{kn}) \text{vec}(X_n D_k^C \Lambda_k^C) \\
= (I_M \otimes H_{kn})(I_M \otimes (X_n D_k^C)) \text{vec}(\Lambda_k^C) \\
= (I_M \otimes H_{kn})(I_M \otimes (X_n D_k^C)) \alpha_k \quad (33)
\]

where in the last equality, column-wise Kronecker product (Khatri-Rao product) is used to separate the diagonal elements of \( \Lambda_k^C \) as a column vector \( \alpha_k \). It should be noted that \( K_{kn}^c \) represents the MIMO channel for \( k^{th} \) UUT and \( L_{kn}^c \) represents the temporal code of the \( k^{th} \) UUT. The MIMO channel is not dependent on the symbols \( X_n \), which is expected, so it is deterministic and it repeats itself through the Kronecker product operation because it is assumed to be the same for \( M \) symbol duration. The temporal code actually depends on the transmitted symbols \( X_n \) but since the symbols are assumed to be given, \( L_{kn}^c \) also becomes deterministic. Then,

\[
R_{kn}^c = K_{kn}^c L_{kn}^c \mathbb{E} \left\{ \alpha_k \alpha_k^H \right\} (L_{kn}^c)^H (K_{kn}^c)^H \\
= K_{kn}^c L_{kn}^c R_k^c (L_{kn}^c)^H (K_{kn}^c)^H. \quad (34)
\]

Similarly,

\[
\text{vec}(S_{un}) = \text{vec}(g_{un} j^n_u D_u^S \Lambda_u^S) \\
= (I_M \otimes g_{un}) \text{vec}(j^n_u D_u^S \Lambda_u^S) \\
= (I_M \otimes g_{un})(I_M \otimes (j^n_u D_u^S)) \text{vec}(\Lambda_u^S) \\
= (I_M \otimes g_{un})(I_M \otimes (j^n_u D_u^S)) \beta_u \quad (35)
\]

and

\[
R_{un}^s = K_{un}^S \mathbb{E} \left\{ \beta_u \beta_u^H \right\} (L_{un}^S)^H (K_{un}^S)^H \\
= K_{un}^S \mathbb{E} \left\{ \beta_u \beta_u^H \right\} (L_{un}^S)^H (K_{un}^S)^H. \quad (36)
\]

The last expectation in (36) is over \( j_{un} \) because Sol signals are always assumed to be random, even if \( X_n \) are given. On the other hand, since \( j_{un}^T D_u^S \) is a row vector, \( L_{un}^S = \text{diag}(j_{un}) D_u^S \) and

\[
R_{un}^s = K_{un}^S \mathbb{E} \left\{ \text{diag}(j_{un}) D_u^S \beta_u \text{diag}(j_{un}) D_u^S \right\}^H (K_{un}^S)^H \\
= K_{un}^S \mathbb{E} \left\{ j_{un}^T j_{un}^H \right\} \otimes D_u^S R_u^S (D_u^S)^H (K_{un}^S)^H \\
= K_{un}^S \left( \frac{1}{M} I_M \otimes D_u^S \right) R_u^S (D_u^S)^H (K_{un}^S)^H \\
= \frac{1}{M} K_{un}^S \mathbb{E} \left\{ j_{un}^T j_{un}^H \right\} (K_{un}^S)^H \\
= \gamma_u^S (I_M \otimes g_{un})(I_M \otimes g_{un})^H \\
= \gamma_u^S \frac{I}{M} (I_M \otimes g_{un} g_{un})^H. \quad (37)
\]

The noise is assumed to be uncorrelated both spatially and temporally before the receive beamforming. After the beamforming, the noise covariance matrix becomes \( R_{nn}^n = \sigma_n^2 (I_M \otimes (W_r^H W_r)) \). However, \( W_r \) is chosen such that \( W_r^H W_r = I_{D_r} \) is satisfied, therefore \( R_{nn}^n = \sigma_n^2 I_{M D_r} \) after receive beamforming. Finally, \( R_{kn}^\psi \) in (31) becomes

\[
R_{kn}^\psi = \sum_{k=1}^{K} R_{kn}^c + \sum_{u=1}^{U} R_{un}^s + R_{nn}^n, \quad (38)
\]

which is consistent with (19).

When the transmitted symbols \( X_n \) are known, \( R_{n_1 n_2}^\psi (\hat{m}) \) in (20) can be written as:

\[
R_{n_1 n_2}^\psi = \mathbb{E} \left\{ \text{vec}(\Psi_{n_1}) \ \text{vec}(\Psi_{n_2})^H \right\} \\
= \sum_{k=1}^{K} \mathbb{E} \left\{ \text{vec}(C_{kn_1}) \ \text{vec}(C_{kn_2})^H \right\} \\
+ \sum_{u=1}^{U} \mathbb{E} \left\{ \text{vec}(S_{un_1}) \ \text{vec}(S_{un_2})^H \right\} \\
+ \mathbb{E} \left\{ \text{vec}(N_{n_1}) \ \text{vec}(N_{n_2})^H \right\} . \quad (39)
\]

We can again calculate clutter, Sol and noise parts separately. Let’s define:

\[
R_{kn_{12}}^c \triangleq \mathbb{E} \left\{ \text{vec}(C_{kn_1}) \ \text{vec}(C_{kn_2})^H \right\} , \\
R_{un_{12}}^s \triangleq \mathbb{E} \left\{ \text{vec}(S_{un_1}) \ \text{vec}(S_{un_2})^H \right\} , \\
R_{nn_{12}}^n \triangleq \mathbb{E} \left\{ \text{vec}(N_{n_1}) \ \text{vec}(N_{n_2})^H \right\} . \quad (40)
\]

Following the derivation in (33),

\[
R_{kn_{12}}^c = K_{kn_1}^C L_{kn_1}^C R_k^C (L_{kn_2}^C)^H (K_{kn_2}^C)^H. \quad (41)
\]

On the other hand, Sol symbols are assumed to be uncorrelated among both users and subcarriers, namely,

\[
\mathbb{E} \left\{ j_{u_1 u_2}^H \right\} = \frac{1}{M} I_M [\delta(u_1 - u_2) \delta(n_1 - n_2)]. \quad (42)
\]

Therefore, following the derivation in (37),

\[
R_{un_{12}}^s = K_{un_1}^S \mathbb{E} \left\{ j_{un_1}^H \right\} \otimes D_u^S R_u^S (D_u^S)^H (K_{un_2}^S)^H \\
= \frac{1}{M} \gamma_u^S (I_M \otimes g_{un_1} g_{un_2}). \quad (43)
\]
\[ [E_c^\tau](r_1, r_2) = \sum_{m_1, m_2} \sum_{t_1, t_2} [R_k^\tau]_{(m_1, m_2)} [d_{m_1, m_2}[H_{kn}(r_1, t_1)][H_{kn}]^*_{(r_2, t_2)} \times \mathbb{E}\{[X_n^\tau](t_1, m_1)[X_n](t_2, m_2)[X_n^\tau](t_1, m_1)[X_n^\tau](t_2, m_2)] \} \]

(45)

Then, \((r_1, r_2)\)th element of \(E_c^\tau\) can be calculated as in (45), where the scalar parameter \(d_{m_1, m_2}\) about the Doppler mismatches is defined as:

\[ d_{m_1, m_2} = \delta[D_m^*D_k^\tau]_{(m_1, m_2)}[D_m^*D_k^\tau]_{(m_2, m_2)}. \]

(51)

The expectation \(\mu\) in (45) includes four random variables which are assumed to be zero-mean and iid. Therefore, \(\mu\) is nonzero only for some elements in the summation. Let's call each element of \(X_n\) as \(\tilde{x}\), a zero-mean complex random variable. Then, let's \(\mathbb{E}\{\tilde{x}\} = 1/M\) and \(\mathbb{E}\{\tilde{x}^4\} \equiv \sigma_4\). Then, \(\mu\) can be written as:

\[
\mu = \begin{cases} 
1/M^2, & m_1 = m_2, \quad t_1 = t_2, \quad l_1 = l_2, \quad \delta \neq \delta_1 \\
1/M^2, & m_1 = m_2, \quad t_1 = t_2, \quad l_1 = l_2, \quad \delta = \delta_1 \\
1/M^2, & m_1 \neq m_2, \quad t_1 = t_2, \quad l_1 \neq l_2, \quad \delta = \delta_1 \\
1/M^2, & m_1 \neq m_2, \quad t_1 = t_2, \quad l_1 = l_2, \quad \delta = \delta_1 \\
0, & \text{otherwise}
\end{cases}
\]

(52)

For the SoI part inside \(R_c^\tau\), it is not necessary to begin investigation from the \((t_1, t_2)\)th block because the calculations are easier. Following the (18) and (19), the expectation of SoI part inside \(R_c^\tau\) over \(X_n\) can be found as:

\[
\mathbb{E}\left\{[V_n]^T \left[\sum_{u=1}^{\mu} \frac{\gamma_u}{M} (I_M \otimes g_{sun}) (V_n \otimes I_{D_r}) \right] V_n \right\} = \mathbb{E}\left\{[V_n]^T \left[\sum_{u=1}^{\mu} \frac{\gamma_u}{M} (V_n^H \otimes I_{D_r}) (I_M \otimes g_{sun})^H (V_n \otimes I_{D_r}) \right] \right\} = \mathbb{E}\left\{[V_n]^T \left[\sum_{u=1}^{\mu} \frac{\gamma_u}{M} (V_n^H V_n) \otimes (g_{sun})^H \right] \right\} \approx \sum_{u=1}^{\mu} \frac{\gamma_u}{M} I_{D_r} \otimes (g_{sun}).
\]

(53)

Then, the \((t_1, t_2)\)th block of \(R_c^\tau\) corresponding to \(u\)th user is seen to be written as:

\[
E_{ut, t_2} = \mathbb{E}_{ut, t_2} \approx \delta[t_1 - t_2] \frac{\gamma_u}{M} g_{sun} g_{sun}^H.
\]

(54)

For the noise part, the calculations seem unnecessary. The \((t_1, t_2)\)th block of \(R_c^\tau\) corresponding to noise is \(\delta[t_1 - t_2] \sigma_1^2 I_{D_r}\). Finally, the \((t_1, t_2)\)th block of \(R_c^\tau\) is:

\[
E_{t_1, t_2} = \sum_{k=1}^{K} E_{v_1, v_2} + \sum_{u=1}^{\mu} E_{ut, t_2} + \delta[t_1 - t_2] \sigma_1^2 I_{D_r} \quad \text{for} \quad t_1, t_2 = 1, 2, \ldots, D_r.
\]

(55)

Similar to \(R_c^\tau\), it is computationally burdensome to calculate \(R_n^\tau\) for each transmitted symbol sequence. Therefore, its expectation over \(X_n\) and \(X_{n}^\tau\) can be taken to reduce the computational complexity. In the following calculations, the symbols transmitted in different subcarriers are considered to be uncorrelated, the expectations can be separately calculated for returns from different scatterers and SoIs are assumed to be uncorrelated, the expectations can be separately calculated for returns from different scatterers and SoIs are assumed to be uncorrelated.
be independent from each other, so that the double expectations can be separated. Let’s define
\[
\mathbf{R}^\eta_{n_1 n_2} \triangleq \mathbb{E}_{X_{n_1}} \{ \mathbb{E}_{X_{n_2}} \{ \mathbf{R}^\eta_{n_1 n_2} \} \}. \tag{56}
\]
Following the definition in (26),
\[
\mathbf{R}^\eta_{n_1 n_2} = \sum_{k=1}^K \mathbb{E} \left\{ \mathbf{V}_{n_1} \mathbf{K}^C_{\mathbf{k}n_1} \mathbf{L}^C_{\mathbf{k}n_2} \mathbf{R}^\eta_{\mathbf{k}} (\mathbf{L}^C_{\mathbf{k}n_2})^H (\mathbf{K}^C_{\mathbf{k}n_2})^H \mathbf{V}^H_{n_2} \right\}
\]
\[
+ \mathbb{E} \{ \mathbf{V}_{n_1} \} \{ \mathbf{R}^\eta_{n_1 n_2} + \mathbf{R}^\eta_{n_1 n_2} \} \mathbb{E} \{ \mathbf{V}^H_{n_2} \}
\]
\[
= \sum_{k=1}^K \mathbb{E} \left\{ \mathbf{V}_{n_1} \mathbf{K}^C_{\mathbf{k}n_1} \mathbf{L}^C_{\mathbf{k}n_2} \mathbf{R}^\eta_{\mathbf{k}} (\mathbf{L}^C_{\mathbf{k}n_2})^H (\mathbf{K}^C_{\mathbf{k}n_2})^H \mathbf{V}^H_{n_2} \right\}
\]
\[
+ \delta[n_1 - n_2] \mathbb{E} \{ \mathbf{V}_{n_1} \} \{ \mathbf{R}^\eta_{n_1 n_2} + \mathbf{R}^\eta_{n_1 n_2} \} \mathbf{V}^H_{n_1} \}
\tag{57}
\]
because \( \mathbf{R}^\eta_{n_1 n_2} \) in (43) includes \( \delta[n_1 - n_2] \). Similar to (48), we can define a new matrix
\[
\mathbf{F}^C_{\mathbf{k}n_1} (\mathbf{m}) \triangleq \mathbb{E} \{ \mathbf{V}^H_{n_1} \mathbf{K}^C_{\mathbf{k}n_1} \mathbf{L}^C_{\mathbf{k}n_1} \}
\]
\[
= \mathbb{E} \{ (\mathbf{V}^H_{n_1} \otimes \mathbf{H}_{\mathbf{k}n_1}) (\mathbf{I}_M \otimes (\mathbf{X}_{\mathbf{n}_1} \mathbf{D}^C_{\mathbf{k}})) \}
\]
\[
\approx \mathbb{E} \{ (\mathbf{X}^*_n \mathbf{D}^*_n \otimes \mathbf{H}_{\mathbf{k}n_1}) (\mathbf{I}_M \otimes (\mathbf{X}_{\mathbf{n}_1} \mathbf{D}^C_{\mathbf{k}})) \}, \tag{58}
\]
using the approximation in (50). \( \mathbf{F}^C_{\mathbf{k}n_1} \) is a \( D_r D_t \times M \) matrix. Let’s consider it as a matrix consisting of \( D_r \times 1 \) blocks, where these blocks are indexed by \( t_1 \) and \( m_1 \). Then, \( (t_1, m_1) \)th block of \( \mathbf{F}^C_{\mathbf{k}n_1} \) can be defined as:
\[
f^C_{t_1, m_1} \triangleq [\mathbf{F}^C_{\mathbf{k}n_1}]_{[t_1, m_1]} = \mathbb{E} \{ v_{t_1, m_1} \mathbf{H}_{\mathbf{k}n_1} \mathbf{v}_{n_1, m_1} \} \tag{59}
\]
where \( v_{t_1, m_1} \) and \( \mathbf{v}_{n_1, m_1} \) are defined the same way as in (48). The \( r \)th element of \( D_r \times 1 \) vector \( f^C_{t_1, m_1} \) can be written as:
\[
[f^C_{t_1, m_1}]_{(r_1)} = \mathbb{E} \left\{ \sum_{t_2=1}^{D_t} v_{t_1, m_1} [\mathbf{H}_{\mathbf{k}n_1}]_{[r_1, t_1]} [\mathbf{v}_{n_1, m_1}]_{(t_2)} \right\}
\]
\[
= \sum_{t_2=1}^{D_t} \mathbb{E} \left\{ [\mathbf{X}^*_n \mathbf{D}^*_m]_{[t_1, m_1]} [\mathbf{H}_{\mathbf{k}n_1}]_{[r_1, t_1]} [\mathbf{X}_{\mathbf{n}_1} \mathbf{D}^C_{\mathbf{k}}]_{(t_2, m_1)} \right\}
\]
\[
= \mathbb{E} \{ [\mathbf{X}^*_n \mathbf{m}]_{[t_1, m_1]} \} [\mathbf{D}^*_m]_{[m_1, m_1]} [\mathbf{D}^C_{\mathbf{k}}]_{[m_1, m_1]} [\mathbf{H}_{\mathbf{k}n_1}]_{[r_1, t_1]}
\]
\[
= \frac{1}{M} [\mathbf{D}^*_m \mathbf{D}^C_{\mathbf{k}}]_{[m_1, m_1]} [\mathbf{H}_{\mathbf{k}n_1}]_{[r_1, t_1]} \tag{60}
\]
where the third equality comes from the fact that transmitted symbols from different antennas are uncorrelated, namely \( \mathbb{E} \{ [\mathbf{X}^*_n \mathbf{m}]_{[t_1, m_1]} [\mathbf{X}_{\mathbf{n}_1}]_{(t_2, m_1)} \} = 0 \) for all \( t_2 \neq t_1 \). Then, the vector \( f^C_{t_1, m_1} \) becomes
\[
f^C_{t_1, m_1} = \frac{1}{M} [\mathbf{D}^*_m \mathbf{D}^C_{\mathbf{k}}]_{[m_1, m_1]} [\mathbf{H}_{\mathbf{k}n_1}]_{[r_1, t_1]} \tag{61}
\]
By some algebraic manipulations,
\[
\mathbf{F}^C_{\mathbf{k}n_1} = \frac{1}{M} \text{vec} \{ \mathbf{H}_{\mathbf{k}n_1} \} \mathbf{I}^T \mathbf{D}^*_m \mathbf{D}^C_{\mathbf{k}} \tag{62}
\]
where \( \mathbf{1} \) is the all-ones vector of proper size. Then,
\[
\mathbf{R}^\eta_{n_1 n_2} \approx \begin{cases} 
\mathbf{R}^\eta_{n_1 n_2}, & \text{if } n_1 = n_2 \\
\sum_{k=1}^K \mathbf{F}^C_{\mathbf{k}n_1} \mathbf{R}^\eta_{\mathbf{k}} (\mathbf{F}^C_{\mathbf{k}n_2})^H, & \text{if } n_1 \neq n_2 
\end{cases} \tag{63}
\]