IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 27 September 2023, accepted 21 October 2023, date of publication 25 October 2023, date of current version 1 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3327439

RESEARCH ARTICLE

An ADC-Aware Receiver Design for Multi-User MIMO Underlay System With Strong Cyclostationary Legacy Signal

JOOHEE CHAE AND JOON HO CHO, (Member, IEEE)
Department of Electrical Engineering (EE), Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea

Corresponding author: Joon Ho Cho (jcho@postech.ac.kr)

This work was supported by the Electronics and Telecommunications Research Institute (ETRI) Grant funded by the Information and Communication Technology (ICT) Research and Development Program of The Ministry of Science and ICT (MSIT)/Institute for Information and Communication Technology Planning and Evaluation (IITP) (Development of Adaptive Interference Reduction Technology Based on Civil-Military Shared Frequency Environment) under Grant 2022-0-00024.

ABSTRACT In this paper, we design a receiver (RX) for a multi-user multiple-input multiple-output (MU-MIMO) underlay system in the presence of a strong cyclostationary legacy signal. Each desired signal is chosen to be a single-carrier signal that exhibits cyclostationarity with the same cycle period as the legacy signal. If a conventional RX is used that does not take into account the effect of the analog-to-digital converter (ADC), then the strong interference from the legacy transmitter (TX) may cause quantization noise to dominate the desired signals from the underlay TXs and severely degrade the reception performance. To address this issue together with the cyclostationarity of the signals, we propose an ADC-aware RX that adopts an analog-digital hybrid combiner structure. The analog combiner first suppresses the interference power to a negligible level through the projection onto an approximate null space of the interference correlation matrix. Then, it enhances the total power of desired signals by applying the principal component analysis on the signal correlation matrix. After the output of the analog combiner is sampled and quantized at the ADCs, the digital combiner generates the linear minimum mean-squared error estimates of the symbols transmitted by the underlay TXs. During the symbol estimation, it fully exploits the periodic spectral correlation induced by the cyclostationarity of signals. Simulation results show that our proposed RX effectively maximizes signal-to-interference-plus-noise ratio and, under certain conditions, approaches the performance achievable in the absence of the legacy TX.

INDEX TERMS Cyclostationary signal, hybrid combiner, linear MMSE, MU-MIMO, quantization noise, strong interference, underlay spectrum sharing.

I. INTRODUCTION

The demand for radio spectrum is rapidly growing with widespread use of the variety of wireless systems [1]. Given both the limited availability of frequency resources and the need to effectively support the wireless systems, spectrum sharing that enables different systems to coexist in the same frequency band has emerged as an essential strategy [2], [3], [4]. In this paper, we are interested in non-cooperative spectrum sharing [5], [6], [7], [8] rather than the cooperative one [9], [10].

The associate editor coordinating the review of this manuscript and approving it for publication was Olutayo O. Oyerinde.

In non-cooperative spectrum sharing scenarios, interference suppression is crucial for achieving desirable reception performance [5]. In particular, when an underlay system attempts to utilize a frequency band already in use by a non-cooperative legacy system, the underlay receiver (RX) needs to suppress the interference caused by the legacy transmitter (TX). The task becomes more challenging when the legacy system is designed for high-power and long-range transmission, such as military radar for target detection and ranging [6], [7], [8] or fixed wireless access (FWA) for broadband connectivity in rural areas [11], [12]. In such cases, the underlay RX may be exposed to a strong interfering signal, whose received power is

significantly higher than those of the desired signals from the underlay TXs.

Man-made interfering signals often exhibit cyclostationarity [13], [14], [15], [16]. When the legacy system is a digital communication system employing linear modulation with proper-complex symbols, such as QAM symbols, the interfering signal received by the underlay RX is better modeled as a wide-sense cyclostationary (WSCS) random process rather than a wide-sense stationary (WSS) random process. The most important characteristic of a WSCS signal is that it exhibits non-zero periodic correlation in the frequency domain [14]. By noticing this, [17], [18], [19] have reported communication- and information-theoretic results for the spectrum sharing with WSCS legacy signals. In [17], it is shown that, if an underlay TX transmits a WSCS signal with the same cycle period as a legacy TX, the underlay system can share the spectrum without interfering the legacy system. In [18] and [19], it is also shown that the optimal input achieving the channel capacity in the presence of a WSCS Gaussian interference is a WSCS Gaussian random process with the same cycle period.

In addition to the above results, several studies have developed optimal transmit and receive filters that exploit the non-zero spectral correlation of WSCS signals with the same cycle period [17], [20], [21], [22]. However, they do not take the effect of the analog-to-digital converter (ADC) into account when a strong interference is encountered in spectrum sharing. In practice, the strong interference may cause quantization noise to dominate the desired signals at the ADC and severely degrade the reception performance of the conventional RX [23]. This occurs because the magnitude of an ADC input signal is usually adjusted by a gain controller to fit within the quantization range to prevent clipping and, consequently, the desired signals lose their effective bit resolutions.

In this paper, we design an ADC-aware RX for a multi-user multiple-input multiple-output (MU-MIMO) underlay system in the presence of a strong cyclostationary legacy signal. Motivated by the results in [17], [18], and [19], we assume that the symbol rates of the underlay TXs are matched to that of the legacy TX. We consider frequency-selective channels with a moderate number of antenna elements, rather than a massive number of antenna elements considered in the millimeter wave (mmWave) band. Hence, the design is applicable to the centimeter wave (cmWave) band and other lower frequency bands, rather than the mmWave band. Our main objective is to accurately approximate the linear minimum mean-squared error (L-MMSE) estimates of symbols transmitted by the underlay TXs.

To mitigate the effect of the strong interference on the ADCs, we propose to adopt at the RX a structure similar to a hybrid beamformer [24], [25], [26]. It is referred to as a *hybrid combiner structure* in this paper and consists of an analog combiner operating in radio frequency (RF) and a digital combiner operating in digital baseband. The analog combiner first processes the output signals of low-noise amplifiers (LNAs) from multiple antenna elements by using analog components such as power splitters, phase shifters, variable gain amplifiers (VGAs), and power combiners. Then, it passes the combined signal to multiple RF chains [23], [27], [28], [29], where it is assumed that the number of RF chains is less than the number of antenna elements. After the output baseband signals of the RF chains are sampled and quantized at the ADCs, the digital combiner processes the ADC outputs by using a digital signal processor.

Such a hybrid structure is widely adopted for reducing hardware complexities required to implement massive MIMO beamformers in mmWave band [24], [25], [26]. However, only a few studies have proposed to adopt a hybrid combiner structure for strong interference suppression by noticing the ability to pre-process the received signals before they arrive at the ADCs [30], [31], [32]. In [30], an analog combiner called analog preprocessing network (APN) performs, jointly with a digital combiner, the L-MMSE estimation of transmitted symbols. It is shown that the APN can reduce the dominant quantization noise caused by the strong interference. In [31], the L-MMSE weight matrix for the symbol estimation is decomposed into analog and digital parts, where the analog part approximates pre-whitening on the correlation matrix of the received signal. However, the approaches in [30] and [31] assume that the signals are WSS and the channels are frequency-flat, which imply that the signals have time-invariant correlation functions and does not exhibit any spectral correlation. In [32], a beam design is conducted using an analog beamformer to generate nulls at the angles of arrival (AoAs) of strong interference and to make flat at the other directions. This technique can effectively suppress the interference and can be shown to be applicable for the suppression of strong cyclostationary interference. However, the angular dispersion due to multiple sub-paths within each cluster, typically considered in multipath propagation [33], may lead to unintended power loss in the desired signals.

Our analog combiner is designed to maximize a total signal-to-interference-plus-noise ratio (SINR) of the combined received signals. To maximize the total SINR, it performs the following three tasks. First, it suppresses the interference power to a negligible level by approximating the projection onto the null space of the interference correlation matrix. Second, after the interference suppression, it enhances the total power of desired signals by applying the principal component analysis (PCA) on the signal correlation matrix. Third, it adjusts the magnitudes of the ADC input signals to fit them within the quantization range of the ADCs without clipping. To perform these tasks, we conceptually decompose the analog combining matrix into three matrices, design them separately, and then multiply them to construct a single analog combining matrix.

Our digital combiner is designed to generate the estimates of symbols, given the ADC output signals. We first present a method for obtaining the L-MMSE estimates of

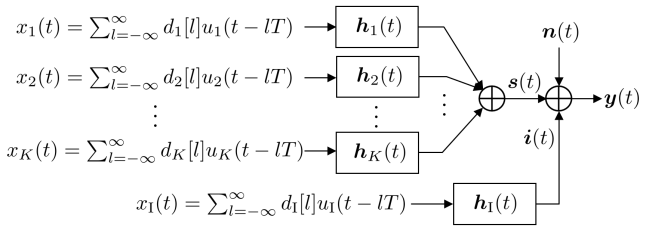

FIGURE 1. Block diagram of MU-MIMO underlay system with K underlay TXs in the presence of a legacy signal.

symbols without considering the quantization noise, and then numerically analyze performance degradation caused by the quantization noise. To exploit the periodic spectral correlation in deriving the L-MMSE solution, we employ the vectorized Fourier transform (VFT) technique. The VFT has been employed for scalar-valued functions [17], [20], [21], and it is shown that the VFT enables the filter design for a scalar-valued WSCS process to be performed by using the conventional design method for an equivalent vector-valued WSS process. In our MU-MIMO system with interference, we extend the VFT to matrix-valued functions. Using this approach, our digital combiner well suppresses the multiple-access interference (MAI) by utilizing the dimensions obtained from both the multiple RF chains and the spectral redundancy of WSCS signals. Simulation results demonstrate that our proposed RX effectively maximizes the SINR and, under certain conditions, approaches the performance achievable in the absence of the legacy TX.

The rest of this paper is organized as follows. In Section II, the signal and system models are described. In Section III, the analog combiner is designed. In Section IV, the digital combiner is designed. In Section V, simulation results are provided. Finally, concluding remarks are offered in Section VI.

II. SIGNAL AND SYSTEM MODELS

In this section, we present signal models in complex baseband and provide an overview of the proposed RX structure.

A. SIGNAL MODELS

There is an MU-MIMO underlay system that attempts to use a frequency band, which is already occupied by a legacy system. To minimize interference to the legacy system, the underlay system is assumed to allocate its uplink to the downlink of the legacy system with low transmission powers of the underlay TXs. In frequency-division duplexing (FDD), the underlay system's uplink frequency band overlaps with the legacy system's downlink band, and in time-division duplexing (TDD), the underlay system's uplink time slot overlaps with the legacy system's downlink slot.

Fig. 1 illustrates the system block diagram in complex baseband. We assume that the legacy system has one TX and the underlay system has K TXs, each equipped with a single antenna. The legacy TX employs linear modulation and emits a signal $x_{\mathrm{I}}(t)$ given by

$$x_{\mathrm{I}}(t) \triangleq \sum_{l=-\infty}^{\infty} d_{\mathrm{I}}[l]u_{\mathrm{I}}(t-lT), \tag{1}$$

where $d_{\mathrm{I}}[l]$ is a sequence of independent and identically distributed (i.i.d.) proper-complex symbols with zero mean and unit variance, $u_{\mathrm{I}}(t)$ is the impulse response of a pulse-shaping filter, and T is the symbol period. We assume that the legacy TX utilizes non-zero excess bandwidth $\beta_{\mathrm{I}} > 0$ defined by the relation [17]

$$B_{\mathrm{I}}T = \frac{1+\beta_{\mathrm{I}}}{2} \tag{2}$$

with bandwidth B_{I} in complex baseband. As well known in [14], $x_{\mathrm{I}}(t)$ is well modeled as a zero-mean proper-complex WSCS random process with cycle period T, i.e., its auto-correlation function is a periodic function with period T.

As mentioned earlier, we apply the results in [17], [18], and [19] and assume that the symbol periods of the underlay TXs are the same with the symbol period T of the legacy TX. Hence, we can model the transmitted signal from the kth underlay TX as

$$x_k(t) \triangleq \sum_{l=-\infty}^{\infty} d_k[l]u_k(t-lT), \tag{3}$$

for $k = 1, 2, \cdots, K$, where $d_k[l]$ is a sequence of i.i.d. proper-complex symbols with zero mean and unit variance and $u_k(t)$ is the impulse response of a pulse-shaping filter. Similarly to the legacy TX, we assume that each kth underlay TX utilizes non-zero excess bandwidth $\beta_k > 0$ defined by the relation

$$B_kT = \frac{1+\beta_{\mathrm{k}}}{2} \tag{4}$$

with bandwidth B_k in complex baseband, so that $x_k(t)$ is also modeled as a zero-mean proper-complex WSCS process with cycle period T, for all k. Throughout this paper, we assume that all the transmitted signals from the underlay TXs have the same bandwidth, i.e., $\beta_k = \beta$ and $B_k = B$, for all k. Furthermore, we also assume that this underlay system partially or fully utilizes the frequency band occupied by the legacy system, i.e., $\beta \le \beta_{\mathrm{I}}$ and $B \le B_{\mathrm{I}}$.

The transmitted signals from the legacy TX and K underlay TXs pass through the frequency-selective channels, and then arrive at the underlay RX with N_{R} antenna elements. Let $\boldsymbol{h}_k(t)$ and $\boldsymbol{h}_{\mathrm{I}}(t)$ be length-N_{R} impulse responses of channels from the kth underlay TX and the legacy TX, respectively. Then, a length-N_{R} received signal $\boldsymbol{y}(t)$ is given by

$$\boldsymbol{y}(t) \triangleq \boldsymbol{s}(t) + \boldsymbol{i}(t) + \boldsymbol{n}(t), \tag{5}$$

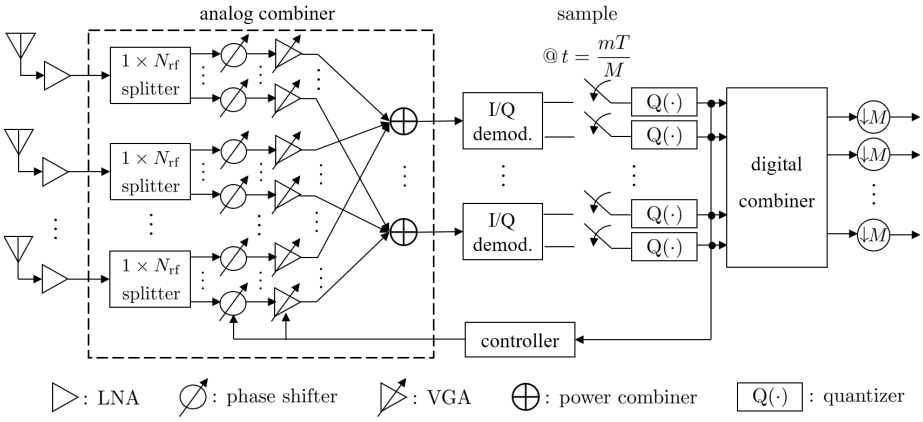

FIGURE 2. Illustration of the proposed RX structure adopting the hybrid combiner structure.

where the first term $\boldsymbol{s}(t)$ denotes the sum of the channel outputs of desired signals given by

$$\boldsymbol{s}(t) \triangleq \sum_{k=1}^{K} \boldsymbol{h}_k(t) * x_k(t) \tag{6a}$$

$$= \sum_{k=1}^{K} \sum_{l=-\infty}^{\infty} d_k[l] \boldsymbol{p}_k(t - lT) \tag{6b}$$

$$= \sum_{l=-\infty}^{\infty} \boldsymbol{P}(t - lT) \boldsymbol{d}[l] \tag{6c}$$

with

$$\boldsymbol{p}_k(t) \triangleq \boldsymbol{h}_k(t) * u_k(t), \tag{7a}$$

$$\boldsymbol{P}(t) \triangleq [\boldsymbol{p}_1(t), \boldsymbol{p}_2(t), \cdots, \boldsymbol{p}_K(t)], \tag{7b}$$

and

$$\boldsymbol{d}[l] \triangleq [d_1[l], d_2[l], \cdots, d_K[l]]^{\mathcal{T}}, \tag{7c}$$

the second term $\boldsymbol{i}(t)$ denotes the channel output of interfering signal given by

$$\boldsymbol{i}(t) \triangleq \boldsymbol{h}_{\mathrm{I}}(t) * x_{\mathrm{I}}(t) \tag{8a}$$

$$= \sum_{l=-\infty}^{\infty} d_{\mathrm{I}}[l] \boldsymbol{p}_{\mathrm{I}}(t - lT) \tag{8b}$$

with

$$\boldsymbol{p}_{\mathrm{I}}(t) \triangleq \boldsymbol{h}_{\mathrm{I}}(t) * u_{\mathrm{I}}(t), \tag{9}$$

and the third term $\boldsymbol{n}(t)$ denotes the length-N_{R} additive white Gaussian noise (AWGN) vector with auto-correlation function $\mathbb{E}\{\boldsymbol{n}(t+\tau)\boldsymbol{n}^{\mathcal{H}}(t)\} = N_0 \boldsymbol{I}_{N_{\mathrm{R}}} \delta(\tau)$. Since we consider the strong interference from the legacy TX, we have $\int_{-\infty}^{\infty} \|\boldsymbol{p}_{\mathrm{I}}(t)\|^2 \, dt \gg \int_{-\infty}^{\infty} \|\boldsymbol{p}_k(t)\|^2 \, dt$, for all $k = 1, 2, \cdots, K$.

B. OVERVIEW OF PROPOSED RX STRUCTURE

Fig. 2 illustrates the structure of the proposed ADC-aware RX adopting the hybrid combiner structure. There are N_{R} antenna elements and N_{rf} RF chains with $N_{\mathrm{R}} > N_{\mathrm{rf}}$. The analog combiner admits the length-N_{R} received signal and combines them to generate a length-N_{rf} signal. An LNA is connected to each mth antenna element, for $m = 1, 2, \cdots, N_{\mathrm{R}}$. Throughout this paper, we assume that all the LNAs operate in the linear region[1]. Then, at each mth antenna element, a power splitter divides the signal into N_{rf} identical signals, and at each branch of the splitter output, a phase shifter and a VGA adjust the phase and magnitude of the signal, respectively [23], [27], [28], [29]. Then, a power combiner connected to each nth RF chain combines the N_{R} output signals from the nth branch of N_{R} power splitters, for $n = 1, 2, \cdots, N_{\mathrm{rf}}$.

At each RF chain, an in-phase and quadrature (I/Q) demodulator generates real and imaginary parts of the complex baseband signal. Then, a sampler conducts M-times oversampling, and then a quantizer conducts uniform quantization within a quantization range. The following digital combiner admits the ADC outputs and then generates K output signals by using a digital signal processor. Finally, after M-times downsampling, the RX generates the length-K decision statistic.

In Fig. 2, it can be seen that the VGAs, which are usually connected to each RF chain before the ADCs [35], are omitted. As will be discussed later, we design the analog combiner to perform the gain adjustment for the ADC input signals. To do this, a controller shown in Fig. 2 periodically updates the amounts of phase shift and gain in the analog combiner that will perform the SINR maximization and the gain control for the ADC input signals.

Fig. 3 illustrates a continuous-time (CT) equivalent model for the proposed RX structure in complex baseband. The length-N_{R} received signal $\boldsymbol{y}(t)$ corresponds to the LNA outputs in complex baseband. In the first block, the RX combines $\boldsymbol{y}(t)$ to generate a length-N_{rf} signal $\boldsymbol{r}(t)$ by using

[1]The strong interference can saturate the LNA, resulting in nonlinear distortion of the received signal. Considering this issue may be an interesting future research direction, but it is beyond the scope of this paper.

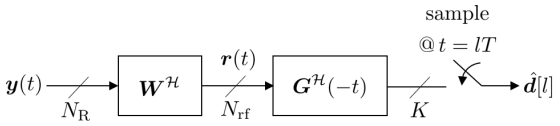

FIGURE 3. CT equivalent model for the proposed RX structure in complex baseband.

an N_R-by-N_{rf} matrix $\boldsymbol{W}$, called analog combining matrix, as

$$\boldsymbol{r}(t) \triangleq \boldsymbol{W}^{\mathcal{H}}\boldsymbol{y}(t). \tag{10}$$

As shown in Fig. 2, the analog combiner adopts phase shifter and VGA for each branch of power splitter output, so that each (m,n)th entry of $\boldsymbol{W}$ can be expressed as a complex number [27], [28], [29], for $m = 1, 2, \cdots, N_R$ and $n = 1, 2, \cdots, N_{rf}$. Here, the length-N_{rf} signal $\boldsymbol{r}(t)$ corresponds to the power combiner outputs in Fig. 2.

In the second block, the RX combines the length-N_{rf} signal $\boldsymbol{r}(t)$ by using an N_{rf}-by-K matrix of LTI filters with impulse response matrix $\boldsymbol{G}(t)$ and then performs uniform sampling with symbol period T to generate a length-K decision statistic $\hat{\boldsymbol{d}}[l]$ as

$$\hat{\boldsymbol{d}}[l] \triangleq \boldsymbol{G}^{\mathcal{H}}(-t) * \boldsymbol{r}(t)\Big|_{t=lT} \tag{11a}$$

$$= \boldsymbol{G}^{\mathcal{H}}(-t) * \boldsymbol{W}^{\mathcal{H}}\boldsymbol{y}(t)\Big|_{t=lT}. \tag{11b}$$

It is noteworthy that the digital combiner is designed in the CT equivalent model, although it actually operates in the discrete-time (DT) domain. For details on this equivalence, see [36, Sec. 4.6]. The quantization is omitted here, because this CT equivalent model is employed for the purpose of derivation for $\boldsymbol{W}$ and $\boldsymbol{G}(t)$.

Given this structure, our objective is to design $\boldsymbol{W}$ and $\boldsymbol{G}(t)$ so that the decision statistic $\hat{\boldsymbol{d}}[l]$ well approximates the transmitted symbol vector $\boldsymbol{d}[l]$. We briefly summarize design criteria for $\boldsymbol{W}$ and $\boldsymbol{G}(t)$ as follows:

- The analog combining matrix $\boldsymbol{W}$ is designed to maximize a total SINR of the analog combiner output and to adjust the magnitude of the signal in each RF chain. The total SINR is denoted by $\mathrm{SINR}_{\mathrm{total}}$ and defined as

$$\mathrm{SINR}_{\mathrm{total}} \triangleq \frac{\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{W}^{\mathcal{H}}\boldsymbol{s}(t)\|^2\right\}\right]}{\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{W}^{\mathcal{H}}\boldsymbol{i}(t)\|^2\right\}\right] + \mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{W}^{\mathcal{H}}\boldsymbol{n}(t)\|^2\right\}\right]} \tag{12}$$

  with $\mathbb{A}[\cdot]$ denoting the time average. That is, the following problem is considered:

$$\begin{aligned} &\underset{\boldsymbol{W}}{\text{maximize}} && \mathrm{SINR}_{\mathrm{total}} \\ &\text{subject to} && \boldsymbol{w}_n^{\mathcal{H}}\boldsymbol{w}_{n'} = 0, \ \forall\, n \neq n', \end{aligned} \tag{13}$$

  where $\boldsymbol{w}_n$ is the nth column of $\boldsymbol{W}$, for $n = 1, 2, \cdots, N_{rf}$. The reason for including the time average in (12) is because we consider WSCS signals having time-varying correlation functions, unlike WSS signals having time-invariant correlation functions [30], [31]. Moreover, the reason for imposing the orthogonal constraint on the column vectors of $\boldsymbol{W}$ is to prevent non-zero noise correlation in different RF chains.
- The digital combining filter $\boldsymbol{G}(t)$ is designed to minimize a total MSE between the decision statistic $\hat{\boldsymbol{d}}[l]$ and the symbol vector $\boldsymbol{d}[l]$, for all $l \in \mathbb{Z}$, given the analog combiner output $\boldsymbol{r}(t)$. The total MSE is denoted by $\mathrm{MSE}_{\mathrm{total}}$ and defined as

$$\mathrm{MSE}_{\mathrm{total}} \triangleq \mathbb{E}\left\{\|\hat{\boldsymbol{d}}[l] - \boldsymbol{d}[l]\|^2\right\}, \tag{14}$$

  where $\hat{\boldsymbol{d}}[l]$ is defined in (11). That is, the following problem is considered:

$$\underset{\boldsymbol{G}(t)}{\text{minimize}} \qquad \mathrm{MSE}_{\mathrm{total}}. \tag{15}$$

  Note that $\mathrm{MSE}_{\mathrm{total}}$ is uniquely defined for all $l \in \mathbb{Z}$, because all the transmitted signals are WSCS with the common cycle period T.

We solve the above two optimization problems in the following two sections. In deriving $\boldsymbol{W}$ and $\boldsymbol{G}(t)$, we assume that the underlay RX has prior information about the channels from the legacy TX and the underlay TXs. This can be justified by considering the following scenario. During the deployment of the underlay system, before any transmissions from K underlay TXs, the underlay RX is exposed to the strong interference from the legacy TX. This initial exposure allows the underlay RX to learn the propagation environment from the legacy TX to the underlay RX. Then, after the underlay TXs start their operation, such a prior information about the channel from the legacy TX can be exploited to obtain the channel information from the underlay TXs to the underlay RX [37].

III. ANALOG COMBINER DESIGN

In this section, we solve the optimization problem (13) to design the analog combining matrix $\boldsymbol{W}$.

A. DECOMPOSITION OF W INTO THREE MATRICES

We consider the following decomposition of $\boldsymbol{W}$ given by

$$\boldsymbol{W} = \boldsymbol{ABD}, \tag{16}$$

where $\boldsymbol{A}$ is an N_R-by-N_{eff} matrix, $\boldsymbol{B}$ is an N_{eff}-by-N_{rf} matrix, and $\boldsymbol{D}$ is an N_{rf}-by-N_{rf} diagonal matrix. The constant N_{eff} can be equal to N_R, N_{rf}, or any natural number between N_R and N_{rf}. By (10) and (16), the analog combiner output $\boldsymbol{r}(t)$ can be rewritten as

$$\boldsymbol{r}(t) = \boldsymbol{W}^{\mathcal{H}}\boldsymbol{y}(t) = \boldsymbol{D}^{\mathcal{H}}\boldsymbol{B}^{\mathcal{H}}\boldsymbol{A}^{\mathcal{H}}\boldsymbol{y}(t). \tag{17}$$

To maximize $\mathrm{SINR}_{\mathrm{total}}$ in (12), we assign the following tasks to the three matrices. First, the matrix $\boldsymbol{A}$ reduces the interference term at the denominator of $\mathrm{SINR}_{\mathrm{total}}$ in (12) to a negligible level, while the noise term is maintained at a consistent level. Second, given $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{y}(t)$, the matrix $\boldsymbol{B}$ enhances the numerator of $\mathrm{SINR}_{\mathrm{total}}$ in (28), while the

noise term is maintained at a consistent level. Third, the diagonal matrix $\boldsymbol{D}$ adjusts the magnitude of each element of $\boldsymbol{B}^{\mathcal{H}}\boldsymbol{A}^{\mathcal{H}}\boldsymbol{y}(t)$. We will impose the semi-unitary constraints on $\boldsymbol{A}$ and $\boldsymbol{B}$ to ensure the uniqueness of the solution and to prevent noise boosting effect. These constraints allow $\boldsymbol{W} = \boldsymbol{ABD}$ to be a semi-orthogonal matrix that satisfies the constraint in (13) for any diagonal matrix $\boldsymbol{D}$. It is noteworthy that, although $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{D}$ are separately designed, the phase shifters and VGAs in the analog combiner are adjusted based on the single matrix $\boldsymbol{W}$.

B. DESIGN OF MATRIX $\boldsymbol{A}$

Our objective is to find an N_{R}-by-N_{eff} semi-unitary matrix $\boldsymbol{A}_{\mathrm{PROP}}$ such that

$$\frac{\mathbb{A}\Big[\mathbb{E}\Big\{\|\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{i}(t)\|^2\Big\}\Big]}{\mathbb{A}\Big[\mathbb{E}\big\{\|\boldsymbol{i}(t)\|^2\big\}\Big]} \le \epsilon, \tag{18}$$

for a pre-defined $\epsilon > 0$, and $\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}} = \boldsymbol{I}_{N_{\mathrm{eff}}}$. That is, for a sufficiently small ϵ, the matrix $\boldsymbol{A}_{\mathrm{PROP}}$ can significantly reduce the received interference power level, while maintaining the noise power level due to the semi-unitary constraint. However, since N_{eff} is not specified yet, the solution that satisfies (18) given $\epsilon > 0$ is not unique. Among the solutions, we find the one with the largest N_{eff}. This choice is reasonable, as will be discussed in the next subsection, in that a larger N_{eff} provides more degrees of freedom for the following N_{eff}-by-N_{rf} matrix $\boldsymbol{B}$ to enhance the total power of desired signals. To find such a solution, we propose an algorithm that approximates the projection onto the null space of the interference correlation as follows.

Consider the eigenvalue decomposition (EVD) of the time-averaged correlation matrix of $\boldsymbol{i}(t)$ as

$$\mathbb{A}\Big[\mathbb{E}\Big\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\Big\}\Big] = \boldsymbol{U}_{\mathrm{I}}\boldsymbol{\Lambda}_{\mathrm{I}}\boldsymbol{U}_{\mathrm{I}}^{\mathcal{H}}, \tag{19a}$$

where $\boldsymbol{U}_{\mathrm{I}}$ is a size-N_{R} unitary matrix and $\boldsymbol{\Lambda}_{\mathrm{I}}$ is a size-N_{R} diagonal matrix having the eigenvalues with

$$\lambda_{\mathrm{I},1} \ge \lambda_{\mathrm{I},2} \ge \cdots \ge \lambda_{\mathrm{I},N_{\mathrm{R}}}. \tag{19b}$$

Then, the left-side of (18) becomes

$$\frac{\mathbb{A}\Big[\mathbb{E}\Big\{\|\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{i}(t)\|^2\Big\}\Big]}{\mathbb{A}\Big[\mathbb{E}\big\{\|\boldsymbol{i}(t)\|^2\big\}\Big]} = \frac{\mathrm{tr}\Big(\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{U}_{\mathrm{I}}\boldsymbol{\Lambda}_{\mathrm{I}}\boldsymbol{U}_{\mathrm{I}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}\Big)}{\mathrm{tr}\Big(\boldsymbol{U}_{\mathrm{I}}\boldsymbol{\Lambda}_{\mathrm{I}}\boldsymbol{U}_{\mathrm{I}}^{\mathcal{H}}\Big)} \tag{20a}$$

$$= \frac{\mathrm{tr}\Big(\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{U}_{\mathrm{I}}\boldsymbol{\Lambda}_{\mathrm{I}}\boldsymbol{U}_{\mathrm{I}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}\Big)}{\sum_{m=1}^{N_{\mathrm{R}}}\lambda_{\mathrm{I},m}}. \tag{20b}$$

Let $\boldsymbol{A}_{\mathrm{eff}}$ be an N_{R}-by-N_{eff} matrix constructed by choosing the eigenvectors corresponding to the smallest N_{eff} eigenvalues, i.e.,

$$\boldsymbol{A}_{\mathrm{eff}} \triangleq \big[\boldsymbol{u}_{\mathrm{I},N_{\mathrm{R}}-N_{\mathrm{eff}}+1}\ \boldsymbol{u}_{\mathrm{I},N_{\mathrm{R}}-N_{\mathrm{eff}}+2}\ \cdots\ \boldsymbol{u}_{\mathrm{I},N_{\mathrm{R}}}\big], \tag{21}$$

Algorithm 1 Line Search Algorithm to Find N_{eff}

Input: N_{R}, N_{rf}, $\boldsymbol{\Lambda}_{\mathrm{I}}$, ϵ
Output: N_{eff}

```
Set N = 0 and f(0) = 1;
while f(N) > ε and N < N_R − N_rf do
    N ← N + 1;
    Compute CCDF f(N) in (23);
end while
return N_eff = N_R − N
```

for some N_{eff} to be determined, where $\boldsymbol{u}_{\mathrm{I},m}$ denotes the mth column vector of $\boldsymbol{U}_{\mathrm{I}}$, for $m = 1, 2, \cdots, N_{\mathrm{R}}$. Then, by replacing $\boldsymbol{A}_{\mathrm{PROP}}$ in (20) with $\boldsymbol{A}_{\mathrm{eff}}$, we have

$$\frac{\mathbb{A}\Big[\mathbb{E}\Big\{\|\boldsymbol{A}_{\mathrm{eff}}^{\mathcal{H}}\boldsymbol{i}(t)\|^2\Big\}\Big]}{\mathbb{A}\Big[\mathbb{E}\big\{\|\boldsymbol{i}(t)\|^2\big\}\Big]} = \frac{\sum_{m=N_{\mathrm{R}}-N_{\mathrm{eff}}+1}^{N_{\mathrm{R}}}\lambda_{\mathrm{I},m}}{\sum_{m=1}^{N_{\mathrm{R}}}\lambda_{\mathrm{I},m}}. \tag{22}$$

Note that $\boldsymbol{A}_{\mathrm{eff}}$ in (21) is a semi-unitary matrix satisfying $\boldsymbol{A}_{\mathrm{eff}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{eff}} = \boldsymbol{I}_{N_{\mathrm{eff}}}$ for any N_{eff}, because $\boldsymbol{U}_{\mathrm{I}}$ is a unitary matrix. To find N_{eff}, we propose a simple line search algorithm as follows, which is also described in **Algorithm 1**.

Let $f(N)$ be a complementary cumulative distribution function (CCDF) of eigenvalues given by

$$f(N) \triangleq 1 - \frac{\sum_{m=1}^{N}\lambda_{\mathrm{I},m}}{\sum_{m=1}^{N_{\mathrm{R}}}\lambda_{\mathrm{I},m}}, \tag{23}$$

for $N = 0, 1, \cdots, N_{\mathrm{R}}$. We perform the iteration to find N_{eff} and begin by initializing $N = 0$ and $f(0) = 1$. At each iteration, N is increased by one, and $f(N)$ is computed as in (23). The process iterates while $f(N) > \epsilon$ and $N < N_{\mathrm{R}} - N_{\mathrm{rf}}$. If either of these two conditions is not met, we terminate the iteration and return $N_{\mathrm{eff}} = N_{\mathrm{R}} - N$. The reason for involving the second condition $N < N_{\mathrm{R}} - N_{\mathrm{rf}}$ is because the analog combiner must pass at least N_{rf} signals to the RF chains.

If we set $\epsilon = 1$, i.e., 0 dB, then $N_{\mathrm{eff}} = N_{\mathrm{R}}$ and $\boldsymbol{A}_{\mathrm{eff}}$ becomes $\boldsymbol{I}_{N_{\mathrm{R}}}$, which implies that the pre-combiner does not perform the strong interference suppression. On the contrary, if we set $\epsilon \approx 0$, e.g. -100 dB, and $f(N_{\mathrm{eff}}) \le \epsilon$ with $N_{\mathrm{eff}} \ge N_{\mathrm{rf}}$, then $\boldsymbol{A}_{\mathrm{eff}}$ best approximates the projection onto the null space of $\mathbb{A}\Big[\mathbb{E}\Big\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\Big\}\Big]$. Therefore, by (20) and (22), we choose the matrix $\boldsymbol{A}_{\mathrm{eff}}$ with $\epsilon \approx 0$ as $\boldsymbol{A}_{\mathrm{PROP}}$.

We now investigate the eigenvalue distribution of the matrix $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$, because it is crucial in determining N_{eff}.

Lemma 1: The matrix $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$ can be rewritten as

$$\mathbb{A}\Big[\mathbb{E}\Big\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\Big\}\Big] = \frac{1}{T}\int_{-B_{\mathrm{I}}}^{B_{\mathrm{I}}}\tilde{\boldsymbol{p}}_{\mathrm{I}}(\xi)\tilde{\boldsymbol{p}}_{\mathrm{I}}^{\mathcal{H}}(\xi)\,d\xi, \tag{24}$$

where $\tilde{\boldsymbol{p}}_{\mathrm{I}}(\xi)$ is an element-wise Fourier transform of $\boldsymbol{p}_{\mathrm{I}}(t)$.

Proof: See Appendix. □

Before analyzing the eigenvalue distribution, we first specify the impulse response $\boldsymbol{h}_{\mathrm{I}}(t)$ of the channel from the legacy TX to the underlay RX because $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$ is

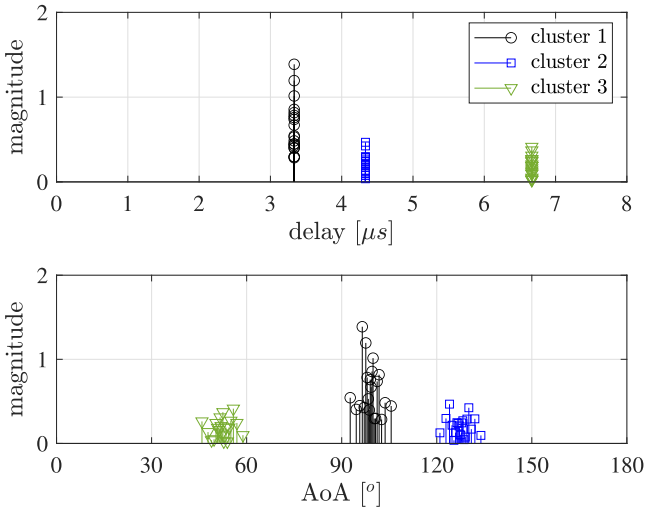

FIGURE 4. Exemplary multipath intensity profile of propagation channel from the legacy TX to the underlay RX in terms of delay and AoA.

constructed by using $\boldsymbol{p}_{\mathrm{I}}(t) = \boldsymbol{h}_{\mathrm{I}}(t) * u_{\mathrm{I}}(t)$. We consider the ray-based specular channel model and assume that the underlay RX is equipped with uniform linear array (ULA) antenna elements that have half-wavelength spacing. Hence, the length-N_{R} impulse response $\boldsymbol{h}_{\mathrm{I}}(t)$ can be written as

$$\boldsymbol{h}_{\mathrm{I}}(t) = \sum_{l=1}^{L_{\mathrm{I}}} \sum_{j=1}^{J_{\mathrm{I}}} \alpha_{\mathrm{I},l,j} \delta\left(t - \tau_{\mathrm{I},l,j}\right) \boldsymbol{e}\left(\theta_{\mathrm{I},l,j}\right), \tag{25}$$

where L_{I} is the number of clusters, J_{I} is the number of sub-paths assumed to be the same across the clusters, $\alpha_{\mathrm{I},l,j}$ is complex gain, $\tau_{\mathrm{I},l,j}$ is delay, and $\boldsymbol{e}(\theta_{\mathrm{I},l,j})$ is the length-N_{R} array response vector corresponding to the AoA $\theta_{\mathrm{I},l,j}$ with the nth element $\left[\boldsymbol{e}\left(\theta_{\mathrm{I},l,j}\right)\right]_n \triangleq e^{-\mathrm{j}\pi(n-1)\cos\theta_{\mathrm{I},l,j}}$, for the lth cluster and jth sub-path.

Here, we consider the case where the transmitted signal $x_{\mathrm{I}}(t)$ from the legacy TX is reflected by only a few distant objects, such as mountains or large buildings, and are received by the underlay RX. This situation can occur when both the legacy TX and the underlay RX are located at elevated positions to enable high-power and long-range transmission and to increase cell coverage, respectively. Fig. 4 illustrates an exemplary multipath intensity profile of such a propagation environment, where $L_{\mathrm{I}} = 3$ and $J_{\mathrm{I}} = 20$. Note that the delays of sub-paths within each cluster are almost the same in the delay domain, but the AoAs are distinct in the space domain. These sub-path parameters are from the rural macro (RMa) environment with non-line-of-sight (NLOS) scenario, which is specified in the technical report on the spatial channel model (SCM) of the 3rd Generation Partnership Project (3GPP) [33]. Hence, we assume that $\tau_{\mathrm{I},l,j} \approx \tau_{\mathrm{I},l}$, for all $j = 1, 2, \cdots, J_{\mathrm{I}}$ and for each lth cluster.

By (25) and the above assumption, the element-wise Fourier transform $\tilde{\boldsymbol{p}}_{\mathrm{I}}(\xi)$ of $\boldsymbol{p}_{\mathrm{I}}(t) = \boldsymbol{h}_{\mathrm{I}}(t) * u_{\mathrm{I}}(t)$ is approximated by

$$\tilde{\boldsymbol{p}}_{\mathrm{I}}(\xi) \approx \tilde{u}_{\mathrm{I}}(\xi) \sum_{l=1}^{L_{\mathrm{I}}} \left(\sum_{j=1}^{J_{\mathrm{I}}} \alpha_{\mathrm{I},l,j} \boldsymbol{e}\left(\theta_{\mathrm{I},l,j}\right) \right) e^{-\mathrm{j}2\pi\xi\tau_{\mathrm{I},l}} \tag{26a}$$

$$= \tilde{u}_{\mathrm{I}}(\xi) \sum_{l=1}^{L_{\mathrm{I}}} \boldsymbol{a}_l e^{-\mathrm{j}2\pi\xi\tau_{\mathrm{I},l}}, \tag{26b}$$

for $\xi \in [-B_{\mathrm{I}}, B_{\mathrm{I}})$, where $\tilde{u}_{\mathrm{I}}(\xi)$ is the Fourier transform of $u_{\mathrm{I}}(t)$ and $\boldsymbol{a}_l \triangleq \sum_{j=1}^{J_{\mathrm{I}}} \alpha_{\mathrm{I},l,j} \boldsymbol{e}(\theta_{\mathrm{I},l,j})$ is a length-N_{R} vector. Note that $\boldsymbol{a}_l$'s are almost linearly independent, because the channel parameters for different clusters are uncorrelated in general. For analytical convenience, let $\tilde{u}_{\mathrm{I}}(\xi) = 1$, for all $\xi \in [-B_{\mathrm{I}}, B_{\mathrm{I}})$. Then, we can further rewrite (24) as

$$\mathbb{A}\left[\mathbb{E}\left\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\right\}\right] = \frac{1}{T}\int_{-B_{\mathrm{I}}}^{B_{\mathrm{I}}} \sum_{l=1}^{L_{\mathrm{I}}} \sum_{l'=1}^{L_{\mathrm{I}}} \boldsymbol{a}_l \boldsymbol{a}_{l'}^{\mathcal{H}} e^{-\mathrm{j}2\pi\xi(\tau_{\mathrm{I},l} - \tau_{\mathrm{I},l'})}\, d\xi \tag{27a}$$

$$= \frac{2B_{\mathrm{I}}}{T} \sum_{l=1}^{L_{\mathrm{I}}} \sum_{l'=1}^{L_{\mathrm{I}}} \boldsymbol{a}_l \boldsymbol{a}_{l'}^{\mathcal{H}} \mathrm{sinc}(2B_{\mathrm{I}}(\tau_{\mathrm{I},l} - \tau_{\mathrm{I},l'})). \tag{27b}$$

In (27b), we can see that the eigenvalue distribution of the matrix $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$ is determined by the values of $\mathrm{sinc}(2B_{\mathrm{I}}(\tau_{\mathrm{I},l} - \tau_{\mathrm{I},l'}))$ for $l \neq l'$. We analyze it by using $B_{\mathrm{I}} = 7.5$ MHz, $N_{\mathrm{R}} = 16$, and the multipath intensity profile shown in Fig. 4. In this scenario, since the differences of cluster delays are larger than 1 μs, the values of $\mathrm{sinc}(2B_{\mathrm{I}}(\tau_{\mathrm{I},l} - \tau_{\mathrm{I},l'}))$ becomes much smaller than one for all combinations of $l \neq l'$. That is, the matrix $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$ becomes the sum of the outer product of almost linearly independent vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. Hence, as can be seen in Fig. 5-(a), there are three dominant eigenvalues in $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$. In Fig. 5-(b), the value N for which the CCDF $f(N)$ of eigenvalues becomes less than -100 dB is 3, so that we have $N_{\mathrm{eff}} = N_{\mathrm{R}} - 3$ in this exemplary scenario.

C. DESIGNS OF MATRICES B AND D

By substituting (16) into (12) with $\boldsymbol{A}_{\mathrm{PROP}}$ satisfying (18) for $\epsilon \approx 0$, the total SINR can be approximated by

$$\mathrm{SINR}_{\mathrm{total}} \approx \frac{\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{D}^{\mathcal{H}}\boldsymbol{B}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)\|^2\right\}\right]}{\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{D}^{\mathcal{H}}\boldsymbol{B}^{\mathcal{H}}\boldsymbol{n}(t)\|^2\right\}\right]}, \tag{28}$$

where $\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}$ is omitted from the denominator because $\boldsymbol{A}_{\mathrm{PROP}}$ is a semi-unitary matrix and $\boldsymbol{n}(t)$ is the AWGN vector. Our objective is now to find an N_{eff}-by-N_{rf} semi-unitary matrix $\boldsymbol{B}_{\mathrm{PROP}}$ such that

$$\boldsymbol{B}_{\mathrm{PROP}} \triangleq \underset{\boldsymbol{B}}{\mathrm{argmax}}\, \mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{B}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)\|^2\right\}\right] \tag{29}$$

and $\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{B}_{\mathrm{PROP}} = \boldsymbol{I}_{N_{\mathrm{rf}}}$. To find $\boldsymbol{B}_{\mathrm{PROP}}$, we consider the PCA as follows.

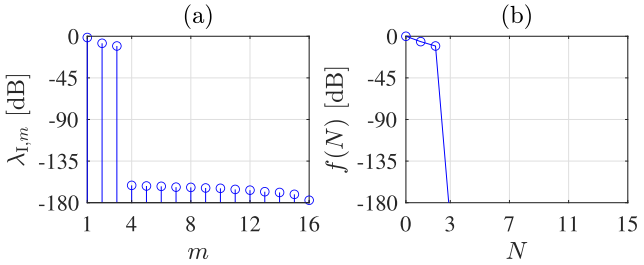

FIGURE 5. Illustration of exemplary (a) eigenvalue distribution of the time-averaged interference correlation matrix $\mathbb{A}[\mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}]$ and (b) CCDF $f(N)$, when $B_{\mathrm{I}} = 7.5$ MHz, $N_{\mathrm{R}} = 16$, and the multipath intensity profile shown in Fig. 4.

Proposition 1: Let an EVD of the time-averaged correlation matrix of $\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)$ be

$$\mathbb{A}\left[\mathbb{E}\left\{\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)\boldsymbol{s}^{\mathcal{H}}(t)\boldsymbol{A}_{\mathrm{PROP}}\right\}\right] = \boldsymbol{U}_{\mathrm{S}}\boldsymbol{\Lambda}_{\mathrm{S}}\boldsymbol{U}_{\mathrm{S}}^{\mathcal{H}}, \tag{30}$$

where $\boldsymbol{U}_{\mathrm{S}}$ is a size-N_{eff} unitary matrix and $\boldsymbol{\Lambda}_{\mathrm{S}}$ is a size-N_{eff} diagonal matrix having the eigenvalues with

$$\lambda_{\mathrm{S},1} \geq \lambda_{\mathrm{S},2} \geq \cdots \geq \lambda_{\mathrm{S},N_{\mathrm{eff}}}. \tag{31}$$

Then, the optimal solution to (29) is given by

$$\boldsymbol{B}_{\mathrm{PROP}} = \left[\boldsymbol{u}_{\mathrm{S},1}\ \boldsymbol{u}_{\mathrm{S},2}\ \cdots\ \boldsymbol{u}_{\mathrm{S},N_{\mathrm{rf}}}\right], \tag{32}$$

where $\boldsymbol{u}_{\mathrm{S},n}$ denotes the nth column vector of $\boldsymbol{U}_{\mathrm{S}}$, for $n = 1, 2, \cdots, N_{\mathrm{eff}}$.

Proof: By (29)–(32), we have

$$\begin{aligned}
&\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)\|^2\right\}\right] \\
&= \mathrm{tr}\left(\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\mathbb{A}\left[\mathbb{E}\left\{\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)\boldsymbol{s}^{\mathcal{H}}(t)\boldsymbol{A}_{\mathrm{PROP}}\right\}\right]\boldsymbol{B}_{\mathrm{PROP}}\right) && \text{(33a)} \\
&= \mathrm{tr}\left(\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{U}_{\mathrm{S}}\boldsymbol{\Lambda}_{\mathrm{S}}\boldsymbol{U}_{\mathrm{S}}^{\mathcal{H}}\boldsymbol{B}_{\mathrm{PROP}}\right) && \text{(33b)} \\
&= \sum_{n=1}^{N_{\mathrm{rf}}}\lambda_{\mathrm{S},n}. && \text{(33c)}
\end{aligned}$$

According to the definition of trace and Rayleigh's inequality [34], $\boldsymbol{B}_{\mathrm{PROP}}$ in (32) maximizes the objective function in (29) by satisfying $\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{B}_{\mathrm{PROP}} = \boldsymbol{I}_{N_{\mathrm{rf}}}$, given any N_{rf}. Therefore, the conclusion follows. □

It can be shown that, if $\boldsymbol{A}_{\mathrm{PROP}}$ satisfies the inequality in (18) with $\epsilon \approx 0$, then the objective function in (29) can be replaced by $\mathbb{A}[\mathbb{E}\{\|\boldsymbol{B}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{y}(t)\|^2\}]$. This implies that, under the condition (18) with $\epsilon \approx 0$, the matrix $\boldsymbol{B}_{\mathrm{PROP}}$ can also be obtained by using the statistical information of the received signals.

After obtaining $\boldsymbol{A}_{\mathrm{PROP}}$ and $\boldsymbol{B}_{\mathrm{PROP}}$, the total SINR can be approximated by

$$\mathrm{SINR}_{\mathrm{total}} \approx \frac{\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{D}^{\mathcal{H}}\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{s}(t)\|^2\right\}\right]}{\mathbb{A}\left[\mathbb{E}\left\{\|\boldsymbol{D}^{\mathcal{H}}\boldsymbol{n}(t)\|^2\right\}\right]}, \tag{34}$$

where $\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}$ is again omitted from the denominator because $\boldsymbol{B}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}^{\mathcal{H}}\boldsymbol{A}_{\mathrm{PROP}}\boldsymbol{B}_{\mathrm{PROP}} = \boldsymbol{I}_{N_{\mathrm{rf}}}$ and $\boldsymbol{n}(t)$ is the AWGN vector. Now, we design the diagonal matrix $\boldsymbol{D}$ given the matrices $\boldsymbol{A}_{\mathrm{PROP}}$ and $\boldsymbol{B}_{\mathrm{PROP}}$. Our objective is to adjust the magnitudes of the ADC input signals using $\boldsymbol{D}$. To do this, the nth diagonal entry $d_{n,n}$ of $\boldsymbol{D}$ is tuned so that the average magnitude of the sampler output observed over a previous time window, e.g., 1 msec, becomes the maximum output level of quantizer. Then, a certain amount of power backoff is applied to prevent signal clipping at the quantizer. The amount of the power backoff can be determined based on the instantaneous normalized power (INP) of each analog combiner output signal. Besides this method, other gain control algorithms can be utilized to meet their system requirements [35], [38]. A key feature of our analog combiner is that it can perform the gain control as well as pre-combining, eliminating the need for VGAs connected to the RF chains as shown in Fig. 2. In Section V, our simulations will be conducted under the assumption that the signal's peak magnitude fits the quantization range without any signal clipping [30], [31]. We denote the matrix $\boldsymbol{D}$ that achieves this as $\boldsymbol{D}_{\mathrm{PROP}}$. Therefore, our proposed analog combining matrix $\boldsymbol{W}_{\mathrm{PROP}}$ is given by

$$\boldsymbol{W}_{\mathrm{PROP}} \triangleq \boldsymbol{A}_{\mathrm{PROP}}\boldsymbol{B}_{\mathrm{PROP}}\boldsymbol{D}_{\mathrm{PROP}}. \tag{35}$$

A key contribution of our proposed analog combiner is to mitigate the dominant quantization noise caused by the strong interference from the legacy TX. This advantage of our analog combiner will be numerically analyzed in terms of the reception performance, after the digital combiner is designed in the next section.

IV. DIGITAL COMBINER DESIGN

In this section, we solve the optimization problem (15) to design the impulse response matrix $\boldsymbol{G}(t)$.

A. FREQUENCY-DOMAIN EXPRESSION OF TOTAL MSE AND DERIVATION OF SOLUTION

Recall that our digital combining filter is designed in our CT equivalent model shown in Fig. 3, although it operates in the DT domain after oversampling. The equivalence of sampling after CT filtering, and downsampling after DT filtering with oversampling can be found in [36, Sec. 4.6].

The impulse response matrix $\boldsymbol{G}(t)$ is designed to find $\hat{\boldsymbol{d}}[l]$ in (11) minimizing $\mathrm{MSE}_{\mathrm{total}} = \mathbb{E}\{\|\hat{\boldsymbol{d}}[l] - \boldsymbol{d}[l]\|^2\}$ in (14). We derive the L-MMSE solution to this problem, given an analog combiner output $\boldsymbol{r}(t) = \boldsymbol{W}^{\mathcal{H}}\boldsymbol{y}(t)$ with a semi-orthogonal matrix $\boldsymbol{W}$. To do this, we express $\mathrm{MSE}_{\mathrm{total}}$ in the frequency domain by applying the VFT technique to matrix-valued functions. The definition of the VFT for a scalar-valued function can be found in [20, *Definition 5*].

Definition 1: Let $\boldsymbol{X}(t)$ be a matrix-valued time function with an entry-by-entry Fourier transform $\tilde{\boldsymbol{X}}(\xi)$. Given bandwidth B [Hz] in complex baseband and symbol rate $1/T$

[symbols/Hz], the VFT $\bar{\boldsymbol{X}}(f)$ of $\boldsymbol{X}(t)$ is defined as

$$\bar{\boldsymbol{X}}(f) \triangleq \begin{bmatrix} \tilde{\boldsymbol{X}}\left(f - \frac{N_{\mathrm{f}}-1}{2T}\right) \\ \tilde{\boldsymbol{X}}\left(f - \frac{N_{\mathrm{f}}-3}{2T}\right) \\ \vdots \\ \tilde{\boldsymbol{X}}\left(f + \frac{N_{\mathrm{f}}-1}{2T}\right) \end{bmatrix}, \tag{36}$$

for $f \in \mathcal{F}$, where $N_{\mathrm{f}} \triangleq 2\lceil\beta\rceil+1$ and $\mathcal{F}$ is the Nyquist interval defined as

$$\mathcal{F} \triangleq \left\{ f : -\frac{1}{2T} \le f \le \frac{1}{2T} \right\}. \tag{37}$$

Using *Definition 1*, we define $N_{\mathrm{R}}N_{\mathrm{f}}$-by-$K$ matrix $\bar{\boldsymbol{P}}(f)$ and $N_{\mathrm{rf}}N_{\mathrm{f}}$-by-$K$ matrix $\bar{\boldsymbol{G}}(f)$, respectively, as the VFTs of the N_{R}-by-K matrix-valued function $\boldsymbol{P}(t)$ and N_{rf}-by-K matrix-valued function $\boldsymbol{G}(t)$. Moreover, we denote a length-$N_{\mathrm{R}}N_{\mathrm{f}}$ vector $\bar{\boldsymbol{p}}_{\mathrm{I}}(f)$ as the VFT of the length-N_{R} vector-valued function $\boldsymbol{p}_{\mathrm{I}}(t)$. In the following theorem, using these VFTs, we express $\mathrm{MSE}_{\mathrm{total}}$ in the frequency domain.

Theorem 1: The total MSE can be rewritten as

$$\begin{aligned} \mathrm{MSE}_{\mathrm{total}} = K + \int_{\mathcal{F}} \mathrm{tr}\big(&\bar{\boldsymbol{G}}^{\mathcal{H}}(f)\bar{\boldsymbol{R}}(f)\bar{\boldsymbol{G}}(f) \\ &- 2\mathsf{Re}\left\{\bar{\boldsymbol{G}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{P}}(f)\right\}\big)\, df, \end{aligned} \tag{38}$$

where $\bar{\boldsymbol{R}}(f)$ is an $N_{\mathrm{rf}}N_{\mathrm{f}}$-by-$N_{\mathrm{rf}}N_{\mathrm{f}}$ correlation matrix given by

$$\begin{aligned} \bar{\boldsymbol{R}}(f) \triangleq\ & \frac{1}{T}\left(\bar{\boldsymbol{W}}^{\mathcal{H}}\left(\bar{\boldsymbol{P}}(f)\bar{\boldsymbol{P}}^{\mathcal{H}}(f) + \bar{\boldsymbol{p}}_{\mathrm{I}}(f)\bar{\boldsymbol{p}}_{\mathrm{I}}^{\mathcal{H}}(f)\right)\bar{\boldsymbol{W}}\right) \\ &+ N_0\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{W}}, \end{aligned} \tag{39}$$

for $f \in \mathcal{F}$, and $\bar{\boldsymbol{W}}$ is an $N_{\mathrm{R}}N_{\mathrm{f}}$-by-$N_{\mathrm{rf}}N_{\mathrm{f}}$ block-diagonal matrix given by

$$\bar{\boldsymbol{W}} \triangleq \boldsymbol{I}_{N_{\mathrm{f}}} \otimes \boldsymbol{W} \tag{40a}$$

$$= \begin{bmatrix} \boldsymbol{W} & \boldsymbol{O}_{N_{\mathrm{R}}\times N_{\mathrm{rf}}} & \cdots & \boldsymbol{O}_{N_{\mathrm{R}}\times N_{\mathrm{rf}}} \\ \boldsymbol{O}_{N_{\mathrm{R}}\times N_{\mathrm{rf}}} & \boldsymbol{W} & \cdots & \boldsymbol{O}_{N_{\mathrm{R}}\times N_{\mathrm{rf}}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{O}_{N_{\mathrm{R}}\times N_{\mathrm{rf}}} & \boldsymbol{O}_{N_{\mathrm{R}}\times N_{\mathrm{rf}}} & \cdots & \boldsymbol{W} \end{bmatrix}. \tag{40b}$$

Proof: See Section IV-B. □

By using the property of trace and applying the Wiener filter theory [39], the L-MMSE solution $\bar{\boldsymbol{G}}_{\mathrm{LMMSE}}(f)$ that minimizes $\mathrm{MSE}_{\mathrm{total}}$ in (38) is given by

$$\bar{\boldsymbol{G}}_{\mathrm{LMMSE}}(f) \triangleq \bar{\boldsymbol{R}}^{-1}(f)\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{P}}(f), \tag{41}$$

for each $f \in \mathcal{F}$. Note that the entry-by-entry Fourier transform of the L-MMSE solution $\boldsymbol{G}_{\mathrm{LMMSE}}(t)$ to the problem (15) can be easily obtained from $\bar{\boldsymbol{G}}_{\mathrm{LMMSE}}(f)$. This is because the VFT is an operation that divides the overall frequency band into the disjoint sub-bands with bandwidth $1/T$ centered around the Nyquist interval.

A key feature of this digital combiner is its ability to combine the analog combiner output $\boldsymbol{r}(t)$ by utilizing the dimensions N_{rf} from multiple RF chains and N_{f} from spectral redundancy. Since the received signal is a WSCS signal with cycle period $1/T$, there exists non-zero periodic correlation between frequency components spaced by $1/T$ [16]. As can be seen in *Theorem 1* and its solution (41), the VFT for matrix-valued functions facilitates the design of the digital combiner that fully exploits such spectral correlation.

B. PROOF OF THEOREM 1

In this subsection, we provide a detailed proof of *Theorem 1*. Unlike [21] that only considers a single RX antenna, we propose a new derivation for $\mathrm{MSE}_{\mathrm{total}}$ in our MU-MIMO system with interference, given the analog combiner output $\boldsymbol{r}(t)$.

We first rewrite the length-K decision statistic $\hat{\boldsymbol{d}}[l]$ as

$$\hat{\boldsymbol{d}}[l] = \int_{-\infty}^{\infty} \boldsymbol{G}^{\mathcal{H}}(t - lT)\boldsymbol{r}(t)\, dt \tag{42a}$$

$$= \sum_{l'=-\infty}^{\infty} \boldsymbol{C}((l-l')T)\boldsymbol{d}[l'] + \sum_{l''=-\infty}^{\infty} \boldsymbol{c}_{\mathrm{I}}((l-l'')T)d_{\mathrm{I}}[l''] + \boldsymbol{n}[l], \tag{42b}$$

where

$$\boldsymbol{C}(t) \triangleq \int_{-\infty}^{\infty} \boldsymbol{G}^{\mathcal{H}}(\tau - t)\boldsymbol{W}^{\mathcal{H}}\boldsymbol{P}(\tau)\, d\tau \tag{43a}$$

and

$$\boldsymbol{c}_{\mathrm{I}}(t) \triangleq \int_{-\infty}^{\infty} \boldsymbol{G}^{\mathcal{H}}(\tau - t)\boldsymbol{W}^{\mathcal{H}}\boldsymbol{p}_{\mathrm{I}}(\tau)\, d\tau \tag{43b}$$

are cross-correlation functions between $\boldsymbol{G}(t)$ and $\boldsymbol{W}^{\mathcal{H}}\boldsymbol{s}(t)$, and between $\boldsymbol{G}(t)$ and $\boldsymbol{W}^{\mathcal{H}}\boldsymbol{i}(t)$, respectively, and

$$\boldsymbol{n}[l] \triangleq \int_{-\infty}^{\infty} \boldsymbol{G}^{\mathcal{H}}(t - lT)\boldsymbol{W}^{\mathcal{H}}\boldsymbol{n}(t)\, dt \tag{43c}$$

is the Gaussian noise component with zero mean and variance N_0P_{N}.

Since $d_k[l]$ and $d_{\mathrm{I}}[l]$ are the i.i.d. proper-complex symbols with zero mean and unit variance for all k and l, we can rewrite $\mathrm{MSE}_{\mathrm{total}}$ as

$$\mathrm{MSE}_{\mathrm{total}} = \mathbb{E}\left\{\|\hat{\boldsymbol{d}}[l] - \boldsymbol{d}[l]\|^2\right\} \tag{44a}$$

$$= \mathbb{E}\left\{\|\boldsymbol{d}[l]\|^2\right\} - \mathrm{tr}\big(2\mathsf{Re}\{\boldsymbol{d}[l]\hat{\boldsymbol{d}}^{\mathcal{H}}[l]\}\big) + \mathbb{E}\left\{\|\hat{\boldsymbol{d}}[l]\|^2\right\} \tag{44b}$$

$$= K - \mathrm{tr}\left(2\mathsf{Re}\left\{\boldsymbol{C}(0)\right\}\right) + \mathrm{tr}\left(\boldsymbol{Q}\right) + N_0P_{\mathrm{N}}, \tag{44c}$$

where $\boldsymbol{Q}$ is a K-by-K matrix given by

$$\boldsymbol{Q} \triangleq \mathbb{E}\{\hat{\boldsymbol{d}}[l]\hat{\boldsymbol{d}}^{\mathcal{H}}[l]\} \tag{45a}$$

$$= \sum_{l=-\infty}^{\infty} \left\{\boldsymbol{C}(lT)\boldsymbol{C}^{\mathcal{H}}(lT) + \boldsymbol{c}_{\mathrm{I}}(lT)\boldsymbol{c}_{\mathrm{I}}^{\mathcal{H}}(lT)\right\}. \tag{45b}$$

Now, we express $\mathrm{tr}(\boldsymbol{Q})$ in the frequency domain.

Proposition 2: The third term $\mathrm{tr}(\boldsymbol{Q})$ in (44c) can be rewritten as

$$\mathrm{tr}(\boldsymbol{Q}) = \frac{1}{T}\int_{\mathcal{F}} \mathrm{tr}\left(\bar{\boldsymbol{G}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{P}}(f)\bar{\boldsymbol{P}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}\bar{\boldsymbol{G}}(f) + \bar{\boldsymbol{G}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{p}}_{\mathrm{I}}(f)\bar{\boldsymbol{p}}_{\mathrm{I}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}\bar{\boldsymbol{G}}(f)\right) df. \tag{46}$$

Proof: Let $c_{k,k}(t)$ and $c_{\mathrm{I},k}(t)$ be the (k,k)th entry of $\boldsymbol{C}(t)$ and the kth element of $\boldsymbol{c}_{\mathrm{I}}(t)$, respectively. Then, by using the definition of trace and Poisson summation formula, we can write the third term $\mathrm{tr}(\boldsymbol{Q})$ in (44c) as

$$\mathrm{tr}(\boldsymbol{Q}) = \sum_{l=-\infty}^{\infty}\left\{\sum_{k=1}^{K}|c_{k,k}(lT)|^2 + |c_{\mathrm{I},k}(lT)|^2\right\} \tag{47a}$$

$$= \frac{1}{T}\sum_{l=-\infty}^{\infty}\left\{\sum_{k=1}^{K}\tilde{v}_{k,k}\left(\frac{l}{T}\right) + \tilde{v}_{\mathrm{I},k}\left(\frac{l}{T}\right)\right\}, \tag{47b}$$

where $\tilde{v}_{k,k}(\xi)$ and $\tilde{v}_{\mathrm{I},k}(\xi)$ are the Fourier transforms of $|c_{k,k}(t)|^2$ and $|c_{\mathrm{I},k}(t)|^2$, respectively. Define $\tilde{\boldsymbol{G}}(\xi)$ and $\tilde{\boldsymbol{P}}(\xi)$ be the entry-by-entry Fourier transforms of $\boldsymbol{G}(t)$ and $\boldsymbol{P}(t)$, respectively. Then, using the definition of trace and convolution theorem, we have

$$\sum_{k=1}^{K}\tilde{v}_{k,k}\left(\frac{l}{T}\right) = \int_{-\infty}^{\infty}\mathrm{tr}\left(\tilde{\boldsymbol{G}}^{\mathcal{H}}(\xi)\boldsymbol{W}^{\mathcal{H}}\tilde{\boldsymbol{P}}(\xi)\tilde{\boldsymbol{P}}^{\mathcal{H}}\left(\xi-\frac{l}{T}\right)\boldsymbol{W}\tilde{\boldsymbol{G}}\left(\xi-\frac{l}{T}\right)\right)d\xi \tag{48a}$$

and

$$\sum_{k=1}^{K}\tilde{v}_{\mathrm{I},k}\left(\frac{l}{T}\right) = \int_{-\infty}^{\infty}\mathrm{tr}\left(\tilde{\boldsymbol{G}}^{\mathcal{H}}(\xi)\boldsymbol{W}^{\mathcal{H}}\tilde{\boldsymbol{p}}_{\mathrm{I}}(\xi)\tilde{\boldsymbol{p}}_{\mathrm{I}}^{\mathcal{H}}\left(\xi-\frac{l}{T}\right)\boldsymbol{W}\tilde{\boldsymbol{G}}\left(\xi-\frac{l}{T}\right)\right)d\xi. \tag{48b}$$

By substituting (48a) and (48b) into (47b), the summation over l leads to spectrum folding with interval $1/T$. Since the integration interval can be divided into sub-intervals of length $1/T$, the term $\mathrm{tr}(\boldsymbol{Q})$ can be expressed in terms of the integration over the Nyquist interval $\mathcal{F}$ by using the VFT in *Definition 1* and (40). Therefore, the conclusion follows. □

Similarly to the above derivation, we can write the second term $\mathrm{tr}(\boldsymbol{C}(0))$ in (44c) as

$$\mathrm{tr}(\boldsymbol{C}(0)) = \int_{\mathcal{F}}\mathrm{tr}\left(\bar{\boldsymbol{G}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{P}}(f)\right)df \tag{49}$$

and the fourth term P_{N} in (44c) as

$$P_{\mathrm{N}} = \int_{\mathcal{F}}\mathrm{tr}\left(\bar{\boldsymbol{G}}^{\mathcal{H}}(f)\bar{\boldsymbol{W}}^{\mathcal{H}}\bar{\boldsymbol{W}}\bar{\boldsymbol{G}}(f)\right)df. \tag{50}$$

Therefore, by substituting (46), (49), and (50) into (44c), we finally have (38) in *Theorem 1*.

As shown in Fig. 2, the digital combiner processes a length-N_{rf} oversampled DT signal with finite bit resolution, so that there is non-zero quantization noise in the digital combiner input. If the analog combiner fails to suppress the strong interference, there is dominant quantization noise that may lead to a severe degradation of the reception performance. Hence, its performance depends on how the analog combining matrix $\boldsymbol{W}$ is designed. In the next section, we will numerically investigate the reception performance of the proposed RX.

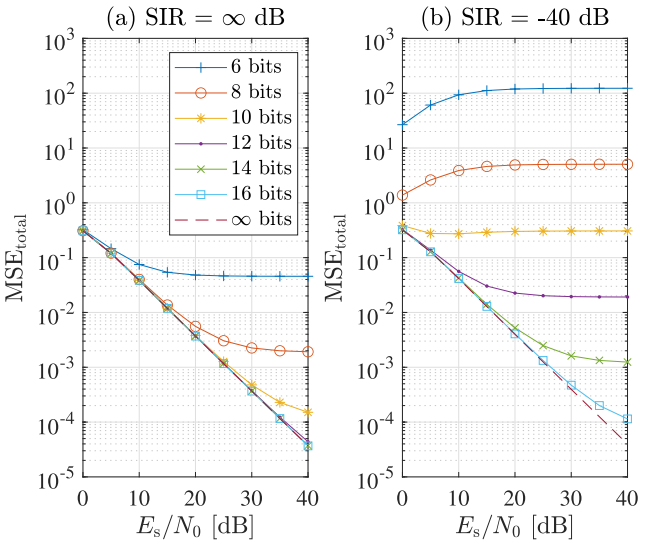

FIGURE 6. Total MSE as a function of E_{s}/N_0 for different numbers of ADC bits, when the RX does not adopt the hybrid combiner structure with $K = 4$ and $N_{\mathrm{R}} = N_{\mathrm{rf}} = 16$.

V. SIMULATION RESULTS

In this section, simulation results are provided. Throughout the simulations, we consider $K = 4$, $N_{\mathrm{R}} = 16$, and, if the hybrid combiner structure is adopted, $N_{\mathrm{rf}} = 8$. We assume that $u_k(t)$ and $u_{\mathrm{I}}(t)$ are the square-root raised cosine (SRRC) shaping filter with excess bandwidth $\beta_k = \beta_{\mathrm{I}} = 0.5$, for all k. The TXs employ linear modulation with symbol rate $1/T = 10$ MHz and bandwidth $B_k = B_{\mathrm{I}} = 7.5$ MHz in complex baseband, for all k. Hence, this underlay system fully utilizes the frequency band occupied by the legacy system. The channel parameters for the underlay TXs are based on the 3GPP SCM's urban micro (Umi) environment with NLOS scenario [33], and those for the legacy TX is based on our exemplary multipath intensity profile shown in Fig. 4. We define post-combined signal-to-noise ratio (SNR) denoted by E_{s}/N_0. This metric is the average SNR multiplied by N_{R}, and it is used to evaluate the performance after removing the effect of array gain.

First, we investigate the reception performance of an RX that does not adopt the hybrid combiner structure for strong interference suppression. This RX directly connects all N_{R} antenna elements to $N_{\mathrm{rf}} = N_{\mathrm{R}}$ RF chains and performs only the digital combining with $\boldsymbol{G}_{\mathrm{LMMSE}}(t)$ given $\boldsymbol{r}(t) = \boldsymbol{y}(t) = \boldsymbol{s}(t)+\boldsymbol{i}(t)+\boldsymbol{n}(t)$. Fig. 6 illustrates the total MSE as a function of E_{s}/N_0, for different numbers of ADC bits. In Fig. 6-(a),

there is no legacy TX, i.e., signal-to-interference ratio (SIR) is ∞ dB, but in Fig. 6-(b), there is a strong interference from the legacy TX with SIR $= -40$ dB. In Fig. 6-(a) with SIR $= \infty$ dB, it can be seen that the total MSE with infinite precision, denoted by "∞ bits", can almost be achieved when using ADCs with 12 bits or higher, and that using 10-bit ADCs result in a slight performance degradation. Hence, the total MSE shown in Fig. 6-(a) serves as the achievable lower bound for each number of ADC bits, given K and N_{R}. However, in Fig. 6-(b) with SIR $= -40$ dB, it can be seen that the dominant quantization noise caused by the strong interference results in severe performance degradation. For example, the total MSE achievable with 8-bit ADCs in Fig. 6-(a) can be outperformed by using at least 14-bit ADCs in Fig. 6-(b). This implies that, in the presence of strong interference, directly connecting all antenna elements to RF chains requires significantly higher ADCs' power consumption, which grows exponentially with the number of bits [30]. Moreover, it requires higher computational complexity for digital signal processors, because it needs to process the received signals from all antenna elements.

Then, we investigate the results from using the hybrid combiner structure. We mainly consider three RXs according to their combinations of analog and digital combiners. Note that the following three RXs have a commonality in that their digital combiners generate the L-MMSE estimates of symbols given the ADC output signals, which are differently derived depending on how the analog combiner is designed. First, the RX in Fig. 6 is considered and will be denoted by "no $\boldsymbol{W}$" in figures. Second, an RX adopting the analog combiner but maximizing only the total power of desired signals [24] is considered and will be denoted by "$\boldsymbol{W}_{\epsilon=1}$" in figures. The sub-script $\epsilon = 1$ means that, as described in Section III-B, this analog combiner does not suppress the strong interference with $\boldsymbol{A}_{\mathrm{eff}} = \boldsymbol{I}_{N_{\mathrm{R}}}$. Third, our proposed RX is considered and will be denoted by "$\boldsymbol{W}_{\mathrm{PROP}}$" in figures. For the first two baseline RXs, we present the performance in both cases with and without the strong interference from the legacy TX. Hence, the RX "no $\boldsymbol{W}$" with SIR $= \infty$ dB presents the achievable performance bound shown in Fig. 6-(a). Moreover, the RX "$\boldsymbol{W}_{\epsilon=1}$" with SIR $= \infty$ dB presents another performance bound achievable under the hybrid combiner structure.

To proceed, we introduce a new metric called *effective step size*, denoted by Δ_{eff}, and define it as

$$\Delta_{\mathrm{eff}} \triangleq \frac{2}{2^q - 1}\sqrt{1 + \frac{1}{\mathrm{SINR}_{\mathrm{total}}}}, \tag{51}$$

where q denotes the number of ADC bits. The conventional *step size* is determined by the quantization range, e.g., $[-1, 1)$ in our simulation, and the number of ADC bits q, and defined as $2/(2^q - 1)$. However, the quantization precision experienced by $\boldsymbol{W}^{\mathcal{H}}\boldsymbol{s}(t)$ depends on the total SINR after analog combining. Hence, the metric Δ_{eff} in (51) enables us to numerically analyze the quantization noise effect in the presence of the residual interference at the analog combiner output.

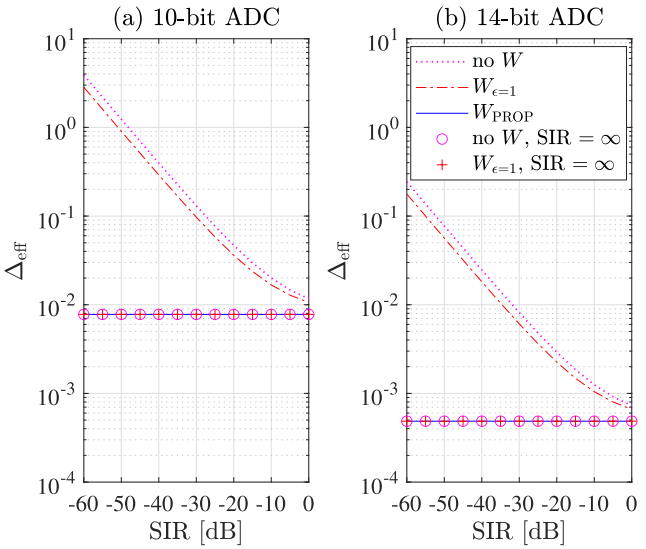

FIGURE 7. Effective step size experienced by desired signals as a function of SIR, when (a) 10-bit ADCs and (b) 14-bit ADCs are employed, and $E_{\mathrm{s}}/N_0 = 40$ dB.

Fig. 7 illustrates the effective step size as a function of SIR, when $E_{\mathrm{s}}/N_0 = 40$ dB. It can be seen that the effective step size in Fig. 7-(b) employing 14-bit ADCs is smaller than in Fig. 7-(a) employing 10-bit ADCs. In both Figs. 7-(a) and (b), it can be seen that the two baseline RXs suffer from a drastic increase in step size for low SIR. However, since our proposed analog combiner reduces the interference power to a negligible level and improves the total SINR, its step size is almost similar to the two lower bounds regardless of SIR. Hence, our proposed RX successfully mitigates the dominant quantization noise caused by the strong interference from the legacy TX.

Fig. 8 illustrates effective step size and total MSE as a function of the number of ADC bits, when SIR $= -40$ dB, and $E_{\mathrm{s}}/N_0 = 40$ dB. In Fig. 8-(a), it can be seen that the effective step size decreases as the number of ADC bits increases. Again, for the two RXs without strong interference suppression at the analog combiner, we can observe a drastically larger effective step size. On the other hand, the ADC outputs of the proposed RX suffer from the quantization noise due to only the finite resolution inherent to the ADCs. In Fig. 8-(b), it can be seen that reception performance is improved as the effective size decreases, and it converges to performance with infinite resolution as the number of ADC bits increases. For the two baseline RXs, we can observe a severe degradation in the total MSE due to the drastically large effective step size. On the other hand, our proposed RX nearly approaches the two lower bounds achievable in the absence of the legacy TX. The slight differences in the total MSE occurs because our analog combiner consumes some degrees of freedom for interference suppression, while the analog combiner $\boldsymbol{W}_{\epsilon=1}$ in the absence of the legacy TX fully

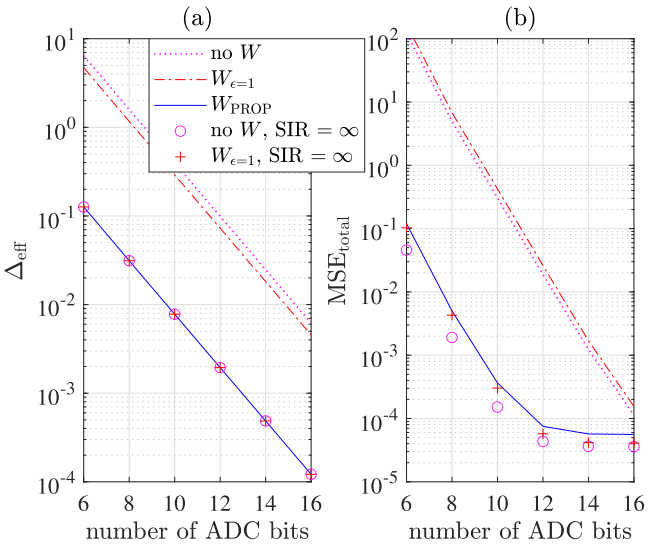

FIGURE 8. (a) Effective step size and (b) total MSE as a function of number of ADC bits, when SIR = −40 dB and E_s/N_0 = 40 dB.

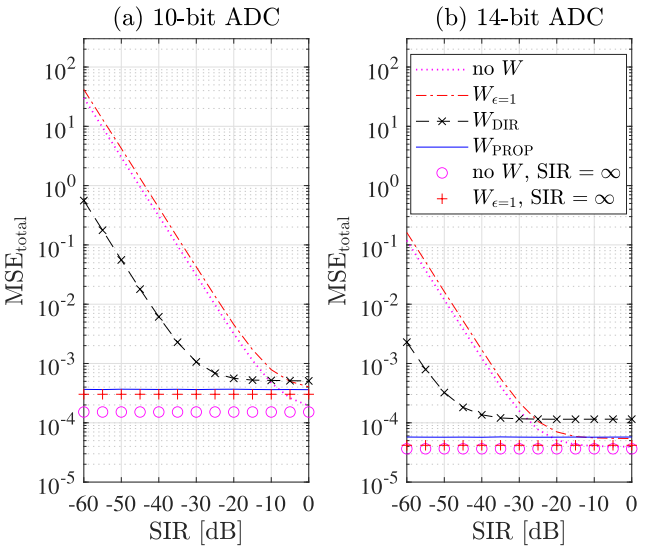

FIGURE 10. Total MSE as a function of SIR, when (a) 10-bit ADCs and (b) 14-bit ADCs are employed, and E_s/N_0 = 40 dB.

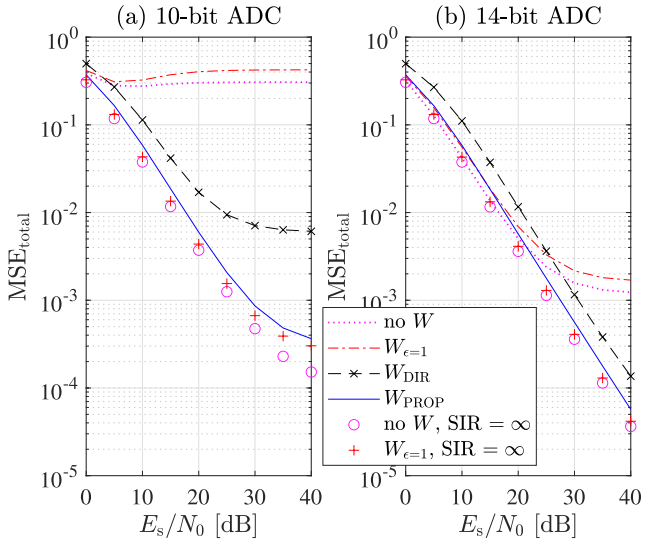

FIGURE 9. Total MSE as a function of E_s/N_0, when (a) 10-bit ADCs and (b) 14-bit ADCs are employed, and SIR = −40 dB.

utilizes dimensions to maximize the total power of desired signals.

Fig. 9 illustrates total MSE as a function of E_s/N_0, when SIR = −40 dB. Here, we introduce and compare a new combination of analog and digital combiners. This analog combiner performs nulling in the space domain towards the direction of the strong interference [32], denoted by "$\boldsymbol{W}_{\mathrm{DIR}}$", and the digital combiner performs L-MMSE filtering. Unlike [32], this analog combiner also maximizes the total power of desired signals after nulling the strong interference. It can be seen in Fig. 9-(a) that, when 10-bit ADCs are employed, there is a severe degradation for the two baseline RXs. Moreover, we can observe performance degradation when the analog combiner suppresses the strong interference by using the direction-based approach. This is because the angular dispersion shown in Fig. 4 requires more dimensions than $N_{\mathrm{R}} - N_{\mathrm{rf}}$ to generate nulls at the AoAs of the strong interference and leads to unintended power loss in the desired signals. On the other hand, in both Figs. 9-(a) and (b), our proposed RX approaches the MSE lower bounds in the absence of strong interference only with a slight performance degradation.

Fig. 10 illustrates total MSE as a function of SIR, when E_s/N_0 = 40 dB. It can be seen that, again, when the strong interference suppression fails, the dominant quantization noise causes severe performance degradation for low SIR. For both two baseline RXs, as SIR approaches to 0 dB, their performance approach to their achievable lower bounds in the absence of the legacy TX. On the other hand, our proposed RX well approximates the null-space projection onto the interference correlation matrix, so that the performance remains nearly close to the interference-free lower bounds regardless of SIR.

Fig. 11 illustrates total MSE as a function of E_s/N_0 and excess bandwidth β, when (a) β = 0.5 and (b) E_s/N_0 = 40 dB. Here, we compare the performance of the RXs based on whether the digital combiner is designed to fully exploit the periodic spectral correlation of the WSCS signals or not. The proposed digital combiner is denoted by "$\boldsymbol{G}_{\mathrm{LMMSE}}(t)$" and the digital combiner that does not exploit the spectral correlation is denoted by "no VFT" in the figure. The digital combiner "no VFT" is designed under the assumption that the signals are WSS. It can be seen that fully exploiting the spectral correlation of WSCS signals is crucial for the MAI suppression in the MU-MIMO system. In Fig. 11-(a), there is a severe performance degradation for "no VFT" due to failure in MAI suppression. In Fig. 11-(b), the difference in $\mathrm{MSE}_{\mathrm{total}}$ increases as β increases. This is because, as spectral redundancy increases with higher β [16], failure to exploit this redundancy leads to further decrease of MAI suppression

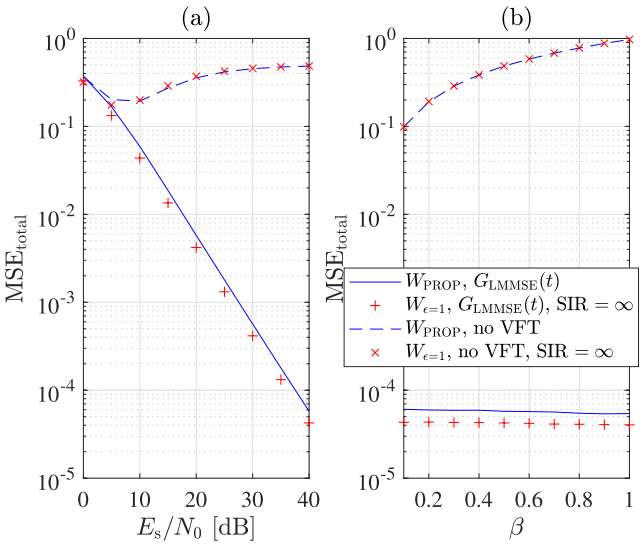

FIGURE 11. Total MSE as a function of (a) E_s/N_0 and (b) excess bandwidth β, when 14-bit ADCs are employed, SIR $= -40$ dB, (a) $\beta = 0.5$, and (b) $E_s/N_0 = 40$ dB.

capability. On the other hand, our proposed RX that fully exploits such spectral redundancy can successfully suppress the MAI.

VI. CONCLUSION AND FUTURE WORK

In this paper, we design an MU-MIMO underlay RX in the presence of a strong WSCS legacy signal. The dominant quantization noise at the ADC caused by the strong interference drastically degrade the reception performance. To mitigate strong interference, we adopt the hybrid combiner structure consisting of analog and digital combiners. In our analog combiner design, we first suppress the strong interference by approximating null-space projection onto the interference correlation matrix. Then, we enhance the total SINR by performing the PCA on the desired signal correlation matrix. Finally, we adjust the magnitudes of the ADC input signals. In our digital combiner design, we present the derivation for L-MMSE estimation of symbols. To exploit the periodic spectral correlation of WSCS signals, we extend the VFT to matrix-valued functions. Simulation results confirm that, under certain conditions, our proposed RX can approach the reception performance achievable in the absence of the legacy TX. Since we assume that the channels from the legacy TX and underlay TXs are known, the channel estimation under the proposed hybrid combiner structure may be an interesting future research direction. To further reduce hardware costs, considering the hybrid structure that omits the VGAs in the analog combiner or that employs a partially-connected hybrid structure can also be a future research topic.

APPENDIX.

In this appendix, we provide the proof of *Lemma 1*.

Since the signal $\boldsymbol{i}(t)$ is a WSCS process with cycle period T, the time average is conducted for a finite interval T, so that

$$\mathbb{A}\left[\mathbb{E}\left\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\right\}\right] = \frac{1}{T}\int_0^T \mathbb{E}\{\boldsymbol{i}(t)\boldsymbol{i}^{\mathcal{H}}(t)\}\,dt \tag{52a}$$

$$= \frac{1}{T}\int_0^T \sum_{l=-\infty}^{\infty} \boldsymbol{p}_{\mathrm{I}}(t-lT)\boldsymbol{p}_{\mathrm{I}}^{\mathcal{H}}(t-lT)\,dt \tag{52b}$$

$$= \frac{1}{T}\int_{-\infty}^{\infty} \boldsymbol{p}_{\mathrm{I}}(t)\boldsymbol{p}_{\mathrm{I}}^{\mathcal{H}}(t)\,dt \tag{52c}$$

$$= \frac{1}{T}\int_{-B_{\mathrm{I}}}^{B_{\mathrm{I}}} \tilde{\boldsymbol{p}}_{\mathrm{I}}(\xi)\tilde{\boldsymbol{p}}_{\mathrm{I}}^{\mathcal{H}}(\xi)\,d\xi, \tag{52d}$$

where (52b) comes from the definition of $\boldsymbol{i}(t)$ in (8) with i.i.d. symbols $d_{\mathrm{I}}[l]$ and (52d) comes from the Parseval's relation. Therefore, the conclusion follows.

REFERENCES

[1] N. Kaabouch and W.-C. Hu, *Handbook of Research on Software-Defined and Cognitive Radio Technologies for Dynamic Spectrum Management*. Hershey, PA, USA: IGI Global, 2014.
[2] A. M. Voicu, L. Simic, and M. Petrova, "Survey of spectrum sharing for inter-technology coexistence," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1112–1144, 2nd Quart., 2019.
[3] L. Zhang, Y.-C. Liang, and M. Xiao, "Sharing for Internet of Things: A survey," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 132–139, Jun. 2019.
[4] W. S. H. M. W. Ahmad, N. A. M. Radzi, F. S. Samidi, A. Ismail, F. Abdullah, M. Z. Jamaludin, and M. N. Zakaria, "5G technology: Towards dynamic spectrum sharing using cognitive radio networks," *IEEE Access*, vol. 8, pp. 14460–14488, 2020.
[5] Y. Noam and A. J. Goldsmith, "Blind null-space learning for MIMO underlay cognitive radio with primary user interference adaptation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1722–1734, Apr. 2013.
[6] N. Nartasilpa, A. Salim, D. Tuninetti, and N. Devroye, "Communications system performance and design in the presence of radar interference," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 4170–4185, Sep. 2018.
[7] Y. Li, L. Zheng, M. Lops, and X. Wang, "Interference removal for radar/communication co-existence: The random scattering case," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4831–4845, Oct. 2019.
[8] C. D'Andrea, S. Buzzi, and M. Lops, "Communications and radar coexistence in the massive MIMO regime: Uplink analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 19–33, Jan. 2020.
[9] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Phys. Commun.*, vol. 4, no. 1, pp. 40–62, Mar. 2011.
[10] J. B. Evans. (2016). *Shared Spectrum Access for Radar and Communications (SSPARC) DARPA*. [Online]. Available: http://www.darpa.mil/program/shared-spectrum-access-for-radar-and-communications
[11] W. Webb, "Broadband fixed wireless access as a key component of the future integrated communications environment," *IEEE Commun. Mag.*, vol. 39, no. 9, pp. 115–121, Sep. 2001.
[12] A. Lappalainen, Y. Zhang, and C. Rosenberg, "Planning 5G networks for rural fixed wireless access," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 1, pp. 441–455, Mar. 2023.
[13] W. Gardner and L. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 1, pp. 4–14, Jan. 1975.
[14] W. A. Gardner Ed., *Cyclostationarity in Communications and Signal Processing*. New York, NY, USA: IEEE Press, 1994.
[15] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, Apr. 2006.
[16] A. Napolitano, *Cyclostationary Processes and Time Series: Theory, Applications, and Generalizations*. San Diego, CA, USA: Academic, 2019.
[17] Y. H. Yun and J. H. Cho, "An optimal orthogonal overlay for a cyclostationary legacy signal," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1557–1567, May 2010.

[18] B. W. Han and J. H. Cho, "Capacity of second-order cyclostationary complex Gaussian noise channels," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 89–100, Jan. 2012.
[19] J. Yeo, B. W. Han, J. H. Cho, and J. S. Lehnert, "Capacity of an orthogonal overlay channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6111–6124, Nov. 2015.
[20] J. H. Cho, "Joint transmitter and receiver optimization in additive cyclostationary noise," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3396–3405, Dec. 2004.
[21] J. H. Cho and W. Gao, "Continuous-time equivalents of Welch bound equality sequences," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3176–3185, Sep. 2005.
[22] J. Chae, J. Choi, J. Kim, and J. H. Cho, "A low-complexity widely-linear MMSE equalizer for DFT-spread OFDM with frequency-domain spectrum shaping," *IEEE Trans. Wireless Commun.*, early access. [Online]. Available: https://ieeexplore.ieee.org/document/10236978
[23] H. Krishnaswamy and L. Zhang, "Analog and RF interference mitigation for integrated MIMO receiver arrays," *Proc. IEEE*, vol. 104, no. 3, pp. 561–575, Mar. 2016.
[24] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
[25] A. F. Molisch, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.
[26] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3060–3097, 4th Quart., 2018.
[27] S. Mondal, R. Singh, A. I. Hussein, and J. Paramesh, "A 25–30 GHz fully-connected hybrid beamforming receiver for MIMO communication," *IEEE J. Solid-State Circuits*, vol. 53, no. 5, pp. 1275–1287, May 2018.
[28] S. Mondal and J. Paramesh, "Power-efficient design techniques for mm-wave hybrid/digital FDD/full-duplex MIMO transceivers," *IEEE J. Solid-State Circuits*, vol. 55, no. 8, pp. 2011–2026, Aug. 2020.
[29] T. Gong, N. Shlezinger, S. S. Ioushua, M. Namer, Z. Yang, and Y. C. Eldar, "RF chain reduction for MIMO systems: A hardware prototype," *IEEE Syst. J.*, vol. 14, no. 4, pp. 5296–5307, Dec. 2020.
[30] V. Venkateswaran and A.-J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.
[31] W. Zhang, Y. Jiang, B. Zhou, and D. Hu, "Hybrid interference mitigation using analog prewhitening," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6595–6605, Oct. 2021.
[32] K. Wu, J. A. Zhang, X. Huang, Y. J. Guo, D. N. Nguyen, A. Kekirigoda, and K.-P. Hui, "Analog-domain suppression of strong interference using hybrid antenna array," *Sensors*, vol. 22, no. 6, p. 2417, Mar. 2022.
[33] *Technical Specification Group Radio Access Network; Study on Channel Model for Frequencies From 0.5 to 100 GHz*, Standard 3GPP TR 38.901 v17.0.0, Release 17, 3rd Generation Partnership Project, Sophia Antipolis, France, Mar. 2022.
[34] G. Strang, *Linear Algebra and Its Applications*, 4th ed. New York, NY, USA: Academic, 1976.
[35] F. Sun, J. Singh, and U. Madhow, "Automatic gain control for ADC-limited communication," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, pp. 1–5, Dec. 2010.
[36] R. W. Heath Jr., *Introduction to Wireless Digital Communication: A Signal Processing Perspective*, 1st ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2017.
[37] A. Kumar, I. B. Franco de Almeida, N. Franchi, and G. Fettweis, "A deep neural network based environment sensing in the presence of jammers," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–7.
[38] T. J. Shan and T. Kailath, "Adaptive algorithms with an automatic gain control feature," *IEEE Trans. Circuits Syst.*, vol. 35, no. 1, pp. 122–127, Jan. 1988.
[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

JOOHEE CHAE received the B.S. degree in electronic engineering from Sogang University, Seoul, South Korea, in 2017, and the M.S. degree in electrical engineering from the Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2019, where he is currently pursuing the Ph.D. degree. His research interests include multiuser communications and MIMO signal processing.

JOON HO CHO (Member, IEEE) received the B.S. degree (summa cum laude) in electrical engineering from Seoul National University, Seoul, South Korea, in 1995, and the M.S.E.E. and Ph.D. degrees in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 1997 and 2001, respectively.

From 2001 to 2004, he was an Assistant Professor with the University of Massachusetts at Amherst. Since July 2004, he has been with the Pohang University of Science and Technology (POSTECH), Pohang, South Korea, where he is currently an Associate Professor with the Department of Electrical Engineering. His research interests include waveform design, multiuser communications, MIMO signal processing, channel measurement and modeling, and information theory.

Dr. Cho served as an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He served as the Vice President for Admissions and Student Affairs of POSTECH.