High Dynamic Range mmWave Massive MU-MIMO with Householder Reflections

Victoria Palhares, Gian Marti, Oscar Castañeda, and Christoph Studer

Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland
e-mail: vmenescal@ethz.ch, gimarti@ethz.ch, caoscar@ethz.ch, and studer@ethz.ch

Abstract—All-digital massive multiuser (MU) multiple-input multiple-output (MIMO) at millimeter-wave (mmWave) frequencies is a promising technology for next-generation wireless systems. Low-resolution analog-to-digital converters (ADCs) can be utilized to reduce the power consumption of all-digital basestation (BS) designs. However, simultaneously transmitting user equipments (UEs) with vastly different BS-side receive powers either drown weak UEs in quantization noise or saturate the ADCs. To address this issue, we propose *high dynamic range (HDR) MIMO*, a new paradigm that enables simultaneous reception of strong and weak UEs with low-resolution ADCs. HDR MIMO combines an adaptive analog spatial transform with digital equalization: The spatial transform focuses strong UEs on a subset of ADCs in order to mitigate quantization and saturation artifacts; digital equalization is then used for data detection. We demonstrate the efficacy of HDR MIMO in a massive MU-MIMO mmWave scenario that uses Householder reflections as spatial transform.

I. INTRODUCTION

Next-generation wireless systems are expected to combine communication at millimeter-wave (mmWave) frequencies with massive multiuser (MU) multiple-input multiple-output (MIMO) technology [2]. While communication at mmWave frequencies provides access to large, contiguous, and available portions of the frequency spectrum, massive MU-MIMO provides beamforming gains that combat the high path-loss at mmWave frequencies and enables communication with multiple user equipments (UEs) in the same time-frequency resource.

The deployment of all-digital massive MIMO basestations (BSs), i.e., BSs where each antenna has a dedicated radio-frequency (RF) chain and dedicated data converters, comes at the cost of increased digital circuit complexity, power consumption, and interconnect data rates [3], [4]. To alleviate these drawbacks, one can use low-resolution analog-to-digital converters (ADCs) in the uplink [5]. However, low-resolution ADCs do not quantize high dynamic range signals properly. In a scenario where one UE is significantly stronger than the other UEs, weak UE signals either drown in quantization noise, if the BS adjusts the ADC input gains to the strongest UE, or the receive signals are distorted due to ADC saturation caused by the strongest UE, if the BS adjusts the ADC input gains to the weakest UEs. Since quantization and saturation effects

are, in general, irreversible, digital equalization cannot fully mitigate such distortions [6].

A. Contributions

We propose *high dynamic range (HDR) MIMO*, a new paradigm that deals with receive signals of high dynamic range in all-digital BS architectures that are equipped with low-resolution ADCs. HDR MIMO utilizes one or multiple adaptive analog spatial transforms that focus either the energy of the strongest UE or the signal dimension with the highest receive power on a pair of ADCs (for the in-phase and quadrature components) per transform; digital equalization is then used to detect the transmitted data of the strong UE as well as of the weak UEs. For the adaptive analog spatial transforms, we propose two distinct Householder reflections (HRs): (i) strongest UE isolation (HR-ISO), which isolates the energy of the strongest UE on dedicated ADCs, and (ii) maximum power isolation (HR-MAX), which focuses the signal dimension with the highest receive power on dedicated ADCs. We demonstrate the efficacy of HDR MIMO by simulating a mmWave massive MU-MIMO uplink system with an all-digital BS that utilizes low-resolution ADCs and channel vectors from a commercial ray tracer.

B. Related Work

Past work on interference or jammer mitigation has led to the development of methods that perform spatial nulling of strong signals; see, e.g., [7]–[10]. For example, spatial projection onto the orthogonal subspace (POS, for short) of the jammer channel has been proposed for jammer mitigation in [7], [8]. Since the content of the interference or jamming signal is typically of no interest, it can be nulled and discarded. In contrast, HDR MIMO assumes that all signal sources are of interest and enables the simultaneous detection of weak and strong received signals. Furthermore, if one is interested in analyzing interferer or jammer signals in the context of jammer mitigation, then HDR MIMO provides means to recover such signals jointly with detecting the legitimate UE data.

The use of analog transforms to perform spatial filtering for systems with low-resolution BS designs has been proposed in [9], [10] for jammer mitigation, and in [11] for strong adjacent channel interference mitigation. In [9], the authors propose HERMIT, which uses an adaptive analog transform to

remove most of the jammer's energy prior to sampling by the ADCs. In [10], the authors propose beam-slicing, a nonadaptive analog transform that focuses the jammer energy onto few ADCs. In [11], the authors propose a hybrid beamformer with spatial filtering that adaptively mitigates interference. In contrast to these methods, we propose the use of analog spatial HRs that focus the energy of the strongest UE or the signal dimension with the highest receive power onto a subset of the available ADCs without discarding any signals, thus enabling the simultaneous recovery of strong and weak UE signals.

In MU communication, one of the challenges is the near-far problem, i.e., if a UE is physically much closer to the BS than another UE, then the BS-side receive power of the close UE is often much stronger than that of the other UEs. Power control at the UE side is typically used to confine the dynamic range at the receiver side [12]–[15], reducing the necessity of HDR MIMO. If, however, there exist UEs, rogue transmitters, interferers, or jammers in the same or nearby frequency bands that do *not* adhere to such power control mechanisms, then HDR MIMO still enables reliable data transmission.

C. Notation

Upper case and lower case bold symbols denote matrices and column vectors, respectively; uppercase calligraphic letters denote sets. We use $[\mathbf{a}]_i$ for the ith element of vector $\mathbf{a}$, $[\mathbf{A}]_{(i,j)}$ for the element in the ith row and jth column of $\mathbf{A}$, and $\mathbf{a}_j$ for the jth column of $\mathbf{A}$. The $M \times M$ identity matrix is $\mathbf{I}_M$ and the M-dimensional unit vector is $\mathbf{e}_i$, where the ith entry is one and the rest is zero. The superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the complex conjugate, transpose, and Hermitian transpose, respectively. A diagonal matrix with the vector $\mathbf{a}$ on its main diagonal is $\mathrm{diag}(\mathbf{a})$. The Euclidean norm of $\mathbf{a}$ is $\|\mathbf{a}\|$. The eigenvalue decomposition of a symmetric matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$ is $\mathbf{A} = \mathbf{L}\boldsymbol{\Lambda}\mathbf{L}^H = \sum_{m=1}^{M} \lambda_m \boldsymbol{\ell}_m \boldsymbol{\ell}_m^H$, where λ_m, $m = 1, \ldots, M$, are the eigenvalues, which are sorted in descending order, and $\boldsymbol{\ell}_m$ are the corresponding unit-length eigenvectors. The operators $\Re\{a\}$ and $\Im\{a\}$ extract the real and imaginary part of $a \in \mathbb{C}$, respectively. The complex-valued sign (or phase) of $a \in \mathbb{C}$ is $\mathrm{sign}(a) = a/|a|$, which we define to be 1 if $a = 0$. The floor function $\lfloor a \rfloor$ yields the greatest integer less than or equal to $a \in \mathbb{R}$. The expectation operator is $\mathbb{E}[\cdot]$. We use j as the imaginary unit.

II. PREREQUISITES

A. System Model

We consider the uplink of a mmWave massive MU-MIMO system with U single-antenna UEs transmitting data to an all-digital BS with B antennas. We consider frequency-flat channels with the following baseband input-output relation:

$$\mathbf{y} = \mathbf{G}\mathbf{D}\mathbf{s} + \mathbf{n}. \tag{1}$$

Here, $\mathbf{y} \in \mathbb{C}^B$ is the unquantized BS-side receive vector, $\mathbf{G} \in \mathbb{C}^{B \times U}$ is the uplink channel matrix, $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_U) \in \mathbb{R}^{U \times U}$ is the power control matrix, $\mathbf{s} \in \mathcal{X}^U$ is the transmit symbol vector with entries taken from a constellation $\mathcal{X}$, and $\mathbf{n} \in \mathbb{C}^B$ models noise as an i.i.d. circularly-symmetric
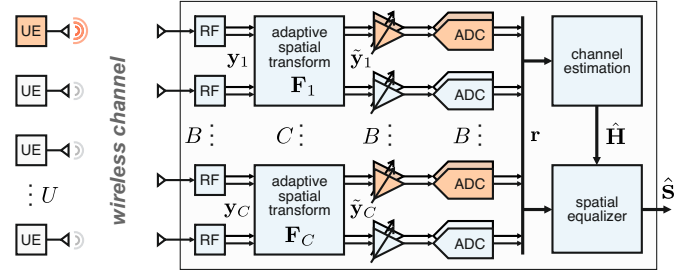

Fig. 1. HDR MIMO receiver architecture: The RF chains, adaptive analog spatial transforms $\mathbf{F}_c$, $c = 1, \ldots, C$, automatic gain controllers (AGCs), and ADCs are divided into C antenna clusters. In a scenario in which the BS-side receive power of one UE (highlighted in orange) is much stronger than that of the other UEs, each adaptive analog spatial transform focuses the energy of the strongest UE on two AGCs and two ADCs per cluster (highlighted in orange); this prevents saturation of the remaining ADCs in the same cluster. The transmitted UE data can then be recovered by a spatial equalizer.

complex Gaussian random vector with variance N_0 per entry. The constellation $\mathcal{X}$ is normalized so that $\mathbb{E}[|[\mathbf{s}]_u|^2] = 1$, $u = 1, \ldots, U$. We define the effective channel matrix as $\mathbf{H} = \mathbf{G}\mathbf{D}$. We assume that the columns of $\mathbf{H}$ are sorted in descending order with respect to the column norms $\|\mathbf{h}_u\|$, $u = 1, \ldots, U$. The matrices $\mathbf{G}$ and $\mathbf{D}$ are sorted accordingly.

In what follows, we consider a scenario in which the BS-side receive power of the first UE is much stronger than that of the other UEs. We model such a situation by assuming that all UEs except for the first UE adhere to power control, which is expressed through the per-UE gains $d_2, \ldots, d_U$ of the power control matrix $\mathbf{D}$ (cf. Section IV-A for the details on the used power control strategy). For the first UE, the gain d_1 can be expressed by the dynamic range ρ (in decibels) of the BS-side receive power between the strongest and weakest UEs as

$$\rho = 10 \log_{10}\left(\frac{d_1^2 \|\mathbf{g}_1\|^2}{d_U^2 \|\mathbf{g}_U\|^2}\right). \tag{2}$$

As a result, the squared Euclidean norm of the strongest UE's channel vector $\mathbf{h}_1$ is ρ decibels larger than the squared norm of the weakest UE's channel vector $\mathbf{h}_U$. Here, $\mathbf{h}_1$ and $\mathbf{h}_U$ correspond to the first and last columns of the effective channel matrix $\mathbf{H}$, respectively.

To estimate the channel matrix $\mathbf{H}$, we assume block fading where the UEs transmit orthogonal pilots $\mathbf{S}_T \in \mathbb{C}^{U \times K}$ for the duration of $K \geq U$ time slots. Each column of $\mathbf{S}_T$ contains the pilots of all UEs transmitted during one training time slot. The (unquantized) receive vectors in the training phase are modeled as $\mathbf{Y}_T = \mathbf{H}\mathbf{S}_T + \mathbf{N}_T$, where each column in the receive matrix $\mathbf{Y}_T \in \mathbb{C}^{B \times K}$ and the noise matrix $\mathbf{N}_T \in \mathbb{C}^{B \times K}$ corresponds to one training time slot. We estimate the effective channel matrix $\mathbf{H}$ using the least-squares estimator $\hat{\mathbf{H}} = \mathbf{Y}_T \mathbf{S}_T^H (\mathbf{S}_T \mathbf{S}_T^H)^{-1}$. In what follows, we define $\underline{\mathbf{h}} = \mathbf{h}_1$ as the channel vector associated with the first (strongest) UE and $\hat{\underline{\mathbf{h}}}$ as its estimate.

B. High Dynamic Range (HDR) MIMO Receiver

We now describe the HDR MIMO BS architecture depicted in Fig. 1, which is able to deal with receive signals of high

dynamic range. In order to prevent weak UEs from drowning in quantization noise or strong UEs from saturating all ADCs, we introduce adaptive analog spatial transforms prior to the automatic gain controllers (AGCs) and ADCs. We focus on linear transforms represented by the matrices $\mathbf{F}_c \in \mathbb{C}^{S \times S}$, $c = 1, \ldots, C$, where each matrix operates on a cluster of $S = B/C$ antennas with C such antenna clusters in total.[1] These analog spatial transforms implement

$$\tilde{\mathbf{y}}_c = \mathbf{F}_c \mathbf{y}_c, \ c = 1, \ldots, C, \tag{3}$$

where $\mathbf{y}_c \in \mathbb{C}^S$ is the cth cluster of the receive vector $\mathbf{y} = \left[\mathbf{y}_1^\mathrm{T}, \ldots, \mathbf{y}_C^\mathrm{T}\right]^\mathrm{T} \in \mathbb{C}^B$ and $\tilde{\mathbf{y}}_c \in \mathbb{C}^S$ is the corresponding output of the adaptive analog spatial transform. The transformed receive vector is, therefore,

$$\tilde{\mathbf{y}} = \mathbf{F}\mathbf{y}, \tag{4}$$

where $\tilde{\mathbf{y}} = \left[\tilde{\mathbf{y}}_1^\mathrm{T}, \ldots, \tilde{\mathbf{y}}_C^\mathrm{T}\right]^\mathrm{T} \in \mathbb{C}^B$ and $\mathbf{F} = \mathrm{diag}(\mathbf{F}_1, \ldots, \mathbf{F}_C)$ is a $B \times B$ block diagonal matrix.

Note that if no structural constraints are imposed on the clusters of the spatial transform matrix $\mathbf{F}$, then (4) would have to be computed as a matrix-vector product consisting of B^2/C complex-valued multiplications. However, the specific structure of the per-cluster transforms $\mathbf{F}_c$, $c = 1, \ldots, C$, proposed in Section III, enables simpler analog hardware with fewer complex-valued multiplications.

C. Analog-to-Digital Conversion

As we will show in Section III, the adaptive analog spatial transform ensures that only a subset of its outputs contains strong signals. One can then utilize the AGCs (cf. Fig. 1) to individually adjust the input gains for each pair of ADCs with the goal of minimizing saturation and quantization artifacts. In order to model such ADC artifacts, we follow [9], [10] and apply a uniform midrise quantizer to each entry of the real and imaginary parts of the transformed receive vector $\tilde{\mathbf{y}}$ in (4). The scalar function $Q(\cdot)$ that models each ADC is given by

$$Q(x) = \begin{cases} \Delta\lfloor\frac{x}{\Delta}\rfloor + \frac{\Delta}{2}, & \text{if } |x| \leq \Delta 2^{q-1}, \\ \frac{\Delta}{2}(2^q - 1)\frac{x}{|x|}, & \text{if } |x| > \Delta 2^{q-1}, \end{cases} \tag{5}$$

where Δ is the quantizer's step size and q is the number of quantization bits. The distortions introduced by this model are then taken into account by the spatial equalizer (cf. Fig. 1).

To design the spatial equalizer, we employ Bussgang's decomposition [17], where the effect of the quantizer on a real-valued zero-mean random variable (RV) x is modeled as

$$Q(x) = \gamma x + d, \tag{6}$$

with the so-called Bussgang gain

$$\gamma = \frac{\mathbb{E}[Q(x)x]}{\mathbb{E}[x^2]}. \tag{7}$$

In (6), the choice of γ as in (7) ensures that the distortion d is uncorrelated with x and is zero mean with variance

$$D = \mathbb{E}[d^2] = \mathbb{E}[Q(x)^2] - \gamma^2\,\mathbb{E}[x^2]. \tag{8}$$

In order to minimize the distortion caused by saturation and quantization, one can either pick an optimal step size Δ while fixing the input variance or pick an optimal input variance while fixing the step size. We force the variance of the input x to one and select an optimal step size Δ that minimizes the mean squared error (MSE) between the quantizer's output $Q(x)$ and its input x [18]. To enforce unit variance at the inputs of all ADCs, we need to adjust the AGC settings, which we model by a diagonal gain control matrix $\boldsymbol{\Omega} = \mathrm{diag}(\omega_1, \ldots, \omega_B)$ with

$$\omega_b = \sqrt{\frac{2}{[\mathbf{C}_{\tilde{\mathbf{y}}}]_{(b,b)}}}, \quad b = 1, \ldots, B. \tag{9}$$

Here, the covariance matrix $\mathbf{C}_{\tilde{\mathbf{y}}}$ is given by

$$\mathbf{C}_{\tilde{\mathbf{y}}} = \mathbb{E}[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^\mathrm{H}] = \mathbf{F}\,\mathbb{E}[\mathbf{y}\mathbf{y}^\mathrm{H}]\,\mathbf{F}^\mathrm{H} = \mathbf{F}\mathbf{C}_{\mathbf{y}}\mathbf{F}^\mathrm{H}. \tag{10}$$

By multiplying the gain control matrix $\boldsymbol{\Omega}$ to the transformed receive vector $\tilde{\mathbf{y}}$, the entries of $\boldsymbol{\Omega}\tilde{\mathbf{y}}$ have unit variance per real dimension. The resulting AGC and ADC model is

$$\mathbf{r} = \mathcal{Q}(\tilde{\mathbf{y}}) = Q(\Re\{\boldsymbol{\Omega}\tilde{\mathbf{y}}\}) + jQ(\Im\{\boldsymbol{\Omega}\tilde{\mathbf{y}}\}). \tag{11}$$

By applying Bussgang's decomposition (6) individually to each entry of (11), we obtain the following linearized input-output relation:

$$\mathbf{r} = Q(\Re\{\boldsymbol{\Omega}\tilde{\mathbf{y}}\}) + jQ(\Im\{\boldsymbol{\Omega}\tilde{\mathbf{y}}\}) \tag{12}$$

$$= \gamma\Re\{\boldsymbol{\Omega}\tilde{\mathbf{y}}\} + \mathbf{d}_r + j(\gamma\Im\{\boldsymbol{\Omega}\tilde{\mathbf{y}}\} + \mathbf{d}_j) \tag{13}$$

$$= \gamma\boldsymbol{\Omega}(\Re\{\tilde{\mathbf{y}}\} + j\Im\{\tilde{\mathbf{y}}\}) + \mathbf{d} \ = \gamma\boldsymbol{\Omega}\tilde{\mathbf{y}} + \mathbf{d}. \tag{14}$$

Here, we define $\mathbf{d} = \mathbf{d}_r + j\mathbf{d}_j$, where $\mathbf{d}_r$ and $\mathbf{d}_j$ are the real and imaginary parts of the distortions caused by $\mathcal{Q}(\tilde{\mathbf{y}})$, respectively.

D. Spatial Equalization

The linearized input-output relation in (14) can now be used to design a linear minimum mean squared error (LMMSE)-type equalizer, which generates estimates for the transmitted symbols as $\hat{\mathbf{s}} = \mathbf{W}\mathbf{r}$. In the equalization step, we do not assume perfect knowledge of the channel matrix. Therefore, using the estimate $\hat{\mathbf{H}}$ of the channel matrix $\mathbf{H}$, we define our LMMSE-type equalization matrix as

$$\mathbf{W} = \frac{1}{\gamma}\hat{\mathbf{H}}^\mathrm{H}\mathbf{F}^\mathrm{H}\boldsymbol{\Omega}\bigg(\boldsymbol{\Omega}\mathbf{F}\hat{\mathbf{H}}\hat{\mathbf{H}}^\mathrm{H}\mathbf{F}^\mathrm{H}\boldsymbol{\Omega}$$

$$+ N_0\boldsymbol{\Omega}\mathbf{F}\mathbf{F}^\mathrm{H}\boldsymbol{\Omega} + \frac{2D}{\gamma^2}\mathbf{I}_B\bigg)^{-1}, \tag{15}$$

where we approximate the distortion covariance matrix as

$$\mathbf{C}_{\mathbf{d}} = \mathbb{E}[\mathbf{d}\mathbf{d}^\mathrm{H}] \approx 2D\,\mathbf{I}_B, \tag{16}$$

with D from (8).

---

[1]In practice, the number S of antennas per cluster should not be too large as implementing large analog spatial transforms is challenging [16].

III. ADAPTIVE ANALOG SPATIAL TRANSFORMS

We are now ready to describe the specifics of suitable adaptive analog spatial transforms $\mathbf{F}_c$, $c = 1, \ldots, C$, which we use to focus strong receive signals onto a subset of the ADCs. Specifically, we present two HR matrices designed for different purposes: HR-ISO, which isolates the energy of the strongest UE on dedicated ADCs, and HR-MAX, which isolates the signal dimension with the highest receive power on dedicated ADCs.

A. Householder Reflections

An HR is carried out by multiplying a vector by a matrix $\mathbf{Q}_{\mathbf{v}}$ that reflects the vector with respect to a subspace described by its normal vector $\mathbf{v}$ [19], [20].

Definition 1. *Let $\mathbf{v} \in \mathbb{C}^M$ be a nonzero vector. Then, the HR matrix $\mathbf{Q}_{\mathbf{v}}$ is defined as follows:*

$$\mathbf{Q}_{\mathbf{v}} = \mathbf{I}_M - 2\frac{\mathbf{v}\mathbf{v}^{\mathrm{H}}}{\|\mathbf{v}\|^2}. \tag{17}$$

HR matrices are (i) unitary so that the statistics of an i.i.d Gaussian random vector remain unaffected and (ii) Hermitian symmetric, i.e., $\mathbf{Q}_{\mathbf{v}}^{\mathrm{H}}\mathbf{Q}_{\mathbf{v}} = \mathbf{Q}_{\mathbf{v}}\mathbf{Q}_{\mathbf{v}} = \mathbf{I}_M$.

By replacing $\mathbf{F}_c$ in (3) with $\mathbf{Q}_{\mathbf{v}_c}$ as in (17), we obtain a transformed receive vector in the cth antenna cluster as

$$\tilde{\mathbf{y}}_c = \mathbf{Q}_{\mathbf{v}_c}\mathbf{y}_c = \mathbf{y}_c - 2\frac{\mathbf{v}_c\mathbf{v}_c^{\mathrm{H}}}{\|\mathbf{v}_c\|^2}\mathbf{y}_c, \tag{18}$$

where $\mathbf{v}_c \in \mathbb{C}^S$. This operation requires only one inner-product calculation $\left(\mathbf{v}_c^{\mathrm{H}}/\|\mathbf{v}_c\|\right)\mathbf{y}_c$ followed by subtracting a scaled version of $\mathbf{v}_c/\|\mathbf{v}_c\|$ from $\mathbf{y}_c$. Thus, the number of real- and complex-valued multiplications required to transform all C clusters decreases from B^2/C to only $2B + C$, which can be implemented more efficiently, e.g., using analog multiplication circuitry originally developed for beamforming [21]. We reiterate that the statistics of the noise vector $\mathbf{n}$ in (1) remain unaffected by (18) since the matrices $\mathbf{Q}_{\mathbf{v}_c}$, $c = 1, \ldots, C$, are unitary. Suitable choices for the vector $\mathbf{v}_c$ are discussed next.

B. Strongest UE Isolation (HR-ISO)

As the first choice for the HR matrix, we wish to isolate the energy of the strongest UE on the ith output of the adaptive analog spatial transform. To this end, we need the following result; the proof is given in Appendix A.

Lemma 1. *Let $\mathbf{a} \in \mathbb{C}^M$ be a nonzero vector. Then, the vector*

$$\hat{\mathbf{v}} = \mathbf{a} + \|\mathbf{a}\|\,\mathrm{sign}([\mathbf{a}]_i)\mathbf{e}_i \tag{19}$$

is a solution of the following optimization problem:

$$\hat{\mathbf{v}} \in \arg\max_{\tilde{\mathbf{v}} \in \mathbb{C}^M} |\mathbf{e}_i^{\mathrm{H}}\mathbf{Q}_{\tilde{\mathbf{v}}}\mathbf{a}|^2. \tag{20}$$

We can use Lemma 1 to isolate the energy of the strongest UE on the first pair of AGCs/ADCs of each antenna cluster by setting $i = 1$ and $\mathbf{a} = \hat{\underline{\mathbf{h}}}_c$, which corresponds to solving

$$\hat{\mathbf{v}}_c^{\mathrm{HR-ISO}} \in \arg\max_{\tilde{\mathbf{v}}_c \in \mathbb{C}^S} |\mathbf{e}_1^{\mathrm{H}}\mathbf{Q}_{\tilde{\mathbf{v}}_c}\hat{\underline{\mathbf{h}}}_c|^2, \tag{21}$$

where $\hat{\underline{\mathbf{h}}}_c = \left[[\hat{\underline{\mathbf{h}}}]_{(c-1)S+1}, \ldots, [\hat{\underline{\mathbf{h}}}]_{cS}\right]^{\mathrm{T}} \in \mathbb{C}^S$ is the estimated channel vector associated with the cth cluster of the strongest UE. According to Lemma 1, a vector that solves (21) is

$$\hat{\mathbf{v}}_c^{\mathrm{HR-ISO}} = \hat{\underline{\mathbf{h}}}_c + \|\hat{\underline{\mathbf{h}}}_c\|\,\mathrm{sign}([\hat{\underline{\mathbf{h}}}_c]_1)\mathbf{e}_1. \tag{22}$$

Therefore, the HR matrix that focuses the energy of the strongest UE on the first per-cluster AGCs and ADCs is $\mathbf{Q}_{\hat{\mathbf{v}}_c^{\mathrm{HR-ISO}}}$. Note that $\hat{\mathbf{v}}_c^{\mathrm{HR-ISO}}$ is the standard HR vector used for QR decompositions applied to $\hat{\underline{\mathbf{h}}}_c$ [20, Sec. 5.2.1].

C. Maximum Power Isolation (HR-MAX)

As the second choice for the HR matrix, we wish to isolate the signal dimension with the highest receive power per antenna cluster on the ith output of the adaptive analog spatial transform. To this end, we need the following result; the proof is given in Appendix B.

Lemma 2. *Let $\mathbf{a} \in \mathbb{C}^M$ be a zero-mean random vector with covariance matrix $\mathbf{C}_{\mathbf{a}} = \mathbb{E}\left[\mathbf{a}\mathbf{a}^{\mathrm{H}}\right]$. Furthermore, let $\boldsymbol{\ell}_1$ be the eigenvector associated with the largest eigenvalue λ_1 of the covariance matrix $\mathbf{C}_{\mathbf{a}}$. Then, the vector*

$$\hat{\mathbf{v}} = \boldsymbol{\ell}_1 + \mathrm{sign}([\boldsymbol{\ell}_1]_i)\mathbf{e}_i \tag{23}$$

is a solution of the following optimization problem:

$$\hat{\mathbf{v}} \in \arg\max_{\tilde{\mathbf{v}} \in \mathbb{C}^M} \mathbb{E}\left[|\mathbf{e}_i^{\mathrm{H}}\mathbf{Q}_{\tilde{\mathbf{v}}}\mathbf{a}|^2\right]. \tag{24}$$

We can now use Lemma 2 to isolate the signal dimension with the highest receive power of antenna cluster c on the first pair of AGCs/ADCs by setting $i = 1$ and $\mathbf{a} = \mathbf{y}_c$ with $\mathbf{C}_{\mathbf{a}} = \mathbf{C}_{\mathbf{y}_c} = \mathbb{E}\left[\mathbf{y}_c\mathbf{y}_c^{\mathrm{H}}\right]$, which corresponds to solving

$$\hat{\mathbf{v}}_c^{\mathrm{HR-MAX}} \in \arg\max_{\tilde{\mathbf{v}}_c \in \mathbb{C}^S} \mathbb{E}\left[|\mathbf{e}_1^{\mathrm{H}}\mathbf{Q}_{\tilde{\mathbf{v}}_c}\mathbf{y}_c|^2\right]. \tag{25}$$

According to Lemma 2, a vector that solves (25) is

$$\hat{\mathbf{v}}_c^{\mathrm{HR-MAX}} = \boldsymbol{\ell}_1 + \mathrm{sign}([\boldsymbol{\ell}_1]_1)\mathbf{e}_1, \tag{26}$$

where $\boldsymbol{\ell}_1$ is the eigenvector associated with the largest eigenvalue λ_1 of the covariance matrix $\mathbf{C}_{\mathbf{y}_c}$. Therefore, the HR matrix that focuses the signal dimension with the highest receive power on the first per-cluster pair of AGCs and ADCs is $\mathbf{Q}_{\hat{\mathbf{v}}_c^{\mathrm{HR-MAX}}}$.

IV. SIMULATION RESULTS

We now demonstrate the efficacy of HDR MIMO in a mmWave massive MU-MIMO scenario where one UE has significantly stronger BS-side receive power than the other UEs.

A. Simulation Details

We consider an uplink scenario in which $U = 32$ single-antenna UEs transmit data to a BS with $B = 256$ antennas arranged as a uniform linear array (ULA) with half-wavelength antenna spacing. The BS is at a height of $25\,\mathrm{m}$ and the UEs at 1.5m; both the BS and the UEs are equipped with omnidirectional antennas. The system operates at a carrier frequency of $60\,\mathrm{GHz}$ and a bandwidth of $40\,\mathrm{MHz}$. The UEs transmit 16-QAM symbols. We limit the BS-side receive power
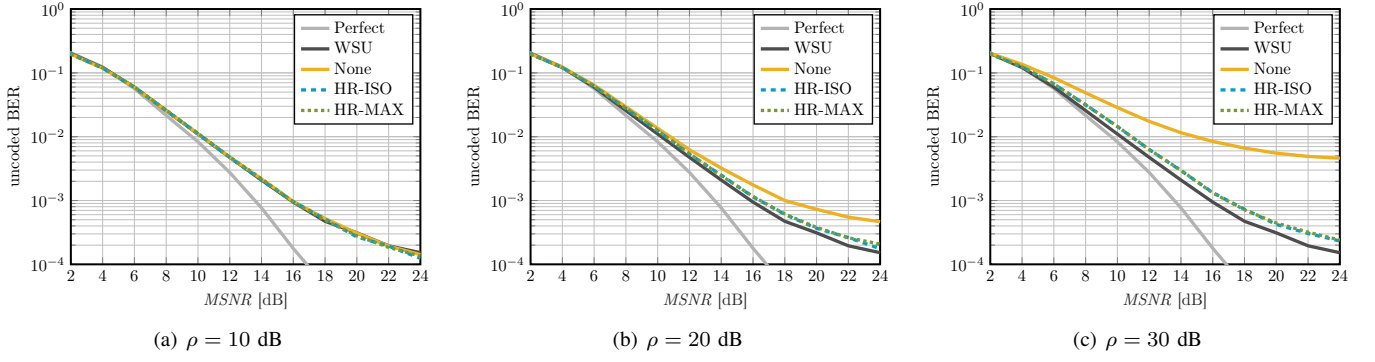
Fig. 2. Uncoded BER of HDR MIMO and baselines when varying the dynamic range ρ. In these results, $q = 3$ bit and $C = 32$ clusters.
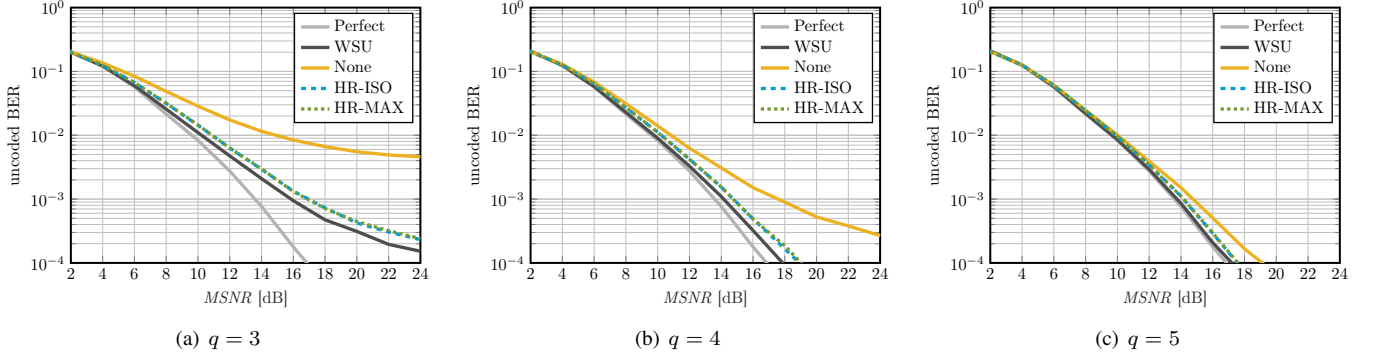

Fig. 3. Uncoded BER of HDR MIMO and baselines when varying the number of quantization bits q. In these results, $\rho = 30$ dB and $C = 32$ clusters.
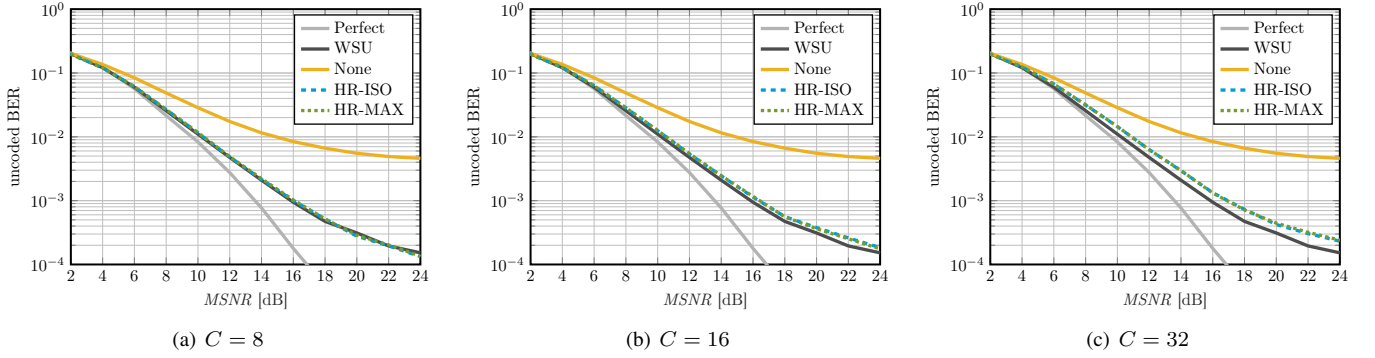

Fig. 4. Uncoded BER of HDR MIMO and baselines when varying the number of antenna clusters C. In these results, $\rho = 30$ dB and $q = 3$ bit.

between the weakest and the second strongest UE by 6 dB using the power control scheme in [22, Sec. II-B], which determines the per-UE gains d_u, $u = 1, \ldots, U$, from (2). In this scheme, strong UEs that exceed the dynamic range limit have to transmit with lower power while weak UEs transmit at their original power. We use mmWave channel vectors generated with Remcom's Wireless InSite [1] for 30 351 UE positions in an area of $150\,\text{m} \times 200\,\text{m}$. For every channel realization, we select U UE channel vectors uniformly at random. We define the median receive signal-to-noise ratio (MSNR) as follows:

$$MSNR = \frac{U \, \text{median}\{\|\mathbf{h}_u\|^2, u = 1, \ldots, U\}}{B N_0}. \tag{27}$$

For the AGC matrix in (9) and (10) and for HR-MAX in (25), we estimate the covariance matrix $\mathbf{C_y}$ as $\mathbf{C_y} \approx \frac{1}{\tau} \mathbf{Y}_T \mathbf{Y}_T^{\text{H}}$. Here, $\mathbf{Y}_T$ is obtained through pilot-based training (cf. Section II-A), where we use orthogonal pilot sequences of length $K = U$ taken from a properly scaled Hadamard matrix.

B. Baseline Algorithms

We compare the proposed HDR MIMO methods HR-ISO and HR-MAX to three baselines. The first baseline is called "Perfect" and considers the same simulation setup as for HDR MIMO, but uses a receiver with infinite-resolution ADCs. Therefore, it does not require any spatial transform. The second baseline is called "WSU" (short for "without strongest UE") and considers a different simulation setup in which all UEs are power-controlled

to a dynamic range of 6 dB and no spatial transform is used. The third baseline is called "None" and assumes the same simulation setup as for the HDR MIMO methods but without any spatial transform. Thus, we have $\mathbf{F} = \mathbf{I}_B$ for all baselines.

C. Simulation Results

We evaluate the performance of HDR MIMO and the baseline algorithms in terms of uncoded bit error rate (BER). We vary three parameters from our system: the dynamic range $\rho = \{10\,\text{dB}, 20\,\text{dB}, 30\,\text{dB}\}$ from (2), the number of ADC quantization bits $q = \{3, 4, 5\}$ from (5), and the number of antenna clusters $C = \{8, 16, 32\}$. When varying one of these parameters, we fix all of the others to $\rho = 30\,\text{dB}$, $q = 3\,\text{bit}$, and $C = 32$ clusters.

In Fig. 2, we show results varying the dynamic range ρ. We observe that "Perfect," which has infinite ADC resolution, outperforms all of the other presented methods which utilize 3-bit ADC resolution. As we increase the dynamic range ρ, the MSNR gap between "WSU" and the proposed HDR MIMO methods (HR-ISO and HR-MAX) increases. However, even with a 1000× stronger BS-side receive power between the weakest and strongest UEs, our proposed methods maintain an MSNR gap of only 1 dB at 0.1% BER between the HDR MIMO methods and "WSU". Furthermore, the proposed HDR MIMO methods demonstrate significant performance gains compared to "None," especially for scenarios in which the BS-side receive power of the strongest UE far exceeds that of the other UEs (see Fig. 2(c)). Hence, the benefits of HDR MIMO are more pronounced with increasing BS-side receive power of the strongest UE.

In Fig. 3, we show results varying the number of quantization bits q. In Fig. 3(a), we observe a similar behavior as in Fig. 2, where "Perfect" outperforms all of the other presented methods due to its infinite ADC resolution. As we increase the number of quantization bits q, the MSNR gap between "Perfect," "WSU," "None," and the proposed HDR MIMO methods decreases, almost closing the MSNR gap in Fig. 3(c). This observation indicates that the benefits of HDR MIMO are less pronounced if the ADC resolution increases. Although with 5-bit ADCs the HDR MIMO methods approach a nearly equivalent BER compared to the use of infinite-resolution ADCs, it also presents a smaller advantage when compared "None." Thus, HDR MIMO is suitable for systems with low-resolution ADCs.

In Fig. 4, we show results varying the number of antenna clusters C, and observe a similar behavior as the one from Fig. 2 and Fig. 3(a), where "Perfect" outperforms all of the other presented methods due to its infinite ADC resolution. As we increase the number of antenna clusters C, the MSNR gap between "WSU" and the HDR MIMO methods also increases. However, even when we have four times as many clusters as in Fig. 4(a), we can still maintain a relatively small MSNR gap of only 1 dB at 0.1% BER between the HDR MIMO methods and "WSU" (Fig. 4(c)). Furthermore, the HDR MIMO methods demonstrate significant performance gains compared to "None," especially in scenarios with fewer antenna clusters, as shown in Fig. 4(a). Hence, the benefit of HDR MIMO improves with fewer antenna clusters (which is equivalent to more antennas per cluster), even though it is disadvantageous for hardware implementation [16].

Finally, we compare the BER performance between the two adaptive analog spatial transforms: HR-ISO and HR-MAX. In Figs. 2, 3, and 4, we find that the BER is nearly identical. In conclusion, both HDR MIMO methods are beneficial in systems featuring high dynamic range UEs and low-resolution ADCs. However, from a complexity perspective, HR-ISO is the preferable option as it does not require the calculation of the covariance matrix $\mathbf{C}_{\mathbf{y}_c}$ and an eigenvalue decomposition.

V. Conclusions

We have proposed HDR MIMO, a novel approach that consists of an adaptive analog spatial transform which enables one to focus the energy of the strongest UE or the signal dimension with the highest receive power on a few ADCs in order to mitigate saturation and quantization artifacts caused by high dynamic range UEs in systems with low-resolution ADCs. Simulation results have shown that HDR MIMO greatly outperforms systems without spatial transforms. Furthermore, we have shown that (i) as the strongest UE BS-side receive power increases and (ii) as the quantization resolution decreases, the more pronounced the benefits of HDR MIMO become. Besides that, we have shown that, although transforming large antenna clusters is favorable, small analog spatial transforms that are more hardware friendly only slightly deteriorate the effectiveness of HDR MIMO.

Appendix A
Proof of Lemma 1

The objective function of (20) is bounded by

$$|\mathbf{e}_i^{\text{H}} \mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{a}|^2 \overset{(a)}{\leq} \|\mathbf{e}_i\|^2 \|\mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{a}\|^2 \overset{(b)}{=} \|\mathbf{a}\|^2. \tag{28}$$

Here, (a) follows from the Cauchy–Schwarz inequality and (b) from the fact that $\mathbf{Q}_{\tilde{\mathbf{v}}}$ is unitary. We now show that the HR matrix $\mathbf{Q}_{\hat{\mathbf{v}}}$ with the vector $\hat{\mathbf{v}}$ from (19) achieves the upper bound in (28) with equality. To this end, notice that the Cauchy–Schwarz inequality holds with equality if and only if $\mathbf{e}_i$ and $\mathbf{Q}_{\hat{\mathbf{v}}} \mathbf{a}$ are collinear, i.e., if $\alpha \mathbf{e}_i = \mathbf{Q}_{\hat{\mathbf{v}}} \mathbf{a}$ holds for some $\alpha \in \mathbb{C}$. From (17), it follows that

$$\alpha \mathbf{e}_i = \mathbf{a} - 2 \frac{\hat{\mathbf{v}} \hat{\mathbf{v}}^{\text{H}} \mathbf{a}}{\|\hat{\mathbf{v}}\|^2}. \tag{29}$$

From (29), after some algebraic manipulations, we arrive at

$$\hat{\mathbf{v}} = \frac{(\mathbf{a} - \alpha \mathbf{e}_i) \|\hat{\mathbf{v}}\|^2}{2 \hat{\mathbf{v}}^{\text{H}} \mathbf{a}}, \tag{30}$$

where we assume that $\hat{\mathbf{v}}^{\text{H}} \mathbf{a} \neq 0$, i.e., $\hat{\mathbf{v}}$ and $\mathbf{a}$ are not orthogonal. By inserting $\hat{\mathbf{v}}$ from (30) into (29), the following must hold:

$$\alpha \mathbf{e}_i = \mathbf{a} - 2 \frac{(\mathbf{a} - \alpha \mathbf{e}_i)(\mathbf{a} - \alpha \mathbf{e}_i)^{\text{H}} \mathbf{a}}{\|\mathbf{a} - \alpha \mathbf{e}_i\|^2}. \tag{31}$$

Since the factor $\|\hat{\mathbf{v}}\|^2 / (2 \hat{\mathbf{v}}^{\text{H}} \mathbf{a})$ from (30) cancels out, its specific choice does not matter (assuming that it is nonzero). Therefore, we assume $\|\hat{\mathbf{v}}\|^2 / (2 \hat{\mathbf{v}}^{\text{H}} \mathbf{a}) = 1$ in what follows.

In order to determine the optimal value of α, we first assume that $\mathbf{a} \neq \alpha \mathbf{e}_i$. By rearranging terms in (31), we obtain the following necessary optimality condition

$$\|\mathbf{a} - \alpha \mathbf{e}_i\|^2 = 2(\mathbf{a} - \alpha \mathbf{e}_i)^{\mathrm{H}} \mathbf{a}, \qquad (32)$$

which can be simplified to

$$|\alpha|^2 = \|\mathbf{a}\|^2 - \alpha^* [\mathbf{a}]_i + \alpha [\mathbf{a}]_i^*. \qquad (33)$$

This equation has two solutions $\hat{\alpha} = \pm \|\mathbf{a}\| \operatorname{sign}([\mathbf{a}]_i)$. By inserting $\hat{\alpha}$ into (31), both the left and right sides are equal, proving that the two solutions are indeed solutions to (31). For reasons given below, we pick the solution

$$\hat{\alpha} = -\|\mathbf{a}\| \operatorname{sign}([\mathbf{a}]_i). \qquad (34)$$

Let us now inspect the two excluded cases $\hat{\mathbf{v}}^{\mathrm{H}} \mathbf{a} = 0$ and $\mathbf{a} = \alpha \mathbf{e}_i$. If $\hat{\mathbf{v}}^{\mathrm{H}} \mathbf{a} = 0$, then the vectors $\hat{\mathbf{v}}$ and $\mathbf{a}$ are orthogonal. With (19), this implies that

$$(\mathbf{a} + \|\mathbf{a}\| \operatorname{sign}([\mathbf{a}]_i) \mathbf{e}_i)^{\mathrm{H}} \mathbf{a} = \|\mathbf{a}\|^2 + \|\mathbf{a}\| |[\mathbf{a}]_i| = 0, \qquad (35)$$

which is impossible as the vector $\mathbf{a}$ was assumed to be nonzero; this is the reason why we picked the solution in (34). If $\mathbf{a} = \alpha \mathbf{e}_i$, then according to (19), $\hat{\mathbf{v}} = 2\alpha \mathbf{e}_i$. By plugging this vector into (20), the bound in (28) is achieved with equality.

In summary, by combining (30) and (34), we see that a vector $\hat{\mathbf{v}}$ that maximizes (20) is given by (19). ∎

APPENDIX B
PROOF OF LEMMA 2

With the covariance matrix $\mathbf{C_a}$, we can rewrite the optimization problem in (24) as

$$\hat{\mathbf{v}} \in \arg \max_{\tilde{\mathbf{v}} \in \mathbb{C}^M} \mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{C_a} \mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{e}_i. \qquad (36)$$

We can now bound the objective of (36) as follows:

$$\max_{\tilde{\mathbf{v}} \in \mathbb{C}^M} \mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{C_a} \mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{e}_i \overset{(c)}{\leq} \max_{\mathbf{z}, \|\mathbf{z}\|=1} \mathbf{z}^{\mathrm{H}} \mathbf{C_a} \mathbf{z} \overset{(d)}{=} \lambda_1. \qquad (37)$$

Here, (c) holds as $\mathbf{Q}_{\tilde{\mathbf{v}}} \mathbf{e}_i$ is a specific unit-norm vector and taking the maximum over arbitrary unit-norm vectors $\mathbf{z} \in \mathbb{C}^M$ bounds the objective. The equality (d) follows from the definition of the spectral norm, which is equal to the largest eigenvalue λ_1 of the covariance matrix $\mathbf{C_a}$ of $\mathbf{a}$ [23, Ex. 5.6.6]. We now show that the HR matrix $\mathbf{Q}_{\hat{\mathbf{v}}}$ with $\hat{\mathbf{v}}$ from (23) achieves the upper bound in (37) with equality.

We start by rewriting (36) as follows:

$$\hat{\mathbf{v}} \in \arg \max_{\tilde{\mathbf{v}} \in \mathbb{C}^M} \sum_{m=1}^{M} \lambda_m |\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\tilde{\mathbf{v}}} \boldsymbol{\ell}_m|^2. \qquad (38)$$

Here, the expression $|\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\tilde{\mathbf{v}}} \boldsymbol{\ell}_1|^2$ with $\tilde{\mathbf{v}} = \hat{\mathbf{v}}$ from (23) is bounded by

$$|\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_1|^2 \overset{(e)}{\leq} \|\mathbf{e}_i\|^2 \|\mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_1\|^2 \overset{(f)}{=} 1, \qquad (39)$$

where (e) follows from the Cauchy–Schwarz inequality and (f) from the fact that $\mathbf{Q}_{\hat{\mathbf{v}}}$ is unitary. From Lemma 1, we know that

using $\hat{\mathbf{v}}$ from (23) leads to equality in (e), i.e., $|\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_1|^2 = 1$. This conclusion together with the objective from (38) leads to

$$\lambda_1 \underbrace{|\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_1|^2}_{1} + \sum_{m=2}^{M} \lambda_m \underbrace{|\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_m|^2}_{\geq 0} \geq \lambda_1, \qquad (40)$$

since the eigenvalues λ_m, $m = 2, \ldots, M$ are nonnegative. By combining (37) and (40), we conclude that the objective function in (38) is both lower and upper bounded by λ_1. Thus, with $\hat{\mathbf{v}}$ from (23), we have that

$$\sum_{m=1}^{M} \lambda_m |\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_m|^2 = \lambda_1 |\mathbf{e}_i^{\mathrm{H}} \mathbf{Q}_{\hat{\mathbf{v}}} \boldsymbol{\ell}_1|^2 = \lambda_1, \qquad (41)$$

which completes the proof. ∎

REFERENCES

[1] Remcom, Inc. (2023) Wireless InSite. [Online]. Available: https://www.remcom.com/wireless-insite-em-propagation-software

[2] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[3] S. Dutta, C. N. Barati, D. Ramirez, A. Dhananjay, J. F. Buckwalter, and S. Rangan, "A case for digital beamforming at mmWave," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 756–770, Feb. 2020.

[4] S. H. Mirfarshbafan, A. Gallyas-Sanhueza, R. Ghods, and C. Studer, "Beamspace channel estimation for massive MIMO mmWave systems: Algorithm and VLSI design," *IEEE Trans. Circuits Syst. I*, vol. 67, no. 12, pp. 5482–5495, Sep. 2020.

[5] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[6] B. C. Geiger, C. Feldbauer, and G. Kubin, "Information loss in static nonlinearities," in *IEEE Int. Symp. Wireless Commun. Syst.*, Nov. 2011, pp. 799–803.

[7] H. Subbaram and K. Abend, "Interference suppression via orthogonal projections: a performance analysis," *IEEE Trans. Antennas Propag.*, vol. 41, no. 9, pp. 1187–1194, Sep. 1993.

[8] Q. Yan, H. Zeng, T. Jiang, M. Li, W. Lou, and Y. T. Hou, "MIMO-based jamming resilient communication in wireless networks," in *IEEE Conf. Comput. Commun. (INFOCOM)*, May 2014, pp. 2697–2706.

[9] G. Marti, O. Castañeda, S. Jacobsson, G. Durisi, T. Goldstein, and C. Studer, "Hybrid jammer mitigation for all-digital mmWave massive MU-MIMO," in *Asilomar Conf. Signals, Syst., Comput.*, Nov. 2021, pp. 93–99.

[10] G. Marti, O. Castañeda, and C. Studer, "Jammer mitigation via beam-slicing for low-resolution mmWave massive MU-MIMO," *IEEE Open J. Circuits Syst.*, vol. 2, pp. 820–832, Dec. 2021.

[11] M. A. Alaei, S. Golabighezelahmad, P.-T. de Boer, F. E. van Vliet, E. A. M. Klumperink, and A. B. J. Kokkeler, "Interference mitigation by adaptive analog spatial filtering for MIMO receivers," *IEEE Trans. Microw. Theory Tech.*, vol. 69, no. 9, pp. 4169–4179, May 2021.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, 2005.

[13] M. Chiang, P. Hande, T. Lan, and C. W. Tan, *Power Control in Wireless Cellular Networks*. Now Foundations and Trends, 2008.

[14] V. Saxena, G. Fodor, and E. Karipidis, "Mitigating pilot contamination by pilot reuse and power control schemes for massive MIMO systems," in *IEEE Veh. Technol. Conf. Spring (VTC-Spring)*, May 2015, pp. 1–6.

[15] E. Björnson, J. Hoydis, and L. Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*. Now Foundations and Trends, 2017.

[16] Z. M. Enciso, S. Hadi Mirfarshbafan, O. Castañeda, C. J. Schaefer, C. Studer, and S. Joshi, "Analog vs. digital spatial transforms: A throughput, power, and area comparison," in *IEEE Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2020, pp. 125–128.

[17] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Elec., Cambridge, MA, USA, Tech. Rep. 216, Mar. 1952.

[18] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[19] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*. SIAM, 1997.

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.

[21] E. Naviasky, L. Iotti, G. LaCaille, B. Nikolić, E. Alon, and A. M. Niknejad, "A 71-to-86-GHz 16-element by 16-beam multi-user beamforming integrated receiver sub-array for massive MIMO," *IEEE J. Solid-State Circuits*, vol. 56, no. 12, pp. 3811–3826, Dec. 2021.

[22] H. Song, T. Goldstein, X. You, C. Zhang, O. Tirkkonen, and C. Studer, "Joint channel estimation and data detection in cell-free massive mu-mimo systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4068–4084, Jun. 2022.

[23] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2013.