User Selection and Multiuser Widely Linear Precoding for One-Dimensional Signalling

Majid Bavand, *Member, IEEE*, and Steven D. Blostein, *Senior Member, IEEE*

***Abstract*—Emergence of ultradense networks in 5G communications, Internet of things, and eHealth devices prompts us to develop new communication techniques that can support a large number of low data rate devices. In particular, it has been shown that when the data are real-valued and the observation is complex-valued, widely linear (WL) estimation can be employed in lieu of linear estimation to improve the performance. With these motivations, we study user selection and transmit precoding in multiuser communication systems assuming transmitted signals are one-dimensionally modulated. A closed-form solution for widely linear maximum signal-to-leakage-and-noise ratio precoding is obtained. We also investigate the design of WL maximum ratio transmission, WL zero-forcing, and WL minimum mean square error precoding techniques. Furthermore, to enable an increased number of communication devices, a user selection algorithm compatible with widely linear processing of one-dimensionally modulated signals is proposed. The proposed user selection algorithm can potentially double the number of simultaneously selected users compared to that of conventional user selection methods.**

***Index Terms*—Broadcast channels, co-channel interference, multiuser communications, scheduling, semi orthogonal user selection, transmit beamforming, widely linear processing.**

I. Introduction

ADVANCES in wireless communications, in conjunction with advances in electronics, are paving the way for emergence of technologies enabling the Internet of things (IoT), pervasive ubiquitous eHealth, dense user deployments, smart cities, vehicular networks, and others, either as proprietary systems or as part of 5th generation (5G) mobile networks [1]. Communicating wirelessly to a large numbers of low-power devices requires power efficiency to be a paramount consideration. One-dimensionally (1D) modulated signals such as binary phase shift keying (BPSK) are therefore of interest to reliably support these emerging systems with massive numbers of low-data-rate devices [1]. It is similarly assumed that the base station or access point could easily be equipped with multiple antennas while the large numbers of users are limited to a single antenna due to physical constraints on equipment size, cost, power supply, and computational capabilities [2]. Consequently, to achieve the required low latencies and high throughput for 5G networks a downlink transmitter is tasked with transmitting different data streams to multiple devices simultaneously to achieve spatial multiplexing gain. Application scenarios that are particularly relevant include *dense deployments* such as HD Wi-Fi stadiums [3], airports, shopping malls or large crowded public spaces where multicasting information to large numbers of users is critical, whether it be related to sporting events, schedule updates, alerts, etc. Other scenarios include broadcasting information to large numbers of smart devices for software updating, control, infrastructure monitoring or security-related purposes. In addition to power efficient BPSK reception, the proposed combination of one-dimensional (M-PAM) modulations and WL processing may offer improved spectral efficiency potentially competitive with 2-D modulations (M-QAM) by its ability to support a greater density of simultaneous users.

Widely linear (WL) processing of complex-valued signals, originally introduced in [4] and later resurrected in [5] in the context of minimum mean square error estimation, refers to the superposition of linear filtering of the observation and linear filtering of its complex conjugate, or equivalently, superposition of linear filtering of real and imaginary parts [6]. The latter representation is known as the composite real representation. When the distribution of the estimand (signal of interest) is improper, i.e., not circularly symmetric, widely linear estimation will improve mean square error (MSE) estimation, whether or not the observation is improper [5]. Since its resurrection by Picinbono and Chevalier [5], WL processing has been applied to communication systems when improper signal constellations or improper noise is encountered in several contexts including point-to-point MIMO channels [7], MIMO broadcast with QoS constraints [6], interference alignment [8], wireless energy harvesting [9], and cooperative relaying [10].

When the data is one-dimensionally modulated, implying an improper distribution, widely linear estimation has been applied to receive beamforming in the context of multiple-antenna communications [11]–[14]. In contrast to linear receive beamforming, i.e., spatial filtering of the received signal, widely linear receive beamforming is the superposition of beamforming of the received signal and beamforming of the complex conjugate of the received signal. While WL processing has been applied to noncircularly symmetric complex distributions in general [14] with a focus on adaptive and robust beamforming, less attention has been paid to precoder design for one-dimensionally

Manuscript received November 29, 2017; revised April 6, 2018 and August 6, 2018; accepted September 12, 2018. Date of publication October 1, 2018; date of current version December 14, 2018. This work was supported by Natural Sciences and Engineering Research Council of Canada Discovery Grant 05061-2014. Part of this work is accepted to be presented at 52nd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, October 2018. The review of this paper was coordinated by Dr. X. Wang. *(Corresponding author: Majid Bavand.)*

The authors are with the Department of Electrical and Computer Engineering, Queen's University, Kingston ON K7L 3N6, Canada (e-mail: m.bavand@queensu.ca; steven.blostein@queensu.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2018.2872881

modulated signals. From one perspective, widely linear precoding can be viewed as the superposition of linear precoding of a modulated signal and linear precoding of its complex conjugate [15]. From a different perspective, widely linear precoding can be viewed as linear precoding of the modulated signal in conjunction with widely linear estimation of the received signal [16]. This paper adopts the latter perspective since the modulated signal is real-valued (1-D).

The task of transmit precoding is to reduce the effect of co-channel interference which arises in wireless broadcast channels due to spatial multiplexing. Four basic approaches to linear transmit precoding that are well researched in the last decades are: (i) transmit matched filtering or maximum ratio transmission (MRT) precoding which maximizes the desired signal portion of the received signal at each receiver [17], (ii) transmit zero-forcing (channel inversion) precoding that nulls the interference at each receiver [17], [18], (iii) transmit minimum mean square error or regularized channel inversion precoding that minimizes sum of mean square errors (MSE) of users [17]–[19], and (iv) maximum signal to leakage and noise ratio (MSLNR) precoding [20]. In this paper, we develop widely linear counterparts of the above linear precoding techniques for the case of one-dimensional signalling. The motivation is to achieve high user capacity and overall throughput, which in turn addresses the dense deployment application scenarios described above.

More recent precoding techniques have become more context specific. For example, the symbol-by-symbol constructive interference based precoding matrix is obtained for each symbol duration rather than for the channel coherence time [21], [22]. In other words, the symbol-level precoding matrix not only depends on channel information but also on the users' transmitted symbols. This approach results in a higher complexity compared to precoding techniques that do not consider individual symbols in each transmission. Other precoding approaches exploit constant envelope modulations [23]. Pilot contamination precoding, proposed for massive MIMO and multi-cell systems address non-orthogonality of pilots using statistical CSI information [24] as well as rate balancing [25]. Accounting for imperfect channel state estimation, while worthy for future investigation, is outside the scope of this contribution.

Selecting the right sets of simultaneous transmissions can opportunistically leverage transmit precoding to achieve further throughput gains to improve spectral efficiency [26]. Therefore, the choice of best user subset, which depends on precoding method, is critical. When the number of available users is large and data is available for transmission to all users, the transmitter can select a subset of users with sufficiently *good* channel conditions for simultaneous transmission. Proper user selection also reduces co-channel interference. In principle, the optimal user subset that maximizes throughput can be found by brute-force search over all possible user subsets, although with prohibitive computational complexity when the number of available users is large.

Emerging applications such as IoT and eHealth traffic require support of a large number of low-data-rate users. Most existing multiuser linear precoding methods for multi input single output (MISO) systems can only support at most as many users as the number of transmit antennas, M [17], [18], [27]. Low-complexity user selection techniques that account for multiuser interference communications also face this restriction to at most M users [2], [26], [28], [29]. One such algorithm is semi-orthogonal user selection (SUS) [26] that tries to incrementally select users with large channel gain that are also nearly orthogonal to the channels of the other selected users. Ideally, SUS channels are mutually orthogonal at any given time and have the largest gains among all available channels. In SUS, the user with the strongest channel among available users is first selected. Channels less orthogonal to the one selected are then removed from consideration. This process is repeated until either there are no more available channels or until the number of selected channels reaches the number of transmit antennas, M.

More recent approaches to user selection are similarly restricted to M simultaneous users. User selection for an interference alignment scenario is proposed in [30] along with sum rate simulation results. In [31], the product of effective channel eigenvalues is proposed as a user selection metric based on principal angles between subspaces. In [32] and [33] user selection is proposed for linear and nonlinear MMSE precoding with computation per iteration proportional to the total number of available users. In [32], knowledge of interference is exploited in a greedy suboptimal approach based on dirty paper coding and sum rate results are presented. In [33], knowledge of channel gains and angle of departure are employed for hybrid beamforming mm-wave systems. In [34], attainability of subspace-based user selection with performance fairness aided by round robin scheduling is investigated based on sum rate maximization.

In contrast to previous approaches, geometric user selection (GUS) algorithm in [35] is capable of overloading the system with more simultaneous users than transmit antennas. The GUS algorithm, however, is derived with the assumption of minimum probability of error (MPE) precoding and required a series of approximations resulting in suboptimality. As a result, GUS should not be expected to perform well in tandem with general linear and widely linear precoding techniques. Slow saturation rate of the number of selected users as the total number of available users increases is another shortcoming of GUS. Moreover, increase in the number of selected users as the number of antennas increases is very slow in GUS. Therefore, as shown in [35], GUS may not always select more users compared to other existing user selection methods. In this paper, inspired by the merits of the semi-orthogonal user selection (SUS) algorithm [26], a semi-orthogonal user selection method for one-dimensional modulation (SUSOM) is developed. It is shown that SUSOM is potentially able to double the number of selected users, a claim that cannot be made for GUS. In other words, a transmitter with M transmit antennas is shown to be capable of supporting at most $2M$ simultaneous users.

The paper therefore contributes the following to densely deployed multi-user, multi-antenna downlink systems with one-dimensional signalling: (i) a generic user selection method called semi-orthogonal user selection for one-dimensional modulations that is capable of overloading the system with more users than the number of transmit antennas, (ii) widely linear precoding counterparts of conventional linear precoding techniques with capability of better supporting overloaded systems with more users than the number of transmit antennas.

The rest of the paper is organized as follows. Section II introduces our system model. Section III proposes semi-orthogonal user selection for 1-D modulation with overloading capability of twice as large as the number of transmit antennas. Section IV redesigns MRT, ZF, MMSE, and MSLNR precoding techniques by assuming WL estimation at the receiver rather than linear estimation for 1-D signalling. Closed-form solutions for the WL-MRT and WL-ZF precoders are obtained by using complex-domain analysis and closed-form solutions for the WL-MMSE and WL-MSLNR precoders are obtained by analysis of the composite real representation. Numerical results are presented in Section V. Finally, conclusions are drawn in Section VI.

The following mathematical notations are used throughout the paper. Boldface upper case and lower case letters denote matrices and vectors, respectively. The superscripts *, $(\cdot)^T$, and $(\cdot)^H$ denote the conjugate, transpose, and conjugate transpose, respectively. The eigenvector corresponding to the maximum eigenvalue is denoted by $\mathbf{v}_{\max}$. $\|\cdot\|_2$ denotes the ℓ_2-norm. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ represent real and imaginary parts of complex parameters, respectively. $\mathrm{E}[\cdot]$ denotes the expected value and $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix.

II. System Model

We consider a multiuser multiple-input single-output wireless broadcast channel with an M-antenna transmitter and K single-antenna users. The transmitter is assumed to simultaneously send independent pulse amplitude modulated (PAM) signals to all users using the same carrier frequency and bandwidth. In low-pass vector space representation, the one-dimensionally modulated signal of a user can be described by a real-valued scalar which is the projection of the low-pass representation of the signal over the basis function defined as $f(t) = g(t)/\sqrt{E_g}$, where $g(t)$ is the low-pass real-valued pulse shaping signal with power E_g. Therefore, the PAM signal of user k can be represented by

$$s_k(l_k) \in \left\{(2l_k - 1 - L_k)d\sqrt{E_g} | 1 \le l_k \le L_k\right\},\ 1 \le k \le K, \tag{1}$$

where d is a scale factor used to achieve a desired transmitted signal power or in other words, half the distance between adjacent signal amplitudes in signal constellation [36]. The modulation order of user k is denoted by L_k, i.e., the number of constellation points in the pulse amplitude modulated signal of user k is L_k. This also implies that different users may not necessarily employ the same modulation order. The distance between adjacent signal constellation points is $2d\sqrt{E_g}$. Given l_k, the power of the signal is $s_k^2(l_k)$. Consequently, the average power of the modulated signal of user k is $\sigma_{s_k}^2 = \frac{L_k^2-1}{3}d^2E_g$. Using an $M \times 1$ precoding vector $\mathbf{u}_k$ to encode the symbol transmitted to user k, $1 \le k \le K$, the transmitted signal from the M-antenna array is then given by

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{u}_k s_k = \mathbf{U}\mathbf{s}, \tag{2}$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_K]$, $\mathbf{s} = [s_1, \ldots, s_K]^T$, and $s_k = s_k(l_k)$. Therefore, the transmit power is expressed by

$$\mathrm{E}\left[\|\mathbf{x}\|_2^2\right] = \mathrm{Tr}\left(\mathbf{U}\mathbf{R}_s\mathbf{U}^H\right), \tag{3}$$

where it is assumed that the input signals are mutually independent with covariance matrix $\mathbf{R}_s = \mathrm{E}[\mathbf{s}\mathbf{s}^T] = \mathrm{diag}(\sigma_{s_1}^2, \ldots, \sigma_{s_K}^2)$.

Assuming a fading channel with additive white Gaussian noise (AWGN), the received signal r_k at user k is given by

$$r_k = \mathbf{h}_k\mathbf{x} + z_k, \quad 1 \le k \le K, \tag{4}$$

where the additive noise z_k is a circularly symmetric complex Gaussian (CSCG) random variable with zero mean and variance $\sigma_{z_k}^2$, and the $1 \times M$ vector $\mathbf{h}_k$ is the channel between the M antennas of the transmitter and the single antenna of user k. It should be noted that we use the same notation for vector representation as [19], [26], i.e., both row and column vectors are represented using lower-case boldface. The entries of $\mathbf{h}_k$ follow an independent identically distributed (i.i.d.) CSCG distribution with zero mean and unit variance. This channel model is valid for narrowband (frequency non-selective) systems if the transmit and receive antennas are in non line-of-sight rich-scattering environments with sufficient antenna spacing [29], [37]. Equivalently, (4) can be represented in vector form by

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{5}$$

where $\mathbf{r} = [r_1, \ldots, r_K]^T$, $\mathbf{H} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T$, and the noise $\mathbf{z} = [z_1, \ldots, z_K]^T$ has zero mean and covariance matrix $\mathbf{R}_z = \mathrm{diag}(\sigma_{z_1}^2, \ldots, \sigma_{z_K}^2)$.

The received signal at each user is passed through a filter which simply scales it by a weight. Therefore, the processed signal at the receiver of user k is represented as a function of the transmit precoding matrix $\mathbf{U}$ and the receive filtering coefficient w_k by

$$\begin{aligned} y_k = w_k r_k &= w_k\mathbf{h}_k\mathbf{U}\mathbf{s} + w_k z_k = \sum_{j=1}^{K} w_k\mathbf{h}_k\mathbf{u}_j s_j + z_k' \\ &= w_k\mathbf{h}_k\mathbf{u}_k s_k + w_k\mathbf{h}_k\mathbf{U}_{\bar{k}}\mathbf{s}_{\bar{k}} + z_k', \quad 1 \le k \le K, \end{aligned} \tag{6}$$

where z_k' is also a CSCG noise term[1] with variance $\sigma_{z_k'}^2 = \sigma_{z_k}^2 w_k w_k^*$, $\mathbf{s}_{\bar{k}} = [s_1, \ldots, s_{k-1}, s_{k+1}, \ldots, s_K]^T$, and $\mathbf{U}_{\bar{k}} = [\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_K]$. Equivalently, the processed signals at the receivers can be represented in vector form by

$$\mathbf{y} = \mathbf{W}\mathbf{H}\mathbf{U}\mathbf{s} + \mathbf{z}', \tag{7}$$

where $\mathbf{W} = \mathrm{diag}(w_1, \ldots, w_K)$ and $\mathbf{y} = [y_1, \ldots, y_K]^T$, and $\mathbf{z}' = \mathbf{W}\mathbf{z}$ and has zero mean and covariance matrix $\mathbf{R}_{z'} = \mathbf{W}\mathbf{R}_z\mathbf{W}^H$.

Since the focus of this paper is on precoder design, it is assumed that $\mathbf{W}$ is known at the transmitter. For example, when the structure of the receivers are required to be simple with no receive filtering, $\mathbf{W}$ can be assumed to be the identity matrix. Moreover, it is assumed that perfect channel state information between transmitter and all users is available at the transmitter in

[1] Affine transformation preserves properness (circular symmetry) of a random variable [38].

TABLE I
SEMI-ORTHOGONAL USER SELECTION FOR ONE-DIMENSIONAL MODULATION

Initialization:
$i = 1$.
$\mathcal{A} = \{1, \cdots, K_T\}$.
$\mathcal{S} = \emptyset$.
Main Body of Algorithm:
while $i \leq 2M$ and $\mathcal{A} \neq \emptyset$ **do**
1) $\pi_i = \underset{k \in \mathcal{A}}{\operatorname{argmax}} \frac{\|\tilde{\mathbf{h}}_k\|_2}{\sigma_{z_k}}$ where $\tilde{\mathbf{h}}_k$ is given by (9).
2) $\mathcal{S} \leftarrow \mathcal{S} \cup \{\pi_i\}$.
3) $\mathbf{e}_{\pi_i} = \tilde{\mathbf{h}}_{\pi_i}$
4) $\mathcal{A} \leftarrow \mathcal{A} \setminus \{\pi_i\}$.
5) $\mathcal{A} \leftarrow \mathcal{A} \setminus \{\forall j \in \mathcal{A} | \mathrm{dist}(\mathbf{h}_\mathrm{j}, \mathbf{e}_{\pi_\mathrm{i}}) > \alpha\}$
6) $i \leftarrow i + 1$.
end while

order to focus on the precoding methods rather than on the effect of channel estimation. This information could be obtained, for example, by using uplink feedback of pilot-based estimation at the receivers or by assuming time division duplex (TDD) systems.

III. SEMI-ORTHOGONAL USER SELECTION FOR ONE-DIMENSIONAL MODULATION

User selection algorithms are coupled with a chosen precoding technique as well as a chosen modulation. In this section, we design a user selection algorithm for one dimensionally modulated signals. To the best of our knowledge, the geometric user selection (GUS) algorithm proposed in [35] is the only existing user selection algorithm (based on the notion of interference avoidance) that is designed for one-dimensionally modulated signals. Although the computational complexity of the GUS algorithm is very low, the number of selected users, K, may not be as large as possible, particularly when the number of available users, K_T, is sufficiently large. This prompts us to devise a user selection algorithm with better performance.

Obviously, when modulated signals are complex-valued, channels $\mathbf{h}_k$ and $\mathbf{h}_j$, $j \neq k$, are considered to be orthogonal if $\mathbf{h}_k \mathbf{h}_j^H = 0$, as is assumed in SUS [26]. However, when the transmitted signals are one-dimensionally modulated on a real basis function, the notion of orthogonality may be relaxed to [36]:

$$\Re\left\{\mathbf{h}_k \mathbf{h}_j^H\right\} = 0. \tag{8}$$

In [39], this type of orthogonality is introduced as semi-orthogonality. Therefore, SUS can be refashioned as in Table I to incorporate the above notion of orthogonality for one-dimensionally modulated signals.

We term the proposed algorithm in Table I as semi-orthogonal user selection for one-dimensional modulation (SUSOM). In SUSOM, at first, the set of available channels is initialized by all available channels and the set of selected channels is set to be empty. In Step 1, π_i is the index of the channel of user k with the strongest effective channel to noise ratio defined as $\frac{\|\tilde{\mathbf{h}}_k\|_2}{\sigma_{z_k}}$, where i is the iteration counter of the SUSOM algorithm. The effective channel of user k, $\tilde{\mathbf{h}}_k$, is defined as the component of $\mathbf{h}_k$ semi-orthogonal to the subspace spanned by the selected channels:

$$\tilde{\mathbf{h}}_k \triangleq \mathbf{h}_k - \sum_{j=1}^{i-1} \frac{\Re\left\{\mathbf{h}_k \mathbf{e}_{\pi_j}^H\right\}}{\|\mathbf{e}_{\pi_j}\|_2^2} \mathbf{e}_{\pi_j}. \tag{9}$$

It should be remarked that to find the component of a channel that is orthogonal to the linear subspace spanned by the previously selected channels, the projection of that channel on any element of the linear subspace includes a real operator [39]. In Step 2, the channel with index π_i is added to the set of selected channels. In Step 3, the orthogonal component of the selected channel, $\tilde{\mathbf{h}}_{\pi_i}$, is saved in $\mathbf{e}_{\pi_i}$, to be used later in Step 5 and in the following iterations of SUSOM algorithm. In Step 4, the selected channel is removed from the set of available channels. In Step 5, all the available channels in $\mathcal{A}$ that have distance "dist" greater than a predetermined threshold $\alpha, 0 \leq \alpha < 1$, are removed from the set of available channels $\mathcal{A}$. The distance "dist" which is defined as

$$\mathrm{dist}(\mathbf{h}_\mathrm{j}, \mathbf{e}_{\pi_\mathrm{i}}) \triangleq \frac{\left|\Re\left\{\mathbf{h}_\mathrm{j} \mathbf{e}_{\pi_\mathrm{i}}^\mathrm{H}\right\}\right|}{\|\mathbf{h}_\mathrm{j}\|_2 \|\mathbf{e}_{\pi_\mathrm{i}}\|_2}, \tag{10}$$

measures the orthogonality of $\mathbf{h}_j$ and $\mathbf{e}_{\pi_i}$. In other words, Step 5 results in semi-orthogonality of the selected channels [26]. It should be remarked that if the real operator in the distance (10) was absent, (10) would indicate the cosine of the principal angles between $\mathbf{h}_j$ and $\mathbf{e}_{\pi_i}$ [40]. Ideally, $\mathrm{dist}(\mathbf{h}_\mathrm{j}, \mathbf{e}_{\pi_\mathrm{i}})$ should be zero, i.e., $\mathbf{h}_j$ and $\mathbf{e}_{\pi_i}$ should be orthogonal.

In SUS, the number of selected users can be at most as large as the number of transmit antennas M. On the other hand, in SUSOM, since the notion of orthogonality is relaxed to only consider the real part, the number of selected users may be greater than M. Therefore, we have the following claim:

Claim 1: If the orthogonality is defined as (8), the maximum number of channels that are mutually orthogonal is $2M$.

Proof: See Appendix. ■

Remark 1: It should be remarked that the complexity order of both SUSOM and SUS is $K_T M^3$ [26], [29]. The SUSOM algorithm is similar to SUS algorithm except for definition of orthogonality which results in a different metric and different number of selected users. In other words, although two algorithms may not have exactly the same flop counts, their orders of complexity are the same [26]

IV. WIDELY LINEAR PROCESSING

Widely linear (WL) processing was considered by Picinbono and Chevalier in the context of mean square error estimation of complex-valued data [5]. In general, when data s and observation r are both improper and complex, widely linear estimation of s is given by $\hat{s} = wr + vr^*$, i.e., by superposition of linear estimates of observation r and its complex conjugate r^*. In case of real-valued data and complex-valued observation, it is known that $v = w^*$ and therefore $\hat{s} = 2\Re\{wr\}$, i.e., the estimation is given by the real part of the output of a linear estimator [5].

It is expected that calculating w by optimizing a metric based on $\Re\{wr\}$ provides additional degrees of freedom compared to optimizing a metric based on wr. In other words, with fewer restrictions placed on the imaginary part of the output, $\Im\{wr\}$, WL processing is expected to be capable of providing additional

degrees of freedom (DoF) for increased throughput by enabling an increase in the number of users or for increased reliability.

Revisiting the system model introduced in Section II, the widely linear symbol decision can be made as [36], [41]

$$\hat{s}_k = \begin{cases} s_k(1) & y_k^R \le \Re\{w_k \mathbf{h}_k \mathbf{u}_k s_k(1) + w_k \mathbf{h}_k \mathbf{u}_k d\sqrt{E_g}\} \\ & \Re\{w_k \mathbf{h}_k \mathbf{u}_k s_k(l_k) - w_k \mathbf{h}_k \mathbf{u}_k d\sqrt{E_g}\} \\ & < y_k^R \le \\ s_k(l_k) & \Re\{w_k \mathbf{h}_k \mathbf{u}_k s_k(l_k) + w_k \mathbf{h}_k \mathbf{u}_k d\sqrt{E_g}\}; \\ & 2 \le l_k \le L_k - 1 \\ s_k(L_k) & y_k^R > \Re\{w_k \mathbf{h}_k \mathbf{u}_k s_k(L_k) - w_k \mathbf{h}_k \mathbf{u}_k d\sqrt{E_g}\} \end{cases}, \tag{11}$$

where the superscript ^{R} denotes the real part, i.e., $x^R = \Re\{x\}$. Since y_k^R is considered in calculating the receive filter or the transmit precoder rather than y_k, the processing is considered as being widely linear rather than linear.

A. WL Maximum Ratio Transmission Precoding

Maximum ratio transmission (MRT) or matched filtering is the transmit counterpart of maximum ratio combining at the receiver [17]. MRT intends to maximize the received signal to noise ratio by matching the transmit precoding vector of each user to its channel. Since MRT does not consider co-channel interference, it may only perform close to optimally in noise-limited channels and in single-user communications. Considering widely linear processing, the SNR at the kth user, $1 \le k \le K$, can be defined as the ratio of the power of the real part of the desired signal at receiver k to the power of the real part of the post-processing noise, i.e.,

$$\mathrm{SNR}_k = \frac{\mathrm{E}\left[|\Re\{w_k \mathbf{h}_k \mathbf{u}_k s_k\}|^2\right]}{\mathrm{E}\left[|\Re\{z'_k\}|^2\right]} = \frac{2\sigma_{s_k}^2 \left(\Re\{w_k \mathbf{h}_k \mathbf{u}_k\}\right)^2}{\sigma_{z'_k}^2}. \tag{12}$$

Then the MRT precoding problem can be formulated by maximizing the SNR at the receiver subject to a constraint on the transmit power as

$$\max_{\mathbf{u}_k} \ \mathrm{SNR}_k \tag{13a}$$

$$\text{subject to } \sigma_{s_k}^2 \|\mathbf{u}_k\|_2^2 \le \tau_k, \tag{13b}$$

for $1 \le k \le K$. In (13), τ_k, the power constraint on the transmitted signal to user k, should satisfy $\sum_{k=1}^K \tau_k = \tau$, where τ is the total transmit power constraint. To calculate the values of τ_ks, a power allocation strategy such as equal power allocation or sum rate maximizing water-filling power allocation can be employed [26]. Using the Cauchy-Schwarz inequality we have

$$|\Re\{\mathbf{h}'_k \mathbf{u}_k\}|^2 \le |\mathbf{h}'_k \mathbf{u}_k|^2 \le |\mathbf{h}'_k|^2 |\mathbf{u}_k|^2, \tag{14}$$

where $\mathbf{h}'_k = w_k \mathbf{h}_k$. The equality holds if and only if $\mathbf{u}_k$ is in the direction of $\mathbf{h}'^H_k$. Therefore, we have $\mathbf{u}_k = \alpha \mathbf{h}'^H_k$, where

$$\alpha = \frac{\sqrt{\tau_k}}{\sigma_{s_k} |\mathbf{h}'_k|}, \tag{15}$$

using the constraint (13b). Therefore, the solution to the WL MRT problem is

$$\mathbf{u}_{k\,\mathrm{MRT}} = \frac{\sqrt{\tau_k}}{\sigma_{s_k}} \frac{\mathbf{h}'^H_k}{|\mathbf{h}'_k|}, \tag{16}$$

or in matrix form

$$\mathbf{U}_{\mathrm{MRT}} = \mathbf{H}'^H \mathbf{\Lambda}, \tag{17}$$

where $\mathbf{H}' = \mathbf{W}\mathbf{H} = [\mathbf{h}'^T_1, \ldots, \mathbf{h}'^T_K]^T$, and $\mathbf{\Lambda} = \mathrm{diag}(\frac{\sqrt{\tau_1}}{\sigma_{s_1}|\mathbf{h}'_1|}, \ldots, \frac{\sqrt{\tau_K}}{\sigma_{s_K}|\mathbf{h}'_K|})$. Interestingly, (17) is identical to MRT with linear processing in the broadcast channel [42]. In other words, widely linear processing has no advantage compared to linear processing when the transmitter uses MRT to transmit one-dimensionally modulated signals.

B. WL Zero-Forcing Precoding

Next, we consider zero-forcing (ZF) precoding also known as channel inversion [17]. In zero-forcing, it is assumed that the received signals are interference free, i.e., the received signal at receiver k is free of interference caused by the signal transmitted to user $j \in \{1, \ldots, K\}/\{k\}$. In other words, zero interference imposes the following constraint on the precoding matrix:

$$\mathbf{W}\mathbf{H}\mathbf{U} = \mathbf{\Lambda}, \tag{18}$$

where non-negative real-valued diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_K})$. Imposing this constraint results in $\sigma_{s_k}^2 \lambda_k$ as the average power of the received signal at receiver k. Since ZF precoding only considers the effect of interference but not noise, it may only perform close to optimally in interference-limited channels. From (18) it can be seen that ZF not only forces interference to be zero, but also it makes the power received by each user to be fixed (and not necessarily equal).

When the transmitter sends one-dimensionally modulated signals to the users and estimation of the received signal is performed only over the real part of the received signal (11), widely linear processing can be employed which results in relaxing (18) to

$$\Re\{\mathbf{W}\mathbf{H}\mathbf{U}\} = \mathbf{\Lambda}. \tag{19}$$

Thus, as an extension to linear ZF precoding [17], widely linear zero-forcing precoding is formulated by minimizing the total transmit power subject to the interference-free constraint of (19) as

$$\min_{\mathbf{U}} \ \mathrm{Tr}(\mathbf{U}\mathbf{R}_s\mathbf{U}^H)$$

$$\text{subject to } \Re\{\mathbf{W}\mathbf{H}\mathbf{U}\} = \mathbf{\Lambda}. \tag{20}$$

To solve this problem we first rewrite it in the following form:

$$\min_{\mathbf{u}_1, \ldots, \mathbf{u}_K} \ \sum_{k=1}^K \sigma_{s_k}^2 \mathbf{u}_k^H \mathbf{u}_k$$

$$\text{subject to } \Re\{\mathbf{H}'\mathbf{u}_k\} - \sqrt{\lambda_k}\mathbf{e}_k = \mathbf{0}, \quad 1 \le k \le K, \tag{21}$$

where $\mathbf{e}_k$ is the kth standard basis vector in K-dimensional Euclidean space and $\mathbf{H}'$ was introduced in Section IV-A.

Accordingly, the Lagrangian is given by

$$L(\mathbf{u}_1,\ldots,\mathbf{u}_K,\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_K) = \sum_{k=1}^{K}\sigma_{s_k}^2\,\mathbf{u}_k^H\mathbf{u}_k + \sum_{k=1}^{K}\boldsymbol{\mu}_k^T\left(\Re\{\mathbf{H}'\mathbf{u}_k\} - \sqrt{\lambda_k}\mathbf{e}_k\right) \quad (22)$$

where $\boldsymbol{\mu}_k$, $1\le k\le K$, are the nonnegative Lagrange multipliers. Using Wirtinger calculus [43]–[45] to take the derivative of the Lagrangian (22) with respect to the complex-valued precoding vectors $\mathbf{u}_k$, $1\le k\le K$, and writing the KKT conditions[2] results in the following equations for the stationary points of (21):

$$\sigma_{s_k}^2\,\mathbf{u}_k^* + \frac{1}{2}\mathbf{H}'^T\boldsymbol{\mu}_k = \mathbf{0}, \quad 1\le k\le K \quad (23a)$$

$$\Re\{\mathbf{H}'\mathbf{u}_k\} - \sqrt{\lambda_k}\mathbf{e}_k = \mathbf{0}, \quad 1\le k\le K, \quad (23b)$$

where (23a) is the stationarity condition and (23b) is a primal feasibility condition [46]. From (23a), and for $1\le k\le K$

$$\mathbf{u}_k = -\frac{1}{2\sigma_{s_k}^2}\mathbf{H}'^H\boldsymbol{\mu}_k. \quad (24)$$

Substituting (24) into (23b) yields the corresponding Lagrange multipliers

$$\boldsymbol{\mu}_k = -2\sigma_{s_k}^2\sqrt{\lambda_k}\left[\Re\left\{\mathbf{H}'\mathbf{H}'^H\right\}\right]^{-1}\mathbf{e}_k. \quad (25)$$

Using (25) in (24) and concatenating the obtained ZF precoding vectors to form a matrix, the widely linear ZF precoder is obtained as

$$\mathbf{U}_{\text{ZF}} = \mathbf{H}'^H\left[\Re\left\{\mathbf{H}'\mathbf{H}'^H\right\}\right]^{-1}\boldsymbol{\Lambda}. \quad (26)$$

Widely linear ZF precoding suffers from the lack of a constraint on the transmit power at the expense of a fixed received power, which makes the total transmit power depend on the channel characteristics. To overcome this shortcoming of WL-ZF, a simple heuristic approach is to introduce a scaling factor γ to normalize $\mathbf{U}$ and constrain the total transmit power to τ [17]. In other words, setting $\gamma^2\,\text{Tr}(\mathbf{U}_{\text{ZF}}\mathbf{R}_s\mathbf{U}_{\text{ZF}}^H) = \tau$, which results in

$$\gamma = \sqrt{\frac{\tau}{\text{Tr}(\mathbf{H}'^H[\Re\{\mathbf{H}'\mathbf{H}'^H\}]^{-1}\boldsymbol{\Lambda}^2\mathbf{R}_s[\Re\{\mathbf{H}'\mathbf{H}'^H\}]^{-1}\mathbf{H}')}}. \quad (27)$$

Using the scaling factor (27) to normalize $\mathbf{U}_{\text{ZF}}$ of (26), results in the normalized WL-ZF precoding matrix

$$\mathbf{U}_{\text{ZF}}^{\text{Norm}} = \frac{\sqrt{\tau}\mathbf{H}'^H[\Re\{\mathbf{H}'\mathbf{H}'^H\}]^{-1}\Lambda}{\sqrt{\text{Tr}(\mathbf{H}'^H[\Re\{\mathbf{H}'\mathbf{H}'^H\}]^{-1}\boldsymbol{\Lambda}^2\mathbf{R}_s[\Re\{\mathbf{H}'\mathbf{H}'^H\}]^{-1}\mathbf{H}')}}. \quad (28)$$

It should be remarked that in general, using this approach or other power allocation approaches such as water-filling results in diagonal matrix $\Re\{\mathbf{WHU}\}$ not being necessarily equal to $\boldsymbol{\Lambda}$ as required by (19). In other words, imposing a constraint on the transmit power, e.g., by using regularization factor of (27), would not necessarily be compatible with the receive power constraint imposed by (19). Although, transmit power allocation may distort the receive power constraint, it would not distort the interference-free transmission inherent to ZF precoding.

C. WL Minimum Mean Square Error Precoding

In this section, we consider minimizing the sum of mean square errors (MSE) of users by constraining the transmit power. Using the estimate $\hat{\mathbf{s}} = \Re\{\mathbf{y}\}$, the sum MSE between the estimated signals and the desired signals can be written as

$$\text{MSE} = \text{E}\left[\|\hat{\mathbf{s}}-\mathbf{s}\|_2^2\right] = \text{Tr}(\Re\{\mathbf{H}'\mathbf{U}\}\mathbf{R}_s\Re\{\mathbf{U}^T\mathbf{H}'^T\}) - 2\,\text{Tr}(\Re\{\mathbf{H}'\mathbf{U}\}\mathbf{R}_s) + \text{Tr}\left(\frac{1}{2}\mathbf{R}_{z'}+\mathbf{R}_s\right), \quad (29)$$

where the factor $1/2$ stems from considering the real part of a complex noise, i.e., because $\hat{\mathbf{s}} = \Re\{\mathbf{y}\}$. Unfortunately, a similar Wirtinger calculus approach of Section IV-B results in $\mathbf{U}$ and $\mathbf{U}^*$ of WL-MMSE being coupled in such a way that a closed-form or a semi closed-form solution would not be possible to obtain.

Widely linear minimum mean square error (MMSE) precoding is optimized using complex augmentation [14] which is equivalent to the following two bijective transformations from the complex field to the real field:

$$\mathbf{U} \xrightarrow{\mathcal{T}_1} \bar{\mathbf{U}} = \begin{bmatrix}\Re\{\mathbf{U}\}\\ \Im\{\mathbf{U}\}\end{bmatrix} \quad (30a)$$

$$\mathbf{H}' \xrightarrow{\mathcal{T}_2} \tilde{\mathbf{H}}' = \begin{bmatrix}\Re\{\mathbf{H}'\} & -\Im\{\mathbf{H}'\}\end{bmatrix}. \quad (30b)$$

Using these transformations the sum MSE in (29) can be equivalently expressed as

$$\text{MSE} = \text{Tr}\left(\tilde{\mathbf{H}}'\bar{\mathbf{U}}\mathbf{R}_s\bar{\mathbf{U}}^T\tilde{\mathbf{H}}'^T\right) - 2\,\text{Tr}\left(\tilde{\mathbf{H}}'\bar{\mathbf{U}}\mathbf{R}_s\right) + \text{Tr}\left(\frac{1}{2}\mathbf{R}_{z'}+\mathbf{R}_s\right). \quad (31)$$

The total transmit power (3) can also be rewritten as

$$\text{E}\left[\|\mathbf{x}\|_2^2\right] = \text{Tr}\left(\bar{\mathbf{U}}\mathbf{R}_s\bar{\mathbf{U}}^T\right). \quad (32)$$

Having the MSE (31) and the transmit power (32), in a fashion similar to the linear MMSE precoder [17], the widely linear MMSE precoder for the broadcast channel is obtained by solving

$$\min_{\bar{\mathbf{U}}}\ \text{MSE} \quad (33a)$$

$$\text{subject to}\ \ \text{Tr}\left(\bar{\mathbf{U}}\mathbf{R}_s\bar{\mathbf{U}}^T\right)\le\tau. \quad (33b)$$

It should be noted that (33b) is not necessarily an active constraint [47].

The Lagrangian function corresponding to the optimization problem (33) is

$$L\left(\bar{\mathbf{U}},\mu\right) = \text{Tr}\left(\tilde{\mathbf{H}}'\bar{\mathbf{U}}\mathbf{R}_s\bar{\mathbf{U}}^T\tilde{\mathbf{H}}'^T\right) - 2\,\text{Tr}\left(\tilde{\mathbf{H}}'\bar{\mathbf{U}}\mathbf{R}_s\right) + \text{Tr}\left(\frac{1}{2}\mathbf{R}_{z'}+\mathbf{R}_s\right) + \mu\left(\text{Tr}\left(\bar{\mathbf{U}}\mathbf{R}_s\bar{\mathbf{U}}^T\right)-\tau\right), \quad (34)$$

[2]Employing $\Re\{\mathbf{H}'_k\mathbf{u}_k\} = \frac{\mathbf{H}'_k\mathbf{u}_k+\mathbf{u}_k^H\mathbf{H}'^H_k}{2}$, $\nabla_{\mathbf{u}_k}\mathbf{u}_k^H\mathbf{u}_k = \mathbf{u}_k^*$, $\nabla_{\mathbf{u}_k}\boldsymbol{\mu}_k^T\mathbf{H}'_k\mathbf{u}_k = \mathbf{H}'^T_k\boldsymbol{\mu}_k$, $\nabla_{\mathbf{u}_k}\boldsymbol{\mu}_k^T\mathbf{u}_k^H\mathbf{H}'^H_k = \mathbf{0}$, we have $\nabla_{\mathbf{u}_k}L(\mathbf{u}_1,\ldots,\mathbf{u}_K,\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_K) = \sigma_{s_k}^2\,\mathbf{u}_k^* + \frac{1}{2}\mathbf{H}'^T\boldsymbol{\mu}_k$.

TABLE II
DUAL ASCENT ALGORITHM FOR FINDING WL-MMSE PRECODING MATRIX IN BROADCAST CHANNELS

Initialize μ.
$l = 1$.
repeat
  Compute $\bar{\mathbf{U}}$ using (37).
  Update μ using (38).
  $l \leftarrow l + 1$.
until $\bar{\mathbf{U}}$ converges

where μ is the non-negative Lagrange multiplier. Therefore, the Lagrange dual function of (33) is $g(\mu) = \min_{\bar{\mathbf{U}}} L(\bar{\mathbf{U}}, \mu)$, and hence the corresponding dual problem is

$$\max_{\mu} \; g(\mu)$$
$$\text{subject to } \mu \geq 0. \tag{35}$$

Inherently, the dual problem is a convex optimization problem with respect to μ. To solve the dual problem, similar to [48]–[50], we take a dual ascent approach which minimizes the Lagrangian (34) and maximizes the dual function alternatingly. Using typical matrix derivative properties, available for example in Appendix E. of [51], the gradient of the Lagrangian (34) with respect to $\bar{\mathbf{U}}$ is

$$\nabla_{\bar{\mathbf{U}}} L(\bar{\mathbf{U}}, \mu) = 2\tilde{\mathbf{H}}'^{T}\tilde{\mathbf{H}}'\bar{\mathbf{U}}\mathbf{R}_s - 2\tilde{\mathbf{H}}'^{T}\mathbf{R}_s + 2\mu\bar{\mathbf{U}}\mathbf{R}_s. \tag{36}$$

By setting (36) to zero, the minimizer could be obtained as

$$\begin{aligned}\bar{\mathbf{U}}_{\text{MMSE\_Iter}} &= \left(\tilde{\mathbf{H}}'^{T}\tilde{\mathbf{H}}' + \mu\mathbf{I}_{2M}\right)^{-1}\tilde{\mathbf{H}}'^{T} \\ &= \tilde{\mathbf{H}}'^{T}\left(\tilde{\mathbf{H}}'\tilde{\mathbf{H}}'^{T} + \mu\mathbf{I}_{K}\right)^{-1},\end{aligned} \tag{37}$$

where the second equality is the result of the matrix inversion lemma [51], i.e., $(\mathbf{A} + \mathbf{XRY})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{X}(\mathbf{R}^{-1} + \mathbf{YA}^{-1}\mathbf{X})^{-1}\mathbf{YA}^{-1}$. From (37), it is obvious that $K \leq 2M$, otherwise the channels would not be linearly independent[3]. Similar to [50], the dual ascent algorithm summarized in Table II is proposed to solve the WL-MMSE precoding problem. To maximize the dual function, in each iteration of the proposed algorithm the Lagrange multiplier μ is updated in a way that it moves in the direction of its steepest ascent, or derivative, as

$$\mu^{l+1} = \left[\mu^{l} + \delta_{\mu}^{l}\left(\text{Tr}\left(\bar{\mathbf{U}}\mathbf{R}_s\bar{\mathbf{U}}^{T}\right) - \tau\right)\right]^{+} \tag{38}$$

where $[\cdot]^{+} = \max\{0, \cdot\}$, l denotes the iteration number, and δ_{μ}^{l} indicates the sequence of positive scalar step sizes for μ [49], [50].

It should be remarked that, besides the dual ascent approach, an alternative approach to address the WL-MMSE precoding problem is from the perspective of regularized zero-forcing [18]. In this approach, similar to (37), the precoding matrix is written as

$$\bar{\mathbf{U}} = \tilde{\mathbf{H}}'^{T}\left(\tilde{\mathbf{H}}'\tilde{\mathbf{H}}'^{T} + \mu\mathbf{I}_{K}\right)^{-1}. \tag{39}$$

When μ approaches zero, the precoding matrix $\bar{\mathbf{U}}$ approaches $\tilde{\mathbf{H}}'^{T}(\tilde{\mathbf{H}}'\tilde{\mathbf{H}}'^{T})^{-1}$ or by using the inverse of transformations of (30), $\mathbf{U}$ approaches $\mathbf{H}'^{H}\left[\Re\{\mathbf{H}'\mathbf{H}'^{T}\}\right]^{-1}$. In other words, μ is a regularization factor of zero-forcing precoding and hence the name regularized zero-forcing. When μ approaches infinity, $\bar{\mathbf{U}}$ approaches $\tilde{\mathbf{H}}'^{T}$ or by using the inverse of transformation (30b), $\mathbf{U}$ approaches $\mathbf{H}'^{H}$. Similar to [18] and also receive MMSE beamforming, regularization factor can be set as the inverse of SNR which can be defined as $\gamma = \frac{\tau}{K\sigma_z^2/2}$, assuming that $\mathbf{R}_z = \sigma_z^2\mathbf{I}_K$, i.e., equal noise variance for all users. This would result in the following WL regularized zero-forcing or WL-MMSE precoding matrix:

$$\bar{\mathbf{U}}_{\text{MMSE}} = \tilde{\mathbf{H}}'^{T}\left(\tilde{\mathbf{H}}'\tilde{\mathbf{H}}'^{T} + \frac{1}{\gamma}\mathbf{I}_{K}\right)^{-1}. \tag{40}$$

In this scenario, when SNR increases and the channel becomes interference-limited, the precoding matrix approaches that of zero-forcing and when SNR decrease and the channel becomes noise-limted, the precoding matrix approaches that of MRT.

D. WL Maximum Signal to Leakage and Noise Ratio Precoding

So far, we have developed widely linear MRT, ZF, and MMSE precoding. Similar to linear MRT, ZF, and MMSE precoders (excluding the regularized zero-forcing approach) [17], it can be seen that their widely linear counterparts also do not consider the effect of receiver's additive noise in calculating the precoding vectors. A conventional performance criterion in communication systems which also reflects the effect of additive noise is the signal to interference and noise ratio (SINR). However, finding precoding vectors by maximizing SINR of each user is a prohibitively complex problem and does not lead to a closed-form solution [52], [53]. On the other hand, signal to leakage and noise ratio (SLNR) is a metric that not only considers the effect of noise but also its maximization results in a closed-form solution for the precoding vectors [20]. In a broadcast channel with linear precoding, the power of the leakage of user k, $1 \leq k \leq K$, is defined as the expected total power of the signal transmitted to user k that is leaked to other users' receivers, i.e.,

$$\begin{aligned}\text{Leakage}_k &= \sum_{\substack{j=1 \\ j\neq k}}^{K}\text{E}\left[\left|w_j\mathbf{h}_j\mathbf{u}_k s_k\right|^2\right] \\ &= \sum_{\substack{j=1 \\ j\neq k}}^{K}\sigma_{s_k}^2\,\mathbf{u}_k^H\mathbf{h}_j'^{H}\mathbf{h}_j'\mathbf{u}_k = \sigma_{s_k}^2\,\mathbf{u}_k^H\mathbf{H}_{\bar{k}}'^{H}\mathbf{H}_{\bar{k}}'\mathbf{u}_k,\end{aligned} \tag{41}$$

where $\mathbf{H}_{\bar{k}}' = [\mathbf{h}_1'^{H}, \ldots, \mathbf{h}_{k-1}'^{H}, \mathbf{h}_{k+1}'^{H}, \ldots, \mathbf{h}_K'^{H}]^{H}$ and $\mathbf{h}_j'$s were introduced in Section IV-A. Consequently, the SLNR of user k is defined as the ratio between the expected power of the desired part of the received signal of that user and the combined expected noise and leakage powers [20]:

$$\frac{\sigma_{s_k}^2\,\mathbf{u}_k^H\mathbf{h}_k'^{H}\mathbf{h}_k'\mathbf{u}_k}{\sigma_{s_k}^2\,\mathbf{u}_k^H\mathbf{H}_{\bar{k}}'^{H}\mathbf{H}_{\bar{k}}'\mathbf{u}_k + \sigma_{z_k'}^2}. \tag{42}$$

Considering one-dimensional modulation combined with widely linear processing, the SLNR expression can be revised

[3]Linearly dependent channels indicate a degraded broadcast channel.

to accommodate only the effective part of the powers on the estimation of the received signals:

$$\mathrm{SLNR}_k = \frac{\sigma_{s_k}^2 \left(\Re\{\mathbf{h}_k' \mathbf{u}_k\}\right)^2}{\sigma_{s_k}^2 \Re\left\{\mathbf{u}_k^H \mathbf{H}_{\bar{k}}'^H\right\} \Re\left\{\mathbf{H}_{\bar{k}}' \mathbf{u}_k\right\} + \frac{\sigma_{z_k'}^2}{2}}. \tag{43}$$

To maximize the SLNR (43), we again have to use the linear transformations in (30) to decouple $\mathbf{U}$ and $\mathbf{U}^*$ in the optimization problem. Using these transformations, SLNR (43) can be rewritten as

$$\mathrm{SLNR}_k = \frac{\sigma_{s_k}^2 \left(\tilde{\mathbf{h}}_k' \bar{\mathbf{u}}_k\right)^2}{\sigma_{s_k}^2 \sum_{\substack{j=1 \\ j\neq k}}^{K} \left(\tilde{\mathbf{h}}_j' \bar{\mathbf{u}}_k\right)^2 + \frac{\sigma_{z_k'}^2}{2}} = \frac{\sigma_{s_k}^2 \bar{\mathbf{u}}_k^T \tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k' \bar{\mathbf{u}}_k}{\sigma_{s_k}^2 \bar{\mathbf{u}}_k^T \tilde{\mathbf{H}}_{\bar{k}}^T \tilde{\mathbf{H}}_{\bar{k}} \bar{\mathbf{u}}_k + \frac{\sigma_{z_k'}^2}{2}}$$

$$= \frac{\bar{\mathbf{u}}_k^T \tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k' \bar{\mathbf{u}}_k}{\bar{\mathbf{u}}_k^T \left(\tilde{\mathbf{H}}_{\bar{k}}^T \tilde{\mathbf{H}}_{\bar{k}} + \frac{\sigma_{z_k'}^2}{2\sigma_{s_k}^2 \|\bar{\mathbf{u}}_k\|_2^2} \mathbf{I}_{2M}\right) \bar{\mathbf{u}}_k}. \tag{44}$$

Therefore, the WL maximum SLNR (MSLNR) precoding problem can be stated as

$$\max_{\bar{\mathbf{u}}_k} \ \mathrm{SLNR}_k \tag{45a}$$

$$\text{subject to } \sigma_{s_k}^2 \|\bar{\mathbf{u}}_k\|_2^2 = \tau_k, \tag{45b}$$

where SLNR_k is given by (44). Using the constraint (45b) in (44) yields

$$\mathrm{SLNR}_k = \frac{\bar{\mathbf{u}}_k^T \tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k' \bar{\mathbf{u}}_k}{\bar{\mathbf{u}}_k^T \left(\tilde{\mathbf{H}}_{\bar{k}}^T \tilde{\mathbf{H}}_{\bar{k}} + \frac{\sigma_{z_k'}^2}{2\tau_k} \mathbf{I}_{2M}\right) \bar{\mathbf{u}}_k}, \tag{46}$$

which is a generalized Rayleigh quotient.[4] Therefore, optimization (45) is equivalent to maximization of a generalized Rayleigh quotient as

$$\operatorname*{argmax}_{\bar{\mathbf{u}}_k} \ \frac{\bar{\mathbf{u}}_k^T \tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k' \bar{\mathbf{u}}_k}{\bar{\mathbf{u}}_k^T \left(\tilde{\mathbf{H}}_{\bar{k}}^T \tilde{\mathbf{H}}_{\bar{k}} + \frac{\sigma_{z_k'}^2}{2\tau_k} \mathbf{I}_{2M}\right) \bar{\mathbf{u}}_k} \tag{47a}$$

$$\text{subject to } \sigma_{s_k}^2 \|\bar{\mathbf{u}}_k\|_2^2 = \tau_k. \tag{47b}$$

Therefore, the standard solution to the above optimization problem [54] yields the WL-MSLNR precoding vector as

$$\bar{\mathbf{u}}_{k_{\mathrm{MSLNR}}} = \frac{\sqrt{\tau_k}}{\sigma_{s_k}} \mathbf{v}_{\max}(\tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k', \mathbf{Q}_k'), \tag{48}$$

where $\mathbf{Q}_k' = \tilde{\mathbf{H}}_{\bar{k}}'^T \tilde{\mathbf{H}}_{\bar{k}}' + \frac{\sigma_{z_k'}^2}{2\tau_k} \mathbf{I}_{2M}$ and $\mathbf{v}_{\max}(\tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k', \mathbf{Q}_k')$ is the normalized eigenvector corresponding to the largest generalized eigenvalue of $\tilde{\mathbf{h}}_k'^T \tilde{\mathbf{h}}_k'$ and $\mathbf{Q}_k'$.

[4]As can be seen in [54], if $\mathbf{A}$ and $\mathbf{B}$ are Hermitian matrices and $\mathbf{x}$ is a non-zero vector, $\frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}}$ is called the generalized Rayleigh quotient. The maximum value of the generalized Rayleigh quotient is given by the maximum generalized eigenvalue of $\mathbf{A}$ and $\mathbf{B}$, i.e., $\lambda_{\max}(\mathbf{A}, \mathbf{B})$ and is achieved when $\mathbf{x}$ is proportional to the corresponding generalized eigenvector of $\mathbf{A}$ and $\mathbf{B}$. It should be noted that the generalized eigenequation of $\mathbf{A}$ and $\mathbf{B}$ is given by $\mathbf{A}\mathbf{x} = \lambda \mathbf{B}\mathbf{x}$ or equivalently by $\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$ when $\mathbf{B}$ is invertible.

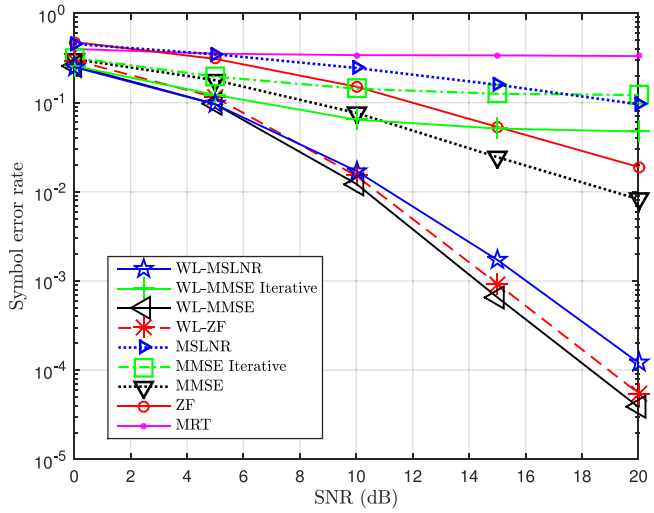

Fig. 1. Average symbol error rates of users for $M = 4$ transmit antennas and $K = 4$ users with 4-PAM modulation. The MRT, ZF, and MMSE iterative precoding methods are given [17], MMSE precoding is given in [18], and MSLNR precoding is given in [20].

V. Numerical Results

A. WL Precoding

We consider a multiple-input single-output broadcast channel (BC) with a 4-antenna transmitter and four single-antenna users. The transmitter is assumed to send independent 4-PAM signals to the users simultaneously and at the same carrier frequency. The channel gains are assumed to be quasi static and follow a Rayleigh distribution with unit variance. In other words, each element of the channel is generated as a zero-mean and unit-variance i.i.d. CSCG random variable. Since our focus is on comparing transmit precoding methods rather than on the effects of channel estimation, we assume that perfect CSI of all channels is available at the transmitter [18], [20]. At the receiver, an i.i.d. Gaussian noise is added to the received signal. All simulations are performed over 10,000 different channel realizations and at each channel realization a block of 1,000 symbols is transmitted to each user. The above set up is used for all of the following simulations unless indicated otherwise.

Fig. 1 compares the average symbol error rates of linear MRT, ZF, MMSE, iterative MMSE, and MSLNR precoding and their widely linear counterparts. As is expected, all the proposed widely linear precoding methods substantially outperform their linear counterparts. Moreover, it can be seen that the best performances are achieved by WL-MMSE, WL-ZF, and WL-MSLNR. Maximum ratio transmission, which can be considered as both a linear and a widely linear processing technique, does not exhibit good performance in high SNRs, as expected. Similar to iterative MMSE [17], as SNR increases, iterative WL-MMSE quickly reaches an error floor and does not show a promising performance. Compared to MMSE, WL-MMSE shows a gain of about 9.2 dB at the error probability of 8.25×10^{-3}, which demonstrates substantial performance improvement of WL processing compared to that of linear processing. It is known that the high SNR and low SNR performances of the investigated

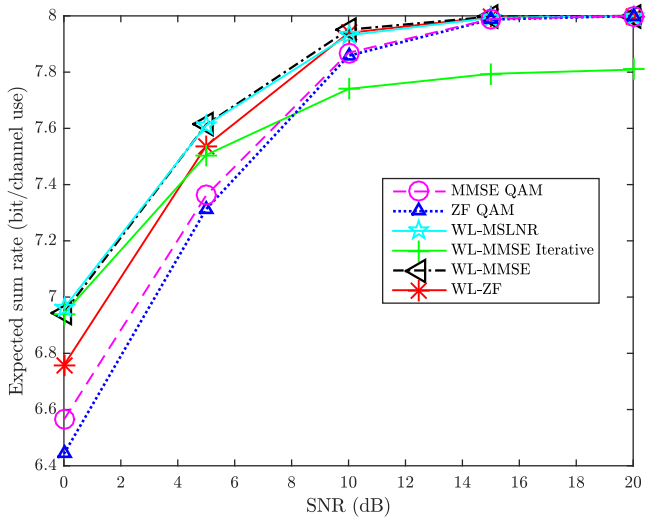

Fig. 2. Average sum rates when $M = 4$ and $K = 2$ with 16-QAM modulated signals for ZF-QAM and MMSE-QAM and when $M = 4$ and $K = 4$ with 4-PAM modulated signals for the simulated widely linear precoding methods.

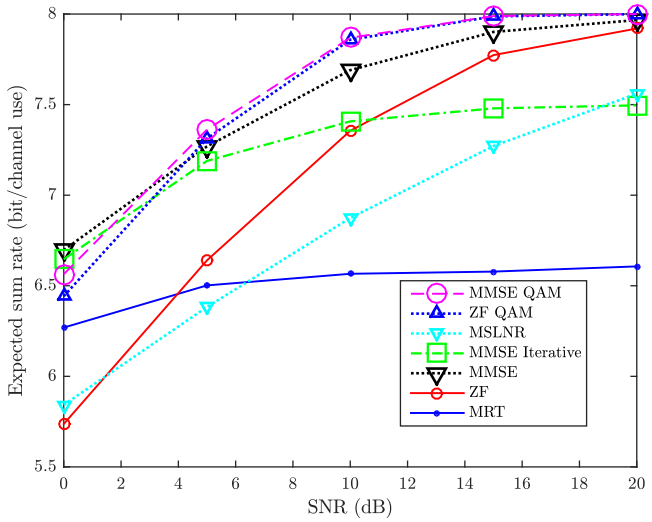

Fig. 3. Average sum rates when $M = 4$ and $K = 2$ with 16-QAM modulated signals for ZF-QAM and MMSE-QAM and when $M = 4$ and $K = 4$ with 4-PAM modulated signals for the other simulated linear precoding methods.

linear precoding methods are different [17]. At higher SNRs, the best performance is achieved from top to bottom by MMSE, ZF, MSLNR, iterative MMSE, and MRT. These results are consistent with the results obtained in [17]. In a similar fashion, if widely linear methods are compared to one another, at higher SNRs from best to worst the order of performance is WL-MMSE, WL-ZF, WL-MSLNR, iterative WL-MMSE, and MRT. An interesting remark on these comparisons is that widely linear processing substantially benefits MSLNR precoding. Linear MSLNR precoding, which did not perform well in this simulation setting, is significantly improved by using widely linear processing which indicates the importance of availability of additional degrees of freedom in design of MSLNR precoder.

In Fig. 1, it was shown that WL precoding of 1D modulated signals outperforms linear precoding of 1D modulated signals. It would also be instructive to compare the performance of WL precoding of 1D modulated signals with that of linear precoding of two-dimensionally modulated signals. Fig. 2 depicts the system throughput (expected sum rate) of four users with 4-PAM modulation employing the proposed WL precoding methods and the system throughput of two users with 16-QAM modulation employing linear precoding methods. Generally in link level simulations, user throughput is computed by considering the number of successfully decoded bits of the transport block of the user [55]. Since we do not employ channel coding and parity checks, transport block size only depends on the modulation order of each user and consequently system throughput depends on the number of users and their modulation order. By assuming gray mapping between information bits and modulated symbols and that one modulated symbol is transmitted per channel use, user throughput is computed numerically using the ratio of the number of successfully estimated bits to the number of transmitted bits multiplied by the modulation order of the user (number of bits per modulated symbol). Consequently, system throughput is the summation of users' throughput, assuming concurrent transmission to users. Theoretically, four users with 4-PAM modulation and two users with 16-QAM modulation, both achieve a maximum sum rate of 8 bits/channel use. Therefore, it is very interesting to observe that WL-ZF and WL-MMSE precoding of four 4-PAM modulated users effectively exploit the relaxed orthogonality and outperform ZF and MMSE precoding of two 16-QAM modulated users. Widely linear MSLNR precoding also outperforms both linear ZF and MMSE precoding, at all simulated SNRs. The fact that iterative WL-MMSE seems to be unable to achieve the system throughput of 8 bits/channel use is also consistent with our findings in Fig. 1 which exhibits the error floor of iterative WL-MMSE precoding.

For a more comprehensive study, we also present Fig. 3 which depicts the system throughput of linear precoding of four users with 4-PAM modulation and ZF and MMSE precoding of two users with 16-QAM modulation. At higher simulated SNRs, the system throughput achieved by MMSE and ZF precoding of two 16-QAM modulated users is higher than that of any other combination of modulation and precoding. At lower simulated SNRs, the system throughput of MMSE precoding of four 4-PAM modulated users is higher than that of any other combination of the investigated modulation and precoding. This can be explained by better performance of a lower order modulation such as 4-PAM, compared to a higher order modulation such as 16-QAM in a noise-limited communication scenario. As expected from Fig. 1, MRT, iterative MMSE, and MSLNR precoding methods do not perform as well as other precoding methods. By comparing Figs. 2 and 3, it becomes clear that, at all SNRs simulated, the proposed widely linear precoding methods achieve higher throughput compared to their linear counterparts.

B. User Selection

In this section, we evaluate the performance of the proposed SUSOM algorithm of Table I. In Fig. 4, the performance of SUSOM is compared to that of the SUS algorithm of [26] and the GUS algorithm of [35], for $M = 2, 8, 16$ transmit antennas

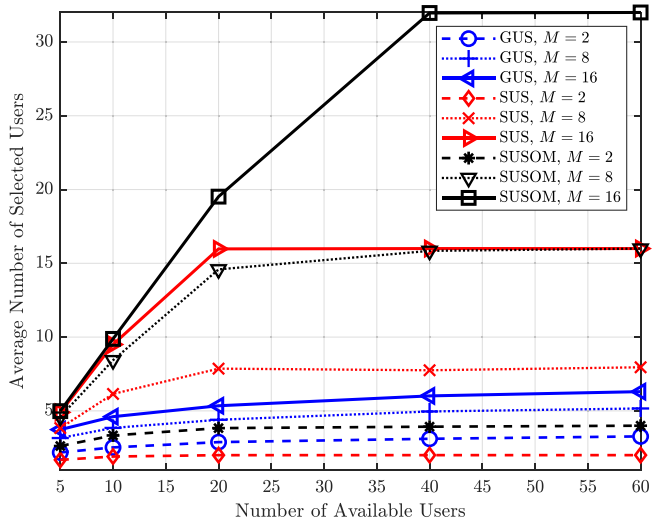

Fig. 4. Average numbers of selected users with SUS [26], GUS [35], and SUSOM vs. total number of available users for $M = 2, 8, 16$ transmit antennas.

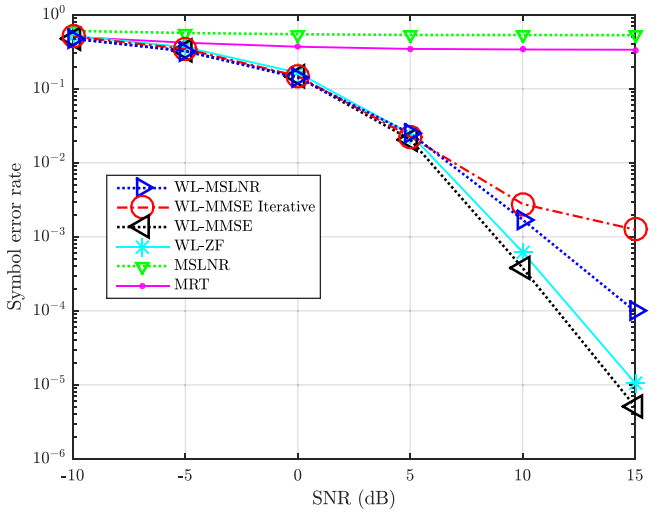

Fig. 5. Average symbol error rates of users for $M = 4$ and $K_T = 100$ available users assuming the SUSOM algorithm is employed and all users transmit 4-PAM signals.

when one receive antenna is employed at each user. Each curve is averaged over 1,000 different channel realizations. It can be observed in Fig. 4 that as the number of available users increases, the number of selected users for all three algorithms increases. As can be seen, the SUS algorithm could select at most M users for all three cases of $M = 2, 8, 16$ transmit antennas when K is large enough. On the other hand, the proposed SUSOM algorithm can select up to at most $2M$ users, i.e., twice of that of the SUS algorithm. The GUS algorithm is also expected to saturate at $2M$ users. Although it becomes clear from Fig. 4 that a major drawback of the GUS algorithm is its slow saturation rate as the number of available users increases. Another disadvantage of the GUS algorithm, as can be seen in Fig. 4, is its slow increase in the number of selected users as the number of antennas increases. This indicates that the GUS algorithm does not always find the best set of users compared to SUSOM, despite the fact that it can potentially select more users than SUS for instance when $M = 2$. It is interesting to observe that even before saturation, SUSOM outperforms both GUS and SUS. For example, when $M = 8$ and there are 20 users available, the average number of selected users is 14.59 with SUSOM, 4.4 with GUS, and 7.86 with SUS.

Fig. 5 compares the average symbol error rates of MRT, MSLNR, WL-ZF, WL-MMSE, iterative WL-MMSE, and WL-MSLNR precoding when the SUSOM algorithm is employed. It is assumed that at first the SUSOM algorithm selects a set of users out of $K_T = 100$ available users and then the transmitter broadcasts information to the selected users using the above precoding methods. From Fig. 4 it is known that using the SUSOM algorithm the average number of selected users is 7.96 when $K_T = 100$, i.e., more than the number of transmit antennas (system is overloaded). Therefore, we do not perform linear ZF, MMSE, and iterative MMSE for these cases, since they are not designed to work on overloaded systems in their presented form. We also do not display the error probabilities of the above precoding methods combined with GUS and SUS algorithms, since GUS and SUS algorithms select different numbers of users compared to SUSOM, and therefore comparing the error probabilities in such a case would not bear a meaningful interpretation. However, if we compare Fig. 5 and Fig. 1, it can be observed that the SUSOM algorithm not only selects more users than that in Fig. 1, but also all the investigated precoding methods under SUSOM achieve a better symbol error rate compared to those of Fig. 1.

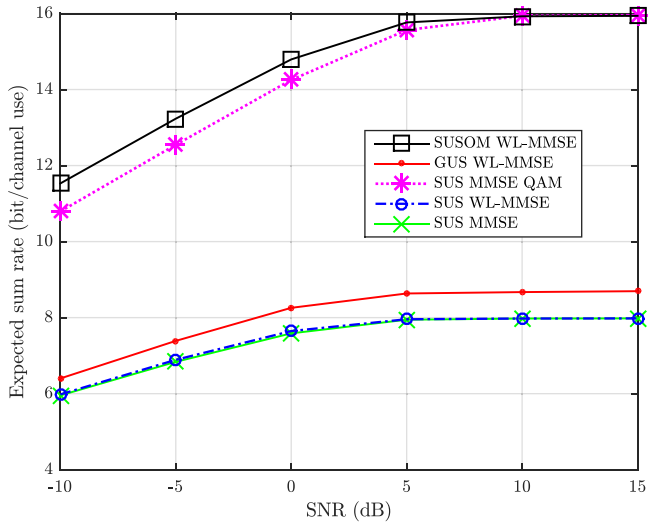

Fig. 6. Average sum rates when $M = 4$ and $K_T = 100$ with 16-QAM modulated signals for SUS MMSE QAM and when $M = 4$ and $K_T = 100$ with 4-PAM modulated signals for the other user selection and precoding methods. The MMSE precoding is given in [18], the SUS algorithm is given in [26], and the GUS algorithm is given in [35].

As observed in Fig. 5, since the error probability alone is not a good indicator of the performance when there are different numbers of users in the system, Fig. 6 is provided to gain more insight into the relative performances of the user selection algorithms. In Fig. 6, the system throughput of 4-PAM modulated users for combinations of SUS with MMSE, SUS

with WL-MMSE, GUS with WL-MMSE, and SUSOM with WL-MMSE are presented. In addition, the system throughput of 16-QAM modulated users with SUS algorithm and MMSE precoding is also plotted for comparison. As can be seen, at all simulated SNRs, the combination of SUSOM with 4-PAM and WL-MMSE achieves the highest throughput. It can be seen in Fig. 6 that as SNR increases, the achievable throughput approaches limits determined by the average numbers of selected users and the order of modulation. For this example, at higher SNRs the achievable throughput of both SUSOM with 4-PAM and SUS with 16-QAM are 16 bits/channel use.

VI. Conclusion

In this paper, we proposed a widely linear (WL) transmit precoding design for one-dimensionally modulated signals in a standard broadcast communication channel. Closed-form solutions for the precoders of the WL-MRT and the WL-ZF were obtained by using complex-domain analysis and closed-form solutions of WL-MMSE and WL-MSLNR were obtained by analysis of the composite real representation. It was shown that WL-ZF and WL-MMMSE precoders can properly operate even if the number of users is twice as large as the number of transmit antennas, as opposed to linear ZF and MMSE precoders which can only support as many users as the number of transmit antennas. We also developed a user selection algorithm, compatible with widely linear precoding, that can select twice as many users as the number of transmit antennas. It has been shown that WL precoding outperforms linear precoding. Moreover, it has been shown that widely linear precoding in conjunction with the proposed semi-orthogonal user selection algorithm for one-dimensional modulation (SUSOM) also outperforms linear precoding in conjunction with semi-orthogonal user selection algorithm (SUS). In summary, higher throughput may be achieved and the number of users can be increased in a multiuser communication system by using widely linear processing of one-dimensionally modulated signals compared to linear processing of both one-dimensionally and two-dimensionally modulated signals.

Appendix
Proof of Claim 1

Let us consider the following transformation from the complex field to the real field:

$$\mathbf{h}_k \in \mathbb{C}^{1\times M} \xrightarrow{\mathcal{T}_1} \bar{\mathbf{h}}_k = \begin{bmatrix} \Re\{\mathbf{h}_k\} & \Im\{\mathbf{h}_k\} \end{bmatrix} \in \mathbb{R}^{1\times 2M}, \quad 1 \le k \le K. \tag{49}$$

Then, it is obvious that

$$\Re\{\mathbf{h}_k \mathbf{h}_j^H\} = \bar{\mathbf{h}}_k \bar{\mathbf{h}}_j^T. \tag{50}$$

If we define the $K \times 2M$ matrix $\bar{\mathbf{H}}$ as

$$\bar{\mathbf{H}} = \begin{bmatrix} \bar{\mathbf{h}}_1 \\ \vdots \\ \bar{\mathbf{h}}_K \end{bmatrix}, \tag{51}$$

then we have the following lemma:

Lemma 1: All K channels are mutually orthogonal if and only if the $K \times K$ matrix $\bar{\mathbf{H}}\bar{\mathbf{H}}^T$ is a full rank diagonal matrix.

Proof: If $\bar{\mathbf{H}}\bar{\mathbf{H}}^T$ is a full rank diagonal matrix, then it is equivalently represented by $\mathrm{diag}(\|\bar{\mathbf{h}}_1\|_2^2, \ldots, \|\bar{\mathbf{h}}_K\|_2^2)$. In other words, $\bar{\mathbf{h}}_k \bar{\mathbf{h}}_j^T = 0$, $1 \le j \ne k \le K$, i.e., the channels are mutually orthogonal. On the other hand, $\|\bar{\mathbf{h}}_k\|_2 \ne 0$ with probability one and if the channels are mutually orthogonal then $\bar{\mathbf{h}}_k \bar{\mathbf{h}}_j^T = 0$, $1 \le j \ne k \le K$, which results in $\bar{\mathbf{H}}\bar{\mathbf{H}}^T$ being a full rank diagonal matrix. ■

On the other hand, it is known that $\mathrm{rank}(\bar{\mathbf{H}}\bar{\mathbf{H}}^T) = \mathrm{rank}(\bar{\mathbf{H}}) \le \min(K, 2M) \le 2M$. Therefore, the maximum number of mutually orthogonal channels is $K = 2M$.

References

[1] "Emerging communication technologies enabling the Internet of things," *Rohde & Schwarz White Paper*, Sep. 2016.

[2] S. Huang, H. Yin, J. Wu, and V. C. M. Leung, "User selection for multiuser MIMO downlink with zero-forcing beamforming," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3084–3097, Sep. 2013.

[3] N. D. Mickulicz, U. Drolia, P. Narasimhan, and R. Gandhi, "Zephyr: First-person wireless analytics from high-density in-stadium deployments," in *Proc. IEEE 17th Int. Symp. World Wireless, Mobile Multimedia Netw.*, Jun. 2016, pp. 1–10.

[4] W. Brown and R. Crane, "Conjugate linear filtering," *IEEE Trans. Inf. Theory*, vol. 15, no. 4, pp. 462–465, Jul. 1969.

[5] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[6] C. Hellings, M. Joham, and W. Utschick, "QoS feasibility in MIMO broadcast channels with widely linear transceivers," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1134–1137, Nov. 2013.

[7] F. Sterle, "Widely linear mmse transceivers for MIMO channels," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4258–4270, Aug. 2007.

[8] A. Medra and T. N. Davidson, "Widely linear interference alignment precoding," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2014, pp. 464–468.

[9] M. Bavand and S. D. Blostein, "Widely linear multiuser simultaneous information and power transfer with one-dimensional signaling," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2016, pp. 1–6.

[10] D. Darsena, G. Gelli, I. Iudice, and F. Verde, "Widely-linear transceiver design for amplify-and-forward MIMO relaying," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Sep. 2016, pp. 1–5.

[11] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 746–756, Mar. 2004.

[12] P. Chevalier, J.-P. Delmas, and A. Oukaci, "Optimal widely linear MVDR beamforming for noncircular signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2009, pp. 3573–3576.

[13] S. Tan, L. Xu, S. Chen, and L. Hanzo, "Iterative soft interference cancellation aided minimum bit error rate uplink receiver beamforming," in *Proc. 63rd IEEE Veh. Technol. Conf.*, May 2006, pp. 17–21.

[14] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[15] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[16] P. Xiao and M. Sellathurai, "Improved linear transmit processing for single-user and multi-user MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1768–1779, Mar. 2010.

[17] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[18] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[19] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.
[20] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1711–1721, May 2007.
[21] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.
[22] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1396–1404, Mar. 2009.
[23] F. Liu, C. Masouros, P. V. Amadori, and H. Sun, "An efficient manifold algorithm for constructive interference based constant envelope precoding," *IEEE Signal Process. Lett.*, vol. 24, no. 10, pp. 1542–1546, Oct. 2017.
[24] L. Li, A. Ashikhmin, and T. Marzetta, "Pilot contamination precoding for interference reduction in large scale antenna systems," in *Proc. 51st Allerton Conf. Commun., Control, Comput.*, Oct. 2013, pp. 226–232.
[25] D. Neumann, A. Gründinger, M. Joham, and W. Utschick, "Rate-balancing in massive MIMO using statistical precoding," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2015, pp. 226–230.
[26] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
[27] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
[28] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath Jr, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, Sep. 2006.
[29] J. Mao, J. Gao, Y. Liu, and G. Xie, "Simplified semi-orthogonal user selection for MU-MIMO systems with ZFBF," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 42–45, Feb. 2012.
[30] G. Gupta and A. K. Chaturvedi, "User selection in MIMO interfering broadcast channels," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1568–1576, May 2014.
[31] S. Nam, J. Kim, and Y. Han, "A user selection algorithm using angle between subspaces for downlink MU-MIMO systems," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 616–624, Feb. 2014.
[32] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.
[33] N. D. Dao and Y. Sun, "User-selection algorithms for multiuser precoding," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3617–3622, Sep. 2010.
[34] R. Rajashekar and L. Hanzo, "User selection algorithms for block diagonalization aided multiuser downlink mm-wave communication," *IEEE Access*, vol. 5, pp. 5760–5772, Mar. 2017.
[35] M. Bavand and S. D. Blostein, "Modulation-specific multiuser transmit precoding and user selection for one-dimensional signalling," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10946–10961, Jun. 2017.
[36] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2001.
[37] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[38] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[39] J. Grob, G. Trenkler, and S.-O. Troschke, "On semi-orthogonality and special class of matrices," *Linear Algebra Its Appl.*, vol. 289, pp. 169–182, Mar. 1999.
[40] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2450–2454, Sep. 2002.
[41] S. Chen, A. Livingstone, H.-Q. Du, and L. Hanzo, "Adaptive minimum symbol error rate beamforming assisted detection for quadrature amplitude modulation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1140–1145, Apr. 2008.
[42] E. Bjornson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.
[43] D. Brandwood, "A complex gradient operator and its application in adaptive array theory," *Proc. IEEE*, vol. 130, no. 1, pp. 11–16, Feb. 1983.
[44] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.
[45] J. Eriksson, E. Ollila, and V. Koivunen, "Essential statistics and tools for complex random variables," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5400–5408, Oct. 2010.
[46] S. Boyd, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[47] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.
[48] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Breda, Netherlands: Now Publishers Inc, 2011.
[49] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
[50] N. Mokari, M. R. Javan, and K. Navaie, "Cross-layered resource allocation in OFDMA systems for heterogenous trafic with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 1011–1017, Feb. 2010.
[51] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.
[52] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
[53] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
[54] J. R. Schott, *Matrix Analysis for Statistics*. New York, NY, USA: Wiley-Interscience, 1997.
[55] "3gpp TS 38.214 v15.2.0: Physical layer procedures for data (release 15)," Technical Specification Group Radio Access Network, Jun. 2018.

Majid Bavand (S'10–M'18) received the B.Sc. degree from Amirkabir University of Technology, Tehran, Iran, and the M.Sc. degree from Tarbiat Modares University, Tehran, Iran, both in electrical engineering, and the Ph.D. degree in electrical and computer engineering from Queen's University, Kingston, ON, Canada. He has worked as wireless R&D and Senior R&D Scientist with Mapsted Corp., Toronto, ON, Canada, working on architectural design of indoor navigation and positioning systems. He is now a RAN System Designer at Ericsson, Ottawa, Canada, working on algorithm development for LTE and 5G NR.

His research interests include signal processing and wireless communications, focusing on transmit and receive beamforming, radio resource management, and energy efficiency in addition to application of machine intelligence in localization. He is the recipient of the Best Student Paper runner-up award in the 27th Biennial Symposium on Communications.

Steven D. Blostein (S'82–M'88–SM'96) received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 1983, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1985 and 1988, respectively. Since 1988, he has been on the Faculty with the Department of Electrical and Computer Engineering, Queens University, Kingston, ON, Canada, and currently holds the position of Professor. From 2004 to 2009, he was the Department Head. He has also been a consultant to industry and government in the areas of image compression, target tracking, radar imaging, and wireless communications.

His current interests include wireless communications systems, including detection and estimation, signal processing, energy efficiency, MIMO, dynamic access, and dense deployments. He was Chair of IEEE Kingston Section, Chair of the Biennial Symposium on Communications, Publications Chair IEEE ICASSP, Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and served on numerous Technical Program Committees for the IEEE Communications Society conferences. He is a registered professional engineer in Ontario.