Low Complexity Dirty Paper Coding for MU-MIMO Channels

Zhibin Zou, and Aveek Dutta
Department of Electrical and Computer Engineering
University at Albany SUNY, Albany, NY 12222 USA
{zzou2, adutta}@albany.edu

***Abstract*—Dirty Paper Coding (DPC) is considered as the optimal precoding which achieves capacity for the Gaussian Multiple-Input Multiple-Output (MIMO) broadcast channel (BC). However, to find the optimal precoding order, it needs to repeat $N!$ times for N users as there are $N!$ possible precoding orders. This extremely high complexity limits its practical use in modern wireless networks. In this paper, we show the equivalence of DPC and the recently proposed Higher Order Mercer's Theorem (HOGMT) precoding [1] in 2-D (spatial) case, which provides an alternate implementation for DPC. Furthermore, we show that the proposed implementation method is linear over the permutation operator when permuting over multi-user channels. Therefore, we present a low complexity algorithm that optimizes the precoding order for DPC with beamforming, eliminating repeated computation of DPC for each precoding order. Simulations show that our method can achieve the same result as conventional DPC with $\approx$30 dB lower complexity for N=10 users.**

***Keywords*—Non-linear Precoding, MU-MIMO, Dirty Paper Coding (DPC), Beamforming, Precoding Orders Optimization.**

I. Introduction

Precoding is a very well investigated area, which can cancel interference if the CSI is available at the transmitter [2]. DPC is a non-linear precoding that achieves optimal interference-free transmission by subtracting the potential interference at the transmitter [3], which is well investigated for MU-MIMO channels [4]. In multi-user information theory literature, the downlink MU-MIMO channel is modeled as MIMO Gaussian broadcast channel (BC) [5], where the sum-rate capacity grows linearly with the number of spatial-domain degrees of freedom [6]. DPC is proven to achieve capacity for MIMO BC channels [7], [8], [9]. However, practical implementation of DPC has the great challenge of *very high computational complexity*. At the same time, the power allocation problem is studied in the beamforming literature from linear methods [10], [11] to nonlinear DPC [12]. However, these approaches only solve the problem for a *fixed DPC precoding order*. An inherent problem with DPC is that for every order, the interference coupling matrix has a different structure. Thus finding the optimum precoding order remains a combinatorial problem that is prohibitive, even for moderate numbers of users. A low complexity but sub-optimal method to achieve this has been shown in [13].

HOGMT precoding [1] is the first method, which is capable of cancelling spatial, temporal and joint spatio-temporal interference in multi-user non-stationary channels. This is achieved by transmitting signals on independent flat-fading subchannels (eigenfunctions) in an eigen-domain. As a joint spatio-temporal precoding method for multi-user non-stationary channels, HOGMT generally analyzes a 4-D channel tensors. However, if time dimension at the transmitter and the receiver are both collapsed, as it would be LTI channels, it will operate on a 2-D MU-MIMO channel matrix to cancel spatial interference only, which is exactly the same as in DPC for MU-MIMO channels [4, Chapter 13].

In this paper, we prove the equivalence between DPC and 2-D HOGMT, since both ensure interference-free communication. 2-D HOGMT precoding is implemented by SVD decomposition [1], which is a linear process. Therefore, the equivalence provides an alternate linear implementation for DPC. Furthermore, we show that the SVD decomposition of a permuted matrix can be obtained by directly permuting the decomposed components (Lemma 1). This property does not hold for the LQ decomposition, commonly used in implementing DPC, because the decomposed triangular matrix cannot preserve its structure after permutation. This difference suggests that conventional method based on LQ decomposition needs to repeat the DPC for each permutation of the channel matrix (or the precoding order), while the proposed alternate method requires *only one DPC computation* for any arbitrary order, followed by *permuting the decomposed components* to find the optimal order, avoiding unnecessary iterations.

The contributions of this paper are summarized as follows:

- We show the equivalence between DPC and 2-D HOGMT precoding with effective channel gains, and give an alternate implementation for DPC.
- We give a general beamforming optimization method by designing a diagonal matrix according to a desired criteria and under certain constraints.
- We show the difference between SVD and LQ decomposition under the permutation operation and give a low complexity algorithm for the optimal precoding order.

II. Background & Preliminaries

A. DPC for MIMO Broadcast Channels

Consider a BC channel with N transmit antennas and N single-antenna users (MU-MISO), $\mathbf{H} \in \mathbb{C}^{N\times N}$, where the received signal $\mathbf{y} \in \mathbb{C}^{N\times 1}$ is given by,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{1}$$

where, $\mathbf{x} \in \mathbb{C}^{N\times 1}$ is the precoded signal and $\mathbf{v} \in \mathbb{C}^{N\times 1}$ is AWGN. Computationally, DPC performs LQ decomposition followed by a series of Gram-Schmidt processes [4]. The channel matrix, $\mathbf{H}$ is decomposed as $\mathbf{H}$=$\mathbf{LQ}$ where, $\mathbf{L} \in \mathbb{C}^{N\times N}$ and $\mathbf{Q}^{N\times N}$ is a triangular and unitary matrix, respectively. Let $\tilde{\mathbf{x}}$=$[\ x_1, \ldots, x_N\ ]^T$ denote the precoded signal for $\mathbf{s}$=$[\ s_1, \ldots, s_N\ ]^T$ to cancel the effect of $\mathbf{L}$. By transmitting $\mathbf{x}$=$\mathbf{Q}^H\tilde{\mathbf{x}}$, the effect of $\mathbf{Q}$ is cancelled and (1) is rewritten as,

$$\begin{aligned}\mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{v} = \mathbf{L}\mathbf{Q}\mathbf{Q}^H\tilde{\mathbf{x}} + \mathbf{v} \\ &= \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{N1} & l_{N2} & \cdots & l_{NN} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_N \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}\end{aligned} \tag{2}$$

Therefore, for n^{th} user, there is no interference from users n'>n and the interference from users n'<n is cancelled by the Gram–Schmidt process as,

$$\tilde{x}_n = s_n - \sum_{n'=1}^{n-1} \frac{l_{nn'}}{l_{nn}} \tilde{x}_{n'} \quad \text{where,} \quad \tilde{x}_1 = s_1 \tag{3}$$

Substituting (3) in (2), the received signal is given by,

$$\mathbf{y} = \mathbf{D_L}\mathbf{s} + \mathbf{v} \tag{4}$$

where, $\mathbf{D_L}$= $\mathrm{diag}(\mathbf{L})$ and l_{nn} is the channel gain for user n.

Multi-antenna users case (MU-MIMO): For multi-antenna user case, each row in (2) corresponds to one antenna instead of one user and then each user would incorporate multiple rows as well as multiple elements $\tilde{\mathbf{x}}_n$ in (3). For notational simplicity, we use the expression of the single-antennas user case as it does not affect the underlying theory in this paper.

B. HOGMT Precoding

HOGMT precoding [1] cancels the spatial, temporal and joint spatio-temporal interference in a 4-D double-selective channel, modeled as in [14],

$$\mathbf{H}(t,\tau) = \begin{bmatrix} h_{1,1}(t,\tau) & \cdots & h_{1,u'}(t,\tau) \\ \vdots & \ddots & \\ h_{u,1}(t,\tau) & & h_{u,u'}(t,\tau) \end{bmatrix} \tag{5}$$

where, $h_{u,u'}(t,t')$ is the multi-user time-varying impulse response. Then the received signal is given by,

$$r(u,t) = \iint k_H(u,t;u',t') s(u',t')\ du'\ dt' + v(u,t) \tag{6}$$

where, $v(u,t)$ is AWGN, $s(u,t)$ is the data symbol and $k_{u,u'}(t,t')$=$h_{u,u'}(t,t-t')$ is the 4-D channel kernel [15], [16].

HOGMT decomposition is the first method to decompose a 4-D channel kernels as follows,

$$k_H(u,t;u',t') = \sum\nolimits_{n=1}^{N} \sigma_n \psi_n(u,t)\phi_n(u',t') \tag{7}$$

with orthonormal properties as in (8),

$$\begin{aligned}\langle \psi_n(u,t), \psi^*_{n'}(u,t)\rangle = \delta_{nn'} \\ \langle \phi_n(u,t), \phi^*_{n'}(u,t)\rangle = \delta_{nn'}\end{aligned} \tag{8}$$

Both (7) and (8) show that the 4-D channel kernel is decomposed into jointly orthogonal subchannels (eigenfunctions). Then the precoded signal $x(u,t)$ based on HOGMT is derived by combining the jointly orthogonal eigenfunctions with the desired coefficients x_n as,

$$x(u,t) = \sum_{n=1}^{N} x_n \phi^*_n(u,t) \text{ where, } x_n = \frac{\langle s(u,t), \psi_n(u,t)\rangle}{\sigma_n} \tag{9}$$

Transmitting $x(u,t)$ over the channel, the received signal is directly the combination of data signal and noise without complementary post-coding step as $r(u,t) = s(u,t) + v(u,t)$. It shows that HOGMT precoding can achieves interference-free communication for multi-user non-stationary channels.

III. Equivalence of DPC and HOGMT Precoding

DPC achieves capacity for MU-MIMO BC channels but is a non-linear precoding with impractical complexity. On the contrary, HOGMT achieves the same interference-free communication for multidimensional non-stationary channels and has a linear implementation. If there exists equivalence between them, then we can use it as an alternate implementation for DPC for practical system implementation.

Theorem 1. *(DPC and 2-D HOGMT precoding with effective channel gains are mathematically equivalent)*

Given a channel matrix $\mathbf{H}$ *with entries* $h(u,u')$ *and data symbols* $\mathbf{s}$=$[s_1, ..., s_N]^T$*, the 2-D HOGMT precoded signal is,*

$$x(u) = \sum_{n}^{N} x_n \phi^*_n(u) \text{ where, } x_n = \frac{\langle s(u), \psi_n(u)\rangle}{\sigma_n} \tag{10}$$

where σ_n*,* $\phi_n(u)$ *and* $\psi_n(u)$ *are given by 2-D HOGMT decomposition as* $h(u,u')$=$\sum_n^N \sigma_n \psi_n(u)\phi_n(u')$ *[1].*

Then the DPC precoded signal is given by

$$x(u) = \sum_{n}^{N} x_n \phi^*_n(u) \text{ where, } x_n = \frac{\langle l(u), s(u), \psi_n(u)\rangle}{\sigma_n} \tag{11}$$

where, $l(u)$ *is the continuous diagonal element of* $\mathbf{D_L}$ *in* (4).

Proof. Let $\mathbf{x}$=$\mathbf{Ws}$, where $\mathbf{W}$ is the precoding matrix. Then we can write,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} = \mathbf{LQWs} + \mathbf{v} \tag{12}$$

Substituting (4) in (12), we have $\mathbf{W}$=$\mathbf{Q}^H\mathbf{L}^{-1}\mathbf{D_L}$. Decomposing $\mathbf{L}$ by SVD as $\mathbf{L} = \mathbf{U_L}\mathbf{\Sigma}\mathbf{V}_L^H$ we get,

$$\mathbf{W} = \mathbf{Q}^H\mathbf{V_L}\mathbf{\Sigma_L}^{-1}\mathbf{U_L}^H\mathbf{D_L} \tag{13}$$

Meanwhile, the SVD of $\mathbf{H}$ can be also represented by the SVD of $\mathbf{L}$ as,

$$\begin{aligned}\mathbf{H} &= \mathbf{LQ} = (\mathbf{U_L}\mathbf{\Sigma_L}\mathbf{V_L}^H)\mathbf{Q} \\ &= \mathbf{U_L}\mathbf{\Sigma_L}(\mathbf{Q}^H\mathbf{V_L})^H \equiv \mathbf{U\Sigma V}^H\end{aligned} \tag{14}$$

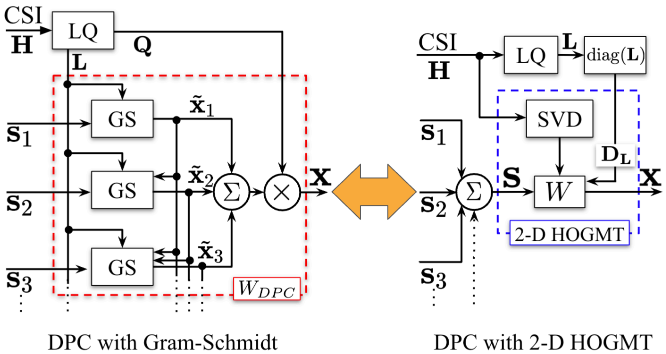


Figure 1: A low-complexity implementation of DPC

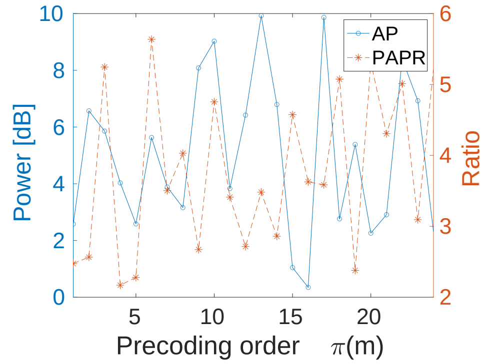


Figure 2: AP and PAPR for different precoding orders

Therefore we have the following equivalence,

$$\mathbf{U} \equiv \mathbf{U_L},\ \ \mathbf{\Sigma} \equiv \mathbf{\Sigma_L} \text{ and } \mathbf{V} \equiv \mathbf{Q}^H\mathbf{V_L} \tag{15}$$

Substituting (15) in (13) and noting that $\mathbf{Q}$ is unitary,

$$\mathbf{W} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{D_L} \tag{16}$$

Then the transmitted symbol $\mathbf{x}$ is given by,

$$\mathbf{x}{=}\mathbf{W}\mathbf{s}{=}\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{D_L}\mathbf{s} \tag{17}$$

Note that the precoded symbol for user u is the u^{th} row of $\mathbf{x}$. By expanding (17), we have

$$x_u{=}\sum_n^N v_{un}\underbrace{\sigma_n^{-1}\sum_u^N l_u s_u u_{un}^*}_{x_n} \tag{18}$$

where, v_{un} and u_{un} are the elements of $\mathbf{V}$ and $\mathbf{U}$ respectively and σ_n is n^{th} diagonal element of $\mathbf{\Sigma}$.

Now, rewriting (18) using two arbitrary continuous-time complex functions, $\phi_n^*(u)$ and $\psi_n^*(u)$ we get,

$$x(u) = \sum_n^N x_n\phi_n^*(u) \text{ where, } x_n = \frac{\langle l(u), s(u), \psi_n(u)\rangle}{\sigma_n} \tag{19}$$

where, $x(u)$, $s(u)$, and $l(u)$ is the continuous form of $\mathbf{x}$, $\mathbf{s}$ and $\{l_u\}$, respectively.

Meanwhile, the continuous form of SVD of $\mathbf{H}$, yields the two eigenfunctions, ϕ and ψ according to the the 2-D HOGMT decomposition in (20) by collapsing time dimension in (7),

$$k(u, u') = \Sigma_n\sigma_n\phi_n(u)\psi(u') \tag{20}$$

Therefore, we have the 2-D form of (9) as,

$$x(u) = \sum_n^N x_n\phi_n^*(u),\ \text{where, } x_n = \frac{\langle s(u), \psi_n(u)\rangle}{\sigma_n} \tag{21}$$

Therefore, observing the similarity of (19) and (21) we find that DPC is mathematically same as 2-D HOGMT precoding after scaling by the effective gain, $l(u)$. $\square$

Figure 1 illustrates the equivalence shown in Theorem 1 to provide an alternate implementation of DPC using HOGMT, as in (16). Note that the non-linearity of DPC is due to the iterative feedback required by the Gram–Schmidt process as shown in Figure 1 Therefore, because of the equivalence and the linear implementation of 2-D HOGMT precoding by SVD provides significant computational advantage in practical implementation of DPC in MU-MIMO channels.

A. Beamformer optimization

From Theorem 1, the beamformer is obtained by designing an optimal pre-equalizer, $b(u)$ for the u^{th} user. Then x_n in (11) can be expressed as

$$x_n = \frac{\langle b(u), l(u), s(u), \psi_n(u)\rangle}{\sigma_n} \tag{22}$$

Specifically, if $b(u) = 1/l(u)$, then the DPC implemented using (11) is numerically equal to 2-D HOGMT in (10). Let $k(u) = b(u)l(u)$, which is the desired effective gain, and denote the discrete form of $k(u)$ as diagonal matrix $\mathbf{K}$, then (16) with beamformer is given by,

$$\mathbf{W} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{K} \tag{23}$$

Then the beamformer design is to find a diagonal matrix $\mathbf{K}$ to replace $\mathbf{D_L}$ in (16). The optimal $\mathbf{K}$ is obtained by the objective function $f(\cdot)$ under the power constraint P as follows:

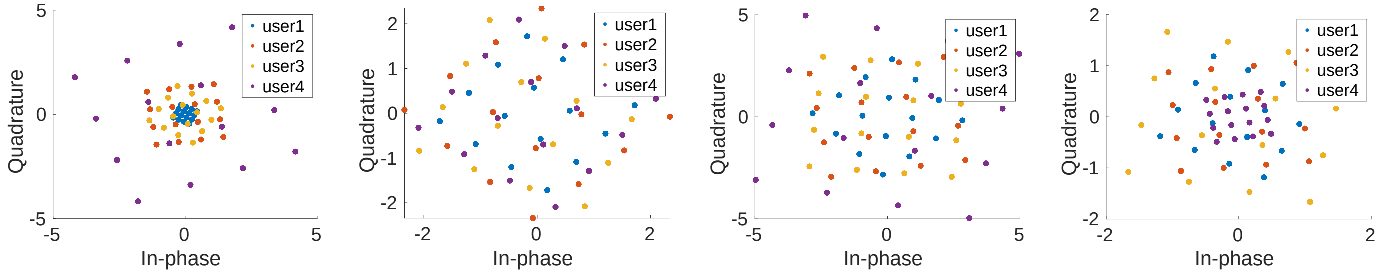
$$\begin{aligned} \arg\max_{\mathbf{K}} \quad & f(\mathbf{K}) \\ \text{s.t.} \quad & tr(\mathbf{W}\mathbf{W}^H) \leq P \end{aligned} \tag{24}$$

IV. Precoding order optimization for DPC

A. Optimum precoding order

DPC treats each user as one layer and iteratively precodes on previous layers by treating the interference from previous layers as *dirty*. This process is widely termed as *writing on dirty paper*. Since each user channel is different, the order of these layers (users) affects the precoded signal [13].

Figure 2 shows the Average Power (AP) and Peak-to-Average Power Ratio (PAPR) of DPC with 4 users for different precoding orders. The total number of precoding orders is $4!{=}24$. We observe that AP and PAPR are varies with each precoding order. Figure 3 shows the constellation diagram for DPC with minima AP, maximal AP, minimal PAPR and



(a) Maximal PAPR order, $m = 6$ (b) Minimal PAPR order, $m = 4$ (c) Maximal AP order, $m = 13$ (d) Minimal AP order, $m = 16$

Figure 3: 16-QAM constellation of 4-user DPC with different precoding order, $\pi(m)$ and power constraints

maximal PAPR precoding orders, where the scatter plot of each layer is *writing* on previous layers. The gap between scatter plots of layers is largest in figure 3a and smallest in figure 3b, as the corresponding precoding order gives the maximal PAPR and minimal PAPR respectively. Figure 3c has the maximal boundary for scatter plots while figure 3d has the minimal boundary, which suggests maximal AP and minimal AP respectively.

The complexity of searching the optimal DPC order is known to be $O(N^3N!)$ [17]. The number of possible DPC precoding order is $N!$ for a N-user case. Now, for each order, DPC is repeated and then the precoded signal is compared based on the given constraint criteria to find the optimal order, which is extremely expensive computationally.

However, from Theorem 1 and Figure 1, we find that the precoded matrix consists of the components from the SVD and LQ decomposition only. The relation of precoded matrix W and the permutation of channel matrix H is given by Lemma1.

Lemma 1. *Let $\mathbf{H}_{\pi(m)}$ be the permutation, by order $\pi(m)$ of a given channel matrix with SVD, $\mathbf{H}{=}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. Then the SVD of $\mathbf{H}_{\pi(m)}$ is given by permuting $\mathbf{U}$ by the same order $\pi(m)$,*

$$\mathbf{H}_{\pi(m)} = \mathbf{U}_{\pi(m)}\mathbf{\Sigma}\mathbf{V}^H \tag{25}$$

Proof. Given a matrix $\mathbf{G}$ consisting of basis vectors $\{\mathbf{e}_n\}$,

$$\mathbf{G} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_N \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \tag{26}$$

if $\{\mathbf{e}_n\}$ is rearranged by a given order $\pi(m)$, we get the permutation operator $\mathbf{G}_{\pi(m)}$. Then permutation of a matrix by order $\pi(m)$ can seen as the multiplication by $\mathbf{G}_{\pi(m)}$. Thus the permuted matrix $\mathbf{H}_{\pi(m)}$ is given by

$$\mathbf{H}_{\pi(m)} = \mathbf{G}_{\pi(m)}\mathbf{H} = \mathbf{G}_{\pi(m)}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{27}$$

Note that $(\mathbf{G}_{\pi(m)}\mathbf{U})(\mathbf{G}_{\pi(m)}\mathbf{U})^H{=}\mathbf{G}_{\pi(m)}\mathbf{U}\mathbf{U}^H\mathbf{G}_{\pi(m)}^H = \mathbf{G}_{\pi(m)}\mathbf{I}\mathbf{G}_{\pi(m)}^H = \mathbf{I}$, implying $(\mathbf{G}_{\pi(m)}\mathbf{U})$ is an unitary matrix. Let $\mathbf{U}_{\pi(m)} = \mathbf{G}_{\pi(m)}\mathbf{U}$, then (27) can be rewritten as

$$\mathbf{H}_{\pi(m)} = \mathbf{U}_{\pi(m)}\mathbf{\Sigma}\mathbf{V}^H \tag{28}$$

□

B. Permutation on SVD and LQ decomposition

Lemma1 shows that the SVD of the permuted matrix $\mathbf{H}_{\pi(m)}$ can be represented by the linear combination of permutation operator $\mathbf{G}_{\pi(m)}$ and SVD of $\mathbf{H}$, i.e.,

$$\text{SVD}(\mathbf{H}_{\pi(m)}) = \mathbf{G}_{\pi(m)}\text{SVD}(\mathbf{H}) \tag{29}$$

This shows the linearity of SVD with respect to permutation. However, it is not the same for LQ decomposition, as the triangular matrix $\mathbf{L}$ is unable to maintain its triangular structure after permutation and since, $\mathbf{G}_{\pi(m)}\mathbf{L}$ is not a triangular matrix the LQ decomposition of a permuted matrix can not be obtained by permuting the decomposed component. Therefore,

$$\text{LQ}(\mathbf{H}_{\pi(m)}) \neq \mathbf{G}_{\pi(m)}\text{LQ}(\mathbf{H}) \tag{30}$$

(30) shows the non-linearity of LQ decomposition with respect to permutation, which requires conventional DPC to repeat LQ decomposition for each permutation of channel matrix $\mathbf{H}$ to find the optimal order for precoding .

From Lemma 1, we find that permutation of $\mathbf{H}$ only changes the order of the elements of $\mathbf{U}$. Therefore, using (23), the precoding matrix under permutation, $\mathbf{W_m}$ is,

$$\mathbf{W}_m = \mathbf{V}\Sigma^{-1}\mathbf{U}_{\pi(m)}^H\mathbf{K} \tag{31}$$

Thus the optimal order can be obtained by an one-time DPC for an arbitrary order using 2-D HOGMT and then comparing each order by permuting the unitary matrix $\mathbf{U}$ and data signal $\mathbf{s}$. Thus we have the Theorem 2.

Theorem 2. *The precoding order for DPC with effective channel gains $\mathbf{K}$ is obtained by permuting the diagonal elements of $\mathbf{K}$.*

Proof. The DPC precoding order is optimized as follows,

$$\begin{aligned} \underset{\{\pi(m)\}}{\arg\min} \quad & g(\mathbf{x}_m) \\ \text{s.t.} \quad & \mathbf{H}_{\pi(m)}\mathbf{W}_m = \mathbf{K}_{\pi(m)} \end{aligned} \tag{32}$$

where $\mathbf{x}_m = \mathbf{W}_m\mathbf{s}_{\pi(m)}$, and $\mathbf{s}_{\pi(m)} = \mathbf{G}_{\pi(m)}\mathbf{s}$ is the permutation of data signal $\mathbf{s}$ by order $\pi(m)$. $g(\cdot)$ is the objective function according to the given criteria such as minimal AP, minimal PAPR, etc.

Remark 1: (Diagonal permutation) Given a permutation operator $\mathbf{G}_{\pi(m)}$, which permutes the rows of matrix by the order $\pi(m)$, for a diagonal matrix $\mathbf{D}$, permuting the diagonal elements on the diagonal direction by the order $\pi(m)$ is,

$$\mathbf{D}_{\pi(m)} = (\mathbf{G}_{\pi(m)}^H \mathbf{D} \mathbf{G}_{\pi(m)}) \tag{33}$$

Then the precoded signal is given by,

$$\begin{aligned} \mathbf{x}_m &= \mathbf{W}_m \mathbf{S}_{\pi(m)} \\ &= \mathbf{V}\Sigma^{-1}(\mathbf{G}_{\pi(m)}\mathbf{U})^H \mathbf{K}\mathbf{G}_{\pi(m)}\mathbf{s} \\ &= \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H \underbrace{(\mathbf{G}_{\pi(m)}^H \mathbf{K}\mathbf{G}_{\pi(m)})}_{\text{Diagonal permutation}} \mathbf{s} \\ &= \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H \mathbf{K}_{\pi(m)}\mathbf{s} \end{aligned} \tag{34}$$

where $\mathbf{K}_{\pi(m)}$ is a diagonal matrix having the same elements of $\mathbf{K}$ with diagonal entries ordered by $\pi(m)$. □

Theorem 2 shows that the solution of optimal precoding order with respect to arbitrary objective function $g(\cdot)$ can be obtained by looping over all precoding orders, where for each precoding order, the proposed method can avoid repeating the decomposition by simply permuting a diagonal matrix.

C. Convergence of beamforming and precoding orders for the same strategy

Specifically, if the criteria of the beamforming is the optimal power allocation, the beamforming solution already achieves minimal power precoding order as shown in Corollary 1.

Corollary 1. *The optimal power allocation strategy achieves minimal power precoding order.*

Proof. In Theorem 1, given power constraint $\text{tr}(\mathbf{W}\mathbf{W}^H) \leq P$, the precoding matrix $\mathbf{W}$ with optimal power allocation beamforming is obtained by

$$\mathbf{W} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H \mathbf{K} \tag{35}$$

where, $\mathbf{K} = \text{diag}(k_1, \ldots, k_N)$, and k_n is the effective gain for user n and can be designed by water-filling algorithm as

$$k_n = \sqrt{p_n \lambda_n}, \quad \text{where}, \quad p_n = \left(\mu - \frac{1}{\lambda_n}\right)^+ \tag{36}$$

where μ is a constant to ensure power constraint, and $\lambda_n = \sigma_n^2$ is n^{th} eigenvalue where σ_n is the n^{th} diagonal element of Σ. $(x)^+$ is defined as $\max(x, 0)$.

On the other hand, to find the optimal precoding order with respect to minimal power, set $g(\mathbf{x}){=}\mathbb{E}\{|\mathbf{x}|^2\}$, then (32),

$$\begin{aligned} \underset{\{\pi(m)\}}{\arg\min} \quad & \mathbb{E}\{|\mathbf{x}_m|^2\} \\ \text{s.t.} \quad & \mathbf{H}_{\pi(m)}\mathbf{W}_m = \mathbf{K}_{\pi_m} \end{aligned} \tag{37}$$

Substituting (34) in (37), we have

$$\underset{\{\pi(m)\}}{\arg\min} \quad \sum_n^N \frac{k_{n,\pi(m)}^2}{\lambda_n} \tag{38}$$

where, $k_{n,\pi(m)}$ is n^{th} diagonal element of $\mathbf{K}_{\pi(m)}$. Thus, the optimal order in (37) is obtained by simply permuting $\mathbf{K}$.

(38) suggests that the order $\{\pi(m)\}$ ensures $\{k_{n,\pi(m)}\}$ has the same magnitude order as $\{\lambda_n\}$, achieves the optimal solution. Meanwhile, as k_n in the original $\mathbf{K}$ given by (36) has the positive relation with λ_n, it is ranked by the same order as λ_n. which is the solution of the optimization (38). Thus the original order of k_n is already optimal. □

The beamforming optimization with optimal power allocation strategy target to maximize the energy efficiency. The precoding order optimization with the minimal power criteria target to minimize the signal power with the effective gain unchanged as in (37), which also maximize the energy efficiency. They are reasonable to converge to the same solution. Thus the equivalence shown in lemma 2 validate the correctness of the proposed precoding order optimization.

D. Low-complexity DPC with optimum ordering

The procedure to implement the proposed DPC with precoding order optimization is given in Algorithm 1. First, decompose the CSI by SVD as $\mathbf{H}{=}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ using Theorem 1 and then constitute $\mathbf{W}$ by the decomposed components with the channel effective gain $\mathbf{K}$ as in (23), where $\mathbf{K}$ is obtained by solving (24). To find the optimal precoding order, we collect the all permutation orders $\{\pi(m)\}_{m=1}^M$, where for N users, there are $M{=}N!$ precoding orders. For the precoding order $\pi(m)$, the corresponding precoded signal $\mathbf{x}_m{=}\mathbf{W}_m\mathbf{s}_{\pi(m)}$ is obtained from (34), which only diagonally permutes $\mathbf{K}$ by order $\pi(m)$ as $\mathbf{K}_{\pi(\mathbf{m})} = \mathbf{G}_{\pi(\mathbf{m})}^{\mathbf{H}}\mathbf{K}\mathbf{G}_{\pi(\mathbf{m})}$. The precoded signal with optimal precoding order is obtained by looping over all orders and comparing the corresponding precoded signal by the decision function $g(\cdot)$. In practice, $g(\cdot)$ is based on desired criteria such as minimal AP or minimal PAPR.

Algorithm 1: Low-complexity DPC with optimum ordering

Input: Data $\mathbf{s}$, CSI $\mathbf{H}$, beamforming optimization function $f(\cdot)$, power limitation P, decision function $g(\cdot)$;
Output: Precoded signal $\mathbf{x}$;
Decompose $\mathbf{H}$ by SVD decomposition $[\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \text{SVD}(\mathbf{H})$;
Solve $\arg\max_{\mathbf{K}} f(\mathbf{K})$, under the constraint $\text{tr}(\mathbf{W}\mathbf{W}^H) \leq P$, where $\mathbf{W} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{K}$ to get $\mathbf{K}$;
Collect diagonal permutation orders $\{\pi(\cdot)\}$ of $\mathbf{K}$;
Initialize the order index $m = 1$;
Permute $\mathbf{K}$ by order $\pi(m)$ to get $\mathbf{K}_{\pi(m)}$;
Initialize $\mathbf{x} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{K}_{\pi(m)}$;
while $m < \textit{size}(\{\pi(\cdot)\})$ **do**
    $m = m + 1$;
    Update $\mathbf{K}_{\pi(m)}$ by permuting $\mathbf{K}$ by order $\pi(m)$;
    $\mathbf{x}_m = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{K}_{\pi(m)}$;
    Update $\mathbf{x}$ by decision function $\mathbf{x} = g(\mathbf{x}_m, \mathbf{x})$;
end

V. Results

Conventional DPC implementation needs to repeat DPC for each precoding order, whose complexity is $\mathcal{O}(N^3 N!)$ for N users case. In contrast the proposed algorithm requires

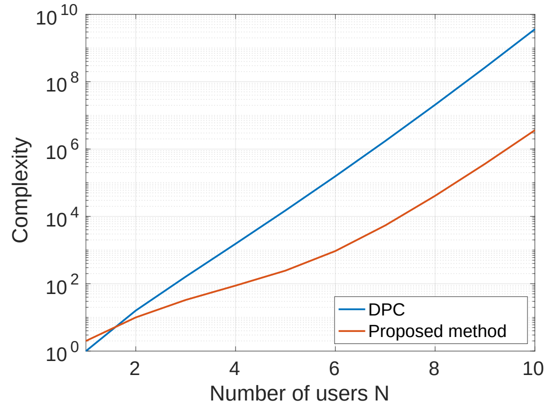


Figure 4: Complexity of conventional DPC and Algorithm 1

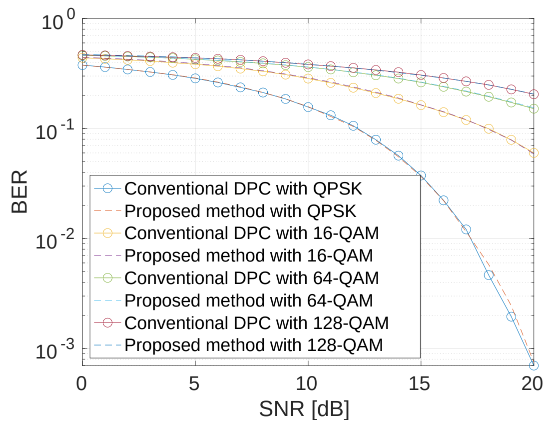


Figure 5: Equivalence of conventional DPC and Algorithm 1 for QPSK, 16-QAM, 64-QAM and 128-QAM modulation

a one-time DPC and then search for the optimal precoding order by just permuting the channel gain matrix that has a computational complexity of $\mathcal{O}(N^3+N!)$. Figure 4 shows that for multi-user cases, i.e., $N \geq 2$, the proposed methods always achieve lower complexity than conventional DPC. For users number $N = 10$, the complexity ratio of conventional DPC over proposed method is $\approx$30 dB, which is very encouraging.

We validate the equivalence of the proposed methods and conventional DPC with perfect CSI at the transmitter using MATLAB simulations. The number of the transmitter antennas and the users are both N=10, where each user is equipped with one antenna. The coefficient of the channel matrix is generated by standard Gaussian distribution. The effective channel gain of both methods are normalized. Figure 5 compares the BER of conventional DPC and proposed method for QPSK, 16-QAM, 64-QAM and 128-QAM schemes. It is evident that the proposed method achieves the same result as conventional DPC, supporting the theoretical equivalence discussed earlier.

VI. Conclusion

In this paper, we show the equivalence of DPC and 2-D HOGMT precoding for MU-MIMO channels and give an alternate low-complexity implementation based on SVD decomposition to replace the iterative method based on Gram-Schmidt processes and LQ decomposition. Then we show the difference between SVD decomposition and LQ decomposition with respect to permutation, where an unitary matrix after permutation is still unitary but a triangular matrix cannot maintain its structure under the same permutation. This difference suggests that conventional implementation needs to repeat DPC for different precoding orders while the proposed method just needs one-time DPC and then search the optimal precoding order by permuting a component of the precoding matrix. Simulations show that our method is able to achieve the same BER performance as DPC but with less complexity. For $N = 10$ users case, the proposed method achieves near 30 dB lower complexity.

References

[1] Z. Zou, M. Careem, A. Dutta, and N. Thawdar, "Unified characterization and precoding for non-stationary channels," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 5140–5146.

[2] M. Vu and A. Paulraj, "MIMO Wireless Linear Precoding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 86–105, 2007.

[3] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[4] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley Publishing, 2010.

[5] Y. Ma, A. Yamani, N. Yi, and R. Tafazolli, "Low-complexity mu-mimo nonlinear precoding using degree-2 sparse vector perturbation," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 497–509, 2016.

[6] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[7] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[8] J. Lee and N. Jindal, "High SNR Analysis for MIMO Broadcast Channels: Dirty Paper Coding Versus Linear Precoding," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4787–4792, 2007.

[9] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.

[10] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, 1998.

[11] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual sinr constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, 2004.

[12] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2324–2334, 2005.

[13] T. Michel and G. Wunder, "Transmitter and precoding order optimization for nonlinear downlink beamforming," 2007. [Online]. Available: https://arxiv.org/abs/cs/0702122

[14] P. Almers, E. Bonek, A. Burr, N. Czink, m. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz, A. Molisch, C. Oestges, and H. Özcelik, "Survey of Channel and Radio Propagation Models for Wireless MIMO Systems." *EURASIP J. Wireless Comm. and Networking*, vol. 2007, 01 2007.

[15] G. Matz, "On non-wssus wireless fading channels," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2465–2478, 2005.

[16] F. Hlawatsch and G. Matz, *Wireless Communications Over Rapidly Time-Varying Channels*, 1st ed. USA: Academic Press, Inc., 2011.

[17] Y. Mao and B. Clerckx, "Beyond Dirty Paper Coding for Multi-Antenna Broadcast Channel With Partial CSIT: A Rate-Splitting Approach," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6775–6791, 2020.