COOPERATIVE MULTI-CELL MIMO DOWNLINK PRECODING FOR FINITE-ALPHABET INPUTS

Wenhao Wu* *Kun Wang** *Zhi Ding** *Chengshan Xiao*†

*Department of Electrical and Computer Engineering,
University of California, Davis, CA 95616

†Department of Electrical and Computer Engineering,
Missouri University of Science and Technology, Rolla, MO 65409

*{wnhwu,kunwang,zding}@ucdavis.edu, †xiaoc@mst.edu

ABSTRACT

This work studies the design of linear precoders for cooperative multi-cell MIMO downlink systems with finite alphabet inputs. Traditionally, multi-cell MIMO downlink precoder designs rely on Gaussian input assumption, which may lead to performance loss when the true inputs admit discrete non-Gaussian symbols. In this work, we present optimized precoders by maximizing weighted sum rate (of finite-alphabet-input) under a set of per-base station power constraints. Specifically, we propose a simple gradient algorithm for general multi-cell MIMO downlink channel and a block diagonalization gradient algorithm while supporting interference cancellation.

***Index Terms*—** Multi-cell downlink , linear precoding, finite-alphabet input, gradient descent, dual decomposition.

1. INTRODUCTION

Cooperative transmission in multi-cell MIMO downlink channel (MC-MIMO-DLC) has been a topic that attracted growing interest in recent years, as a promising technology for cellular interference management. Among various approaches, cooperative precoding has shown substantial rate improvement when given channel state information (CSI) and when data streams are available to all base stations (BSs) [1]. An MC-MIMO-DLC resembles MIMO broadcasting channel (BC) except for the stricter per-BS (or per-antenna) power constraints (rather than the sum-power constraint). Although the uplink-downlink duality and capacity-reaching dirty paper coding (DPC) scheme can be generalized [2, 3], their implementation difficulty has made suboptimal precoder design such as linear precoders more favorable.

While existing linear precoder design approaches for MC-MIMO-DLC are mostly based on Gaussian input assumption, practical communication systems have input signals from finite-alphabet constellation. This difference leads to considerable gap between the actual rate and the capacity. Such mismatch can often degrade the performance of Gaussian input based methods. Noting this key discrepancy and mismatch, our linear precoder designs for MC-MIMO-DLC shall specifically target finite-alphabet inputs.

There exist several linear precoder design algorithms for maximizing weighted sum rate (WSR) in MC-MIMO-DLC. Due to the non-convexity of the problem formulation, most of them rely on numerical alternating methods, including mean square error (MSE) receiver/weighting matrix update - precoder optimization [4, 5], and interference/leakage update - precoder optimization [6]. When block diagonalization (BD) constraint can be enforced, the MC-MIMO-DLC becomes simpler and the WSR maximization problem with per-BS power constraints becomes, convex. A dual-decomposition based, cooperative, multi-cell BD precoder design method is proposed in [7] and two other suboptimal solutions are provided in [8]. As a special case, beamforming of MC-MIMO-DLC can be found in [9, 10].

Although the aforementioned studies are all based on Gaussian input assumption, linear precoder optimization for finite alphabet inputs has been investigated for several scenarios including single user MIMO [11, 12], MIMO BC [13], MIMO multiple access channels (MAC) [14], MIMO relay channels [15, 16], MIMO hybrid-ARQ (HARQ) [17], as well as MIMO wiretap channels [18]. Each case demonstrates performance gain over the laissez faire use of their Gaussian based precoder design counterpart.

To the best of our knowledge, there has been scant research on precoder design for MC-MIMO-DLC under finite-alphabet inputs. In this work, we examine the linear finite-alphabet precoder design problem in MC-MIMO-DLC under per-BS power constraint. We first derive a simple gradient-descent precoder optimization algorithm. To avoid the time-consuming Monte-Carlo numerical evaluation, we generalize the rate approximation and its gradient in [19] to MIMO-BC model [13] while incorporating the per-BS constraint. Next, we include the BD restriction to our precoder design scheme and subsequently decompose the WSR objective into single user rate objectives. We enable a nontrivial extension of the alternating optimization algorithm in [12]. This is achieved by applying the dual-decomposition technique [7, 20] to the newly formulated convex power allocation subproblems un-

This material is based on works supported by the National Science Foundation Grants CNS-1147930, ECCS-1307820 and CIF-1321143.

der per-BS power constraints. Our numerical test results show that both our proposed design algorithms outperforms the Gaussian based designs, particularly in the medium SNR regime.

2. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a cellular network consisting of A cells, each having a BS with N_t Tx antennas. The a-th cell has K_a mobile stations (MSs) each equipped with N_r Rx antennas. The total number of MSs is $K = \sum_{a=1}^{A} K_a$ and the total number of TX antennas equals $N = AN_t$. The received signal at the k-th MS is given by

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \sum_{i\neq k}\mathbf{H}_k\mathbf{x}_i + \mathbf{z}_k,\ k=1,\ldots,K \tag{1}$$

in which $\mathbf{x}_k \in \mathbb{C}^{N_t\times 1}$ is the transmitted signal vector intended for the k-th MS. $\mathbf{H}_k \in \mathbb{C}^{N_r\times N}$ is the normalized channel matrix from the virtual BS formed by the set of cell BSs to the k-th MS with $\mathrm{tr}\left(\mathbf{H}_k\mathbf{H}_k^H\right) = N_r$. $\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0},\sigma^2\mathbf{I})$ is receiver noise. We assume a full cooperative scheme where the channel state information (CSI) is perfectly known to all BSs and MSs involved.

The linear precoding is represented as

$$\mathbf{x}_k = \mathbf{P}_k\mathbf{s}_k,\ k=1,\ldots,K \tag{2}$$

where $\mathbf{P}_k \in \mathbb{C}^{N\times N_r}$ denotes the precoder and $\mathbf{s}_k \in \mathbb{C}^{N_r\times 1}$ denotes the finite-alphabet data vector for the k-th MS. Elements in $\mathbf{s}_k$ is independent and uniformly distributed from a constellation of size M, with zero mean and unit variance.

Throughout this paper, we strive to maximize the weighted sum rate as a function of the precoders

$$R(\{\mathbf{P}_k\}) = \sum_{k=1}^{K}\mu_k R_k(\{\mathbf{P}_k\}) \tag{3}$$

where the k-th MS achievable rate $R_k(\{\mathbf{P}_k\})$ is the same as its counterpart in MIMO-BC given by *Proposition 1* in [13]. Let N_t be the normalized maximum transmit power of each BS. The per-BS power constraints are given by

$$\sum_{k=1}^{K}\|\mathbf{B}_a\mathbf{P}_k\|_F^2 = \sum_{k=1}^{K}\mathrm{tr}\left(\mathbf{B}_a\mathbf{P}_k\mathbf{P}_k^H\right) \leq N_t,\ a=1,\ldots,A \tag{4}$$

with Frobenius norm $\|\cdot\|_F$ while $\mathbf{B}_a$ is defined in [7] as

$$\mathbf{B}_a \coloneqq \mathrm{Diag}(\underbrace{0,\ldots,0}_{(a-1)N_t},\underbrace{1,\ldots,1}_{N_t},\underbrace{0,\ldots,0}_{(A-a)N_t}). \tag{5}$$

The challenges from this new formulation beyond existing works are twofold. First, this constrained optimization problem is non-convex, which compels the search for sub-optimal solutions. Second, evaluating the finite-alphabet rate $R_k(\{\mathbf{P}_k\})$ in Eq. (3) requires time-consuming Monte-Carlo statistical sampling for non-trivial model settings. In the following section, we will address these two challenges in the development of our algorithms.

3. PRECODER DESIGN

3.1. Simple Gradient-Descent Precoder (S-GDP)

Gradient descent method is a common way for finding local optima [21]. Our obstacle lies in the need to repeatedly evaluate the rate (for finite-alphabet inputs) and its gradient using computationally costly Monte-Carlo. This is particularly costly for MC-MIMO-DLC as the interferences are also finite-alphabetical. To tackle this problem, we generalize the approximation for non-interference channel in [19] to the interference channel in our problem.

Denote constellation vector index set $\mathbb{I} = \{1,\ldots,M\}$ and its K-ary and $(K-1)$-ary Cartesian products as $\mathbb{A} = \{(p_1,\ldots,p_K)|p_i \in \mathbb{I}\}$ and $\mathbb{A}_k = \{(p_1,\ldots,p_K)|p_i \in \mathbb{I}, i\neq k\}$, respectively. Then the approximation to $R_k(\{\mathbf{P}_k\})$ is given by

$$\hat{R}_k(\{\mathbf{P}_k\}) = \log_2 M + \frac{1}{M^{K-1}}\sum_{\mathbf{m}_k\in\mathbb{A}_k}\log_2\sum_{\mathbf{n}_k\in\mathbb{A}_k}u_{\mathbf{m}_k\mathbf{n}_k} - \frac{1}{M^K}\sum_{\mathbf{m}\in\mathbb{A}}\log_2\sum_{\mathbf{n}\in\mathbb{A}}v_{\mathbf{mn}} \tag{6}$$

in which $u_{\mathbf{m}_k\mathbf{n}_k} = \exp(-\mathbf{c}_{\mathbf{m}_k\mathbf{n}_k}^H\mathbf{c}_{\mathbf{m}_k\mathbf{n}_k}/2\sigma^2)$ and $v_{\mathbf{mn}} = \exp(-\mathbf{d}_{\mathbf{mn}}^H\mathbf{d}_{\mathbf{mn}}/2\sigma^2)$, where $\mathbf{c}_{\mathbf{m}_k\mathbf{n}_k} = \mathbf{H}_k\sum_{j\neq k}\mathbf{P}_j\mathbf{e}_{m_j,n_j}$ and $\mathbf{d}_{\mathbf{mn}} = \mathbf{H}_k\sum\mathbf{P}_j\mathbf{e}_{m_jn_j}$. Here m_j, n_j are the entries of index vectors $\mathbf{m}_k/\mathbf{m}$ and $\mathbf{n}_k/\mathbf{n}$, and $\mathbf{e}_{m_jn_j} = \mathbf{s}^{(m_j)} - \mathbf{s}^{(n_j)}$ denotes the difference between the m_j-th and n_j-th constellation vectors.

The gradient of $\hat{R}_k(\{\mathbf{P}_k\})$ with respect to (w.r.t.) precoder $\mathbf{P}_i$, according to [22], is given by

$$\nabla_{\mathbf{P}_i}\hat{R}_k|_{i\neq k} = -\frac{\mathbf{H}_k^H\mathbf{H}_k}{2\ln 2\sigma^2}\left(\sum_{\mathbf{m}_k\in\mathbb{A}_k}\frac{\sum_{\mathbf{n}_k\in\mathbb{A}_k}u_{\mathbf{m}_k\mathbf{n}_k}\left(\sum_{j\neq k}\mathbf{P}_j\mathbf{e}_{m_jn_j}\right)\mathbf{e}_{m_in_i}^H}{M^{K-1}\sum_{\mathbf{n}_k\in\mathbb{A}_k}u_{\mathbf{m}_k\mathbf{n}_k}} - \sum_{\mathbf{m}\in\mathbb{A}}\frac{\sum_{\mathbf{n}\in\mathbb{A}}v_{\mathbf{mn}}\left(\sum\mathbf{P}_j\mathbf{e}_{m_jn_j}\right)\mathbf{e}_{m_in_i}^H}{M^K\sum_{\mathbf{n}\in\mathbb{A}}v_{\mathbf{mn}}}\right) \tag{7}$$

$$\nabla_{\mathbf{P}_k}\hat{R}_k = -\frac{\mathbf{H}_k^H\mathbf{H}_k}{2\ln 2\sigma^2}\sum_{\mathbf{m}\in\mathbb{A}}\frac{\sum_{\mathbf{n}\in\mathbb{A}}v_{\mathbf{mn}}\left(\sum\mathbf{P}_j\mathbf{e}_{m_jn_j}\right)\mathbf{e}_{m_kn_k}^H}{M^K\sum_{\mathbf{n}\in\mathbb{A}}v_{\mathbf{mn}}}. \tag{8}$$

To enforce the per-BS power constraints during gradient descent, we project the updated precoders to a feasible solution in each iteration. [23] discussed several projection methods. Here we use a straightforward projection $\mathrm{Proj}(\mathbf{P}_k) = \mathbf{TP}_k$ where $\mathbf{T}$ is a N-by-N block-diagonal matrix with its a-th N_t-by-N_t block-diagonal element as:

$$\mathbf{T}_a = \begin{cases}\mathbf{I}_{N_t} & \text{if } \sum_{k=1}^{K}\|\mathbf{B}_a\mathbf{P}_k\|_F^2 \leq N_t;\\ \frac{\sqrt{P}\mathbf{I}_{N_{BS}}}{\sum_{k=1}^{K}\|\mathbf{B}_a\mathbf{P}_k\|_F^2} & \text{otherwise.}\end{cases} \tag{9}$$

Now we are ready to summarize the steps of our Simple Gradient-Descent Precoder (S-GDP) optimization algorithm:

S1 Select a feasible set of initial precoders $\{\mathbf{P}_k\}$.
S2 Evaluate the approximated weighted sum rate (AWSR) $\hat{R}(\{\mathbf{P}_k\}) = \sum_{k=1}^{K} \mu_k \hat{R}_k(\{\mathbf{P}_k\})$ and its gradients $\nabla_{\mathbf{P}_i} \hat{R}(\{\mathbf{P}_k\}) = \sum_{k=1}^{K} \mu_k \nabla_{\mathbf{P}_i} \hat{R}_k(\{\mathbf{P}_k\})$ for $i = 1, \ldots, K$. Determine the step size t with backtracking line search method [21].
S3 For $k = 1, \ldots, K$, evaluate $\mathbf{P}_k^* = \mathbf{P}_k + t\nabla_{\mathbf{P}_k} \hat{R}(\{\mathbf{P}_k\})$. Then get the updated precoders by $\mathbf{P}_k \leftarrow \mathbf{T}\mathbf{P}_k^*$.
S4 Return to S2 until convergence.

3.2. Block Diagonalization - Alternating Gradient-Descent Precoder (B-GDP)

Although the S-GDP algorithm is straightforward, the interference term makes $\hat{R}(\{\mathbf{P}_k\})$ highly non-convex. Thus, we need good initial precoders to achieve good performance. Each MS also needs to decode the finite-alphabet interference for rate benefit. Moreover, computational complexity required for evaluating $\hat{R}_k$ and $\nabla_{\mathbf{P}_k} \hat{R}_k$ grows rapidly following $\mathcal{O}(K^2 N N_r (M^{2K} + N))$ w.r.t K.. However, when $N \geq K N_r$, it is possible to focus on block-diagonalization (BD) precoding scheme that is substantially simpler and performs well in high SNR regime. BD achieves some convex property and reduces computational complexity.

Denote $\mathbf{G}_k = [\mathbf{H}_1^T, \ldots, \mathbf{H}_{k-1}^T, \mathbf{H}_{k+1}^T, \ldots, \mathbf{H}_K^T]^T$ whose null space $\bar{\mathbf{V}}_{\mathbf{G}_k}$ can be derived from the singular vector decomposition (SVD) as $\mathbf{G}_k = \mathbf{U}_{\mathbf{G}_k} \mathbf{\Sigma}_{\mathbf{G}_k} [\mathbf{V}_{\mathbf{G}_k}, \bar{\mathbf{V}}_{\mathbf{G}_k}]^H$. By restricting $\mathbf{P}_k = \bar{\mathbf{V}}_{\mathbf{G}_k} \bar{\mathbf{P}}_k$, the interference-free channel for the k-th MS is $\bar{\mathbf{H}}_k = \mathbf{H}_k \bar{\mathbf{V}}_{\mathbf{G}_k}$. Denote the SVD of $\bar{\mathbf{P}}_k$ as $\bar{\mathbf{P}}_k = \mathbf{U}_{\bar{\mathbf{P}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k} \mathbf{V}_{\bar{\mathbf{P}}_k}^H$. Although *Proposition 2* in [12] is no longer valid due to per-BS power constraints, we can still get the following equivalent channel if we align $\mathbf{U}_{\bar{\mathbf{P}}_k}$ to the right singular vectors of $\bar{\mathbf{H}}_k$, as if the inputs were Gaussian or we were considering a point-to-point MIMO system with finite alphabet inputs [24]:

$$\bar{\mathbf{y}}_k = \mathbf{\Sigma}_{\bar{\mathbf{H}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k} \mathbf{V}_{\bar{\mathbf{P}}_k}^H \mathbf{s}_k + \bar{\mathbf{z}}_k. \tag{10}$$

The approximated rate for the k-th MS is [19]

$$\bar{R}_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2, \mathbf{V}_{\bar{\mathbf{P}}_k}) = \log_2 M - \frac{1}{M} \sum_{m=1}^{M} \log_2 \sum_{n=1}^{M} w_{mn} \tag{11}$$

where $w_{mn} = \exp(\mathbf{e}_{mn}^H \mathbf{W}_k \mathbf{e}_{mn} / 2\sigma^2)$, $\mathbf{e}_{mn} = \mathbf{s}^{(m)} - \mathbf{s}^{(n)}$ and $\mathbf{W}_k = \mathbf{V}_{\bar{\mathbf{P}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 \mathbf{\Sigma}_{\bar{\mathbf{H}}_k}^2 \mathbf{V}_{\bar{\mathbf{P}}_k}^H$. Its gradient w.r.t $\mathbf{W}_k$ is

$$\nabla_{\mathbf{W}_k} \bar{R}_k = \sum_{m=1}^{M} \frac{\sum_{n=1}^{M} w_{mn} (\mathbf{e}_{mn} \mathbf{e}_{mn}^H)}{2 \ln 2\sigma^2 M \sum_{n=1}^{M} w_{mn}}. \tag{12}$$

Similar to [12], we attempt to maximize the AWSR by alternatingly solving

$$\max \bar{R}_k(\mathbf{V}_{\bar{\mathbf{P}}_k} | \mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2), k = 1, \ldots, K; \tag{13}$$

$$\max \sum\nolimits_{k=1}^{K} \mu_k \bar{R}_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 | \mathbf{V}_{\bar{\mathbf{P}}_k}), \tag{14}$$

$$\text{s.t. } \sum\nolimits_{k=1}^{K} \mathrm{tr}(\mathbf{B}_a \bar{\mathbf{V}}_{\mathbf{G}_k} \mathbf{V}_{\bar{\mathbf{H}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 \mathbf{V}_{\bar{\mathbf{H}}_k}^H \bar{\mathbf{V}}_{\mathbf{G}_k}^H) \leq N_t.$$

Problem (13) can be solved with a Stiefel manifold based method [16] in which the line-search direction is

$$\Delta \mathbf{V}_{\bar{\mathbf{P}}_k} = \nabla_{\mathbf{V}_{\bar{\mathbf{P}}_k}} \bar{R}_k - \mathbf{V}_{\bar{\mathbf{P}}_k} (\nabla_{\mathbf{V}_{\bar{\mathbf{P}}_k}} \bar{R}_k)^H \mathbf{V}_{\bar{\mathbf{P}}_k}, \tag{15}$$

$$\nabla_{\mathbf{V}_{\bar{\mathbf{P}}_k}} \bar{R}_k = (\nabla_{\mathbf{W}_k} \bar{R}_k)^H \mathbf{V}_{\bar{\mathbf{P}}_k} \Sigma_{\bar{\mathbf{P}}_k}^2 \Sigma_{\bar{\mathbf{H}}_k}^2. \tag{16}$$

The convex problem (14) can be solved by dual decomposition [7, 20]. Denote $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_A)^T \succeq \mathbf{0}$ as the dual variables. We alternatingly solve the K individual subproblems $\max f_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 | \boldsymbol{\gamma})$, $k = 1, \ldots, K$, where

$$\begin{aligned} f_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2, \boldsymbol{\gamma}) = {} & \mu_k \bar{R}_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 | \mathbf{V}_{\bar{\mathbf{P}}_k}) \\ & - \mathrm{tr}(\mathbf{B}_{\boldsymbol{\gamma}} \bar{\mathbf{V}}_{\mathbf{G}_k} \mathbf{V}_{\bar{\mathbf{H}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 \mathbf{V}_{\bar{\mathbf{H}}_k}^H \bar{\mathbf{V}}_{\mathbf{G}_k}^H) \end{aligned} \tag{17}$$

and $\mathbf{B}_{\boldsymbol{\gamma}} = \sum_{a=1}^{A} \gamma_a \mathbf{B}_a$ with the barrier function - gradient descent - backtracking line search method [16, 21] and solve the primal dual problem $\min g(\boldsymbol{\gamma})$ where

$$g(\boldsymbol{\gamma}) = \sum_{k=1}^{K} \max_{\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2} f_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2, \boldsymbol{\gamma}) + N_t \sum_{a=1}^{A} \gamma_a \tag{18}$$

with gradient descent method. The gradient of $f_k(\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 | \boldsymbol{\gamma})$ and the subgradient of $g(\boldsymbol{\gamma})$ are given, respectively, by

$$\begin{aligned} \nabla_{\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2} f_k = \mathrm{Diag}\Big(& \mathbf{\Sigma}_{\bar{\mathbf{H}}_k}^2 \mathbf{V}_{\bar{\mathbf{P}}_k}^H (\nabla_{\mathbf{W}_k} \bar{R}_k)^H \mathbf{V}_{\bar{\mathbf{P}}_k} \\ & - \mathbf{V}_{\bar{\mathbf{H}}_k}^H \bar{\mathbf{V}}_{\mathbf{G}_k}^H \mathbf{B}_\gamma \bar{\mathbf{V}}_{\mathbf{G}_k} \mathbf{V}_{\bar{\mathbf{H}}_k} \Big) \end{aligned} \tag{19}$$

$$\nabla_{\gamma_a} g = N_t - \sum_{k=1}^{K} \mathrm{tr}(\mathbf{B}_a \bar{\mathbf{V}}_{\mathbf{G}_k} \mathbf{V}_{\bar{\mathbf{H}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2 \mathbf{V}_{\bar{\mathbf{H}}_k}^H \bar{\mathbf{V}}_{\mathbf{G}_k}^H). \tag{20}$$

We now briefly summarize the steps of our BD Alternating Gradient-Descent Precoder (B-GDP) algorithm:

B1 Evaluate $\mathbf{\Sigma}_{\bar{\mathbf{H}}_k}$, $\mathbf{V}_{\bar{\mathbf{H}}_k}$. Choose initial $\boldsymbol{\gamma}^{(0)}$, $\{\mathbf{V}_{\bar{\mathbf{P}}_k}\}$ and $\{\mathbf{\Sigma}_{\bar{\mathbf{P}}_k}^2\}$.
B2 For $k = 1, \ldots, K$, update $\mathbf{\Sigma}_{\bar{\mathbf{H}}_k}^2$ with the barrier function - gradient descent - backtracking line search method. Here we use the logarithmic barrier function with parameter t [16].
B3 Update the dual variable $\boldsymbol{\gamma} \leftarrow (\boldsymbol{\gamma} - \delta \nabla_{\boldsymbol{\gamma}} g)^+$ where δ is a sufficiently small stepsize and $(x)^+ = \max(0, x)$.
B4 Return to B2 until $\{\mathbf{\Sigma}_{\bar{\mathbf{H}}_k}^2\}$ and $\boldsymbol{\gamma}$ converge.
B5 For $k = 1, \ldots, K$, update $\mathbf{V}_{\bar{\mathbf{P}}_k}$ with the Stiefel manifold - backtracking line search method [16]. The line search direction is given in Eq. (15).
B6 Return to B2 until $\{\mathbf{\Sigma}_{\bar{\mathbf{H}}_k}^2\}$, $\{\mathbf{V}_{\bar{\mathbf{P}}_k}\}$ converge.
B7 Return the BD precoders $\mathbf{P_k} = \bar{\mathbf{V}}_{\mathbf{G}_k} \mathbf{V}_{\bar{\mathbf{H}}_k} \mathbf{\Sigma}_{\bar{\mathbf{P}}_k} \mathbf{V}_{\bar{\mathbf{P}}_k}^H$, $k = 1, \ldots, K$.

4. SIMULATION TESTS

To quantify and demonstrate the performance of our precoder designed specifically for finite-alphabet inputs, we compare our proposed S-GDP and B-GDP with Gaussian based designs in [5] (GP1) and [7] (GP2), respectively. Our multi-cell MIMO channel entries are random independent circularly symmetric complex Gaussian (CSCG) variables, in which the variance of the inter-cell channels are ρ^2 times that of intra-cell channels. In all simulation tests we set $A = 2$, $N_r = 2$, $\rho = 0.6$ and $\mu_k = 1/K$. Our input signals range from BPSK, QPSK, to 8-PSK modulation.

In the first example we test SGDP against GP1 with $K_1 = K_2 = 1$ and $N_t = 2$. The initial precoders for GP1 are $\mathbf{P}_k = [\mathbf{0}_{N_t\times N_r}; \ldots; \mathbf{I}_{N_t\times N_r}; \ldots; \mathbf{0}_{N_t\times N_r}]$ with its a-th block as the tall identity matrix, a being the cell covering the k-th MS. We use the results of GP1 as the initial precoders for SGDP. The backtracking line search parameters [21] in S2 are given in Table 1. We illustrate the convergent behavior of GP1 + S-GDP cascade at SNR = 4dB in terms of AWSR/Monte-Carlo WSR in Fig. 1(a). The results show that GP1 converges after 30 iterations, at which point we switch to S-GDP method which converges to a higher WSR after approximately another 20 iterations. The curves of Gaussian/finite-alphabet input WSR of GP1 and the cascaded GP1 + S-GDP method versus SNR are plotted in Fig. 2. Apparently, our GP1 + S-GDP cascade uniformly offers a higher and more consistent WSR than GP1 for finite alphabet inputs tested.

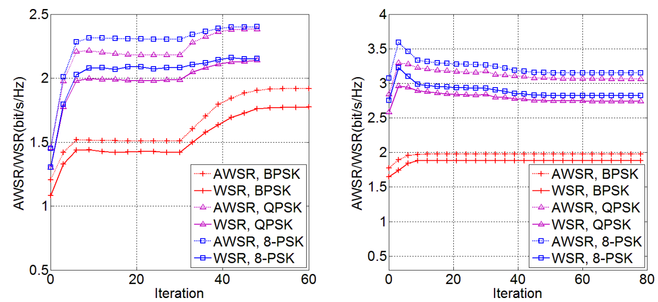
In the second example we test B-GDP against GP2 with $K_1 = K_2 = 2$ and $N_t = 4$. We initialize $\gamma_a = 0.1$, $\Sigma_{\bar{\mathbf{P}}_k} = \sqrt{AN_t/(KN_r)}\mathbf{I}$ and $\mathbf{V}_{\bar{\mathbf{P}}_k}$ as in *Example 1* of [12] with $\omega = \nu = \pi/15$. In B2 the barrier parameter is updated by $t \leftarrow 2t$ with $t_0 = 2$ and $t_{max} = 10^4$. In B3 we set $\delta = 7\times 10^{-4}$. The stopping criterion in B4 is $\|\Delta\gamma\|/\|\gamma\| <= 1\times 10^{-4}$. The backtracking line search parameters in B2 and B5 of B-GDP are also listed in Table 1. The convergence behavior at SNR = 12dB can be seen from Fig. 1(b). Note that the WSR is higher in the beginning when the per-BS power constraints are violated. Upon convergence the maximum power overage at the BSs is 0.0319%. From WSR-SNR results shown in and Fig. 3, we see that for all 3 constellations, B-GDP outperforms GP2 significantly in terms of WSR, especially in the practically important medium SNR regime.

5. CONCLUSION

This work investigates the design optimization of linear precoders in multi-cell MIMO downlink channel (MC-MIMO-DLC) with known finite-alphabet inputs. In our approach, we first simplify the complexity to evaluate the finite-alphabet weighted sum rate (WSR) and its gradient by adopting an approximation to Monte-Carlo sampling and derive a simple gradient descent algorithm (S-GDP) for precoder optimization. Second, by enforcing block diagonalization restrictions whenever possible, we derive an alternating gradient descent precoder optimization algorithm (B-GDP) based on two convex subproblems: One uses Stiefel manifold method and the other utilizes dual decomposition. Our numerical tests demonstrate the convergence behaviors and the performance advantages of both S-GDP and B-DGP algorithms over existing methods given finite-alphabet inputs.

Table 1. Backtracking line search parameters

	S2 S-GDP	B2 B-GDP	B5 B-GDP
α	0.2	0.1	0.8
β	0.5	0.25	0.25



(a) GP1 + S-GDP (b) B-GDP

Fig. 1. Convergence behavior of (a) GP1 + S-GDP and (b) B-GDP

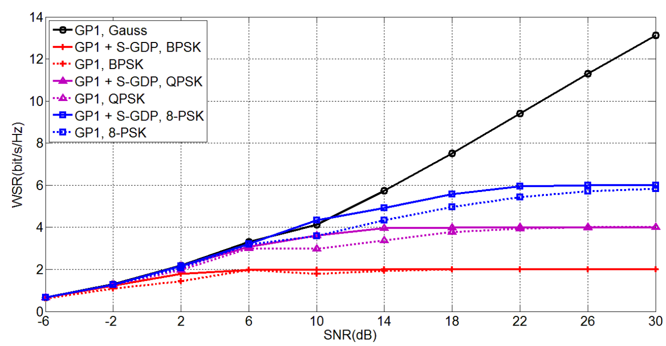


Fig. 2. Comparison between GP1 + S-GDP and GP1.

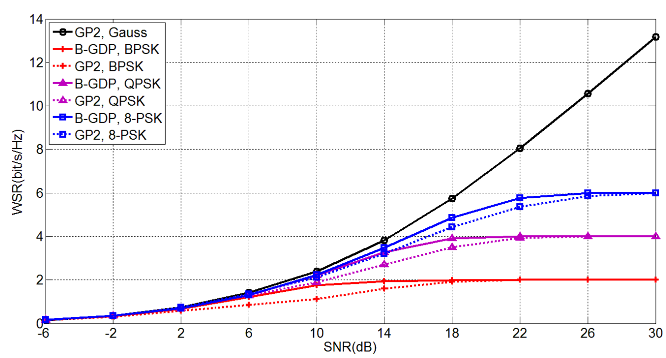


Fig. 3. Comparison between B-GDP and GP2.

6. REFERENCES

[1] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, “Multi-cell MIMO cooperative networks: A new look at interference,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, 2010.

[2] W. Yu and T. Lan, “Transmitter optimization for the multi-antenna downlink with per-antenna power constraints,” *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, 2007.

[3] H. Huh, H. Papadopoulos, and G. Caire, “MIMO broadcast channel optimization under general linear constraints,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2009, pp. 2664–2668.

[4] S. S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, “Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, 2008.

[5] R. Fritzsche and G. P. Fettweis, “Robust sum rate maximization in the multi-cell MU-MIMO downlink,” in *Proc. 2013 IEEE WCNC*, pp. 3180–3184.

[6] K. Wang, X. Wang, W. Xu, and X. Zhang, “Coordinated linear precoding in downlink multicell MIMO-OFDMA networks,” *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4264–4277, 2012.

[7] R. Zhang, “Cooperative multi-cell block diagonalization with per-base-station power constraints,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1435–1445, 2010.

[8] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, “Networked MIMO with clustered linear precoding,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1910–1921, 2009.

[9] L. Venturino, N. Prasad, and X. Wang, “Coordinated linear beamforming in downlink multi-cell wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1451–1461, 2010.

[10] R. Zhang and S. Cui, “Cooperative interference management with MISO beamforming,” *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, 2010.

[11] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, “MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation,” *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070–1084, 2010.

[12] C. Xiao, Y. R. Zheng, and Z. Ding, “Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels,” *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3301–3314, 2011.

[13] Y. Wu, M. Wang, C. Xiao, Z. Ding, and X. Gao, “Linear precoding for MIMO broadcast channels with finite-alphabet constraints,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2906–2920, 2012.

[14] M. Wang, W. Zeng, and C. Xiao, “Linear precoding for MIMO multiple access channels with finite discrete inputs,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3934–3942, 2011.

[15] X. Liang, Z. Ding, and C. Xiao, “On linear precoding of non-regenerative MIMO relays for QAM inputs,” in *Proc. IEEE Int. Conf. Commun.*, June 2012, pp. 2439–2444.

[16] W. Zeng, Y. R. Zheng, M. Wang, and J. Lu, “Linear precoding for relay networks: a perspective on finite-alphabet inputs,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1146–1157, 2012.

[17] X. Liang, C. Zhao, and Z. Ding, “Sequential linear MIMO precoder optimization for hybrid ARQ retransmission of QAM signals,” *IEEE Commun. Lett.*, vol. 15, no. 9, pp. 913–915, 2011.

[18] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, “Linear precoding for finite-alphabet signaling over MIMOME wiretap channels,” *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599–2612, 2012.

[19] W. Zeng, C. Xiao, and J. Lu, “A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs,” *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 38–41, 2012.

[20] D. P. Palomar and M. Chiang, “A tutorial on decomposition methods for network utility maximization,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, 2006.

[21] S. P. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[22] A. Hjørungnes, *Complex-valued matrix derivatives: with applications in signal processing and communications*, Cambridge University Press, 2011.

[23] H. Zhang and H. Dai, “Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2004, no. 2, pp. 222–235, 2004.

[24] M. Payaró and D. P. Palomar, “On optimal precoding in linear vector Gaussian channels with arbitrary input distribution,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2009, pp. 1085–1089.