

A New Approach to Beamformer Design for Massive MIMO Systems Based on k -Regularity

Gilwon Lee, Juho Park, Youngchul Sung[†], and Junyeong Seo

Dept. of Electrical Engineering

KAIST

Daejeon, South Korea 305-701

Email: {gwlee@, jhp@, ysung@ee., and jyseo@}kaist.ac.kr

Abstract—In this paper, a new beamformer design paradigm, named k -regular beamformer, is proposed for massive multiple-input multiple-output (MIMO) transmission systems to achieve most of the gain inherent to a large antenna array without too much complexity. In the proposed k -regular beamforming scheme, each of multiple data streams for MIMO transmission is multiplied by k complex gains and assigned to k out of available N_T transmit antennas, and signals assigned to the same transmit antenna are added and transmitted through the assigned antenna. The proposed k -regular beamformer can implement antenna selection (corresponding to $k = 1$) to optimal eigen-beamforming (corresponding to $k = N_T$) by controlling the parameter k , and thus enables arbitrary trade-off between complexity and performance. Two beamformer design algorithms, the maximum correlation method (MCM) and the projected iterative shrinkage-thresholding algorithm (PISTA), are proposed to design k -regular beamforming matrices. Numerical results show that the proposed k -regular beamformer even with small k significantly improves the rate gain over simple antenna selection and achieves most of the optimal eigen-beamforming performance with far less complexity than that required for optimal eigen-beamforming for massive MIMO transmission.

I. INTRODUCTION

Beyond 4G systems require significant improvement in spectral efficiency over the current maximum of LTE-Advanced attained by 8×8 MIMO transmission. One of the promising innovative technologies to achieve this goal is the massive MIMO technology adopting large-scale antenna arrays [1]–[4]. Employing a large-scale antenna array at the transmitter or receiver provides multiple gains such as rate increase, transmission reliability, and energy efficiency. Whereas the massive MIMO technology can provide such gains, two important practical issues arise regarding the technology in addition to other practical issues such as channel estimation: (I.1) The technology requires high hardware complexity in digital and RF/analog domains in order to support transmission through a large-scale antenna array. Thus, the technology consumes more energy than conventional small-scale MIMO technologies.

(I.2) Although it is possible to employ a large-scale antenna array at the basestation, the size of antenna array at the mobile

station is limited. Consequently, the MIMO channel formed by the massive MIMO technology is likely to be very asymmetric and the number of possible independent data streams for transmission is limited by the number of antennas at the mobile station.

In this paper, we propose a new beamformer design paradigm to handle the above two issues appropriately for the massive MIMO transmission. To illustrate the idea, consider an example of massive MIMO downlink transmission with N_T transmit antennas and M receive antennas, where $N_T \gg M$, and consider the transmit beamformer design to maximize the transmission rate for this MIMO channel. Under the assumption that all M rows of the $M \times N_T$ MIMO channel matrix are independent, the optimal number of independent data streams is M and the optimal transmit beamforming matrix is given by the matrix composed of the M right singular vectors of the channel matrix corresponding to the largest M singular values [5]. To generate the output signals to all N_T transmit antennas, the data vector of M input symbols from the M independent data streams should be multiplied to the $N_T \times M$ transmit beamforming matrix and this requires $N_T M$ multiplications or multipliers depending on whether the beamforming is implemented in digital or analog domains. If the beamforming is implemented in digital domain, we need N_T digital-to-analog converter (DAC) chains. If the beamforming is implemented in analog domain, on the other hand, we need $N_T M$ analog multipliers. In either case, hardware complexity is heavy since N_T itself is large for massive MIMO. A simple and classical solution to this complexity problem is antenna selection [6]–[10]. That is, M transmit antennas are selected out of the N_T available transmit antennas to yield the maximum rate among all possible $\binom{N_T}{M}$ choices and each data stream is assigned to one of the selected M antennas. Although the selection method provides significant complexity reduction, the rate performance of the selection is far inferior to that of the full optimal eigen-beamforming as we shall see later. Thus, in this paper, we propose the k -regular beamforming scheme for massive MIMO to overcome the rate loss of the simple antenna selection method, while keeping the hardware complexity far less than that required full optimal eigen-beamforming. In the proposed k -regular beamforming scheme, each of the M data streams for MIMO transmission

[†] This research was supported in part by the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA (Korea Communications Agency) (KCA-2011-11913-04001). This work was also supported in part by the IT R&D program of MKE/KEIT. [2008-F-004-02, “5G mobile communication systems based on beam division multiple access and relays with group cooperation”]

is multiplied by k complex gains and assigned to k out of the available N_T transmit antennas, and signals assigned to the same transmit antenna are added and transmitted through the assigned antenna. When $k = 1$, the proposed k -regular beamformer is equivalent to the conventional best transmit antenna selection method since it is not rate-optimal to assign two different data streams to the same antenna. (Degrees-of-freedom are lost in this case.) When $k = N_T$, on the other hand, the k -regular beamformer reduces to the full optimal eigen-beamforming. However, when $1 < k < N_T$, the k -regular beamformer resides somewhere in-between the antenna selection method and the full optimal eigen-beamforming, and thus trade-off between complexity and performance is possible through the k -regular beamformer for massive MIMO. The main difficulty in the k -regular beamformer design lies in finding k best antennas for each data stream and corresponding complex gains. If this design problem would be approached by using a combinatorial approach, then complexity of order $\mathcal{O}\left(\binom{N_T}{k}^M\right)$ would be required. To circumvent this difficulty and solve the problem of obtaining the best k -regular beamforming matrix, we propose two beamformer design algorithms: the *maximum correlation method (MCM)* and the *projected iterative shrinkage-thresholding algorithm (PISTA)*. In particular, the proposed PISTA is based on the iterative shrinkage-thresholding algorithm (ISTA) [11] which is an iterative approach to optimization under sparsity constraints. The PISTA avoids combinatorial search for the k -regular beamformer design problem and converges very fast. Numerical results show that the proposed k -regular beamformer significantly improves the rate gain over the simple antenna selection and achieves most of the optimal beamforming rate performance with far less complexity than that required for optimal eigen-beamforming for massive MIMO transmission.

Notation We will make use of standard notational conventions. Vectors and matrices are written in boldface with matrices in capitals. All vectors are column vectors. For matrix \mathbf{A} , \mathbf{A}^T , \mathbf{A}^H and $|\mathbf{A}|$ indicate the transpose, conjugate transpose and determinant of \mathbf{A} , respectively. For vector \mathbf{a} , $\|\mathbf{a}\|_p$ and $[\mathbf{a}]_i$ represent the p -norm and i -th element of \mathbf{a} , respectively. $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means that random vector \mathbf{x} is complex Gaussian distributed with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbb{E}\{\cdot\}$ denotes statistical expectation. \mathbb{C} is the set of complex numbers.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this paper, we consider single-user massive MIMO downlink transmission over a time-invariant MIMO channel. We assume that the transmitter (or basestation) has N_T transmit antennas and the receiver (or mobile station) has M receive antennas. We assume that $N_T \gg M$ for massive MIMO operation and the transmitter transmits M independent data streams which is the maximum for the considered MIMO channel. We further assume that the transmitter uses linear beamforming to transmit the M data streams. Then, the considered MIMO channel model is given by

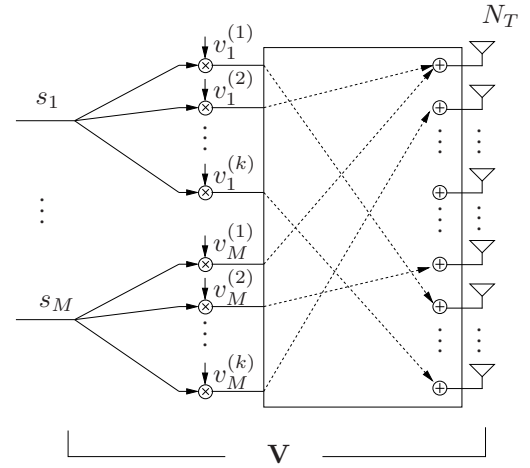


Fig. 1. The k -regular beamformer architecture

$$\mathbf{y} = \mathbf{H}\mathbf{V}\mathbf{s} + \mathbf{n}, \quad (1)$$

where \mathbf{y} is the $M \times 1$ received signal vector at the receiver, \mathbf{H} is the $M \times N_T$ MIMO channel matrix assumed to be known to the transmitter, $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_M]$ is the $N_T \times M$ transmit beamforming matrix, $\mathbf{s} = [s_1, \dots, s_M]^T$ is the $M \times 1$ symbol vector at the transmitter, and $\mathbf{n} \sim \mathcal{N}(0, \mathbf{I})$ is the additive white Gaussian noise vector at the receiver.

In general, the transmit beamforming matrix \mathbf{V} is an $N_T \times M$ full matrix of which elements can all be nonzero. As mentioned already in the introduction, a full transmit beamforming matrix incurs heavy hardware complexity for massive MIMO. Thus, in this paper, we consider a subclass of beamforming matrices, composed of k -regular beamforming matrices defined below:

Definition 1: A beamforming matrix \mathbf{V} is referred to as a k -regular beamforming matrix if each of its column vectors has k nonzero elements and all the rest elements are zero.

Note that the positions of k nonzero elements of each column vector are not predetermined and the k positions are arbitrary for each stream. The schematic diagram for a k -regular beamforming matrix is shown in Fig. 1. As shown in the figure, the connection between the input streams and the transmit antennas forms a bipartite graph and the left side is k -regular. This is why we refer to the beamforming matrix as k -regular. The example in Fig. 2 corresponds to the following 2-regular 6×2 beamforming matrix:

$$\mathbf{V} = \begin{bmatrix} 0 & 0 & 0 & \alpha_2 & \alpha_1 & 0 \\ \beta_1 & 0 & 0 & \beta_2 & 0 & 0 \end{bmatrix}^T. \quad (2)$$

There are several interesting properties of the proposed k -regular beamformer architecture. When $k = N_T$, an N_T -regular beamforming matrix reduces to a conventional full beamforming matrix of which elements can all be nonzero. On the other hand, when $k = 1$ and the assigned transmit antenna for each stream is different, the 1-regular beamformer is equivalent to conventional antenna selection. The complexity advantage of the k -regular beamformer comes from two facts. First, the number of required multiplications or multipliers is kM (not $N_T M$ as the full beamforming matrix case). This

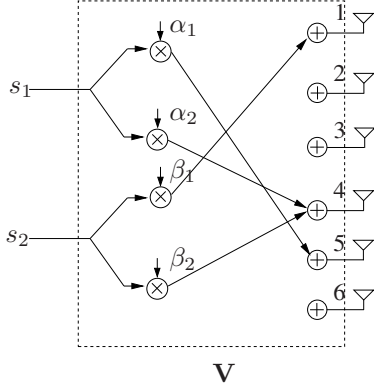


Fig. 2. An example of k -regular beamformer: $N_T = 6$, $M = 2$, and $k = 2$

significantly reduces the number of required analog multipliers if the k -regular beamformer is implemented in analog domain. (Of course, it reduces the number of required multiplications in the digital implementation case.) Second, all transmit antennas are not used for transmission since the k antenna assignment of each stream is arbitrary. When k is small, still many transmit antennas are not used for transmission as we shall see later, and this reduces the number of required DAC chains if the k -regular beamformer is implemented in digital domain. Thus, the parameter k controls the trade-off between complexity and performance.

From here on, we consider the problem of designing the k -regular beamforming matrix to maximize the transmission rate under the scenario of M independent data stream transmission with equal power for each stream. In this case, the covariance matrix of the data symbol vector is given by $\Sigma_s = \mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \frac{P_t}{M}\mathbf{I}$, where P_t is the total transmit power, and the k -regular beamformer design problem is formulated as follows:

$$(P1) \quad \underset{\mathbf{V}}{\text{maximize}} \quad \log \left| \mathbf{I} + \frac{P_t}{M} \mathbf{H} \mathbf{V} \mathbf{V}^H \mathbf{H}^H \right| \quad (3)$$

$$\text{subject to} \quad \|\mathbf{v}_i\|_0 = k, \quad i = 1, \dots, M \quad (4)$$

$$\|\mathbf{v}_i\|_2^2 \leq 1, \quad i = 1, \dots, M, \quad (5)$$

where (4) and (5) are the k -regularity and power constraints for each column of \mathbf{V} , respectively. (P1) is similar to the conventional MIMO precoder design problem but it includes regularity* constraints provided as ℓ_0 -norm constraints.

III. PROPOSED BEAMFORMER DESIGN ALGORITHMS

In this section, we provide two algorithms to design the k -regular beamforming matrix for rate maximization. The first is a two-step suboptimal approach and the second is an iterative algorithm that directly maximizes the rate under the constraints.

A. The maximum correlation method

Without the regularity constraints (4), the optimal transmit beamforming matrix is given by the matrix composed of the

*The regularity constraints (4) can be considered as sparsity constraints when $k < N_T$.

M right singular vectors of \mathbf{H} corresponding to the largest M singular values [5]. Based on this fact, we propose our first method for the k -regular beamformer design. The first method is simply to design the k -regular transmit beamforming matrix to approximate the full optimal eigen-beamforming matrix obtained from singular value decomposition (SVD) of \mathbf{H} under the regularity constraints. Here, we use correlation as the proximity measure and the proposed maximum correlation method (MCM) is given by

$$(P2) \quad \underset{\mathbf{v}_i}{\text{maximize}} \quad |\langle \boldsymbol{\psi}_i, \mathbf{v}_i \rangle| \\ \text{subject to} \quad \|\mathbf{v}_i\|_0 = k, \quad \|\mathbf{v}_i\|_2^2 \leq 1, \quad i = 1, \dots, M, \quad (6)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product operation and $\{\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_M\}$ are the first M right singular vectors of \mathbf{H} .

Proposition 1: The solution to the MCM problem (P2) is given by

$$\mathbf{V} = \left[\frac{\mathcal{P}_k(\boldsymbol{\psi}_1)}{\|\mathcal{P}_k(\boldsymbol{\psi}_1)\|_2}, \frac{\mathcal{P}_k(\boldsymbol{\psi}_2)}{\|\mathcal{P}_k(\boldsymbol{\psi}_2)\|_2}, \dots, \frac{\mathcal{P}_k(\boldsymbol{\psi}_M)}{\|\mathcal{P}_k(\boldsymbol{\psi}_M)\|_2} \right], \quad (7)$$

where $\mathcal{P}_k(\cdot) : \mathbb{C}^{N_T} \rightarrow \mathbb{C}^{N_T}$ is defined by

$$[\mathcal{P}_k(\mathbf{u})]_i = \begin{cases} [\mathbf{u}]_i, & \text{if } i \in \mathcal{K}, \\ 0, & \text{otherwise (o.w.),} \end{cases} \quad (8)$$

and \mathcal{K} is the set of indices of the elements of \mathbf{u} with the first k largest absolute values.

Proof: Suppose that the solution \mathbf{v}_i^* has $\|\mathbf{v}_i^*\|_2^2 < 1$. Let $(\mathbf{v}_i^*)' = \mathbf{v}_i^* / \|\mathbf{v}_i^*\|_2$. Then, $|\boldsymbol{\psi}_i^H (\mathbf{v}_i^*)'| = \frac{|\boldsymbol{\psi}_i^H \mathbf{v}_i^*|}{\|\mathbf{v}_i^*\|_2} > |\boldsymbol{\psi}_i^H \mathbf{v}_i^*|$, which contradicts the assumption. Therefore, $\|\mathbf{v}_i^*\|_2^2 = 1$. Suppose \mathbf{x}_i is a unit-norm vector that satisfies the regularity constraint. Let \mathcal{I} be the set of indices at which the elements of \mathbf{x}_i are nonzero and define

$$[\tilde{\boldsymbol{\psi}}_i]_j = \begin{cases} [\boldsymbol{\psi}_i]_j, & \text{if } j \in \mathcal{I}, \\ 0, & \text{o.w.} \end{cases}$$

Then,

$$\begin{aligned} |\boldsymbol{\psi}_i^H \mathbf{x}_i| &= \sum_{j=1}^{N_T} [\boldsymbol{\psi}_i]_j^* [\mathbf{x}_i]_j = \sum_{j \in \mathcal{I}} [\boldsymbol{\psi}_i]_j^* [\mathbf{x}_i]_j = \sum_{j \in \mathcal{I}} [\tilde{\boldsymbol{\psi}}_i]_j^* [\mathbf{x}_i]_j \\ &= |(\tilde{\boldsymbol{\psi}}_i)^H \mathbf{x}_i| = \|\tilde{\boldsymbol{\psi}}_i\|_2 \cdot \left| \left(\frac{\tilde{\boldsymbol{\psi}}_i}{\|\tilde{\boldsymbol{\psi}}_i\|_2} \right)^H \mathbf{x}_i \right| \leq \|\tilde{\boldsymbol{\psi}}_i\|_2. \end{aligned}$$

The last inequality is by the Cauchy-Schwarz inequality and equality holds when $\mathbf{x}_i = \tilde{\boldsymbol{\psi}}_i / \|\tilde{\boldsymbol{\psi}}_i\|_2$. Therefore, $|\boldsymbol{\psi}_i^H \mathbf{x}_i|$ is maximized when $\|\tilde{\boldsymbol{\psi}}_i\|_2$ is maximized. This is done by choosing \mathcal{I} to be \mathcal{K} of $\boldsymbol{\psi}_i$. ■

B. The projected iterative shrinkage-thresholding algorithm

Although the MCM provides a method for the k -regular beamformer design, this two-step approach is not optimal and furthermore it requires full SVD of the large matrix \mathbf{H} . Thus, in this subsection, we propose an iterative algorithm for (P1) that yields better performance with reduced complexity. Note that the problem (3,4) without (5) is an optimization problem under ℓ_0 -norm constraints. Recently, there has been much progress in this area under the name of compressed sensing.

In the case of compressed sensing, a linear inverse problem is considered and one approach to the inverse problem is that a solution is obtained by minimizing the 2-norm of the residual vector under ℓ_0 -norm (or sparsity) constraints. However, in our case, we should minimize the non-convex rate cost (3) under ℓ_0 -norm regularity constraints (4) and power constraints (5). To solve this complicated problem, we adopt and modify the general framework of the proximal forward-backward splitting method [12], in particular, the iterative shrinkage-thresholding algorithm (ISTA) [11]. First, note that the regularity (or sparsity) constraints in (P1) are not convex and thus it is difficult to tackle the problem in the original form. To circumvent this difficulty, we substitute the ℓ_0 -norm constraints for the k -regularity with ℓ_1 -norm constraints in order to make the constraints convex while maintaining sparsity as in [13]. With this substitution, we have the following relaxed problem for (P1):

$$(P3) \quad \begin{aligned} & \underset{\mathbf{V}}{\text{minimize}} && f(\mathbf{V}) && (9) \\ & \text{subject to} && \|\mathbf{v}_i\|_1 \leq \xi, && i = 1, \dots, M, && (10) \\ & && \|\mathbf{v}_i\|_2^2 \leq 1, && i = 1, \dots, M, && (11) \end{aligned}$$

where $\xi \geq 0$ and $f(\mathbf{V}) := -\log \left| \mathbf{I} + \frac{P_t}{M} \mathbf{H} \mathbf{V} \mathbf{V}^H \mathbf{H}^H \right|$.

In general, a convex optimization problem under an ℓ_1 -norm constraint can be written as

$$\begin{aligned} & \underset{\mathbf{u}}{\text{minimize}} && h(\mathbf{u}) \\ & \text{subject to} && \|\mathbf{u}\|_1 \leq c, \end{aligned} \quad (12)$$

where $h(\mathbf{u})$ is a smooth convex function of \mathbf{u} . By the Lagrange duality the problem (12) is equivalent to the following ℓ_1 -regularized problem with some $\lambda \geq 0$:

$$\underset{\mathbf{u}}{\text{minimize}} \quad h(\mathbf{u}) + \lambda \|\mathbf{u}\|_1. \quad (13)$$

The ISTA is an extension to the classical gradient descent method by proximal regularization and is an efficient iterative method to solve (13) [11]. At each iteration of ISTA, the current solution \mathbf{u}_k is updated as

$$\mathbf{u}_{k+1} = \mathcal{T}_{\lambda\mu}(\mathbf{u}_k - \mu \nabla h(\mathbf{u}_k)), \quad (14)$$

where $\mu > 0$ is the stepsize parameter for gradient descent, and \mathcal{T}_α is the shrinkage operator defined by

$$[\mathcal{T}_\alpha(\mathbf{u})]_i = (|\mathbf{u}_i| - \alpha)_+ \text{sgn}([\mathbf{u}]_i) \quad (15)$$

in the case of real \mathbf{u} . Here, $(r)_+ = r$ if $r \geq 0$ and $(r)_+ = 0$ o.w. In the case of complex \mathbf{u} , the shrinkage operator can be generalized to

$$[\mathcal{T}_\alpha(\mathbf{u})]_i = (|\mathbf{u}_i| - \alpha)_+ e^{j\theta_i}, \quad (16)$$

where θ_i is the phase of $[\mathbf{u}]_i$. That is, at each iteration of ISTA, the current vector is updated by gradient descent first to reduce the cost and then by a shrinkage/soft-threshold step to guarantee sparsity. Note that (P3) without the power constraints can be formulated as an ℓ_1 -regularized optimization problem and the ISTA can be applied to this problem. However, we

TABLE I
THE PROJECTED ISTA FOR k -REGULAR BEAMFORMER DESIGN (R_n IS THE RATE AT THE n -TH ITERATION AND $\mathcal{P}_k(\cdot)$ IS DEFINED IN (8).)

0. (Initialization) Generate \mathbf{V} randomly.
1. (Gradient descent) Update $\mathbf{V} = \mathbf{V} - \mu \nabla f(\mathbf{V})$.
2. (Soft-thresholding) For $i = 1, \dots, M$, $\mathbf{v}_i = \mathcal{T}_{\lambda\mu}(\mathbf{v}_i)$,
3. (Projection of \mathbf{v}_i onto B_1) For $i = 1, \dots, M$, $\mathbf{v}_i = \mathcal{P}_{B_1}(\mathbf{v}_i)$.
4. (Stop criterion) If $\frac{R_n - R_{n-1}}{R_n} \geq \delta$, go to step 1. O.w., go to step 5.
5. (Hard-thresholding) For $i = 1, \dots, M$, update $\mathbf{v}_i = \mathcal{P}_k(\mathbf{v}_i)$.
6. (Power adjusting) For $i = 1, \dots, M$, update $\mathbf{v}_i = \mathbf{v}_i / \ \mathbf{v}_i\ _2$.

have the power constraints (11). Fortunately, each of the power constraints in (P3) yields a convex constraint set which is simply the unit-norm ball B_1 in \mathbb{C}^{N_T} . To incorporate the power constraints (11), we project the output of the ISTA to the unit-norm ball for each column of \mathbf{V} , and the metric projection of vector \mathbf{v}_i onto B_1 is simply given by

$$\mathcal{P}_{B_1}(\mathbf{v}_i) = \begin{cases} \mathbf{v}_i / \|\mathbf{v}_i\|_2, & \text{if } \|\mathbf{v}_i\|_2 > 1, \\ \mathbf{v}_i, & \text{o.w.} \end{cases} \quad (17)$$

In summary, the proposed operation at each update is given by

$$\mathbf{V} = \mathcal{P}_{B_1}^{(M)} \dots \mathcal{P}_{B_1}^{(1)} \circ \mathcal{T}_{\lambda\mu}^{(M)} \dots \mathcal{T}_{\lambda\mu}^{(1)} (\mathbf{V} - \mu \nabla f(\mathbf{V})), \quad (18)$$

where $\mathcal{P}_{B_1}^{(i)}$ is the projection of the i -th column of \mathbf{V} onto the unit-norm ball, $\mathcal{T}_{\lambda\mu}^{(i)}$ is the shrinkage operator for the i -th column of \mathbf{V} , and the gradient is given by

$$\nabla f(\mathbf{V}) = -\frac{P_t}{M} \mathbf{H}^H \left(\mathbf{I} + \frac{P_t}{M} \mathbf{H} \mathbf{V} \mathbf{V}^H \mathbf{H}^H \right)^{-1} \mathbf{H} \mathbf{V}. \quad (19)$$

Note that the proposed algorithm (18) is a combination of metric projections and ISTA. This is why we refer to the algorithm as the projected ISTA (PISTA). In fact, the ISTA itself is a combination of proximal operator[†] and gradient descent. Thus, by viewing the sequential projection $\mathcal{P}_{B_1}^{(M)} \dots \mathcal{P}_{B_1}^{(1)} \circ \mathcal{T}_{\lambda\mu}^{(M)} \dots \mathcal{T}_{\lambda\mu}^{(1)}$ in (18) as a big single “metric” projection, we can consider the proposed PISTA as an extension of the projected gradient method [14]. By using the firmly nonexpansive property of $\mathcal{P}_{B_1}^{(i)}$ and $\mathcal{T}_{\lambda\mu}^{(i)}$, the local convergence of the PISTA can be shown but this is beyond the scope of this paper. (In the next section, we shall see that the PISTA finds a (nearly-)optimal solution very well in the case of $k = 1$.) Since the original ℓ_0 -norm regularity constraints are relaxed to mild ℓ_1 -norm constraints and the PISTA solves (P3), the final output of the PISTA has nonzero elements more than k . Thus, an additional step is necessary to obtain a final solution that satisfies the constraints in (P1). To obtain only k nonzero elements, we pick k elements with the largest absolute values for each column of \mathbf{V} obtained from the PISTA, and this column vector with k nonzero elements is normalized to satisfy the power constraint. The overall algorithm based on the PISTA is summarized in Table I. The complexity of the algorithm mainly comes from the computation of the gradient

[†]A proximal operator is an extension of metric projection and the shrinkage operator is an example of proximal operator [12].

(19), which is order of $O(N_T M^2)$, since the metric projection and the shrinkage operator only require complexity of $O(N_T M)$. It is observed that the number of required iterations for convergence is almost independent of the problem size and it is roughly several tens of iterations.

IV. NUMERICAL RESULTS

In this section, we provide some numerical results to evaluate the performance of the proposed algorithms for k -regular beamformer design. All rates here are the results averaged over 50 independent channel realizations and each element of the MIMO channel matrix was generated i.i.d. according to $\mathcal{CN}(0, 1)$. First, we validated the PISTA by using the available results for $k = 1$. As mentioned already, the case of $k = 1$ corresponds to the best antenna selection. We considered the case of $N_T = 16$ and $M = 4$ so that even brute-force search is possible. Fig. 3 (left) shows the result. It is shown that the PISTA yields almost the same performance as the brute-force search. With this validation, we proceeded to the cases of $k > 1$, which the conventional antenna selection methods cannot handle. Fig. 3 (right) shows the rate performance of the proposed algorithms with respect to P_t when $(N_T, M, k) = (72, 8, 9)$. Here, R_{opt} , R_{PISTA} , R_{MCM} and $R_{8 \times 8}$ denote the rates of optimal eigen-beamforming, PISTA, MCM and $(N_T, M) = (8, 8)$ optimal eigen-beamforming, respectively. It is seen that the MCM is suboptimal indeed and the PISTA outperforms the MCM. Fig. 4 (left) shows the rate versus k for $(N_T, M) = (72, 8)$. It is seen that, although there is some selection gain over $(8, 8)$ MIMO, antenna selection ($k = 1$) is far inferior to the optimal eigen-beamforming but the k -regular beamformer only with small k achieves most of the rate of the optimal eigen-beamforming. Finally, Fig. 4 (right) shows the distribution of number of connections. When $k = 1$ (i.e., for antenna selection), 8 antennas have one connection and $64 = 72 - 8$ (89%) antennas are not connected to signals. As expected, for the k -regular beamformer, still a large portion of antennas are not connected to signals for reasonably small k values. This aspect of the k -regular beamformer can be exploited to reduce the hardware complexity in addition to the reduction in the number of required multiplications or multipliers.

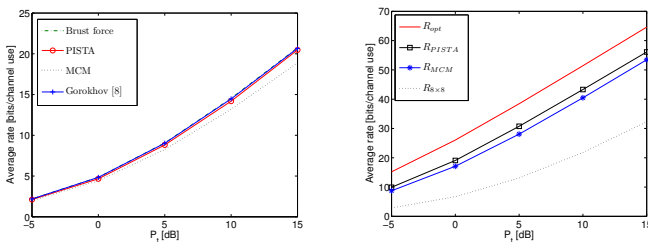


Fig. 3. (Left) Performance of the PISTA for antenna selection: $N_T = 16$, $M = 4$ and $k = 1$ and (right) average rate performance: $N_T = 72$, $M = 8$ and $k = 9$

V. CONCLUSION

In this paper, we have proposed the k -regular beamformer architecture for single-user massive MIMO downlink transmission and the PISTA for k -regular beamformer design.

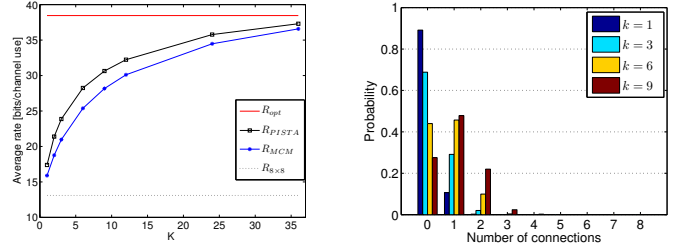


Fig. 4. $N_T = 72$, $M = 8$ and $P_t/1 = 5$ [dB]: (left) rate versus k and (right) distribution of antennas over numbers of connections

The proposed k -regular beamformer can implement antenna selection to full eigen-beamforming depending on the regularity parameter k and provides trade-off between hardware complexity and rate performance. Thus, the proposed k -regular beamformer and the PISTA enable system designers to choose optimal trade-off for their massive MIMO systems based on their hardware constraint and required rate performance. Numerical results show that the proposed k -regular beamformer significantly improves the rate gain over simple antenna selection and achieves most of the optimal eigen-beamforming performance with far less hardware complexity than that required for optimal eigen-beamforming for massive MIMO transmission.

REFERENCES

- [1] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *ArXiv pre-print cs.IT/1201.3210*, Jan. 2012.
- [2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3590 – 3600, Nov. 2010.
- [3] J. Hoydis, S. T. Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?," *ArXiv pre-print cs.IT/1107.1709v2*, Sep. 2011.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *ArXiv pre-print cs.IT/1112.3810*, Dec. 2011.
- [5] I. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585 – 596, Nov.-Dec. 1999.
- [6] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Transactions on Communications*, vol. 47, pp. 1773 – 1776, Dec. 1999.
- [7] D. A. Gore and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2580 – 2588, Oct. 2002.
- [8] A. Gorokhov, "Antenna selection algorithm for MEA transmission systems," in *Proc. ICASSP*, vol. 3, pp. 2857 – 2860, 2002.
- [9] Y. S. Choi, A. F. Molisch, M. Z. Win, and J. H. Winters, "Fast algorithms for antenna selection in MIMO systems," in *Proc. VTC*, vol. 3, pp. 1733 – 1737, Oct. 2003.
- [10] A. F. Molisch, M. Z. Win, Y. S. Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 1759 – 1772, Jul. 2005.
- [11] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, pp. 183 – 202, Mar. 2009.
- [12] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *SIAM Journal on Multiscale Modeling and Simulation*, vol. 4, pp. 1168 – 1200, Nov. 2005.
- [13] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289 – 1306, Apr. 2006.
- [14] H. H. bauschke, R. Burachik, P. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. ch. 17, pp. 345 – 390. Springer, 2011.