# Joint User Selection and Precoding in Multiuser MIMO Systems via Group LASSO

Saba Asaad*, Ali Bereyhi*, Ralf R. Müller*, and Rafael F. Schaefer†,
*Institute for Digital Communications, Friedrich-Alexander Universität Erlangen-Nürnberg
†Information Theory and Applications Chair, Technische Universität Berlin
{saba.asaad, ali.bereyhi, ralf.r.mueller}@fau.de, rafael.schaefer@tu-berlin.de



***Abstract*—Joint user selection and precoding in multiuser MIMO settings can be interpreted as group sparse recovery in linear models. In this problem, a signal with group sparsity is to be reconstructed from an underdetermined system of equations. This paper utilizes this equivalent interpretation and develops a computationally tractable algorithm based on the method of group LASSO. Compared to the state of the art, the proposed scheme shows performance enhancements in two different respects: higher achievable sum-rate and lower interference at the non-selected user terminals.**

***Index Terms*—User selection, precoding, group LASSO, massive MIMO.**

## I. Introduction

Performance gains are often achieved in multiuser massive multiple-input multiple-output (MIMO) systems with a large number of transmit antennas per user [1]. As a result, in dense settings in which the number of available users is comparable to the number of transmit antennas, user selection is required along with downlink beamforming [2]–[5].

The conventional approach for user selection and precoding is to divide them into two separate problems: First, a subset of users is selected; then, the information signals of the selected users are precoded via a classic precoding scheme [5]. Generally, the optimal approach for user selection deals with integer programming. Hence, this problem is often addressed via sub-optimal greedy algorithms [2], [3]. In this work, we deviate from the conventional approach and propose a scheme for joint user selection and downlink beamforming.

### A. User Selection and Precoding as Group Sparsity

Joint user selection and beamforming is interpreted as the problem of constructing a signal with *group sparsity*. To clarify this point, assume a multiuser downlink scenario with $M$ transmit antennas and $K$ users in which we wish to select a subset of $L$ users. A linear precoder in this problem can be seen as a signal with $MK$ entries, such that each block of size $M$ represents an individual beamforming vector. By such a formulation, joint user selection and downlink beamforming with respect to some performance metric, e.g., the achievable sum-rate or mean squared error (MSE), reduces to the problem of finding a signal with group sparsity: A signal of size $MK$ in which only $L$ blocks of size $M$ have non-zero entries.

Following this equivalent interpretation, we employ the generalized least squared error (GLSE) framework for precoding, recently developed in [6]–[8], to formulate joint user selection and precoding as the problem of group sparse recovery in a linear model. A computationally tractable algorithm is then developed based on group least absolute shrinkage and selection operator (LASSO) to address this problem. Our investigations show significant performance enhancements compared to the state of the art.

### B. Notations

Throughout the paper, scalars, vectors, and matrices are represented by non-bold, bold lower case, and bold upper case letters, respectively. The real axis is denoted by $\mathbb{R}$ and the complex plane is shown by $\mathbb{C}$. $\mathbf{H}^{\mathsf{H}}$, $\mathbf{H}^{*}$, and $\mathbf{H}^{\mathsf{T}}$ indicate the Hermitian, complex conjugate, and transpose of $\mathbf{H}$, respectively. $\log(\cdot)$ is the binary logarithm. We denote the statistical expectation by $\mathbb{E}\{\cdot\}$. $\mathrm{diag}(\mathbf{t})$ represents the diagonal matrix constructed from the elements of vector $\mathbf{t}$.

## II. Problem Formulation

Consider a multiuser MIMO system with multiple base stations (BSs) which are equipped with transmit antenna arrays of size $M$. The system is intended to serve $K$ single-antenna user terminals (UTs). For mathematical tractability, we focus on a single BS which aims to transmit information to a group of $L \leq K$ UTs.

### A. System Model

The system operates in the time-division duplexing (TDD) mode. Hence, the uplink and downlink channels are reciprocal. In each coherence time interval, the UTs transmit known training sequences. The BS then utilizes these sequences to estimate the channel state information (CSI).

Let $\mathbf{h}_k \in \mathbb{C}^M$ denote the vector of uplink channel coefficients between UT $k$ and the BS. The signal received by UT $k$ is hence given by

$$y_k = \mathbf{h}_k^{\mathsf{T}} \boldsymbol{x} + z_k \tag{1}$$

where $z_k$ is additive complex Gaussian noise with zero mean and variance $\sigma_k^2$, i.e., $z_k \sim \mathcal{CN}\left(0, \sigma_k^2\right)$, and $\boldsymbol{x}$ is the downlink transmit signal constructed from the information symbols of the selected UTs and the

CSI via linear precoding. As a result, the transmit signal is written as

$$\boldsymbol{x} = \sum_{\ell \in \mathbb{S}} \sqrt{p_\ell} s_\ell \mathbf{w}_\ell. \tag{2}$$

where $\mathbb{S}$, $s_\ell$, $p_\ell$ and $\mathbf{w}_\ell$ are defined as follows:

1) $\mathbb{S} \subseteq \{1, \dots, K\}$ represents the subset of $L$ UTs selected by the BS for downlink transmission.
2) $s_\ell$ is the information symbol of user $\ell$ which is assumed to be zero-mean and unit-variance.
3) $p_\ell$ denotes the power allocated to UT $\ell \in \mathbb{S}$.
4) $\mathbf{w}_\ell$ is the beamforming vector of UT $\ell$.

The transmit power at the BS is restricted. It is hence assumed that $\boldsymbol{x}$ satisfies the power constraint $\mathbb{E}\left\{\boldsymbol{x}^{\mathsf{H}}\boldsymbol{x}\right\} \leq P$ for some non-negative real $P$.

### B. Performance Measure

There are various metrics characterizing the performance of the downlink transmission in this system. One well-known metric is the *weighted average throughput* which is defined as

$$R_{\text{avg}} = \frac{1}{L} \sum_{\ell \in \mathbb{S}} w_\ell R_\ell \tag{3}$$

for some non-negative weights $\{w_\ell\}$ and transmission rates

$$R_\ell = \log\left(1 + \text{SINR}_\ell\right). \tag{4}$$

In (4), $\text{SINR}_\ell$ is defined as

$$\text{SINR}_\ell = \frac{p_\ell |\mathbf{h}_\ell^{\mathsf{T}} \mathbf{w}_\ell|^2}{\sigma_\ell^2 + \sum\limits_{j=1, j\neq \ell}^{K} p_j |\mathbf{h}_\ell^{\mathsf{T}} \mathbf{w}_j|^2}. \tag{5}$$

From signal processing points of view, precoding can be interpreted as *channel inversion*. In this problem, the ultimate aim is to construct the transmit signal such that at a selected UT $\ell$, $\mathbf{h}_\ell^{\mathsf{T}}\boldsymbol{x} = \beta s_\ell$, for some scaling factor $\beta$, and at UT $k$ which has not been selected, we have $\mathbf{h}_k^{\mathsf{T}}\boldsymbol{x} = 0$. The former guarantees channel inversion at the selected UTs which results in minimal post-processing load, and the latter restricts the precoder to have zero leakage at the non-selected UTs.

By this alternative viewpoint, a suitable performance measure is the *residual sum of squares (RSS)* at the UTs defined as

$$\text{RSS} = \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\left\{|\mathbf{h}_k^{\mathsf{T}}\boldsymbol{x} - \beta a_k s_k|^2\right\}, \tag{6}$$

where $a_k = 1$ if UT $k$ is selected and is zero otherwise.

## III. Optimal User Selection and Precoding

Let $\boldsymbol{s} = [s_1, \dots, s_K]^{\mathsf{T}}$ collect the information symbols of all UTs. By defining $p_k = 0$ for those UTs which are not selected, the transmit signal is compactly represented as

$$\boldsymbol{x} = \mathbf{W}\sqrt{\mathbf{P}}\boldsymbol{s}. \tag{7}$$

where $\mathbf{W}$ and $\mathbf{P}$ are defined as follows:

1) $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$ is the beamforming matrix.
2) $\mathbf{P} = \text{diag}\left(\mathbf{p}\right)$ with $\mathbf{p} = [p_1, \dots, p_K]^{\mathsf{T}}$.

The notation $\sqrt{\mathbf{P}}$ moreover denotes a matrix whose entries are the square root of the entries of $\mathbf{P}$. Similarly, the vector of receive signals $\boldsymbol{y} = [y_1, \dots, y_K]^{\mathsf{T}}$ reads

$$\boldsymbol{y} = \mathbf{H}^{\mathsf{T}}\boldsymbol{x} + \boldsymbol{z} \tag{8}$$

where $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K]$ and $\boldsymbol{z} = [z_1, \dots, z_K]^{\mathsf{T}}$.

### A. User Selection and Precoding with Minimum RSS

We design the transmit signal by considering the RSS as the performance measure. In this respect, the optimal approach for joint user selection and precoding is to find $\mathbf{W}$ and $\mathbf{p}$ such that the RSS is minimized and the signal constraints are satisfied. In the sequel, we formulate this approach in a standard form.

*Objective Function:* Following the given representation, the RSS is written as

$$\text{RSS} = \frac{1}{K}\mathbb{E}\left\{\|\mathbf{H}^{\mathsf{T}}\mathbf{W}\sqrt{\mathbf{P}}\boldsymbol{s} - \beta\mathbf{A}\boldsymbol{s}\|^2\right\}, \tag{9}$$

where $\mathbf{A} = \text{diag}\left(a_1, \dots, a_K\right)$. In this formulation, $\mathbf{A}$ is ineffective and can be dropped. To show this, note that for any non-selected UT $k$, $\boldsymbol{x}$ is independent of $s_k$ and hence

$$\begin{aligned} \mathbb{E}\left\{|\mathbf{h}_k^{\mathsf{T}}\boldsymbol{x} - \beta s_k|^2\right\} &= \mathbb{E}\left\{|\mathbf{h}_k^{\mathsf{T}}\boldsymbol{x}|^2\right\} + \beta^2\mathbb{E}\left\{|s_k|^2\right\} && \text{(10a)} \\ &= \mathbb{E}\left\{|\mathbf{h}_k^{\mathsf{T}}\boldsymbol{x}|^2\right\} + \beta^2. && \text{(10b)} \end{aligned}$$

Therefore, we can write

$$\text{RSS} = \frac{1}{K} D\left(\mathbf{W}, \mathbf{p}\right) - \left(1 - \frac{L}{K}\right)\beta^2, \tag{11}$$

where $D\left(\mathbf{W}, \mathbf{p}\right)$ is defined as

$$\begin{aligned} D\left(\mathbf{W}, \mathbf{p}\right) &:= \mathbb{E}\left\{\|\mathbf{H}^{\mathsf{T}}\mathbf{W}\sqrt{\mathbf{P}}\boldsymbol{s} - \beta\boldsymbol{s}\|^2\right\} && \text{(12a)} \\ &= \text{tr}\left\{\mathbf{Q}^{\mathsf{H}}\mathbf{Q}\right\} && \text{(12b)} \end{aligned}$$

with $\mathbf{Q} = \mathbf{H}^{\mathsf{T}}\mathbf{W}\sqrt{\mathbf{P}} - \beta\mathbf{I}_K$. We hence set the objective function to $D\left(\mathbf{W}, \mathbf{p}\right)$.

*Constraints:* There are two main constraints:

1) The number of selected UTs should be less than $L$.
2) The average transmit power is constrained.

Noting that the number of selected UTs in the system is given by the *sparsity* of $\mathbf{p}$, i.e., $\|\mathbf{p}\|_0$, the first constraint is written as

$$\|\mathbf{p}\|_0 \leq L. \tag{13}$$

For the second constraint, we note that

$$\begin{aligned} \mathbb{E}\left\{\boldsymbol{x}^{\mathsf{H}}\boldsymbol{x}\right\} &= \mathbb{E}\left\{\boldsymbol{s}^{\mathsf{H}}\sqrt{\mathbf{P}}\mathbf{W}^{\mathsf{H}}\mathbf{W}\sqrt{\mathbf{P}}\boldsymbol{s}\right\} && \text{(14a)} \\ &\overset{\dagger}{=} \mathbb{E}\left\{\text{tr}\left\{\sqrt{\mathbf{P}}\mathbf{W}^{\mathsf{H}}\mathbf{W}\sqrt{\mathbf{P}}\boldsymbol{s}\boldsymbol{s}^{\mathsf{H}}\right\}\right\} && \text{(14b)} \\ &= \text{tr}\left\{\mathbf{W}\mathbf{P}\mathbf{W}^{\mathsf{H}}\right\} && \text{(14c)} \end{aligned}$$

where $\dagger$ follows the fact that $\mathbb{E}\left\{\boldsymbol{s}\boldsymbol{s}^{\mathsf{H}}\right\} = \mathbf{I}_K$. As a result, the transmit power constraint reads

$$\text{tr}\left\{\mathbf{W}\mathbf{P}\mathbf{W}^{\mathsf{H}}\right\} \leq P. \tag{15}$$

*Optimization Problem:* Considering the objective function and constraints, the jointly optimal approach for user selection and precoding is formulated as

$$\begin{aligned} \min_{\mathbf{W}\in\mathbb{C}^{M\times K},\mathbf{p}\in\mathbb{R}_+^K} \quad & D(\mathbf{W},\mathbf{p}) \qquad (16)\\ \text{subject to} \quad & \mathrm{C}_1 : \|\mathbf{p}\|_0 \leq L,\\ & \mathrm{C}_2 : \mathrm{tr}\left\{\mathbf{W}\mathrm{diag}(\mathbf{p})\mathbf{W}^{\mathsf{H}}\right\} \leq P. \end{aligned}$$

The optimization problem in its initial form is not tractable, since both the objective function and constraints are not convex. We address this issue by converting (16) into a *group selection* problem. We then develop an algorithm based on *group LASSO* to estimate the solution.

## IV. Precoding via Group LASSO

The optimization problem in (16) can be converted into a group selection problem. To show this, let $\mathbf{V} := \mathbf{W}\sqrt{\mathbf{P}}$ be the *overall precoding matrix*. The objective function is rewritten in terms of $\mathbf{V}$ as

$$\begin{aligned} D(\mathbf{W},\mathbf{p}) &= \mathrm{tr}\left\{\left(\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\mathbf{I}_K\right)^{\mathsf{H}}\left(\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\mathbf{I}_K\right)\right\}\\ &= \|\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\,\mathbf{I}_K\|_F^2. \qquad (17) \end{aligned}$$

The power constraint is further given in terms of $\mathbf{V}$ as

$$\mathrm{tr}\left\{\mathbf{V}^{\mathsf{H}}\mathbf{V}\right\} = \|\mathbf{V}\|_F^2 \leq P. \qquad (18)$$

To represent constraint $\mathrm{C}_1$ in terms of $\mathbf{V}$, we note that only the column vectors in $\mathbf{V}$ whose corresponding UT is selected have non-zero entries. This equivalently means that

$$\begin{cases} \|\mathbf{v}_k\| \neq 0 & \text{if UT } k \text{ is selected}\\ \|\mathbf{v}_k\| = 0 & \text{otherwise} \end{cases}, \qquad (19)$$

where $\mathbf{v}_k = \sqrt{p_k}\mathbf{w}_k$ denotes the $k$-th column vector of $\mathbf{V}$. As the result, one can write

$$\|\mathbf{V}\|_{2,0} = \|\mathbf{p}\|_0, \qquad (20)$$

where $\|\mathbf{V}\|_{p,q}$ denotes the $\ell_{p,q}$ norm of $\mathbf{V}$ defined as

$$\|\mathbf{V}\|_{p,q} := \left[\sum_{k=1}^{K}\left(\|\mathbf{v}_k\|_p\right)^q\right]^{1/q}. \qquad (21)$$

From the above derivations, we conclude that the optimal approach for joint user selection and precoding reduces to the following programming:

$$\begin{aligned} \min_{\mathbf{V}\in\mathbb{C}^{M\times K}} \quad & \|\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\,\mathbf{I}_K\|_F^2 \qquad (22)\\ \text{subject to} \quad & \mathrm{C}_1 : \|\mathbf{V}\|_{2,0} \leq L,\\ & \mathrm{C}_2 : \|\mathbf{V}\|_F^2 \leq P. \end{aligned}$$

The optimization in (22) describes a group selection problem in which a matrix with *group sparsity* is to be recovered, i.e., a matrix with a certain fraction of column or row vectors being zero. Such a problem raises in several applications, e.g., distributed compressive sensing and machine leaning [9]–[11]. Group selection in its primitive form is a non-deterministic polynomial time (NP)-hard problem, since it reduces to an integer programming. To address this problem tractably, several suboptimal approaches have been developed in the literature which approximate the solution. Group LASSO is one of the most efficient approaches which relaxes the problem of group selection into a convex programming [12], [13]. In the sequel, we use group LASSO to develop a computationally tractable algorithm for joint user selection and precoding.

### A. A Tractable Algorithm via Group LASSO

Group selection is an extension of the basic *sparse recovery* problem in which a sparse vector is to be recovered from an underdetermined system of equations [14], [15]. Group LASSO extends Tibshirani's regularization approach [16] and convexifies the non-convex $\ell_0$-*norm* with the $\ell_1$-*norm*. This means that constraint $\mathrm{C}_1$ is relaxed as

$$\mathrm{C}_1 : \|\mathbf{V}\|_{2,1} \leq \eta L \qquad (23)$$

for some $\eta$ which regularizes the relaxation. By doing so, the joint user selection and precoding reduces to

$$\begin{aligned} \min_{\mathbf{V}\in\mathbb{C}^{M\times K}} \quad & \|\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\,\mathbf{I}_K\|_F^2 \qquad (24)\\ \text{subject to} \quad & \mathrm{C}_1 : \|\mathbf{V}\|_{2,1} \leq \eta L,\\ & \mathrm{C}_2 : \|\mathbf{V}\|_F^2 \leq P. \end{aligned}$$

This relaxed program represents a group LASSO algorithm which is convex and is posed as a generic linear programming.

### B. An Alternative Formulation via RLS

The joint user selection and precoding scheme in (24) describes *least squares* with side constraints, where the RSS $\|\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\,\mathbf{I}_K\|_F^2$ is minimized subject to some constraints. Following the method of *regularized least-squares (RLS)*, this problem is converted into the following unconstrained optimization[1]

$$\min_{\mathbf{V}\in\mathbb{C}^{M\times K}} \|\mathbf{H}^{\mathsf{T}}\mathbf{V}-\beta\mathbf{I}_K\|_F^2 + \lambda\|\mathbf{V}\|_F^2 + \mu\|\mathbf{V}\|_{2,1} \qquad (25)$$

for some regularizers $\lambda$ a $\mu$. The key features of this algorithm are as follows:

- For given upper bounds on the group sparsity and transmit power of $\mathbf{V}$, there exists a pair of regularizers $\lambda$ and $\mu$, such that the solution to (25) satisfies the constraints. Hence, by tuning $\lambda$ and $\mu$ different constraints are fulfilled.
- Due to its convexity, the problem is tractably solved via generic linear programming. Alternatively, an iterative algorithm based on approximate message passing (AMP) can be developed to find the solution with minimal computational complexity; see [17] for more details on AMP and [18] for its applicatindons to precoding.

Using either the algorithm in (24) or the one in (25), a matrix $\mathbf{V}$ is tractably found which approximates the optimal solution to (22). The beamforming and power allocation matrices are then given by decomposing this matrix as $\mathbf{V} = \mathbf{W}\sqrt{\mathbf{P}}$ for a diagonal $\mathbf{P}$. In the

[1] Alternatively, one could use the method of Lagrange multipliers to conclude the similar unconstrained form.

**Algorithm 1** Joint User Selection and Precoding

**Input:** Channel matrix $\mathbf{H}$, average transmit power $P$ and the number of selected users $L$.

**Set** $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_K]$

$$\mathbf{V} = \mathrm{GroupLASSO}\,(\mathbf{H}, P, L, \beta)$$

**Let** subset $\mathbb{S} \subseteq \{1, \dots, K\}$ contain indices of the column vectors in $\mathbf{V}$ which have the $L$ largest $\ell_2$-norms, i.e., $|\mathbb{S}| = L$ and

$$\|\mathbf{v}_\ell\|^2 \geq \|\mathbf{v}_j\|^2$$

for any $\ell \in \mathbb{S}$ and $j \in \{1, \dots, K\} - \mathbb{S}$.

**Set** $\mathbf{v}_j = 0$ for $j \in \{1, \dots, K\} - \mathbb{S}$, and update $\mathbf{V}$ as

$$\mathbf{V} \leftarrow \frac{\sqrt{P}}{\|\mathbf{V}\|_F}\,\mathbf{V}$$

**Set** $p_k = \|\mathbf{v}_k\|^2$ and $\mathbf{w}_k = \dfrac{\mathbf{v}_k}{\|\mathbf{v}_k\|}$ for $k \in \{1, \dots, K\}$.

**Output:** Beamforming matrix $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$ and power allocation matrix $\mathbf{P} = \mathrm{diag}\,(p_1, \dots, p_K)$.

sequel, we investigate the performance of the proposed approach through some numerical simulations.

## V. Performance Investigation

We study the performance of the proposed approach by simulating some sample scenarios. To jointly precode and select user via group LASSO, Algorithm 1 is used. In this algorithm,

$$\mathbf{V} = \mathrm{GroupLASSO}\,(\mathbf{H}, P, L, \beta) \tag{26}$$

denotes the solution to the minimization in (24) with $\eta = 1$. The algorithm finds first the solution $\mathbf{V}$ to (24), and selects $L$ UTs with strongest precoding vectors while setting the other column vectors zero. It then scales the precoding vectors of the selected users, such that the downlink transmit signal remains $P$.

As a benchmark, we evaluate the performance of maximum ratio transmission (MRT) beamforming with random user selection, and compare it with the performance of Algorthm 1. In this approach, $L$ UTs are selected at random. The precoding vector of selected user $k$ is then set to

$$\mathbf{v}_k = \sqrt{\frac{P}{L}}\,\frac{\mathbf{h}_k^*}{\|\mathbf{h}_k\|}\,. \tag{27}$$

Throughout the simulations the standard Rayleigh model is considered for the fading channel. This means that the entries of $\mathbf{H}$ are generated independently and identically with complex zero-mean and unit-variance Gaussian distribution, i.e.,

$$h_{mk} \sim \mathcal{CN}\,(0, 1) \tag{28}$$

for $m \in \{1, \dots, M\}$ and $k \in \{1, \dots, K\}$.

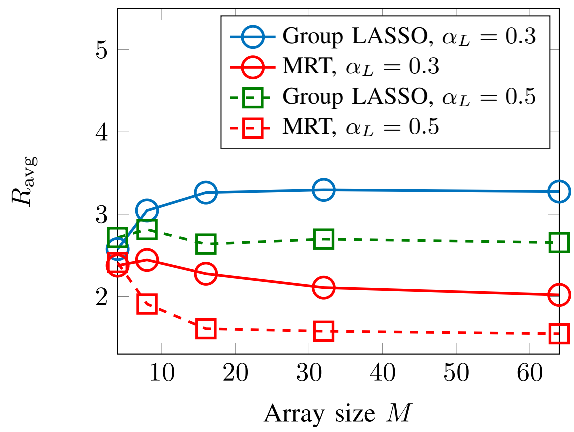


Fig. 1: Average throughput vs. the array size $M$. Here, $P = 1$ and $\sigma_k^2 = 0.1$ for all the UTs. The user load is set to $\alpha_K = 1$, and the scaling factor reads $\beta = 1$.

### A. *Performance Metrics*

To quantify the performance, the following metrics are considered:

1) The weighted average throughput $R_{\mathrm{avg}}$ defined in (3) for uniform wights, i.e., $w_1, \dots, w_K = 1$. This metric determines the average achievable rate per selected UT which is widely used in this literature.
2) The *power leakage* to the non-selected UTs which is given by

$$Q_{\mathrm{Leak}} := \mathbb{E}\left\{ \sum_{k=1, k \notin \mathbb{S}}^{K} |\mathbf{h}_k^{\mathsf{T}} \boldsymbol{x}|^2 \right\} \tag{29a}$$

$$= \sum_{k=1, k \notin \mathbb{S}}^{K} \sum_{\ell \in \mathbb{S}} |\mathbf{h}_k^{\mathsf{T}} \mathbf{v}_\ell|^2. \tag{29b}$$

This metric calculates the total amount of interference at the non-selected UTs from the downlink transmission to the selected UTs.

### B. *Scenario A: Fixed Loads*

We first consider a scenario in which the total number of UTs, as well as the number of selected ones, is a fixed fraction of the transmit array size $M$. More precisely, a downlink transmission scenario is considered in which $K = \lceil \alpha_K M \rceil$ number of users are available and we intend to select $L = \lceil \alpha_L M \rceil$ UTs. Here, $\alpha_K$ and $\alpha_L$ are fixed numbers. For this scenario, both the performance metrics are sketched for fixed transmit power $P$ and noise variance in Fig. 1 and Fig. 2 in terms of the downlink transmit array size $M$.

Fig. 1 shows the weighted average throughput[2] against $M$. Here, $P = 1$ and the noise variances are set to $\sigma_k = 0.1$ for $k \in \{1, \dots, K\}$. Moreover, the scaling factor reads $\beta = 1$. The results are sketched for $\alpha_K = 1$ and two different values of $\alpha_L$; namely, $\alpha_L \in \{0.3, 0.5\}$. As the figure depicts, the proposed approach considerably outperforms the conventional MRT technique. Such an enhancement comes from the

[2]Remember that the average throughput in this case is defined as the sum-rate divided by the number of selected users.

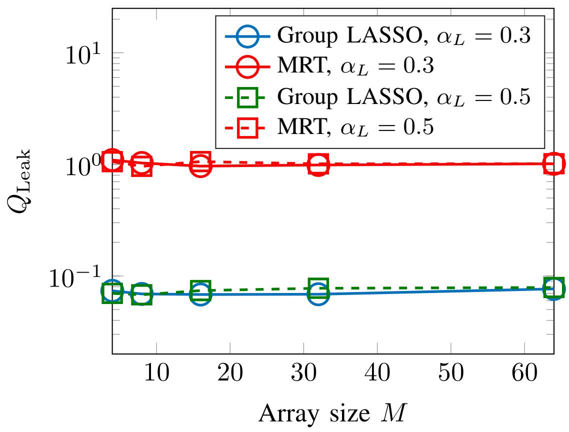


Fig. 2: Power leakage vs. the number of transmit antennas $M$. Here, $P = 1$ and $\sigma_k^2 = 0.1$ for all the UTs. The user load is set to $\alpha_K = 1$, and the scaling factor reads $\beta = 1$.

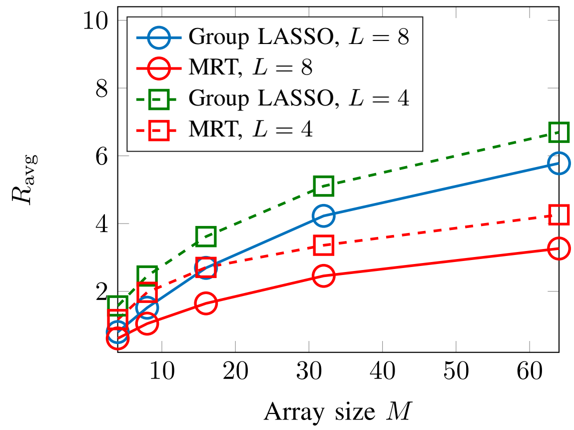


Fig. 3: Average throughput vs. the array size $M$. Here, $P = 1$ and $\sigma_k^2 = 0.1$ for all the UTs. The number of UTs is set to $K = 16$, and the scaling factor reads $\beta = 1$.

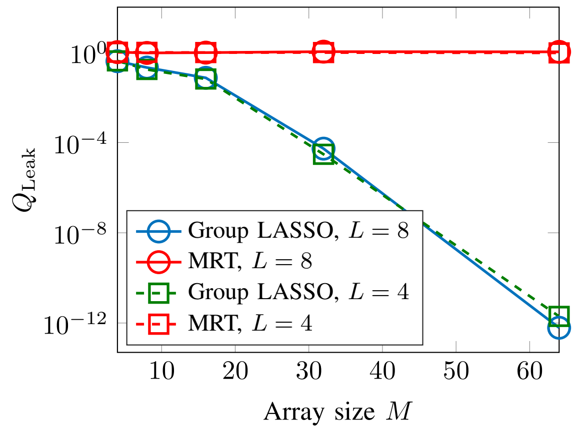


Fig. 4: Power leakage vs. the number of transmit antennas $M$. Here, $P = 1$ and $\sigma_k^2 = 0.1$ for all the UTs. The number of UTs is set to $K = 16$, and the scaling factor reads $\beta = 1$.

joint selection and precoding approach. The convergence of $R_{\text{avg}}$ to a constant in both the techniques follows hardening of the channel in large dimensions for fixed loads [1], [19].

The power leakage for this scenario is plotted in Fig. 2 versus $M$. Here, the parameters are set exactly to the ones considered in Fig. 1. The figure demonstrates the following two observations:

1) The proposed algorithm imposes significantly less interference to the non-selected UTs. This observation comes from the fact that the objective function in (24) contains the power leakage as a penalty term.
2) The power leakage in both techniques converges to a constant value. Such a behavior is naturally following the fact that the loads $\alpha_K$ and $\alpha_L$ are kept fixed.

### *C. Scenario B: Fixed Number of UTs*

As another scenario, we consider a case in which the total number of UTs, as well as the number of selected ones, does not grow with $M$. For this case, we study a settings in which a downlink array of size $M$ is employed to service $L$ users out of $K = 16$ available UTs. Similar to Scenario A, we set $P$ and noise variances to fixed numbers and sketch the average throughput, as well as the power leakage, against the transmit array size $M$ in Fig. 3 and Fig. 4.

In Fig. 3, the average throughput $R_{\text{avg}}$ is sketched against $M$ assuming $\beta = 1$, $P = 1$ and $\sigma_k = 0.1$ for $k \in \{1, \dots, K\}$. The results are given for $L \in \{4, 8\}$. Similar to Scenario A, the figure depicts performance enhancement achieved by using the proposed algorithm based on the group LASSO. In contrast to Scenario A, the throughput in this case grows logarithmically with $M$. Such a behavior follows the fact that in this case, the number of UTs is constant and does not grow with $M$.

Fig. 4 shows the variation of the power leakage against $M$. As the figure demonstrate, in the proposed algorithm, $Q_{\text{Leak}}$ vanishes significantly fast as $M$ grows, such that at $M = 64$ it imposes almost no interference to the non-selected UTs. Such a behavior follows the fact that in the joint approach based on the group LASSO, the beamforming vectors are constructed, such that the power leakage is suppressed at non-selected UTs. For a fixed number of UTs, the suppression is performed more accurately by narrow beamforming towards the selected users, as the array size grows large [20].

## VI. Conclusions

A joint user selection and precoding scheme has been proposed for multiuser MIMO systems based on group LASSO. The scheme depicts performance enhancement in two different aspects: 1) The throughput of the system, defined as the sum-rate divided by the number of active users, shows some gains. 2) The interference imposed by downlink transmission at the non-selected UTs is significantly reduced. For instance, when $L = 8$ UTs are selected out of $K = 16$ users, there is almost zero interference, when the BS is equipped with $M = 64$ antennas. These observations indicate that the proposed scheme is a good candidate for massive MIMO settings.

The current work can be pursued in various directions. For example, considering the RLS-based derivation in (25), an iterative algorithm can be developed via AMP implementing the proposed scheme with low computational complexity. Another direction is to extend the current framework to wiretap settings following the approach in [21]. The work in these directions is currently ongoing.

## REFERENCES


[1] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[2] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: Performance analysis and a simple new algorithm," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3857–3868, 2005.

[3] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," in *Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005.* IEEE, 2005, pp. 628–632.

[4] J. Wang, D. J. Love, and M. D. Zoltowski, "User selection with zero-forcing beamforming achieves the asymptotically optimal sum rate," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3713–3726, 2008.

[5] S. Huang, H. Yin, J. Wu, and V. C. Leung, "User selection for multiuser MIMO downlink with zero-forcing beamforming," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3084–3097, 2013.

[6] A. Bereyhi, M. A. Sedaghat, S. Asaad, and R. R. Müller, "Nonlinear precoders for massive MIMO systems with general constraints," *International ITG Workshop on Smart Antennas (WSA)*, 2017.

[7] A. Bereyhi, M. A. Sedaghat, and R. R. Müller, "Asymptotics of nonlinear LSE precoders with applications to transmit antenna selection," in *International Symposium on Information Theory (ISIT).* IEEE, 2017, pp. 81–85.

[8] A. Bereyhi, M. A. Sedaghat, R. R. Müller, and G. Fischer, "GLSE precoders for massive MIMO systems: Analysis and applications," *arXiv preprint arXiv:1808.01880*, 2018.

[9] S. Sarvotham, D. Baron, M. Wakin, M. F. Duarte, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Asilomar conference on signals, systems, and computers*, 2005, pp. 1537–1541.

[10] A. Bereyhi, S. Haghighatshoar, and R. R. Müller, "Theoretical bounds on MAP estimation in distributed sensing networks," in *IEEE International Symposium on Information Theory (ISIT).* IEEE, 2018, pp. 1385–1389.

[11] F. R. Bach, "Consistency of the group lasso and multiple kernel learning," *Journal of Machine Learning Research*, vol. 9, pp. 1179–1225, 2008.

[12] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[13] W. Deng, W. Yin, and Y. Zhang, "Group sparse optimization by alternating direction method," in *Wavelets and Sparsity XV*, vol. 8858. International Society for Optics and Photonics, 2013, p. 88580R.

[14] D. L. Donoho *et al.*, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[15] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[16] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[17] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *2011 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2011, pp. 2168–2172.

[18] A. Bereyhi, M. A. Sedaghat, and R. R. Müller, "Precoding via approximate message passing with instantaneous signal constraints," in *Proceedings of The International Zurich Seminar on Information and Communication (IZS).* ETH Zurich, 2018, pp. 128–132.

[19] S. Asaad, A. M. Rabiei, and R. R. Müller, "Massive MIMO with antenna selection: Fundamental limits and applications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8502–8516, 2018.

[20] A. Bereyhi, S. Asaad, R. R. Muller, R. F. Schaefer, and A. M. Rabiei, "On robustness of massive MIMO systems against passive eavesdropping under antenna selection," in *2018 IEEE Global Communications Conference (GLOBECOM).* IEEE, 2018, pp. 1–7.

[21] S. Asaad, A. Bereyhi, R. R. Müller, R. F. Schaefer, and H. V. Poor, "Secure regularized zero forcing for multiuser MIMOME channels," *submitted to* $53^{\rm rd}$ *Annual Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019.