Minorization-Maximization Algorithms for Codebook based Downlink Sum-Rate Maximization in TDD Multiuser Large MIMO Broadcast Systems

Sai Subramanyam Thoota

SPC Lab, Department of ECE
Indian Institute of Science

August 5, 2017

Table of contents

1. Motivation
2. Goal & Contributions
3. System Model & Problem Statement
4. Minorization-Maximization Principle
5. Proposed Algorithms
6. Simulation Results
7. Summary & Future Work

Motivation

- Design of downlink (DL) precoding and beamforming schemes for multiuser (MU) multiple input multiple output (MIMO) systems with a large number of antennas at the base station (BS) has attracted significant research interest.
- High control overhead in the uplink (UL) and DL to convey the channel state information (CSI) to the BS, and the precoding matrices to the user equipments (UE), respectively.
- CSI overhead can be avoided using reverse channel training (RCT) in time division duplex (TDD) systems. Pilots transmited in the uplink (UL) by the UEs, which will be used to estimate the DL channel estimates.
- LTE/LTE-A uses codebook based beamforming, which uses a predetermined set of precoding vectors to be chosen during DL transmission. The same scheme can be adopted in the next generation wireless communication systems with a larger codebook size.

Goal & Contributions

Goal

- Multiple data streams are simultaneously transmitted to all users, after precoding each data stream with a beamforming vector selected from a predetermined codebook.
- Goal is to determine the selection of beamforming vectors and power allocation to each user to maximize the achievable sum rate.

Contributions

- The problem is non-convex and combinatorial in nature, which is reformulated to facilitate the use of a minorization-maximization (MM) approach to solve it in a computationally efficient manner.
- Two iterative algorithms are proposed, which involve the use of the MM approach in a nested manner, to obtain a convex objective function, which is solved in closed form.
- In the high SNR regime, lower complexity variants of the algorithms are proposed. The performance of the proposed algorithms is benchmarked against existing approaches.

System Model

- Scenario: Single cell MU-MIMO broadcast system with one BS and K users.
- System Parameters:
  - Number of transmit antennas - N_t
  - Number of receive antennas - N_r
  - Size of the codebook - N
  - Beamforming vectors - $\mathbf{v}_l \in \mathbb{C}^{N_t \times 1}, l = 1, \ldots, N$
  - Channel matrix of the k^{th} user - $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$
  - Codebook $\mathbf{C} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N] \in \mathbb{C}^{N_t \times N}$
- Transmit signal:

$$\mathbf{x} = \sum_{j=1}^{K} \sum_{l=1}^{N} \mathbf{v}_l s_j(l) \tag{1}$$

- Received signal of the k^{th} user:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{w}_k = \mathbf{H}_k \sum_{j=1}^{K} \mathbf{C}\mathbf{s}_j + \mathbf{w}_k, \tag{2}$$

where $\mathbf{w}_k \in \mathbb{C}^{N_r \times 1}$ is the complex additive white Gaussian noise of the k^{th} UE with distribution $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_r})$, and $\mathbf{s}_j = [s_j(1), \ldots, s_j(N)]^T$.

System Model

- The rate achievable for the k^{th} user is given by

$$R_k = \log \left| \mathbf{I}_{N_r} + \mathbf{V}_k^{-1} \mathbf{H}_k \mathbf{C} \boldsymbol{\Phi}_k \mathbf{C}^H \mathbf{H}_k^H \right|, \tag{3}$$

where

$$\mathbf{V}_k = \sigma^2 \mathbf{I}_{N_r} + \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{H}_k \mathbf{C} \boldsymbol{\Phi}_j \mathbf{C}^H \mathbf{H}_k^H \tag{4}$$

is the interference plus noise covariance matrix, and $\boldsymbol{\Phi}_k = \text{diag}\left([P_k(1), P_k(2), ..., P_k(N)]\right)$ is the transmit signal covariance matrix of the k^{th} user.

- The DL sum rate is given by

$$R_{\text{tot}} = \sum_{k=1}^{K} \log \left| \mathbf{I}_{N_r} + \mathbf{V}_k^{-1} \tilde{\mathbf{H}}_k \boldsymbol{\Phi}_k \tilde{\mathbf{H}}_k^H \right|, \tag{5}$$

where $\tilde{\mathbf{H}}_k = \mathbf{H}_k \mathbf{C}$.

Problem Statement

- The objective is to maximize the sum rate R_{tot} given in (5) under a maximum total power constraint. Mathematically, the problem statement is given by

$$\begin{aligned} \underset{\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_K}{\text{maximize}} & \sum_{k=1}^{K} \log \left| \mathbf{I}_{N_r} + \mathbf{V}_k^{-1} \tilde{\mathbf{H}}_k \mathbf{\Phi}_k \tilde{\mathbf{H}}_k^H \right|, \\ \text{s. t.} \quad & \text{Tr}\left(\sum_{k=1}^{K} \mathbf{\Phi}_k \right) \leq P_{\text{max}} \end{aligned} \tag{6}$$

  where P_{max} is the maximum total transmit power allowed at the BS.
- The optimization problem in (6) is nonconvex in $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K$ and combinatorial in nature, which cannot be solved in closed-form.
- We propose two algorithms based on the MM principle, which is briefly explained in the next slide.

Minorization-Maximization Principle

- MM principle proceeds by solving a simple convex optimization problem in place of a complex non-convex optimization problem.
- Surrogate convex function which bounds the objective function either from above (for minimization) or below (for maximization) is computed.
- A function $g(x|x^{(m)})$ is said to minorize a real-valued function $f(x)$ at $x^{(m)}$ if

$$g(x|x^{(m)}) \leq f(x), \forall x \in \mathbb{C}$$
$$g(x^{(m)}|x^{(m)}) = f(x^{(m)})$$

- The algorithm proceeds by maximizing the surrogate function $g(x|x^{(m)})$ and finding the next iterate $x^{(m+1)}$.
- Since $g(x|x^{(m)}) \leq f(x)\ \forall x$, if the maximum of $g(x|x^{(m)})$ is achieved at $x^{(m+1)}$, then $f(x^{(m+1)}) \geq g(x^{(m+1)}|x^{(m)}) \geq g(x^{(m)}|x^{(m)}) = f(x^{(m)})$, i.e., the original objective function either increases or remains unchanged.

Proposed Algorithms

- We propose two algorithms, which we refer to as the square root MM (SMM) and inverse MM (IMM).
- Both algorithms start by executing a preliminary minorization step, and then proceed to solve the resulting optimization problem by two different approaches.
- Preliminary step to find a surrogate function to lower bound the sum rate uses a lemma discussed in the next slide.

Preliminary Step

Lemma 1

For matrices $\mathbf{Z}$, $\mathbf{Y} \succeq 0$, *the non-convex function*

$$f(\mathbf{Z},\mathbf{Y}) = \log\left|\mathbf{Z}^{-1}\mathbf{Y}\right| \tag{7}$$

can be lower bounded by

$$\begin{aligned} f(\mathbf{Z},\mathbf{Y}) \geq -\Big(&\log\left|\mathbf{Z}^{(m)}\right| + \operatorname{Tr}\left(\mathbf{Z}^{(m)^{-1}}\left(\mathbf{Z}-\mathbf{Z}^{(m)}\right)\right)\\ &+\log\left|\mathbf{Y}^{(m)^{-1}}\right| + \operatorname{Tr}\left(\mathbf{Y}^{(m)}\left(\mathbf{Y}^{-1}-\mathbf{Y}^{(m)^{-1}}\right)\right)\Big) \end{aligned} \tag{8}$$

with equality at $\mathbf{Z} = \mathbf{Z}^{(m)}$ *and* $\mathbf{Y} = \mathbf{Y}^{(m)}$.
Proof:

$$f(\mathbf{Z},\mathbf{Y}) = -\log|\mathbf{Z}| + \log|\mathbf{Y}|\,. \tag{9}$$

Here, $\mathbf{Z}$ *and* $\mathbf{Y}$ *are positive semidefinite (p.s.d) matrices, and so the function* f *is convex in* $\mathbf{Z}$ *and* $\mathbf{Y}^{-1}$. *Hence we can bound it from below using the first order Taylor series expansion, resulting in the lower bound given by* (8).

- Let $\mathbf{B}_k \triangleq \sigma^2 \mathbf{I}_{N_r} + \sum_{j=1}^{K} \tilde{\mathbf{H}}_k \mathbf{\Phi}_j \tilde{\mathbf{H}}_k^H$.
- The rate of the k^{th} user can be written as $R_k = \log \left| \mathbf{V}_k^{-1} \mathbf{B}_k \right|$. Applying Lemma 1 to the sum rate objective function, we get the following surrogate optimization problem for (6):

$$\begin{aligned} \{\mathbf{\Phi}_1^{(m+1)}, \ldots, \mathbf{\Phi}_K^{(m+1)}\} = \\ \underset{\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K}{\text{argmax}} \sum_{k=1}^{K} \Bigg\{ - \operatorname{Tr} \Bigg({\mathbf{V}_k^{(m)}}^{-1} \Bigg(\sigma^2 \mathbf{I}_{N_r} + \sum_{\substack{j=1 \\ j \neq k}}^{K} \tilde{\mathbf{H}}_k \mathbf{\Phi}_j \tilde{\mathbf{H}}_k^H \Bigg) \Bigg) \\ - \operatorname{Tr} \Bigg(\mathbf{B}_k^{(m)} \Bigg[\sigma^2 \mathbf{I}_{N_r} + \sum_{j=1}^{K} \tilde{\mathbf{H}}_k \mathbf{\Phi}_j \tilde{\mathbf{H}}_k^H \Bigg]^{-1} \Bigg) \Bigg\}, \quad (10) \\ \text{subject to} \quad \operatorname{Tr} \left(\sum_{k=1}^{K} \mathbf{\Phi}_k \right) \leq P_{\max}, \end{aligned}$$

where m is the iteration index.

- We define some notation before explaining the SMM and IMM algorithms. Let

$$\mathbf{\Phi} \triangleq \text{diag}\left(\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K\right), \tag{11}$$

$$\mathbf{\Psi}_k \triangleq \left[\tilde{\mathbf{H}}_k, \ldots, \tilde{\mathbf{H}}_k\right], k = 1, \ldots, K \tag{12}$$

denote the augmented power allocation and the k^{th} user's channel matrices, respectively. In (12), $\tilde{\mathbf{H}}_k$ is repeated K times. Also, let

$$\mathbf{Q} \triangleq \sum_{k=1}^{K} \text{diag}\left(\tilde{\mathbf{H}}_k^H \mathbf{V}_k^{-1} \tilde{\mathbf{H}}_k, \ldots, \mathbf{0}_N, \ldots, \tilde{\mathbf{H}}_k^H \mathbf{V}_k^{-1} \tilde{\mathbf{H}}_k\right). \tag{13}$$

In the above, the $N \times N$ all zero matrix $\mathbf{0}_N$ is in the k^{th} block diagonal position of $\mathbf{Q}$.

- The first term in (10) can be written as $\text{Tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi}\right)$, where the superscript m denotes the iteration index, and $\mathbf{Q}^{(m)}$ is obtained by substituting $\mathbf{V}_k^{(m)}$ for $\mathbf{V}_k$ in (13).

Square root MM (SMM) Algorithm

- SMM proceeds by working with the square root of the power allocation matrix $\mathbf{\Phi}$.
- Two stages of minorization to bound the objective function from below.
- The second term in (10) can be written as

$$-\sum_{k=1}^{K} \text{Tr}\left(\mathbf{F}_k^{(m)} \left(\sigma^2 \mathbf{I}_{N_r} + \mathbf{\Psi}_k \mathbf{\Phi} \mathbf{\Psi}_k^H \right)^{-1} {\mathbf{F}_k^{(m)}}^H \right), \tag{14}$$

  where $\mathbf{F}_k$ denotes a matrix such that $\mathbf{B}_k = \mathbf{F}_k^H \mathbf{F}_k$ [1].
- The above cost function cannot be directly optimized due to the matrix inversion involved. Hence, we minorize it using the following lemma.

[1] It can be computed, for example, via the Cholesky decomposition of $\mathbf{B}_k$.

Lemma 2

Let $\mathbf{R}$ *denote a diagonal p.s.d. square matrix, and consider the function* $f(\mathbf{R})$ *defined as*

$$f(\mathbf{R}) = -\operatorname{Tr}\left(\mathbf{A}\left(\mathbf{B} + \mathbf{C}\mathbf{R}\mathbf{C}^H\right)^{-1}\mathbf{A}^H\right), \tag{15}$$

where $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{C}$ *are matrices of appropriate dimensions, and* $\mathbf{B}$ *is positive definite, ensuring that* $\mathbf{B} + \mathbf{C}\mathbf{Q}\mathbf{C}^H$ *is invertible. Then, for a given diagonal p.s.d. matrix* $\mathbf{R}^{(m)}$, $f(\mathbf{R})$ *can be lower bounded by*

$$f(\mathbf{R}) \geq g(\mathbf{R}|\mathbf{R}^{(m)}) = -\operatorname{Tr}(\hat{\mathbf{K}}) + \operatorname{Tr}\left(\left(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\mathbf{A}\mathbf{B}^{-1}\mathbf{C} + \mathbf{C}^H\mathbf{B}^{-H}\mathbf{A}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\right)\mathbf{R}^{\frac{1}{2}} - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right), \tag{16}$$

where

$$\hat{\mathbf{X}} \triangleq \mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{(m)\frac{1}{2}}, \hat{\mathbf{Y}} \triangleq \mathbf{I} + \mathbf{R}^{(m)\frac{1}{2}}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{(m)\frac{1}{2}}, \tag{17}$$

$$\hat{\mathbf{K}} \triangleq \mathbf{A}\mathbf{B}^{-1}\mathbf{A}^H + \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}} - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\hat{\mathbf{Y}} + \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1} + \hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H. \tag{18}$$

Also, $g(\mathbf{R}^{(m)}|\mathbf{R}^{(m)}) = f(\mathbf{R}^{(m)})$, *i.e., the lower bound is tight at* $\mathbf{R}^{(m)}$.

Proof of Lemma 2.

Using the Woodbury matrix inversion identity for the inverse term in the function f in (15), we get

$$-\operatorname{Tr}\left(\mathbf{A}\left(\mathbf{B}^{-1}-\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\left(\mathbf{I}+\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right)^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\right)\mathbf{A}^{H}\right)$$
$$=-\operatorname{Tr}\left(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^{H}\right)+\operatorname{Tr}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right), \tag{19}$$

where

$$\mathbf{X}\triangleq\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}},\mathbf{Y}\triangleq\mathbf{I}+\mathbf{R}^{\frac{1}{2}}\mathbf{C}^{H}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}. \tag{20}$$

The function $\operatorname{Tr}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right)$ is jointly convex in $\mathbf{X}$ and $\mathbf{Y}$, and, can be minorized using a first order Taylor series. The complex matrix differential of $\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}$ is computed as follows:

$$\operatorname{Tr}\left(\mathrm{d}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^{H}\right)\right)=\operatorname{Tr}\left(\mathbf{Y}^{-1}\mathbf{X}^{H}\mathrm{d}\mathbf{X}-\mathbf{Y}^{-1}\mathbf{X}^{H}\mathbf{X}\mathbf{Y}^{-1}\mathrm{d}\mathbf{Y}+\mathbf{X}\mathbf{Y}^{-1}\mathrm{d}\mathbf{X}^{H}\right). \tag{21}$$

□

Proof of Lemma 2 contd.

Thus, around the point $(\hat{\mathbf{X}}, \hat{\mathbf{Y}})$, (19) can be lower bounded as

$$\begin{aligned}
&-\operatorname{Tr}(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^H) + \operatorname{Tr}\left(\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^H\right) \\
&\geq -\operatorname{Tr}(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^H) + \operatorname{Tr}\left(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\left(\mathbf{X}-\hat{\mathbf{X}}\right) - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\left(\mathbf{Y}-\hat{\mathbf{Y}}\right) + \hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\left(\mathbf{X}-\hat{\mathbf{X}}\right)^H\right) \\
&= -\operatorname{Tr}(\hat{\mathbf{K}}) + \operatorname{Tr}\left(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}} - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}} + \hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\left(\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right)^H\right) \\
&= -\operatorname{Tr}(\hat{\mathbf{K}}) + \operatorname{Tr}\left(\hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\mathbf{A}\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}} + \mathbf{C}^H\mathbf{B}^{-H}\mathbf{A}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}} - \hat{\mathbf{Y}}^{-1}\hat{\mathbf{X}}^H\hat{\mathbf{X}}\hat{\mathbf{Y}}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{C}^H\mathbf{B}^{-1}\mathbf{C}\mathbf{R}^{\frac{1}{2}}\right),
\end{aligned} \tag{22}$$

where $\hat{\mathbf{K}}$ is as defined in Lemma 2. Grouping all the constant matrices in (22) together, we get (16). □

- Applying Lemma 2 to (14), we get

$$-\operatorname{Tr}\left(\mathbf{W}_{1,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_{2,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right), \tag{23}$$

where

$$\mathbf{W}_{1,k} \triangleq -\left\{\frac{\mathbf{Y}_k^{-1}\mathbf{X}_k^H\mathbf{F}_k\mathbf{\Psi}_k + \mathbf{\Psi}_k^H\mathbf{F}_k^H\mathbf{X}_k\mathbf{Y}_k^{-1}}{\sigma^2}\right\}, \tag{24}$$

$$\mathbf{W}_{2,k} \triangleq \mathbf{Y}_k^{-1}\mathbf{X}_k^H\mathbf{X}_k\mathbf{Y}_k^{-1}, \tag{25}$$

and $\mathbf{S}_k \in \mathbb{C}^{KN\times KN}, \mathbf{X}_k \in \mathbb{C}^{N_r\times KN}$ and $\mathbf{Y}_k \in \mathbb{C}^{KN\times KN}$ are defined as follows:

$$\mathbf{S}_k \triangleq \frac{\mathbf{\Psi}_k^H\mathbf{\Psi}_k}{\sigma^2}, \mathbf{X}_k \triangleq \frac{\mathbf{F}_K\mathbf{\Psi}_k\mathbf{\Phi}^{\frac{1}{2}}}{\sigma^2}, \mathbf{Y}_k \triangleq \mathbf{I}_{KN} + \mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}. \tag{26}$$

- Note that, $\mathbf{W}_{1,k}$ and $\mathbf{W}_{2,k}$ in (24) and (25) are negative and positive semidefinite matrices, respectively.
- The surrogate cost function is not yet amenable to a closed-form solution due to the $\mathbf{W}_{2,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}$ term in (23).

- The following lemma is used to minorize the second term in (23).

Lemma 3

Suppose $\mathbf{R}$ *is a p.s.d. diagonal matrix, and* $\mathbf{A}$ *and* $\mathbf{B}$ *are symmetric p.s.d. square matrices. Then, the function*

$$f(\mathbf{R}) = -\operatorname{Tr}(\mathbf{ARBR}) \tag{27}$$

can be lower bounded by

$$\begin{aligned} f(\mathbf{R}) \geq &-\operatorname{Tr}\left(\mathbf{AR}^{(m)}\mathbf{BR}^{(m)} - \left((\mathbf{B}-\lambda\mathbf{I})\mathbf{R}^{(m)}\mathbf{A} + \mathbf{AR}^{(m)}(\mathbf{B}-\lambda\mathbf{I})\right)\mathbf{R}^{(m)}\right) \\ &-\operatorname{Tr}\left(\left((\mathbf{B}-\lambda\mathbf{I})\mathbf{R}^{(m)}\mathbf{A} + \mathbf{AR}^{(m)}(\mathbf{B}-\lambda\mathbf{I})\right)\mathbf{R}\right) - \lambda\operatorname{Tr}\left(\mathbf{AR}^{2}\right), \end{aligned} \tag{28}$$

where λ *is the largest eigenvalue of* $\mathbf{B}$*. Further, we have equality in* (28) *at* $\mathbf{R} = \mathbf{R}^{(m)}$.

Proof.

We lower bound the function $f(\mathbf{R})$ using λ, the largest eigenvalue of the matrix $\mathbf{B}$, as follows:

$$\begin{aligned} f(\mathbf{R}) &= -\operatorname{Tr}\left(\mathbf{AR}(\mathbf{B}-\lambda\mathbf{I})\mathbf{R} + \lambda\mathbf{AR}^2\right) \\ &= -\operatorname{Tr}\left(\mathbf{ARCR} + \lambda\mathbf{AR}^2\right), \end{aligned} \tag{29}$$

where $\mathbf{C} \triangleq (\mathbf{B}-\lambda\mathbf{I})$. The complex matrix differential of the first term in (29) is

$$\operatorname{Tr}\left(\mathrm{d}(\mathbf{ARCR})\right) = \operatorname{Tr}\left(\mathbf{CRA}(\mathrm{d}\mathbf{R}) + \mathbf{ARC}(\mathrm{d}\mathbf{R})\right). \tag{30}$$

Hence, around the previous iterate $\mathbf{R}^{(m)}$, a lower bound on f can be written as

$$f(\mathbf{R}) \geq -\operatorname{Tr}\left(\mathbf{AR}^{(m)}\mathbf{BR}^{(m)}\right) - \operatorname{Tr}\left(\left(\mathbf{CR}^{(m)}\mathbf{A} + \mathbf{AR}^{(m)}\mathbf{C}\right)\left(\mathbf{R}-\mathbf{R}^{(m)}\right)\right) - \lambda\operatorname{Tr}\left(\mathbf{AR}^2\right). \tag{31}$$

Grouping the constant terms in (31) together and substituting for $\mathbf{C}$, we get (28). □

- Applying Lemma 3 to (23), we get the final lower bound for (14) as follows:

$$-\sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{W}_{1,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_{2,k}^{(m)}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{S}_k\mathbf{\Phi}^{\frac{1}{2}}\right) \geq -\mathrm{Tr}\left(\mathbf{W}_{A}^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_{B}^{(m)}\mathbf{\Phi}\right), \tag{32}$$

where

$$\mathbf{W}_A \triangleq \sum_{k=1}^{K}\left\{\mathbf{W}_{1,k} + \left(\mathbf{S}_k - \lambda_{\max}\left(\mathbf{S}_k\right)\mathbf{I}_{KN}\right)\mathbf{\Phi}^{\frac{1}{2}}\mathbf{W}_{2,k}\right\}, \tag{33}$$

$$\mathbf{W}_B \triangleq \sum_{k=1}^{K}\left\{\lambda_{\max}\left(\mathbf{S}_k\right)\mathbf{W}_{2,k}\right\}, \tag{34}$$

and $\lambda_{\max}(\mathbf{S}_k)$ is the largest eigenvalue of the matrix $\mathbf{S}_k$.

- Combining (32) with $\mathrm{Tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi}\right)$ (above (13)), the optimization problem becomes

$$\{\mathbf{\Phi}^{(m+1)}\} = \underset{\mathbf{\Phi}}{\mathrm{argmax}}\left\{-\mathrm{Tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi} + \mathbf{W}_A^{(m)}\mathbf{\Phi}^{\frac{1}{2}} + \mathbf{W}_B^{(m)}\mathbf{\Phi}\right)\right\} \quad (35)$$
$$\text{subject to} \quad \mathrm{Tr}\left(\mathbf{\Phi}\right) \leq P_{\max}.$$

- The Lagrangian for (35) is given by

$$\sum_{i=1}^{KN}\left([\mathbf{Q}^{(m)}]_{(i,i)}P(i) + [\mathbf{W}_A^{(m)}]_{(i,i)}P(i)^{\frac{1}{2}} + [\mathbf{W}_B^{(m)}]_{(i,i)}P(i)\right) + \eta\left(\sum_{i=1}^{KN}P(i) - P_{\max}\right), \quad (36)$$

where $P(i), i = 1, 2, \ldots, KN$ denotes the diagonal entries of the matrix $\mathbf{\Phi}$.
- By differentiation w.r.t $P(i)$ in (36), we obtain the closed form solution

$$P(i) = \left(\frac{\left[\mathbf{W}_A^{(m)}\right]_{(i,i)}}{2\left(\left[\mathbf{W}_B^{(m)}\right]_{(i,i)} + [\mathbf{Q}^{(m)}]_{(i,i)} + \eta\right)}\right)^2, \quad (37)$$

where η is chosen to satisfy the constraint $\sum_{i=1}^{KN} P(i) = P_{\max}$.

Inverse MM (IMM) Algorithm

- Alternative bounding approach which leads to a different MM procedure for sum rate maximization.
- For convenience, we define the augmented covariance matrix $\tilde{\mathbf{\Phi}}$, the augmented channel matrix $\tilde{\mathbf{\Psi}}_k$ and the matrix $\mathbf{\Xi}_k$ as follows:

$$\tilde{\mathbf{\Phi}} \triangleq \text{diag}\left(\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K, \sigma^2 \mathbf{I}_{N_r}\right) \in \mathbb{R}^{(KN+N_r)\times(KN+N_r)}, \tag{38}$$

$$\tilde{\mathbf{\Psi}}_k \triangleq \left[\tilde{\mathbf{H}}_k, \ldots, \tilde{\mathbf{H}}_k, \mathbf{I}_{N_r}\right], \in \mathbb{C}^{N_r \times (KN+N_r)}, \tag{39}$$

$$\mathbf{\Xi}_k \triangleq \tilde{\mathbf{\Psi}}_k \tilde{\mathbf{\Phi}} \tilde{\mathbf{\Psi}}_k^H, \in \mathbb{C}^{N_r \times N_r}. \tag{40}$$

- In (39), the matrix $\tilde{\mathbf{H}}_k$ is repeated K times. We can rewrite the term inside the square brackets in (10) as $\mathbf{B}_k^{(m)} \mathbf{\Xi}_k^{-1}$.
- Note that $\mathbf{B}_k^{(m)}$ is a p.s.d. matrix.

- In order to develop the IMM procedure, we start with the following proposition.

Proposition 1

Let $\mathbf{R}$ *be an* $m \times n$ *matrix and* $\mathbf{A}$ *be an* $m \times m$ *p.s.d. matrix. We can upper bound the function* $f(\mathbf{U}) \triangleq \text{Tr}\left(\mathbf{A}\left(\mathbf{R}\mathbf{U}\mathbf{R}^H\right)^{-1}\right)$ *as*

$$f(\mathbf{U}) \leq \text{Tr}\left(\mathbf{A}\left(\mathbf{R}\mathbf{U}^{(m)}\mathbf{R}^H\right)^{-1}\mathbf{R}\mathbf{U}^{(m)}\mathbf{U}^{-1}\mathbf{U}^{(m)}\mathbf{R}^H\left(\mathbf{R}\mathbf{U}^{(m)}\mathbf{R}^H\right)^{-1}\right), \tag{41}$$

with equality at $\mathbf{U} = \mathbf{U}^{(m)}$.

- Applying proposition 1 to $\text{Tr}\left(\mathbf{B}_k^{(m)}\mathbf{\Xi}_k^{-1}\right)$ results in the bound given below.

$$\sum_{k=1}^{K}\text{Tr}\left(\mathbf{B}_k^{(m)}\mathbf{\Xi}_k^{-1}\right) \leq \sum_{k=1}^{K}\text{Tr}\left(\mathbf{B}_k^{(m)}\mathbf{\Xi}_k^{(m)^{-1}}\tilde{\mathbf{\Psi}}_k\tilde{\mathbf{\Phi}}^{(m)}\tilde{\mathbf{\Phi}}^{-1}\tilde{\mathbf{\Phi}}^{(m)}\tilde{\mathbf{\Psi}}_k^H\mathbf{\Xi}_k^{(m)^{-1}}\right)$$

$$= \text{Tr}\left(\sum_{k=1}^{K}\tilde{\mathbf{\Phi}}^{(m)}\tilde{\mathbf{\Psi}}_k^H\mathbf{\Xi}_k^{(m)^{-1}}\tilde{\mathbf{\Psi}}_k\tilde{\mathbf{\Phi}}^{(m)}\tilde{\mathbf{\Phi}}^{-1}\right), \tag{42}$$

- Let $\mathbf{Z} \triangleq \sum_{k=1}^{K} \tilde{\mathbf{\Phi}} \tilde{\mathbf{\Psi}}_k^H {\mathbf{\Xi}_k}^{-1} \tilde{\mathbf{\Psi}}_k \tilde{\mathbf{\Phi}}$. Substituting $\mathrm{Tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi}\right)$ (above (13)) and (42) into (6), we get the following surrogate optimization problem:

$$\mathbf{\Phi}^{(m+1)} = \underset{\mathbf{\Phi} \succeq 0}{\mathrm{argmax}} \left\{ -\mathrm{Tr}\left(\mathbf{Q}^{(m)}\mathbf{\Phi} + \mathbf{Z}^{(m)}\tilde{\mathbf{\Phi}}^{-1}\right)\right\} \quad (43)$$
$$\text{subject to } \mathrm{Tr}\left(\mathbf{\Phi}\right) \leq P_{\max},$$

where m is the iteration index. The above cost function is quadratic in $\mathbf{\Phi}$.
- The Lagrangian is given by

$$\sum_{i=1}^{KN} \left(\left[\mathbf{Q}^{(m)}\right]_{(i,i)} P(i) + \left[\mathbf{Z}^{(m)}\right]_{(i,i)} \frac{1}{P(i)} \right) + \eta \left(\sum_{i=1}^{KN} P(i) - P_{\max} \right). \quad (44)$$

- The solution to the surrogate problem is

$$P(i) = \left(\frac{\left[\mathbf{Z}^{(m)}\right]_{(i,i)}}{[\mathbf{Q}^{(m)}]_{(i,i)} + \eta} \right)^{\frac{1}{2}}, \forall i = 1, 2, \ldots, KN. \quad (45)$$

Simulation Results

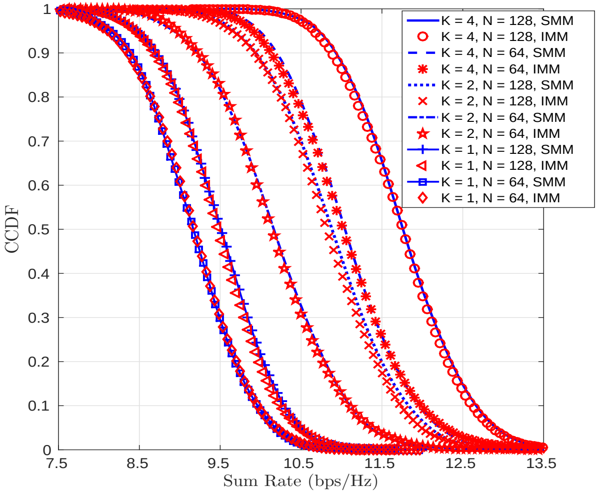


Figure: CCDF comparison between SMM and IMM for SNR $= 0$ dB, $N_t = 16$ and $P_{\max} = 40$ dBm.

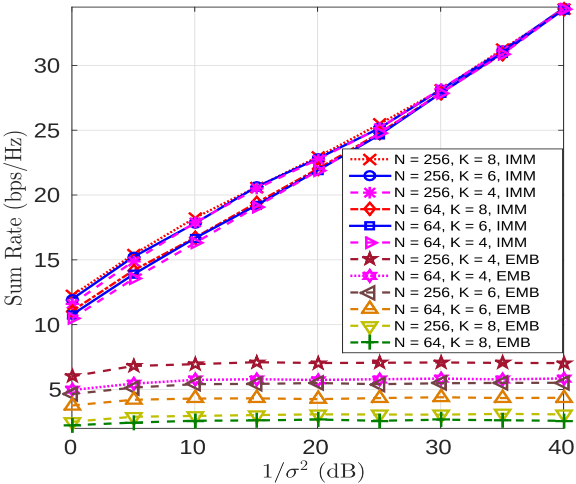
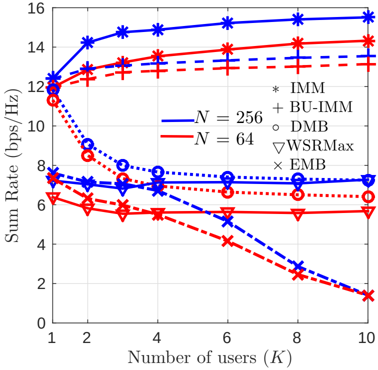
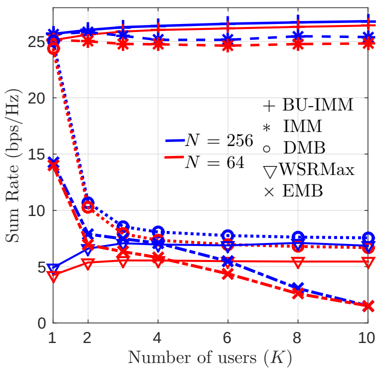


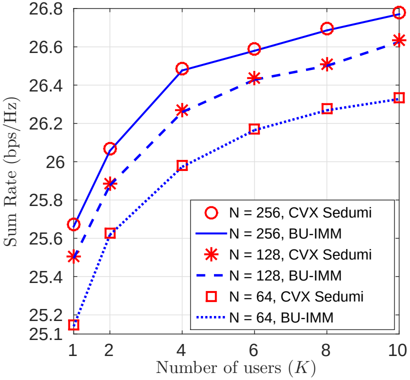
Figure: Sum rate vs. SNR (which is proportional to $1/\sigma^2$), with $N_t = 8$ and $P_{\max} = 40$ dBm.



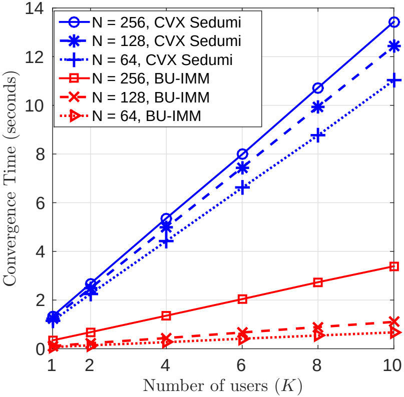
(a) Sum rate vs. K, SNR $= 5$ dB, $P_{\max} = 40$ dBm.



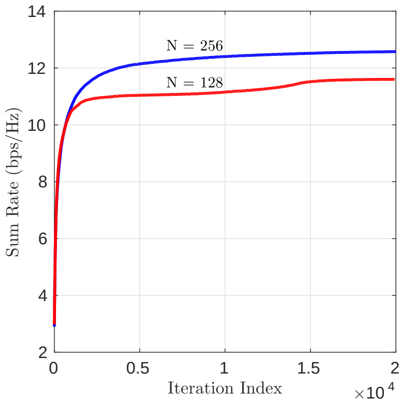
(b) Sum rate vs. K, SNR $= 25$ dB, $P_{\max} = 40$ dBm.



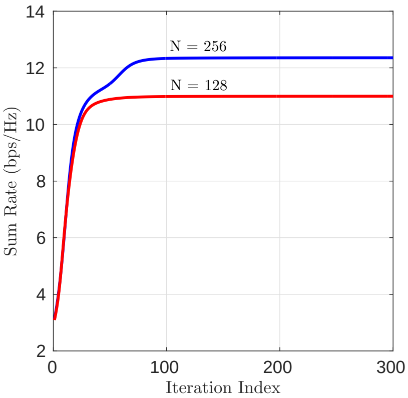
(c) Sum Rate vs K, SNR $=$ 25 dB, $P_{\text{max}} = 40$ dBm.



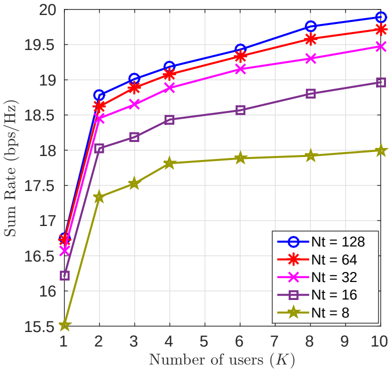
(d) Convergence Time vs K, SNR $=$ 25 dB, $P_{\text{max}} = 40$ dBm.



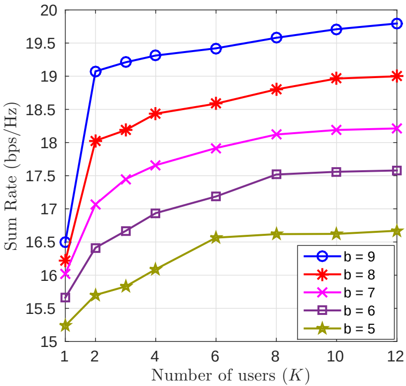
(e) SMM Convergence, $K = 4$, SNR $=$ 0 dB, $P_{\max} = 40$ dBm.



(f) IMM Convergence, $K = 4$, SNR $= 0$ dB, $P_{\max} = 40$ dBm.



(g) Sum rate vs. K, $N = 256$, SNR $=$ 10 dB, $P_{\text{max}} = 40$ dBm.



(h) Sum rate vs. K, $N_t = 16$, SNR $=$ 10 dB, $P_{\text{max}} = 40$ dBm.

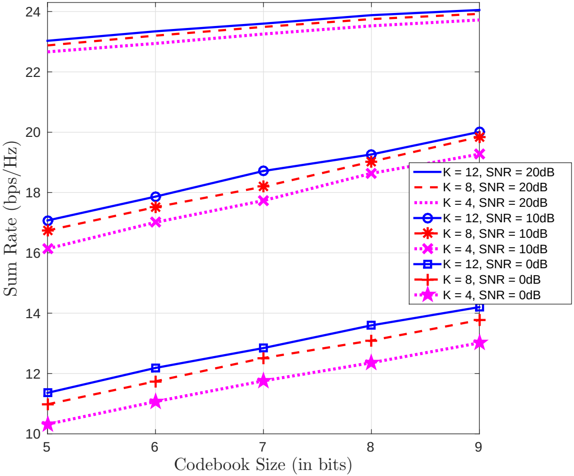


Figure: Sum rate vs. Codebook size (in bits), $N_t = 16$, $P_{\max} = 40$ dBm.

Summary & Future Work

Summary

- Proposed two algorithms, named square root MM (SMM) and inverse MM (IMM), to solve the problem of codebook based DL sum rate maximization in a TDD MU-MIMO broadcast system.
- Algorithms are based on nested application of the MM procedure, and the novelty of the algorithms lies in the choice of the surrogate functions used to bound the objective function.
- Proved the global optimality of the solutions to the surrogate optimization problems of both the SMM and IMM algorithms, and benchmarked the performance with state of the art algorithms.

Future Work

- Design of sum rate optimal codebooks, and to investigate the sum rate performance of hybrid beamforming in massive MIMO and mmwave cellular communication systems.