# MIMO
# Simulation Session

## Part A – User Selection Simulation (80%)

### General Description

In this exercise, we will examine the performance of several transmission and decoding schemes in a multi-antenna wireless network. We assume the base station (BS) has $t$ transmission antennas, a power constraint $P$, and each user (out of $K$) has a single reception antenna.

**Notice**: transmission rate of a single user, according to Shannon, is $R_i = \log(1 + SNR_i)$

### Preliminary Questions

1. What is the maximal number of users the BS can simultaneously transmit to using Zero-Forcing Beam Forming (ZFBF)?
2. What is the SNR observed by user $i$ under ZFBF?
3. How does one generate a zero-mean $\sigma^2$-variance complex Gaussian random variable? You may find this link from Wikipedia helpful.

### Defining A Channel

**Remark:** Each programming language uses a different method to generate Gaussian random variables. Some languages take the standard deviation instead of variance – *be heedful of which is used in your simulation environment*.

In this part, we assume the users are i.i.d. and are indexed from 1 to $t$. We define three types of channels:

1. An **i.i.d complex Gaussian channel** equally distributed – the connection between each receiving antenna and each of the $t$ transmitting antennas is described by a vector of the length $t$ so that each entry in the vector has a real and a complex parts Gaussian distributed with an expected value of 0 and a variance of 0.5 and is i.i.d to all other entries.
2. A **chaotic channel** model – each channel (i.e., row) is characterized by a complex Gaussian distribution with a random expected value uniformly distributed $U[0,1]$ and a variance uniformly distributed $U[0,3]$. The parameters for each channel's distribution are only drawn once.
3. A complete **correlative channel** – all the antennas have the same randomly drawn complex Gaussian distribution with an expected value of 0 and a variance of 1. The is only drawn once.

### Simulation

**Remark:** if your code is designed poorly, expect long running times. Planning your code before writing and executing is highly recommended – especially in a *parallelized* manner. Before running the complete simulation, you should test each part of your code to ensure everything works properly.

Assuming there are $K = 30$ users and the transmission power is $P = 5W$. **Conduct simulations** that receive the number of antennas at the transmitter $t$ as an input (a low value for $t$ is recommended for testing). Repeat each simulation 5,000 times and display the PDF of the total rates per transmission for each of the defined channels above.

Repeat for each of the following cases:

1. The BS chooses a single user randomly and transmits to it at maximum power.
2. The BS chooses to transmit to a single user with the strongest channel (the user with the highest row norm) and transmits to it at maximum power.
3. The BS chooses to transmit to $t$ randomly chosen users using ZFBF with uniform power allocation between the selected users.
4. The BS chooses to transmit to the $t$ strongest users (top $t$ sorted by row norm) using ZFBF. The power allocation is uniform between the selected users.
5. The BS tests all user subgroups of the size $t$, and transmits to the optimal group (e.g., with the highest *rates*) using ZFBF. The power allocation is uniform between the selected users.
6. **(Bonus 10%)** The BS selects $t$ users by applying the SUS algorithm (*see resources for details*) for user selection. It transmits to that group using ZFBF with equal power allocation between the selected users.
7. **(Bonus 10%)** The BS chooses to transmit to a single user with the strongest channel (the user with the highest norm, denoted $\underline{h}_1$). Based on the channel to this user, $\underline{h}_1$, the base station performs a Gram–Schmidt process to create an orthonormal basis. For each orthonormal direction vector, the base station selects the best user to transmit to in that direction. The power allocation is uniform between the selected users.
8. **(Bonus up to 50%)** Any original method to select users and allocate the transmission power. Points will be given based on originality.

## Post-Simulation Questions

1. Provide a short explanation of the advantages and disadvantages of each of these user selection schemes.
2. Compare the user selection fairness results of the **chaotic channel** and **the i.i.d complex Gaussian channel**.

Submit your simulation code and a pdf file, including an elaborated explanation of your results and corresponding graphs to validate it.

### Resources

Wireless communications, Andrea Goldsmith, Stanford University, California, 2005, 9780511841224

Taesang Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," in IEEE Journal on Selected Areas in Communications, vol. 24, no. 3, pp. 528-541, March 2006, doi: 10.1109/JSAC.2005.862421. (SUS algorithm)

J. Kampeas, A. Cohen and O. Gurewitz, "Analysis of Different Approaches to Distributed Multiuser MIMO in the 802.11ac," in IEEE Transactions on Mobile Computing, vol. 19, no. 11, pp. 2548-2561, 1 Nov. 2020, doi: 10.1109/TMC.2019.2931301. (DEMOS and MINOS algorithms)

## Part B – CSI Collection (20%)

### General Description

In this exercise, we focus on how channel state information (CSI) is collected by a receiver. For simplicity, we focus on estimating only real-valued CSI. The main idea is to send a sequence of symbols (named "training symbols") that are agreed upon by both the transmitter and receiver beforehand. Since the training symbols are known, the receiver can estimate (i.e., guess) the channel coefficients. In our settings, the transmitter has $M_t$ antennas and sends $M_t \cdot N$ training symbols (in each "round," it sends $M_t$ training symbols and uses up to $N$ rounds). The transmitter has a power constraint of $P$. The receiver has $M_r$ antennas and estimates the channel, $\boldsymbol{H}$.

There are many techniques to estimate the CSI, and this exercise will focus on the famous Least Squares (LS) method. This exercise uses random symbols instead of the proven-as-optimal orthogonal training symbols. Given a known matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and an observed vector $y \in \mathbb{R}^{m \times 1} = \boldsymbol{A}x + w$ (where $w \in \mathbb{R}^{m \times 1}$ is additive noise), LS solves problems of the following form:

$$\min_{\hat{x} \in \mathbb{R}^n} \|\boldsymbol{A}\hat{x} - y\|_2^2$$

I.e., it finds $\hat{x} \in \mathbb{R}^n$ obtains minimum squared error (MSE). This idea can be extended to matrices and transpose versions. I.e., given a known matrix $\boldsymbol{X} \in \mathbb{R}^{n \times t}$ and an observed *matrix* $\boldsymbol{Y} \in \mathbb{R}^{m \times t} = \boldsymbol{AX} + \boldsymbol{W}$ (where $\boldsymbol{W} \in \mathbb{R}^{m \times t}$ is additive noise), LS solves problems of the following form:

$$\min_{\widehat{\boldsymbol{A}} \in \mathbb{R}^{m \times n}} \left\|\widehat{\boldsymbol{A}}\boldsymbol{X} - \boldsymbol{Y}\right\|_2^2$$

Where $\widehat{\boldsymbol{A}} \in \mathbb{R}^{m \times n}$ achieves MSE and has a closed form expression: $\widehat{\boldsymbol{A}} = \boldsymbol{Y}\boldsymbol{X}^*(\boldsymbol{X}\,\boldsymbol{X}^*)^{-1}$.

### Preliminary Questions

1. What is $\boldsymbol{X}^*(\boldsymbol{X}\,\boldsymbol{X}^*)^{-1}$?
2. Why, under normal circumstances, do we assume that the channel model is complex Gaussian?
3. Formulate the CSI estimation problem as an LS problem. I.e., what are $\boldsymbol{X}$, $\boldsymbol{A}$, $\widehat{\boldsymbol{A}}$ and $\boldsymbol{Y}$ in the CSI estimation problem? What are the mentioned $m$, $n$, and $t$?

### Defining A Channel

Fix $P = 5W$, $M_r = M_t = 4$. Draw a realization of a Gaussian matrix, which would be the channel to estimate, of size $M_r \times M_t$. The channel has zero mean and unit variance. <u>The channel is drawn only once</u>.

### Simulation

**Conduct simulations** that increase $N$ from $M_r$ [1] to $5M_r$. Notice that the total number of training symbols is $M_t \cdot N$. Repeat each simulation $M = 5{,}000$ times and <u>save the per-component mean MSE</u>, given by $\frac{1}{M}\sum_{n=1}^{M} \frac{1}{M_r M_t} \left\|\boldsymbol{H} - \widehat{\boldsymbol{H}}_n\right\|_2^2$ where $\widehat{\boldsymbol{H}}_n$ is the estimation of $\boldsymbol{H}$ in the $n^{th}$ iteration. At the end of the simulation, present a plot of the (empirical) MMSE calculated (<u>in dB</u>) vs. the number of training symbols.

In each iteration, do the following:

1. Generate $M_t \times N$ training symbols, uniformly drawn from $\{-4, -3, \ldots, 3, 4\}$.
2. Normalize the training symbols and allocate power uniformly such that in each round, the total power of the transmitted vector is $P$.
3. Pass the training symbols through the channel, and don't forget to add the zero-mean 0.5-variance additive noise.
4. Estimate $\boldsymbol{H}$, and collect the per-component MSE (to average later).

[1] Theorem from Estimation Theory: Estimations of $M_r \times M_t$ matrices with less than $M_r \cdot M_t$ training symbols are meaningless with high probability.

## Post-Simulation Questions

1. Is there an optimal number of training symbols? ***Hint*:** In our simulation, we assumed $\boldsymbol{H}$ is fixed during the channel estimation process
2. In our simulation, we used random training symbols. The optimal solution (minimal MMSE) uses *orthogonal* training symbols (as Wi-Fi does). Is there any other advantage of using orthogonal (in *rows*) training symbols? ***Hint:*** CDMA

Submit your simulation code and a pdf file, including an elaborated explanation of your results and corresponding graphs to validate it.

### Resources

Wireless communications, Andrea Goldsmith, Stanford University, California, 2005, 9780511841224

M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," in IEEE Transactions on Signal Processing, vol. 54, no. 3, pp. 884-893, March 2006, doi: 10.1109/TSP.2005.863008. (LS-based CSI estimation, optimal set of training symbols)