# Analysis of CRLB for AoA estimation in Massive MIMO systems

Masoud Arash, Hamed Mirghasemi, Ivan Stupia, Luc Vandendorpe
*ICTEAM, Université Catholique de Louvain*
Louvain-la-Neuve, Belgium
{masoud.arash, seyed.mirghasemi, ivan.stupia, luc.vandendorpe}@uclouvain.be

***Abstract*—Massive MIMO systems provide high angular resolution in the next generation of wireless systems. This opportunity can be used to estimate the location of user terminals (UTs) accurately. In this paper, we analyze the Cramer-Rao lower bound ($CRLB$) of planar antenna arrays in Massive MIMO systems for Angle of Arrival (AoA) estimation. With the help of Random Matrix Theory, we prove that for Massive antenna arrays with independent and identically distributed (i.i.d) multipath signals, instantaneous $CRLB$ for AoA estimation converges toward a deterministic value, regardless of channel distribution. In this scenario, $CRLB$ is a function of channel variance instead of instantaneous realizations. Then, antenna selection is studied, and it is shown that using different subsets significantly affects the $CRLB$ of a planar array. Numerical results confirm the convergence of deterministic results and indicate the benefits of antenna selection and the importance of the selection strategy.**

## I. INTRODUCTION

Massive MIMO systems are among the primary candidates for the next generation of wireless systems [1]. By virtue of large number of antennas, many opportunities arise in such systems [2]. In particular, high angular resolution can be used for accurate AoA estimation of UTs [3]. Many applications need real-time information about UTs' locations, such as auto-driving cars and health services [4]. This information can also help reduce the interference in the system by directing beams in the BS to each UT's position, saving power and improving efficiency.

One of the fundamental criteria for assessing an estimator is $CRLB$, which sets a lower bound on the variance of any unbiased estimator [3]. Many works have studied $CRLB$ for AoA estimation in Massive MIMO settings [2], [5], [6]. A common assumption among these studies is to consider one dominant path as the primary carrier of information. In this scenario, for each UT, all of the antennas will have the same channel coefficient. However, in Massive MIMO, the paths that signals take to different antennas might significantly differ from one another [1]. Moreover, based on Dense Multipath Channel (DMC) model [7], a dominant path can be accompanied by many multipath signals. New studies have used multipath signals to extract more information about AoA [2], [4].

In this regard, [8] studied the probability of AoA detection in Massive MIMO i.i.d channels. In [9], we presented an analysis for $CRLB$ of AoA estimation in DMC channels for a linear antenna array. In this work, we extend the analysis by presenting a deterministic form of $CRLB$ for AoA estimation for a planar antenna array under the DMC model in 3D settings.

In the next part, the idea of antenna selection is studied. Antenna selection is motivated by different goals in wireless systems. Initially, it was presented as a solution for hardware shortages, when the number of antennas is more than available RF-chains, and the goal is to take full advantage of available hardware by using them for the best possible set of antennas [10]. Later, regarding growing concerns about energy consumption, antenna selection was revived again, this time to use valuable and limited energy budget in the best possible way to maximize the trade-off between performance and energy consumption [11], [12].

In [9] we presented an antenna selection method to maximize the localization efficiency of a linear array. It has been shown that depending on the antenna selection strategy, the behavior of $CRLB$ with respect to (w.r.t) number of antennas changes and the optimal antenna selection method has been presented. In this paper, we show that in the case of a planar array, when 3D localization is considered, the optimal antenna selection strategy changes. As the proposed antenna selection is for the Multi-User scenario, it is based on minimizing the expected $CRLB$ over the possible range of UT's AoA. The initial optimal subset of antennas is presented, which serves as the starting point of a greedy algorithm. This algorithm finds the best greedy set of antennas when only a limited number of antennas have to be utilized.

*Notation*: Boldface lower case is used for vectors, $\boldsymbol{x}$, and upper case for matrices, $\boldsymbol{X}$. $\boldsymbol{X}^T$, $\boldsymbol{X}^H$ and $\boldsymbol{X}_{m,k}$ denote transpose, hermitian and $(m,k)$th entry of $\boldsymbol{X}$, respectively. $\boldsymbol{I}_K$ is $K \times K$ identity matrix. $\mathbb{E}_x\{.\}$ denotes expectation w.r.t $x$, $j = \sqrt{-1}$, $|.|$ is absolute value, $tr$ is trace operator, $\odot$ is Hadamard product, $\otimes$ is Kronecker product, $diag(\boldsymbol{x})$ is a diagonal matrix with the elements of vector $\boldsymbol{x}$ on the main diagonal and $\xrightarrow{a.s.}$ means Almost Sure convergence.

## II. SYSTEM MODEL

We consider the uplink of a single-cell Multi-User Massive MIMO system with a BS at the center, equipped with $M = M_1 M_2$ antennas. $M_1$ antennas are installed along the $x$ axis and $M_2$ antennas along the $y$ axis. On each axis, adjacent

This work was supported by F.R.S.-FNRS under the EOS research project MUSEWINET (EOS project 30452698).

$$\boldsymbol{CRLB}_{\theta}(k,k) \xrightarrow{a.s.} \frac{v(M_1,2M_1)\cos^2(\theta_k)+v(M_2,2M_2)sin^2(\theta_k)+3v(M_1,M_2)\cos(\theta_k)\sin(\theta_k)}{\rho_k\beta^2(2\sigma_h^2+|\bar{h}_k|^2)v(M_1,M_2)(\frac{v(2M_1,2M_2)}{3}-\frac{3}{4}v(M_1,M_2))\sin^2(\varphi_k)}. \tag{13}$$

antennas are separated by distance $d$ (Fig 1). There are $K$ uniformly distributed single-antenna UTs in the cell. The BS uses pilot signals transmitted by UTs to localize them. The received signal at the BS is

$$\boldsymbol{y} = (\boldsymbol{A}_{Rx} \odot \boldsymbol{H})\boldsymbol{s} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{s}$ is the vector of transmitted pilots and $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma_n^2\boldsymbol{I}_M)$ is additive white Gaussian noise. Also,

$$\boldsymbol{A}_{Rx} = [\boldsymbol{a}_R(\theta_1,\varphi_1) \ \ \dots \ \ \boldsymbol{a}_R(\theta_K,\varphi_K)], \tag{2}$$

contains $K$ columns of $M \times 1$ steering vectors of BS antenna array response, where $\theta_k$ and $\varphi_k$ are $k$th UT's AoA and azimuth for $k \in \{1,\dots,K\}$. The $m$th element of $k$th steering vector is [13]

$$\boldsymbol{a}_R(\theta_k,\varphi_k)_m = \frac{e^{-j\beta\sin(\varphi_k)((m_1-1)\cos(\theta_k)+(m_2-1)\sin(\theta_k))}}{\sqrt{M}}, \tag{3}$$

$$m = (m_2-1)M_1 + m_1,$$
$$m_1 = 1,\dots,M_1, \quad m_2 = 1,\dots,M_2,$$

where $\beta = \frac{2\pi d}{\lambda}$ and $\lambda$ is the wavelength of pilots. $\boldsymbol{H} = [\boldsymbol{h}_1 \ \ \dots \ \ \boldsymbol{h}_K]$ is an $M \times K$ matrix whose $(m,k)$th element, $h_{m,k}$, is the channel coefficient between $k$th UT and $m$th BS antenna. The mean of $\boldsymbol{h}_k$ is equal to the channel coefficient of $k$th UT's dominant path, $\bar{h}_k$. The random part of each elements, $\hat{h}_{m,k}$, accounts for the random effects of aggregated multipath signals at each antenna with limited fourth and eighth moments. The variance of the random part is assumed to be constant and equal to $\sigma_h^2$ for every antenna and user [9]

$$\begin{aligned} &h_{m,k} = h_{m,k}^r + jh_{m,k}^i, \\ &\mathbb{E}\{h_{m,k}^r\} = \mathcal{R}e\{\bar{h}_k\}, \ \ \mathbb{E}\{h_{m,k}^i\} = \mathcal{I}m\{\bar{h}_k\}, \\ &var(h_{m,k}^r) = var(h_{m,k}^i) = \sigma_h^2. \end{aligned} \tag{4}$$

## III. CRLB Analysis

The vector containing desired parameters for estimation is

$$\boldsymbol{\eta} = [\overbrace{\theta_1 \ \ \theta_2 \ \ \dots \ \ \theta_K}^{\boldsymbol{\eta}_\theta} \,|\, \overbrace{\varphi_1 \ \ \dots \ \ \varphi_K}^{\boldsymbol{\eta}_\varphi}]^T. \tag{5}$$

Defining $\hat{\boldsymbol{\eta}}$ as the unbiased estimator of $\boldsymbol{\eta}$, its mean square error is lower bounded as [3]

$$\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\eta}}\{(\boldsymbol{\eta}-\hat{\boldsymbol{\eta}})(\boldsymbol{\eta}-\hat{\boldsymbol{\eta}})^T\} \geq \boldsymbol{CRLB} = \boldsymbol{J}^{-1}, \tag{6}$$

where $\boldsymbol{J}$ is Fisher Information Matrix (FIM) and can be written in block matrix form as [3]

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_{\theta,\theta} & \boldsymbol{J}_{\theta,\varphi} \\ \boldsymbol{J}_{\varphi,\theta} & \boldsymbol{J}_{\varphi,\varphi} \end{bmatrix}, \tag{7}$$

and

$$\boldsymbol{J}_{a,b} = \mathcal{R}e[(\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\eta}_a})^H(\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\eta}_b})], \tag{8}$$

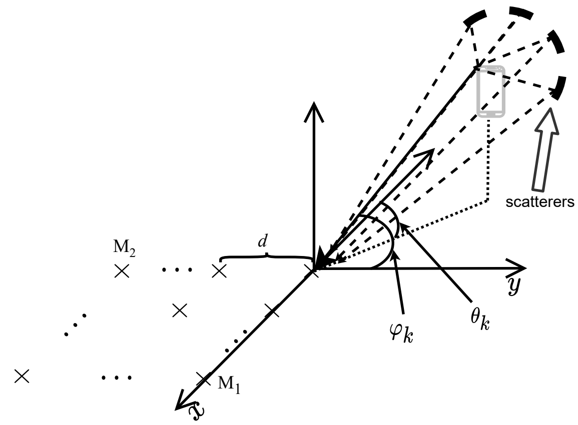


**Figure 1:** Configuration of the antenna array.

in which $\boldsymbol{w} \triangleq (\boldsymbol{A}_{Rx} \odot \boldsymbol{H})\boldsymbol{s}$ and $a,b \in \{\theta,\varphi\}$. Therefore, using block matrix inversion properties, $CRLB$ of $\boldsymbol{\eta}_\theta$ is

$$\boldsymbol{CRLB}_\theta = \frac{\sigma_n^2}{2}(\boldsymbol{J}_{\theta,\theta} - \boldsymbol{J}_{\theta,\varphi}\boldsymbol{J}_{\varphi,\varphi}^{-1}\boldsymbol{J}_{\varphi,\theta})^{-1}. \tag{9}$$

The derivatives of $\boldsymbol{w}$ w.r.t $\boldsymbol{\eta}_\theta$ and $\boldsymbol{\eta}_{\boldsymbol{\varphi}}$ can be written as

$$\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\eta}_\theta} = M(\boldsymbol{\Sigma}_2(\boldsymbol{A}+j\boldsymbol{B})\boldsymbol{C}_\theta - \boldsymbol{\Sigma}_1(\boldsymbol{A}+j\boldsymbol{B})\boldsymbol{S}_\theta)\boldsymbol{S}_\varphi, \tag{10}$$

$$\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\eta}_\varphi} = M(\boldsymbol{\Sigma}_1(\boldsymbol{A}+j\boldsymbol{B})\boldsymbol{C}_\theta + \boldsymbol{\Sigma}_2(\boldsymbol{A}+j\boldsymbol{B})\boldsymbol{S}_\theta)\boldsymbol{C}_\varphi, \tag{11}$$

where

$$\begin{aligned} &\boldsymbol{\Sigma}_1 = \beta(\boldsymbol{I}_{M_2} \otimes diag(0, \frac{1}{M}, \dots, \frac{M_1-1}{M})), \\ &\boldsymbol{\Sigma}_2 = \beta(diag(0, \frac{1}{M}, \dots, \frac{M_2-1}{M}) \otimes \boldsymbol{I}_{M_1}), \\ &\boldsymbol{A} = \bar{\boldsymbol{A}} + \sigma_h\hat{\boldsymbol{A}}, \quad \boldsymbol{B} = \bar{\boldsymbol{B}} + \sigma_h\hat{\boldsymbol{B}}, \\ &\bar{\boldsymbol{A}}_{m,k} = \mathcal{R}e\{\bar{h}_k\boldsymbol{a}_R(\theta_k)_m\boldsymbol{s}_k\}, \bar{\boldsymbol{B}}_{m,k} = \mathcal{I}m\{\bar{h}_k\boldsymbol{a}_R(\theta_k)_m\boldsymbol{s}_k\}, \\ &\hat{\boldsymbol{A}}_{m,k} = \frac{1}{\sigma_h}\mathcal{R}e\{-j\hat{h}_{m,k}s_k(\boldsymbol{a}_R(\theta_k)_m)\}, \\ &\hat{\boldsymbol{B}}_{m,k} = \frac{1}{\sigma_h}\mathcal{I}m\{-j\hat{h}_{m,k}s_k(\boldsymbol{a}_R(\theta_k)_m)\}, \\ &\boldsymbol{S}_a = diag(\sin(a_1),\dots,\sin(a_K)), \ \ a \in \{\theta,\varphi\} \\ &\boldsymbol{C}_a = diag(\cos(a_1),\dots,\cos(a_K)). \ \ a \in \{\theta,\varphi\} \end{aligned} \tag{12}$$

By defining $\rho_k \triangleq \frac{|s_k|^2}{\sigma_n^2}$ and $v(m_1,m_2) = (m_1-1)(m_2-1)$, the following theorem presents a deterministic expression for $\boldsymbol{CRLB}_\theta$ of a planar antenna array.

**Theorem 1.** *In the Massive MIMO settings, $\boldsymbol{CRLB}_\theta$ almost surely converges toward a deterministic diagonal matrix whose $(k,k)$th entry is given by (13), written at the top of this page.*

*Proof.* Replacing (10) and (11) in (8), we have

$$\begin{aligned}\boldsymbol{J}_{\theta,\theta} = M^2\boldsymbol{S}_{\varphi}^T[&\boldsymbol{S}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_1^2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_1^2\boldsymbol{B})\boldsymbol{S}_{\theta}\\ &+\boldsymbol{C}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_2^2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_2^2\boldsymbol{B})\boldsymbol{C}_{\theta}\\ &-\boldsymbol{S}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{B})\boldsymbol{C}_{\theta}\\ &-\boldsymbol{C}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{B})\boldsymbol{S}_{\theta}]\boldsymbol{S}_{\varphi}, \end{aligned} \tag{14}$$

$$\begin{aligned}\boldsymbol{J}_{\varphi,\varphi} = M^2\boldsymbol{C}_{\varphi}[&\boldsymbol{C}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_1^2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_1^2\boldsymbol{B})\boldsymbol{C}_{\theta}\\ &+\boldsymbol{S}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_2^2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_2^2\boldsymbol{B})\boldsymbol{S}_{\theta}\\ &+\boldsymbol{C}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{B})\boldsymbol{S}_{\theta}\\ &+\boldsymbol{S}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{B})\boldsymbol{C}_{\theta}]\boldsymbol{C}_{\varphi}, \end{aligned} \tag{15}$$

$$\begin{aligned}\boldsymbol{J}_{\theta,\varphi} = \boldsymbol{J}_{\varphi,\theta}^T = M^2\boldsymbol{S}_{\varphi}^T[&-\boldsymbol{S}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_1^2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_1^2\boldsymbol{B})\boldsymbol{C}_{\theta}\\ &+\boldsymbol{C}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_2^2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_2^2\boldsymbol{B})\boldsymbol{S}_{\theta}\\ &-\boldsymbol{S}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{B})\boldsymbol{S}_{\theta}\\ &+\boldsymbol{C}_{\theta}^T(\boldsymbol{A}^T\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1\boldsymbol{B})\boldsymbol{C}_{\theta}]\boldsymbol{C}_{\varphi}. \end{aligned} \tag{16}$$

From [9] (lemmas 1-3) we know that for $M \to \infty$

$$\boldsymbol{A}^T\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q\boldsymbol{A}+\boldsymbol{B}^T\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q\boldsymbol{B} \xrightarrow{a.s.} \frac{1}{M}tr(\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q)(2\sigma_h^2+\bar{\boldsymbol{H}}) \tag{17}$$

for $p,q \in \{1,2\}$ and $\bar{\boldsymbol{H}} = diag(|\bar{h}_1|^2,\dots,|\bar{h}_K|^2)$. Also,

$$\begin{aligned}&tr(\boldsymbol{\Sigma}_1^2) = \beta^2(M_1-1)(2M_1-1)/M/6,\\ &tr(\boldsymbol{\Sigma}_2^2) = \beta^2(M_2-1)(2M_2-1)/M/6,\\ &tr(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2) = tr(\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1) = \beta^2(M_1-1)(M_2-1)/M/4.\end{aligned} \tag{18}$$

Replacing (17) and (18) in (14)-(16), we have

$$\begin{aligned}\boldsymbol{J}_{\theta,\theta} \xrightarrow{a.s.} &\beta^2\boldsymbol{S}_{\varphi}^T(\frac{v(M_1,2M_1)}{6}\boldsymbol{S}_{\theta}^2+\frac{v(M_2,2M_2)}{6}\boldsymbol{C}_{\theta}^2\\ &-\frac{v(M_1,M_2)}{2}\boldsymbol{C}_{\theta}\boldsymbol{S}_{\theta})\boldsymbol{S}_{\varphi}(2\sigma_h^2+\bar{\boldsymbol{H}}),\\ \boldsymbol{J}_{\varphi,\varphi} \xrightarrow{a.s.} &\beta^2\boldsymbol{C}_{\varphi}^T(\frac{v(M_1,2M_1)}{6}\boldsymbol{C}_{\theta}^2+\frac{v(M_2,2M_2)}{6}\boldsymbol{S}_{\theta}^2\\ &+\frac{v(M_1,M_2)}{2}\boldsymbol{C}_{\theta}\boldsymbol{S}_{\theta})\boldsymbol{C}_{\varphi}(2\sigma_h^2+\bar{\boldsymbol{H}}),\\ \boldsymbol{J}_{\theta,\varphi} \xrightarrow{a.s.} &\beta^2\boldsymbol{S}_{\varphi}^T(\frac{v(M_2,2M_2)-v(M_1,2M_1)}{6}\boldsymbol{C}_{\theta}\boldsymbol{S}_{\theta}\\ &+\frac{v(M_1,M_2)}{4}(\boldsymbol{C}_{\theta}^2-\boldsymbol{S}_{\theta}^2))\boldsymbol{C}_{\varphi}(2\sigma_h^2+\bar{\boldsymbol{H}}).\end{aligned} \tag{19}$$

By replacing (19) in (9), using the fact that $\boldsymbol{J}_{\theta,\theta}$, $\boldsymbol{J}_{\varphi,\varphi}$, $\boldsymbol{J}_{\theta,\varphi}$ and $\boldsymbol{J}_{\varphi,\theta}$ are diagonal matrices, after some algebraic simplifications, all the elements of $\boldsymbol{CRLB}_{\theta}$ are obtained as (13). □

It is seen that the effects of individual channel realizations are completely removed in the instantaneous $CRLB_{\theta}$. In fact, what matters is the variance of these coefficients when large number of antennas are utilized. Also, (13) shows that in Massive MIMO systems, AoA information is (*almost surely*) never lost due to the contribution of the multipath signals. These results, along with those presented in [9] for linear arrays, provide a complementary theoretical basis for AoA estimation (or refinement) using several path-loss signals with the same order of power, like methods reported in [2], [14].

Now that we have a deterministic value for $\boldsymbol{CRLB}_{\theta}$, we can use it to study how different antennas contribute in $\theta$ estimation and prioritize the antennas with high contribution when only a portion of the total available antennas are utilized.

## IV. Antenna Selection

This section studies the antenna selection for planar antenna array configurations. The expectation of $\boldsymbol{CRLB}_{\theta}$ is minimized w.r.t set of utilized antennas, using a greedy algorithm whose optimal starting point will be presented.

In the following, we assume that $F \leq M$ number of antennas has to be used. $F$ can be either the number of available RF-chains or the optimal point of the trade-off between $CRLB$ and energy consumption. We assume that array dimensions are much smaller than UTs' distances from it, so the difference of received power for different antennas is negligible.

From (9) and Theorem 1 it is seen that $\boldsymbol{CRLB}_{\theta}$ for a planar array is a function of the traces of squared $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$, and their product. When $F$ antennas are utilized, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ will be changed to $\tilde{\boldsymbol{\Sigma}}_1$ and $\tilde{\boldsymbol{\Sigma}}_2$, both with the size of $F \times F$. The diagonal elements of these matrices will be corresponding values of selected antennas from $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. As the antenna selection has to be done before the estimation of $\theta$ (and if $F$ is the number of RF-chains, even before reception of the transmitted pilots by the UTs), there is no apriori knowledge about the $\theta$, except its distribution. With this in mind, we minimize $\mathbb{E}_{\theta}\{tr(\boldsymbol{CRLB}_{\theta})\}$. So, noting that when $F$ antennas are utilized, the $M^{-\frac{1}{2}}$ normalization factor in (3) will be changed to $F^{-\frac{1}{2}}$ and uniform distribution of UTs, we obtain

$$\begin{aligned}\mathbb{E}_{\theta}\{tr(\boldsymbol{CRLB}_{\theta})\} = &\frac{F\sum_{k=1}^{K}(\rho_k(2\sigma_h^2+|\bar{h}_k|^2))^{-1}}{8\beta^2\sin^2(\varphi_k)}\times\\ &\underbrace{\left(\frac{tr(\tilde{\boldsymbol{\Sigma}}_1^2+\tilde{\boldsymbol{\Sigma}}_2^2)}{tr(\tilde{\boldsymbol{\Sigma}}_1^2)tr(\tilde{\boldsymbol{\Sigma}}_2^2)-(tr(\tilde{\boldsymbol{\Sigma}}_1\tilde{\boldsymbol{\Sigma}}_2))^2}\right)}_{U(\mathcal{S})},\end{aligned} \tag{20}$$

where $\mathcal{S}$ is the set of selected antennas according to which the elements of $\tilde{\boldsymbol{\Sigma}}_1$ and $\tilde{\boldsymbol{\Sigma}}_2$ are determined. The only part affected by the selected set of antennas is $U(\mathcal{S})$ (the part inside parentheses), so we just need to minimize it. Minimization of $U(\mathcal{S})$ is a combinatorial optimization problem that we use a greedy algorithm to solve it. $F = 3$ is the minimum possible number of antennas for a planar antenna array (so they form a plane). These first three antennas have to be selected together. After selecting the first three antennas (starting point of the antenna selection), we can proceed with a greedy algorithm that, step by step, selects an antenna that reduces $U(\mathcal{S})$ the most.

**Lemma 1.** *The best choice for first three antennas is* $\mathcal{S}_3 = \{(1,1),(M_1,1),(1,M_2)\}$.

*Proof.* The first choice has to be $(1,1)$th antenna that is the reference point w.r.t which the $\theta$ is measured. Moreover, the other two antennas must have different indices in both dimensions, otherwise they compose a linear array. For notation convenience, we define $M'_1 = M_1 - 1$ and $M'_2 = M_2 - 1$. Let $\bar{\mathcal{S}}_3 = \{(1,1),(M_1 - a, 1+b),(1+c, M_2 - d)\}$, for any set of non-negative integers $\{a,b,c,d\}$ where

$$\begin{aligned} &\{a,c\} \leq M'_1, \quad \{b,d\} \leq M'_2, \\ &a+b+c+d \geq 1, \\ &(a+c, b+d) \neq (M'_1, M'_2), \end{aligned} \tag{21}$$

be a set of three distinctive antennas other than $\mathcal{S}_3$. Evaluating the corresponding $U(\mathcal{S})$ for both $\bar{\mathcal{S}}_3$ and $\mathcal{S}_3$ by using indices of each group in (20), we have

$$\begin{aligned} U(\bar{\mathcal{S}}_3) &= \frac{(M'_1 - a)^2 + b^2 + c^2 + (M'_2 - d)^2}{((M'_1 - a)(M'_2 - d) - bc)^2}, \\ U(\mathcal{S}_3) &= \frac{M'^2_1 + M'^2_2}{(M'_1 M'_2)^2}. \end{aligned} \tag{22}$$

Accordingly,

$$\begin{aligned} &U(\bar{\mathcal{S}}_3)((M'_1 M'_2)^2)((M'_1 - a)(M'_2 - d) - bc)^2 \\ &= (M'_1 M'_2)^2 (M'_1 - a)^2 + (M'_1 M'_2)^2 b^2 + (M'_1 M'_2)^2 c^2 \\ &+ (M'_1 M'_2)^2 (M'_2 - d)^2 \overset{(e)}{>} (M'_1 (M'_2 - d))^2 (M'_1 - a)^2 \\ &+ (M'^2_1 + M'^2_2) b^2 c^2 + ((M'_1 - a) M'_2)^2 (M'_2 - d)^2 \\ &\geq ((M'_2 - d)(M'_1 - a) - bc)^2 (M'^2_1 + M'^2_2) \\ &= U(\mathcal{S}_3)((M'_1 M'_2)^2)((M'_1 - a)(M'_2 - d) - bc)^2, \end{aligned} \tag{23}$$

where $(e)$ follows from the fact that $b \leq M'_2$ and $c \leq M'_1$ and if both are equal either to zero or their maximum, due to (21), either $a$ or $d$ has to be greater than zero. So,

$$U(\bar{\mathcal{S}}_3) > U(\mathcal{S}_3). \tag{24}$$

Based on (24), selecting any other set than $\mathcal{S}_3$ results in higher $U(\mathcal{S})$. Therefore, $\mathcal{S}_3$ is the best set of the first three antennas that minimize $U(\mathcal{S})$. $\square$

Now that the optimal start point for selection has been found, the remaining $F-3$ antennas will be selected using a greedy algorithm. In each step, the antenna that decreases $U(\mathcal{S})$ more than others is added to the $\mathcal{S}$ until all of the required antennas are selected. The algorithm is summarized in Table I. It should be noted that as this selection is a static one, the mentioned algorithm is only needed to run once for the antenna configuration of the system. So, although it is low, the algorithm's computational complexity is not of concern because it is not used in real-time.

In SectionV, it will be shown that although for a $2D$ linear array, the optimal set is composed of furthest antennas from the reference point [9], $\mathcal{S}^*$ of a planar array mostly consists of most separated antennas (whose $x$ and $y$ indices has maximum difference), accompanied by a few numbers of furthest antennas. We compare $\mathcal{S}^*$ with both of these sets separately, in addition to another algorithm that selects the closest antennas to the reference point. The latter's importance is that the behavior of its $CRLB$ w.r.t number of antennas can present some information about designing better antenna arrays when it is not possible to build a larger one.

```
get F
S := S_3
For f := 4 to F do
    For x := 1 to M_1 do
        For y := 1 to M_2 do
            If (x, y) ∉ S  do
                S(f) := (x, y)
                V(x, y) := U(S)
            EndIf
        EndFor
    EndFor
    S(f) := arg min_{(x,y)} V
EndFor
return S
```

**Table I:** Greedy algorithm for optimal antenna selection

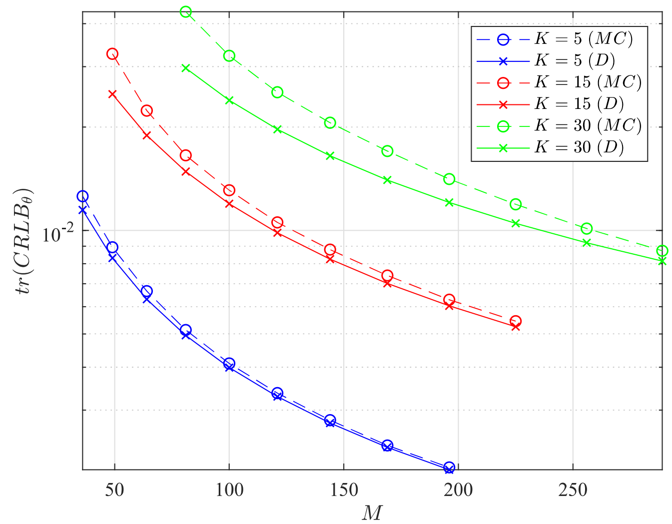


**Figure 2:** Deterministic and MC simulations of $tr(\boldsymbol{CRLB}_\theta)$.

## V. Numerical Results

In this section we verify the analytical results obtained in previous sections using Monte-Carlo (MC) simulations and illustrate the benefits of antenna selection. In the following, $\rho_k = 3$, $\frac{d}{\lambda} = 0.5$, $\theta_k = \{0, \frac{2\pi}{K}, \ldots, \frac{2\pi(K-1)}{K}\}$, $\bar{h}_k = 1$ and $\varphi_k = \pi/3$ for $i \in \{1,\ldots,K\}$. Also, channel coefficients are Gaussian distributed random variables with $\sigma_h^2 = 0.5$.

Fig. 2 shows the values of $tr(\boldsymbol{CRLB}_\theta)$ in which dashed lines are generated by MC simulations of (9), while solid ones are computed using the deterministic expression in (13), when $M_1 = M_2$ for $K = 5, 15, 30$. It should be noted that as the trace of $\boldsymbol{CRLB}_\theta$ is plotted, the gap between deterministic and MC curves is actually the sum of $K$ errors of each almost sure convergence in (13). This explains the seeming increment of the gap between MC and deterministic corresponding curves as $K$ grows. It is seen that (13) completely mimics the

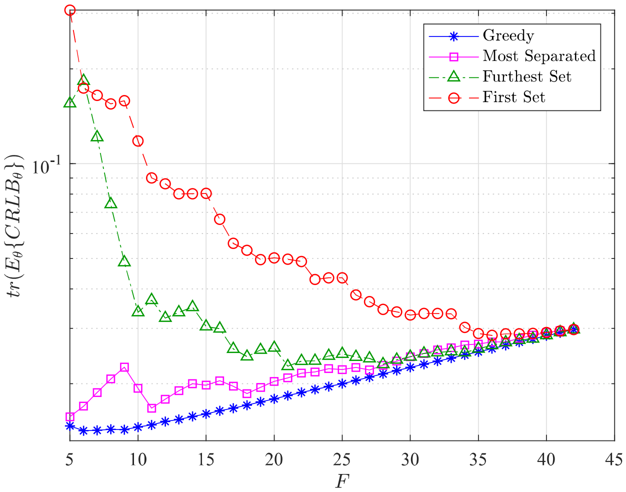


**Figure 3:** $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ for various antenna selection strategies.

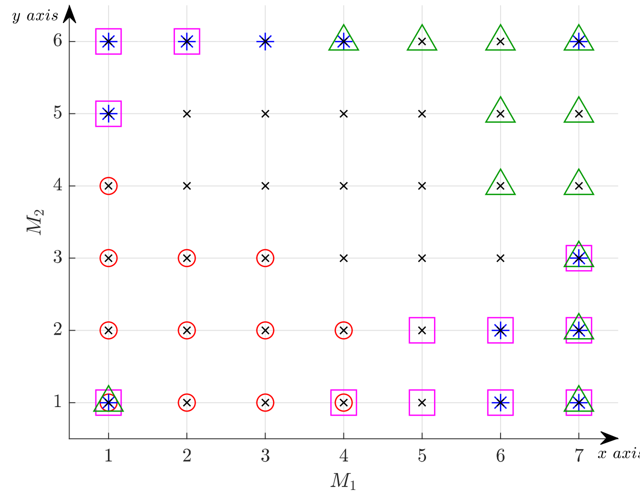


**Figure 4:** Selected antennas of each selection method for $F = 12$.

behavior of actual $\boldsymbol{CRLB}_\theta$ with high accuracy. In the rest of the figures, we only use deterministic formula.

Fig. 3 represents $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ for four different antenna selection algorithms, when $M_1 = 7$, $M_2 = 6$ and $K = 5$. The reason that $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ has increasing behavior for some algorithms stems from the fact that the received power is normalized w.r.t the number of utilized antennas (3). So, for certain $F$, some algorithms have lower $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ compared to the case when all of the $M$ antennas are used (where all curves merge). The scenario where the power is not normalized is also presented in Fig. 5. It is seen that using the greedy algorithm, $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ is significantly reduced. Also, using antenna selection, there will be a global optimal point in which $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ is minimized. In other words, there is an optimal point for the trade-off between number of utilized antennas and $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ that happens before using all of the available antennas. In this regard, the greedy algorithm obtains its minimum for $F = 6$, which is the lowest among all algorithms. Moreover, other algorithms experience several rapid deterioration and refinements, that is due to disordered addition of antennas with low and high contributions in $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$, respectively.

In order to clarify these behaviors, Fig. 4 shows the antenna configurations associated with different algorithms, when $F = 12$, for the same setting as that of Fig. 3. Selected antennas of each method are specified using the same color and marker as Fig. 3. The first set algorithm tries to complete a square in each step. The 12th selected antenna of this method is $(4, 2)$, which has a mild effect in reducing $CRLB$ in Fig. 3, while the previous two antennas, $(4, 1)$ and $(1, 4)$, decreased the $CRLB$ relatively more. Therefore, sudden decreases happen when the added antenna belongs to the most separated set. This shows that in a planar array, antennas inside the array shape do not contribute as high as those on the sides in terms of the $CRLB$. The same phenomenon (with different intensities) happens with the furthest and most separated sets at different points. Finally, the greedy algorithm selects some of the most separated antennas alongside some of the furthest ones (mostly from the former) to achieve the best possible outcome. All in all, Fig. 3 and Fig. 4 show that antennas on the boundary sides of the array have higher contributions than those in the middle, i.e., the further away from the main diagonal an antenna is, the more the contribution it has. Interestingly, we see that the optimal selection strategy creates a collection of four divided sub-arrays for AoA estimation. This effect that four smaller but separated sub-arrays can have better performance in localization rather than one big array has also been reported in [15] for systems that utilize large intelligent surfaces. Our results show that similar effects happen for the Massive MIMO systems, and the best performance is obtained when the sub-arrays vary in size.

Fig. 5 illustrates the percentage of achieved $CRLB_\theta$ versus the percentage of utilized antennas when the $M^{-\frac{1}{2}}$ normalization factor in (3) is removed. $\boldsymbol{CRLB}_\theta(f)$ is the $CRLB$ when $f$ antennas are used. This figure indicates the efficiency of hardware usage when antenna selection is utilized, corresponding to the selection method. It is seen that by using the greedy antenna selection method, more than 80% of the best possible performance is achieved, only by using half of the antennas on the array (i.e., the $\boldsymbol{CRLB}_\theta$ by using half of the total available antennas is only $\frac{1}{0.8} = 1.25$ times the $\boldsymbol{CRLB}_\theta$ when all of antennas are used). This can significantly increase both the hardware and energy efficiency of the system. Therefore, with or without the normalization factor, antenna selection has its advantages, especially with the optimal selection method.

Fig. 6 presents the $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ for $K = 5$, when number of selected antennas is constant, $F = 16$, and the array size increases with $M_1 = M_2 + 2$. This figure shows how different methods appreciate the size of the antenna array. Clearly, the first antenna set will remain constant as it does not care about the total array size. The greedy algorithm and most

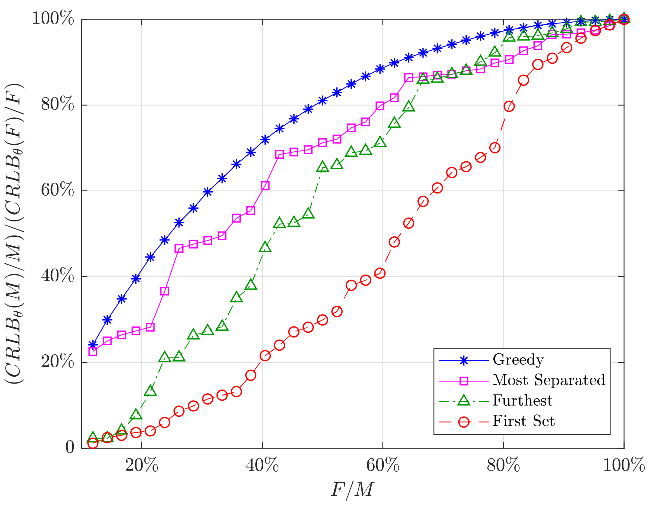


**Figure 5:** Ratio of obtained $CRLB_\theta$ versus ratio of utilized antennas.

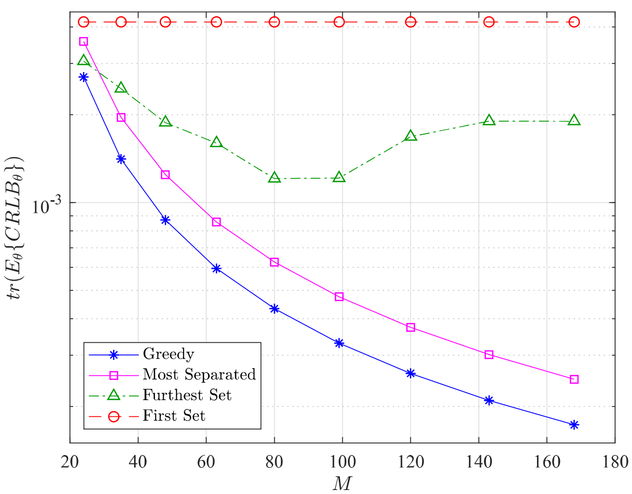


**Figure 6:** $\mathbb{E}_\theta\{tr(\boldsymbol{CRLB}_\theta)\}$ versus $M$ for $F = 16$.

separated have a uni-mode behavior as the size increases. For the furthest set, as long as the array is small enough, it will contain more antennas on the array's boundary sides relative to those near the main diagonal. As the array size grows, the ratio of antennas from the boundary to near diagonal ones will decrease. Eventually, selected antennas will form a square (or a rectangle) shape, in which the ratio will remain the same. When this ratio starts to decrease, the $\boldsymbol{CRLB}_\theta$ rises, and after the formation of the square (or the rectangle), $\boldsymbol{CRLB}_\theta$ remains the same. So, only the greedy and most separated methods always take full advantage of the total array size.

## VI. Conclusion

In this paper, we have investigated the behavior of $CRLB$ for AoA estimation for planar antenna arrays in Multi-User Massive MIMO system by using Random Matrix Theory. We showed that in the asymptotic case, regardless of the distribution of channel coefficients, $CRLB$ almost surely converges toward a deterministic expression that depends on system parameters, namely the channel's variance, instead of its instantaneous realizations. This means that when dominant path is in poor condition, AoA information is not lost and theoretically confirms the results of recent works that use multipath signals to refine their estimation. Moreover, antenna selection for minimizing the $CRLB$ is studied, and the best method to minimize the expected $CRLB$ when a fraction of available antennas have to be used is presented as a greedy algorithm. As a benchmark, other antenna selection methods are also compared with the output of the greedy algorithm. We showed that a greedy antenna selection significantly increases system performance in terms of AoA estimation.

## References


[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE communications magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[2] X. Li, E. Leitinger, M. Oskarsson, K. Åström, and F. Tufvesson, "Massive mimo-based localization and mapping exploiting phase information of multipath components," *IEEE transactions on wireless communications*, vol. 18, no. 9, pp. 4254–4267, 2019.

[3] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave mimo in 5g systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822–1835, 2017.

[4] K. Mahler, W. Keusgen, F. Tufvesson, T. Zemen, and G. Caire, "Tracking of wideband multipath components in a vehicular communication scenario," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 15–25, 2016.

[5] D. Fan, F. Gao, Y. Liu, Y. Deng, G. Wang, Z. Zhong, and A. Nallanathan, "Angle domain channel estimation in hybrid millimeter wave massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8165–8179, 2018.

[6] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Error bounds for uplink and downlink 3d localization in 5g millimeter wave systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 4939–4954, 2018.

[7] D. Astely and B. Ottersten, "The effects of local scattering on direction of arrival estimation with music," *IEEE transactions on Signal Processing*, vol. 47, no. 12, pp. 3220–3234, 1999.

[8] C. A. B. Camargo, G. Fraidenraich, and L. L. Mendes, "Probability of detection of the angle of arrival for massive mimo arrays," *IEEE Access*, vol. 6, pp. 65109–65117, 2018.

[9] M. Arash, H. Mirghasemi, I. Stupia, and L. Vandendorpe, "Localization efficiency in massive mimo systems," *arXiv preprint arXiv:2003.07978*, 2020.

[10] J. M. Romero-Jerez and A. J. Goldsmith, "Performance of multichannel reception with transmit antenna selection in arbitrarily distributed nagakami fading channels," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 2006–2013, 2009.

[11] M. Arash, E. Yazdian, M. S. Fazel, G. Brante, and M. A. Imran, "Employing antenna selection to improve energy efficiency in massive mimo systems," *Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 12, p. e3212, 2017.

[12] H. Li, G. P. Koudouridis, and J. Zhang, "Antenna selection schemes for energy efficiency in distributed antenna systems," in *2012 IEEE International Conference on Communications (ICC)*, pp. 5619–5623, IEEE, 2012.

[13] A. Hu, T. Lv, H. Gao, Z. Zhang, and S. Yang, "An esprit-based approach for 2-d localization of incoherently distributed sources in massive mimo systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 996–1011, 2014.

[14] S. D. Bast and S. Pollin, "Mamimo csi-based positioning using cnns: Peeking inside the black box," *arXiv preprint arXiv:2003.04581*, 2020.

[15] S. Hu, F. Rusek, and O. Edfors, "Beyond massive mimo: The potential of positioning with large intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 7, pp. 1761–1774, 2018.