Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Antenna Selection for Improving Energy Efficiency in XL-MIMO Systems

Marinello, José Carlos; Abrão, Taufik; Amiri, Abolfazl; De Carvalho, Elisabeth; Popovski, Petar

Antenna Selection for Improving Energy Efficiency in XL-MIMO Systems

José Carlos Marinello, Taufik Abrão, Abolfazl Amiri, Elisabeth de Carvalho, Petar Popovski

Abstract—We consider the recently proposed extra-large scale massive multiple-input multiple-output (XL-MIMO) systems, with some hundreds of antennas serving a smaller number of users. Since the array length is of the same order as the distance to the users, the long-term fading coefficients of a given user vary with the different antennas at the base station (BS). Thus, the signal transmitted by some antennas might reach the user with much more power than that transmitted by some others. From a green perspective, it is not effective to simultaneously activate hundreds or even thousands of antennas, since the power-hungry radio frequency (RF) chains of the active antennas increase significantly the total energy consumption. Besides, a larger number of selected antennas increases the power required by linear processing, such as precoding matrix computation, and short-term channel estimation. In this paper, we propose four antenna selection (AS) approaches to be deployed in XL-MIMO systems aiming at maximizing the total energy efficiency (EE). Besides, employing some simplifying assumptions, we derive a closed-form analytical expression for the EE of the XL-MIMO system, and propose a straightforward iterative method to determine the optimal number of selected antennas able to maximize it. The proposed AS schemes are based solely on long-term fading parameters, thus, the selected antennas set remains valid for a relatively large time/frequency intervals. Comparing the results, we find that the genetic-algorithm based AS scheme usually achieves the best EE performance, although our proposed highest normalized received power AS scheme also achieves very promising EE performance in a simple and straightforward way.

Index Terms—Extra large-scale MIMO; Antenna selection; Energy efficiency; Spectral efficiency; Visibility region (VR); Non-stationary; Near-field.

I. INTRODUCTION

In the fifth-generation (5G) networks, massive multiple-input multiple-output (MIMO) is identified as a key technology for achieving large gains in spectral and energy efficiencies [1], [2]. Recently, a new type of very large antenna arrays, which can be integrated into large structures like

J. C. Marinello is with Electrical Engineering Department, Federal University of Technology PR, Cornélio Procópio, PR, Brazil; jcmarinello@utfpr.edu.br.

T. Abrão is with Electrical Engineering Department, State University of Londrina (UEL), Londrina, PR, Brazil; taufik@uel.br.

A. Amiri, E. de Carvalho and P. Popovski are with the Department of Electronic Systems, Technical Faculty of IT and Design; Aalborg University, Denmark; petarp@es.aau.dk.

stadiums, or shopping malls, has been conceived: the so called extra-large scale massive MIMO (XL-MIMO) [3]–[5]. XL-MIMO system is a very promising and recent technology, pointed out as important candidate for sixth-generation (6G) and beyond technologies [6], [7], which is still in its inception, lacking for further elaborated techniques in order to mature the technology. Indeed, due to the large dimension of the antenna array in XL-MIMO systems, different kinds of spatial non-stationarities appear accross the array [3]–[5]; hence, admitting constant long-term fading coefficients between a user and all the antennas of the array is not a valid assumption. This is the main difference between the XL-MIMO scenario and the typical massive MIMO system model assumed in most part of massive MIMO literature. In [8], it is shown through experimental measurements how different regions of an extremely large array see different propagation paths, and in some cases, the terminals might see just a portion of the array, called visibility region (VR). Authors also discuss how the non-stationarity properties of this new scenario change several important design aspects.

In [3] authors seek for mapping users in terms of XL-MIMO array partition, such that the downlink (DL) sum-rate using a truncated zero-forcing (ZF) precoder is maximized. Numerical results show that a properly trained network via deep learning approach solves the problem nearly as well as an optimal mapping algorithm. Hence, increasing the size of current massive MIMO arrays is promising in terms of boosting the spectral efficiency (SE) of the wireless systems.

Since the centralized processing may present very high computational complexity in XL-MIMO arrays, a useful approach is to split the signal processing between subarrays. A subarray-based system architecture for XL-MIMO systems is proposed in [4], where closed-form uplink (UL) SE approximations with linear receivers are derived; the goal is to maximize the system sum achievable SE. Two statistical channel state information (CSI) based greedy user scheduling algorithms are developed, providing improved performance for XL-MIMO systems.

In [5], a simple *non-stationary channel model* is proposed for XL-MIMO systems, and the performance of conjugate beamforming (CB) and ZF in the DL have been investigated considering such channel. The non-stationarities are modeled in a binary fashion, such that each antenna can be visible or not for a specific user, giving rise to the VRs: an area of the massive antenna array concentrating the most of the received user's energy. However, the authors did not consider long-term fading variations between the visible antennas of a given user.

In [9] authors develop procedures for XL-MIMO receivers design. There are two important challenges in designing

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2020.3022708, IEEE Transactions on Vehicular Technology

2

receivers for XL-MIMO systems: increased computational cost of the multi-antenna processing, and how to deal with the variations of user energy distribution over the antenna elements due to the spatial non-stationarities across huge distributed antenna-elements in the 2D or 3D array. Indeed, non-stationarities limit the XL-MIMO system performance. Hence, the authors propose a distributed receiver based on variational message passing that can address both challenges. In the proposed receiver structures, the processing is distributed into local processing units, that can perform most of the complex processing in parallel, before sharing their outcome with a central processing unit. Such designs are specifically tailored to exploit the spatial non-stationarities and require lower computations than linear ZF or minimum mean square error (MMSE) receivers.

In [10], the ZF and regularized ZF schemes operating in XL-MIMO scenarios with a fixed number of subarrays have been emulated using the randomized Kaczmarz algorithm (rKA), deploying non-stationary properties through VRs. Numerical results have shown that, in general, the proposed rKA-based combiner applicable to XL-MIMO systems can considerably decrease computational complexity of the signal detector at the expense of small performance losses. On the other hand, in [11], an expectation propagation detector for XL-MIMO systems has been proposed. In order to reduce complexity, the subarray-based architecture employed distributes baseband data from disjoint subsets of antennas into parallel processing procedures coordinated by a central processing unit. Additionally, authors also propose strategies for further reducing the complexity and overhead of the information exchange between parallel subarrays and the central processing unit to facilitate the practical implementation of the proposed detector.

Recently, to deal with subarrays and channel scatterers in non-stationary XL-MIMO environment, [12] proposed two *channel estimation methods* based on subarray-wise and scatterer-wise near-field non-stationary channel properties. Authors model the multipath channel with the last-hop scatterers under a spherical wavefront and divide the large aperture array into multiple subarrays. The proposed channel estimation methods position the scatterers and perform a mapping between subarrays and scatterers. Hence, the scatterer-wise method simultaneously positions each scatterer and detects its VR to further enhance the positioning accuracy. Moreover, the subarray-wise method can achieve low mean square error (MSE) performance under low-complexity, whereas the scatterer-wise method can accurately arrange the scatterers and determine the non-stationary channel.

In [13], authors propose and validate realistic channel models when employing physically-large arrays, in which non-stationarities and visibility regions are present, as in the XL-MIMO system. The statistical distribution of important channel parameters are found based on measurements. Such contributions are proposed as extensions to the COST 2100 channel model. Besides, key statistical properties of the proposed extensions, e.g., autocorrelation functions, maximum likelihood estimators, and Cramer-Rao bounds, are derived and analyzed. Furthermore, the performance of a spatial modulation massive MIMO system is investigated in [14] under a non-stationary

channel model. Authors show that spatial modulation can outperform typical employed spatial multiplexing transmission in certain scenarios of low correlation among sub-channels, for example under a rich scattering environment.

A novel random access (RA) protocol for crowded XL-MIMO systems is proposed in [15]. Authors have proposed a decentralized and uncoordinated decision rule, which can be evaluated at the users side, for retransmitting or not the RA pilots during the connection stage, taking advantage of the XL-MIMO propagation features. The proposed protocol achieves significant performance improvements in terms of reducing the connection delay and providing access for larger number of devices.

A. Motivation, Contributions and Novelties in Comparison with Existing Works

Current design approaches in telecommunication systems include a global effort in saving energy and reducing pollution [2], [16], [17]. We show in this paper that antenna selection (AS) methods in XL-MIMO systems is a very important issue since the energy expenditure of such systems could be very high if activating the radio frequency (RF) chains of all antennas simultaneously. Besides, some antennas might contribute very little with the system performance due to the non-stationarities and visibility regions, in such a way that the power required to activate their RF chain becomes a burden that severely penalizes the total energy efficiency (EE) of the system. Therefore, the very large number of antennas deployed in the XL-MIMO systems in conjunction with the spatial non-stationarities make the application of AS schemes very important.

The main *contributions* of this work are threefold:

(i) Reformulating the signal to interference plus noise ratio (SINR) performance expressions of [5], considering long-term fading variations across the array and incorporating the maximum transmit power constraint into the expressions for CB and ZF, and finding more compact and comprehensive results, readily applicable for antenna selection procedures.

(ii) Based on the obtained expressions, and on a realistic power consumption model, we evaluate the total EE of the XL-MIMO system. Besides, we propose and compare four low-complexity AS procedures aiming to maximize the total EE of the system, different than [3], [4] which proposed SE-based AS schemes. Our proposed schemes are based solely on the long-term fading parameters, and the obtained solutions remain valid for larger time/frequency intervals.

(iii) Based on our proposed AS schemes, and some simplifying assumptions, we derive approximated closed-form EE expressions, and propose an iterative method for finding the optimal number of selected antennas which maximizes EE. Finally, numerical simulations have validated the proposed performance expressions and compared the different XL-MIMO AS schemes.

AS methods for typical spatially stationary massive MIMO systems [18], [19] is a well investigated topic. However, the

XL-MIMO system is a different scenario. While the spatially stationary model applies for typical cellular systems, where the BS antenna array dimension is much lower than the distance to the users and a single long-term fading coefficient holds for all antennas, significant power variations appear along the XL-MIMO array, due to its large dimension and number of antennas, and proximity with users. The non-stationary XL-MIMO scenario just very recently was introduced in the literature. To the best of our knowledge, this contribution is the first evaluating the EE of the XL-MIMO scenario, showing that AS methods are especially important to improve EE due to the spatial non-stationarities that naturally arise in XL-MIMO systems, proposing long-term fading based AS procedures, and deriving the optimal number of active antennas for this new wireless communication context.

With respect to the existing XL-MIMO literature, we can point out as the *main novelties* of our paper: although our system model and CB and ZF performance expressions are similar to that of [5], authors have considered a binary visibility region model for the XL-MIMO scenario, in which no long-term fading variation occurs for the visible antennas. Besides, performance expressions are dependent of power co-efficients obtained resolving a separated optimization problem for meeting power constraint, and no antenna selection is considered. Differently, we incorporated the power constraint into the performance expressions, arriving at more compact and comprehensive results, readily applicable for AS procedures, and considered long-term fading variations along the array. Besides, AS for XL-MIMO systems has been investigated only in [3], [4] at the moment of writing this paper; however, both works proposed SE-based AS schemes for XL-MIMO systems. Differently, based on only long-term fading coefficients, we propose AS schemes aiming to maximize the XL-MIMO total EE, since this is a very important issue due to the very large number of antennas at the XL-MIMO array, and the non-stationarities and visibility regions which arise in this scenario. Furthermore, the long-term fading approach has the advantages of being simpler than short-term ones, and of providing solutions which remain valid for larger time periods and all subcarriers (if employing a wideband system), reducing the computational complexity of the antenna selection approach and simplifying hardware due to switching and RF chain on-off requirements.

Notations: Boldface lower and upper case symbols represent vectors and matrices, respectively. $\mathbf{I}_N$ denotes the identity matrix of size N, while $\{\cdot\}^T$ and $\{\cdot\}^H$ denote the transpose and the Hermitian transpose operator, respectively. We use $\mathcal{CN}(m, \sigma^2)$ when referring to a circular symmetric complex Gaussian distribution with mean m and variance matrix σ^2. Besides, $\mathrm{tr}(\cdot)$ and $\mathrm{diag}(\cdot)$ are the trace and diagonal matrix operators, respectively, while $[\mathbf{A}]_{i,j}$ holds to the element in the ith row and jth column of matrix $\mathbf{A}$, and $\mathbf{a}_i$ refers to its ith column vector.

II. SYSTEM MODEL

We consider a base station (BS) equipped with a linear XL-MIMO array with M antennas uniformly distributed along a

length of L meters, Fig. 1. In front of the extra-large array structure, K users are randomly distributed in a rectangular area, of length L in the array parallel dimension, and with a distance to the array in the range $[0.1 \cdot L, L]^1$. Since the distances of the users to the antennas is of the same order of the array length L the average received power varies along the XL-MIMO array, and therefore we cannot consider a single long-term fading coefficient for a given user [3], [8]. Instead, we consider a long-term fading coefficient $\beta_{m,k}$ regarding the m-th antenna of the XL-MIMO array and the k-th user, similarly as in [3], [9], [10], [15], given by

$$\beta_{m,k} = q \cdot d_{m,k}^{-\kappa}, \tag{1}$$

in which q is a constant determining the path loss in a reference distance, $d_{m,k}$ is the distance between the m-th antenna of the XL-MIMO array and the k-th user, and κ is the path loss decay exponent. The channel matrix $\mathbf{H} \in \mathbb{C}^{M \times K}$ is thus formed by elements $h_{m,k} = \sqrt{\beta_{m,k}} \cdot \underline{h}_{m,k}$, in which $\underline{h}_{m,k} \sim \mathcal{CN}(0, 1)$, assuming a rich scattering environment as in [4], [5]. If we arrange the long-term fading coefficients of a user in a diagonal matrix:

$$\mathbf{R}_k = \mathrm{diag}([\beta_{1,k}, \beta_{2,k}, \ldots, \beta_{M,k}]) \in \mathbb{R}^{M \times M}, \tag{2}$$

and the elements $\underline{h}_{m,k}$ in a vector $\underline{\mathbf{h}}_k \in \mathbb{C}^{M \times 1}$, we have that each column of $\mathbf{H}$ can be defined as $\mathbf{h}_k = \mathbf{R}_k^{\frac{1}{2}} \underline{\mathbf{h}}_k$ as in [5].
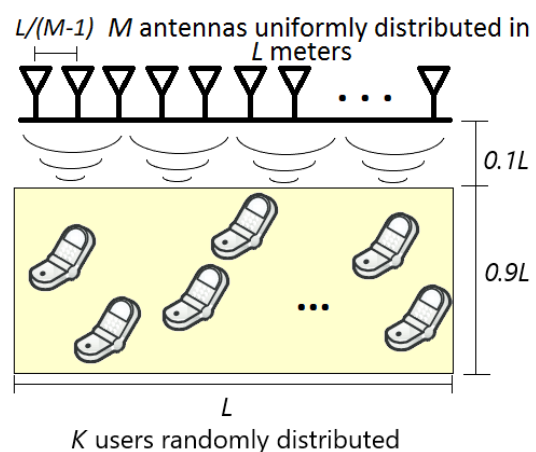
Fig. 1. Illustration of the adopted system model.

In the DL, considering an average received signal-to-noise ratio (SNR) ρ at the users, an average long-term fading coefficient β_{avg} (among all antennas and users' positions), and a uniform power allocation policy for the users, the total transmit power, $P_{\max}$, should satisfy [1]

$$\rho = \frac{P_{\max} \cdot \beta_{\mathrm{avg}}}{\sigma^2}, \tag{3}$$

in which σ^2 is the noise power. Since the channel gain $\beta_{m,k}$ varies significantly along the array, it is more effective to select just the stronger antennas to transmit signal to the k-th user, reducing the number of active antennas, as well the power spent with power-hungry RF chains. We discuss in the

---

[1]In order to guarantee a minimum distance of the users to the XL-MIMO array, as in [10], [16].

next Section different approaches to obtain the set of antennas selected to serve the users, $\mathcal{A}$. For simplicity, we considered $\beta_{\text{avg}} \approx q \cdot L^{-\kappa}$ in our simulations. The signal for user k, s_k, is precoded by $\mathbf{g}_k \in \mathbb{C}^{M \times 1}$ and scaled by $p_k \geq 0$, which adjusts the signal power, before transmission. Considering a similar XL-MIMO system model than [5], the transmit vector $\mathbf{x}$ is the linear combination of the precoded and scaled signal of all the users, *i.e.*,

$$\mathbf{x} = \sum_{k=1}^{K} \sqrt{p_k} \cdot \mathbf{g}_k \cdot s_k. \tag{4}$$

Let $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_K] \in \mathbb{C}^{M \times K}$ be the combined precoding matrix, and $\mathbf{P} = \text{diag}([p_1, p_2, \ldots, p_K]) \in \mathbb{R}^{K \times K}$ be the diagonal matrix of signal powers. The combined precoding matrix $\mathbf{G}$ is normalized to satisfy the power constraint

$$\mathbb{E}[||\mathbf{x}||^2] = \text{tr}(\mathbf{P}\mathbf{G}^H\mathbf{G}) = P_{\max}. \tag{5}$$

The signal received by the k-th user is

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \quad k = 1, 2, \ldots K, \tag{6}$$

in which $n_k \sim \mathcal{CN}(0, \sigma^2)$ is an additive white Gaussian noise (AWGN) sample. Assuming independent Gaussian signaling, *i.e.*, $s_k \sim \mathcal{CN}(0,1)$ and $\mathbb{E}[s_i s_j^*] = 0$, $i \neq j$, the SINR γ_k of the k-th user can be defined as [5]:

$$\gamma_k = \frac{p_k|\mathbf{h}_k^H\mathbf{g}_k|^2}{\sum_{j=1,j\neq k}^{K} p_j|\mathbf{h}_k^H\mathbf{g}_j|^2 + \sigma^2}. \tag{7}$$

We selected the CB and ZF approaches as representative low-complexity linear precoding schemes. The CB precoder matrix is simply defined as

$$\mathbf{G}_{\text{CB}} = \alpha_{\text{CB}}\mathbf{H}, \tag{8}$$

and the ZF precoding matrix is

$$\mathbf{G}_{\text{ZF}} = \alpha_{\text{ZF}}\mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1}, \tag{9}$$

where the scaling factors $\alpha_{\text{CB}} = \sqrt{P_{\max}/\text{tr}(\mathbf{P}\mathbf{H}^H\mathbf{H})}$ and $\alpha_{\text{ZF}} = \sqrt{P_{\max}/\text{tr}(\mathbf{P}(\mathbf{H}^H\mathbf{H})^{-1})}$ ensure that the power constraint (5) is met.

Using (8) in (7), the SINR of the kth user for CB is

$$\gamma_k^{(\text{CB})} = \frac{p_k|\mathbf{h}_k^H\mathbf{h}_k|^2}{\sum_{j=1,j\neq k}^{K} p_j|\mathbf{h}_k^H\mathbf{h}_j|^2 + \frac{\sigma^2}{P_{\max}}\text{tr}(\mathbf{P}\mathbf{H}^H\mathbf{H})}. \tag{10}$$

Similarly, using (9) in (7), the SINR of the kth user for ZF is

$$\gamma_k^{(\text{ZF})} = \frac{p_k P_{\max}}{\sigma^2\text{tr}(\mathbf{P}(\mathbf{H}^H\mathbf{H})^{-1})}. \tag{11}$$

Given the system model presented in this Section in eq. (1)–(11), and the deterministic equivalent analysis of [20], it is presented in [5] the deterministic equivalent of $\gamma_k^{(\text{CB})}$ in (10) as

$$\overline{\gamma}_k^{(\text{CB})} = \frac{p_k(\text{tr}(\mathbf{R}_k))^2}{\sum_{j=1,j\neq k}^{K} p_j\text{tr}(\mathbf{R}_k\mathbf{R}_j) + \frac{\sigma^2}{P_{\max}}\sum_{j=1}^{K} p_j\text{tr}(\mathbf{R}_j)}, \tag{12}$$

and the deterministic equivalent of $\gamma_k^{(\text{ZF})}$ in (11) as

$$\overline{\gamma}_k^{(\text{ZF})} = \frac{p_k P_{\max}}{\sigma^2\sum_{i=1}^{K} p_i\left(\text{tr}(\mathbf{R}_i) - \sum_{j=1,j\neq i}^{K}\frac{\text{tr}(\mathbf{R}_i\mathbf{R}_j)}{\text{tr}(\mathbf{R}_j)}\right)^{-1}}. \tag{13}$$

where $\mathbf{R}_i$ is defined as in (2).

Having found the SINR of the kth user, the spectral efficiency is readily obtained as $\eta_k^s = \log_2(1 + \gamma_k)$. On the other hand, the energy efficiency is [16], [17]

$$\eta_e = \frac{B\sum_{k=1}^{K}\eta_k^s}{\mathcal{P}}, \tag{14}$$

in which B is the system bandwidth, and $\mathcal{P}$ is the total power consumption, discussed in Section II-C.

A. Further Advances in the Performance Expressions

We revisit the performance expressions for non-stationary XL-MIMO discussed in [5], while propose further elaborations to arrive at lean and more comprehensive results. Note that the results of (12) and (13) depend on the signal powers in both numerator and denominators, and such coefficients should be chosen in order to satisfy the power constraint in (5). In the simulation code made available by the authors of [5], they apply the CVX solver of [21] to find a matrix $\mathbf{P}$ satisfying (5). This makes the performance expressions less intuitive, while limiting the application of AS schemes as proposed in Section III of this paper. Hence, in this subsection, we shed light on deriving self-contained closed-form SINR expressions recalling the channel hardening massive MIMO properties. For that, we first rewrite (5) in the following form:

$$\mathbb{E}[||\mathbf{x}||^2] = \text{tr}(\mathbf{P}\mathbf{G}^H\mathbf{G}) = \sum_{k=1}^{K} p_k||\mathbf{g}_k||^2 = P_{\max}. \tag{15}$$

If a uniform power allocation scheme is applied, the following equality holds

$$p_k||\mathbf{g}_k||^2 = \frac{P_{\max}}{K}, \quad k = 1, 2, \ldots K. \tag{16}$$

Hence, when adopting CB, eq. (16) becomes

$$p_k\alpha_{\text{CB}}^2||\mathbf{h}_k||^2 = \frac{P_{\max}}{K}, \quad k = 1, 2, \ldots K, \tag{17}$$

and we have an undetermined system with K equations and $K+1$ variables. By choosing $\alpha_{\text{CB}} = 1$ for simplicity, the p_k coefficients can be obtained for CB as

$$p_k^{(\text{CB})} = \frac{P_{\max}}{K||\mathbf{h}_k||^2}, \quad k = 1, 2, \ldots K. \tag{18}$$

Following similar assumptions as in [5], we have that

$$||\mathbf{h}_k||^2 = \mathbf{h}_k^H\mathbf{h}_k = \underline{\mathbf{h}}_k^H\mathbf{R}_k\underline{\mathbf{h}}_k \xrightarrow{M\to\infty} \text{tr}(\mathbf{R}_k), \tag{19}$$

and a deterministic equivalent of (18) is

$$p_k^{(\text{CB})} = \frac{P_{\max}}{K\text{tr}(\mathbf{R}_k)}, \quad k = 1, 2, \ldots K. \tag{20}$$

Substituting (20) in (12), we arrive at

$$\overline{\gamma}_k^{(\text{CB})} = \frac{\text{tr}(\mathbf{R}_k)}{\sum_{j=1,j\neq k}^{K}\frac{\text{tr}(\mathbf{R}_k\mathbf{R}_j)}{\text{tr}(\mathbf{R}_j)} + \frac{K\sigma^2}{P_{\max}}}. \tag{21}$$

On the other hand, for the case of ZF, (5) becomes

$$
\begin{aligned}
\mathbb{E}[||\mathbf{x}||^2] &= \mathrm{tr}\left(\mathbf{P}\mathbf{G}^H\mathbf{G}\right) = P_{\max}, \\
&= \alpha_{\mathrm{ZF}}^2 \mathrm{tr}\left(\mathbf{P}(\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1}\right) = P_{\max}, \\
&= \alpha_{\mathrm{ZF}}^2 \mathrm{tr}\left(\mathbf{P}(\mathbf{H}^H\mathbf{H})^{-1}\right) = P_{\max}, \\
&= \alpha_{\mathrm{ZF}}^2 \mathrm{tr}\left(\mathbf{P}\mathbf{V}\right) = P_{\max}, \quad (22)
\end{aligned}
$$

in which the matrix $\mathbf{V}$ is a diagonal matrix formed by the main diagonal elements of $(\mathbf{H}^H\mathbf{H})^{-1}$. We can thus rewrite (22) as

$$
\alpha_{\mathrm{ZF}}^2 \sum_{k=1}^{K} p_k [\mathbf{V}]_{k,k} = P_{\max}, \quad (23)
$$

and if a uniform power allocation is employed

$$
\alpha_{\mathrm{ZF}}^2 \, p_k \, [\mathbf{V}]_{k,k} = \frac{P_{\max}}{K}, \quad k = 1, 2, \dots K. \quad (24)
$$

Again, making $\alpha_{\mathrm{ZF}} = 1$, the p_k coefficients can be obtained for the ZF precoding as

$$
p_k^{(\mathrm{ZF})} = \frac{P_{\max}}{K [\mathbf{V}]_{k,k}}, \quad k = 1, 2, \dots K. \quad (25)
$$

Following the analysis of [5, App. A], it can be shown that

$$
[\mathbf{V}]_{k,k} \xrightarrow{M\to\infty} \left(\mathrm{tr}(\mathbf{R}_k) - \sum_{j=1, j\neq k}^{K} \frac{\mathrm{tr}(\mathbf{R}_k\mathbf{R}_j)}{\mathrm{tr}(\mathbf{R}_j)} \right)^{-1}, \quad (26)
$$

and a deterministic equivalent of (25) is

$$
p_k^{(\mathrm{ZF})} = \frac{P_{\max}}{K} \left(\mathrm{tr}(\mathbf{R}_k) - \sum_{j=1, j\neq k}^{K} \frac{\mathrm{tr}(\mathbf{R}_k\mathbf{R}_j)}{\mathrm{tr}(\mathbf{R}_j)} \right), \quad k = 1, \dots K. \quad (27)
$$

Substituting (27) in (13), we arrive at

$$
\overline{\gamma}_k^{(\mathrm{ZF})} = \frac{P_{\max}}{K\sigma^2} \left(\mathrm{tr}(\mathbf{R}_k) - \sum_{j=1, j\neq k}^{K} \frac{\mathrm{tr}(\mathbf{R}_k\mathbf{R}_j)}{\mathrm{tr}(\mathbf{R}_j)} \right). \quad (28)
$$

Equations (21) and (28) show the XL-MIMO DL system performance employing CB and ZF, respectively, as further extensions of eq. (12) and (13) from [5]. This is a first contribution of this manuscript, which serves as basis for the following EE and AS analysis.

Remark 1: Although we have considered $\alpha_{\mathrm{CB}} = \alpha_{\mathrm{ZF}} = 1$ in our analysis, any other choice for these parameters would result in the same expressions, since would affect every numerator and denominator terms in the same way.

Remark 2: The SINR performance expressions presented in [5, Table I] can be seen as particular cases of (21) and (28) when neglecting long-term fading and applying the normalization $\mathrm{tr}(\mathbf{R}_k) = \mathrm{tr}(\boldsymbol{\Theta}_k) = M$ or $\mathrm{tr}(\boldsymbol{\Theta}_k) = D$, where $\boldsymbol{\Theta}_k$ and D are the matrix describing the VR of kth user and the number of visible antennas per user, respectively, as in [5].

B. Antenna Selection Model

Given our deterministic equivalent performance expressions for CB and ZF in eq. (21) and (28), respectively, we can rewrite these expressions considering the activation subset of antennas. Hence, denoting $\mathcal{A}$ as the set containing the indices

of the active antennas, the deterministic equivalent SINR for the CB precoding results

$$
\overline{\gamma}_k^{(\mathrm{CB})} = \frac{\sum_{m\in\mathcal{A}} \beta_{m,k}}{\sum_{j=1, j\neq k}^{K} \frac{\sum_{m\in\mathcal{A}} \beta_{m,k}\beta_{m,j}}{\sum_{m\in\mathcal{A}} \beta_{m,j}} + \frac{K\sigma^2}{P_{\max}}}, \quad (29)
$$

while for the ZF:

$$
\overline{\gamma}_k^{(\mathrm{ZF})} = \frac{P_{\max}}{K\sigma^2} \left(\sum_{m\in\mathcal{A}} \beta_{m,k} - \sum_{j=1, j\neq k}^{K} \frac{\sum_{m\in\mathcal{A}} \beta_{m,k}\beta_{m,j}}{\sum_{m\in\mathcal{A}} \beta_{m,j}} \right). \quad (30)
$$

It is worth to note that, in our formulation, the *activation subset of antennas* is the same for all users, differently from [3], in which each user has its own set of active antennas aiming to maximize the system sum-rate. We justify our formulation since, when aiming to maximize the total energy efficiency, once the power-hungry RF chain of an antenna is active, it is better to take full advantage of it, transmitting signal for all users. It has no significant benefit in defining the activation subset of antennas in a per-user fashion, since the ZF approach is able to eliminate the inter-user interference, while the power increment necessary to compute the precoding vector with a slightly large number of antennas is small if compared to the power to activate the RF chain of the additional antenna, as evinced in the next subsection. Besides, it would result in more complicated performance expressions, probably in terms of short-term fading coefficients, and the dimension of the search space of the AS algorithms would scale with K, becoming considerably more complex and power consuming.

C. Power Consumption Model

We follow the same power consumption model of [16], which is very similar to that in [17], and is a very realistic model. However, as we focus on the DL transmission, we do not consider the UL data rates as well as the UL transmit powers. In the XL-MIMO scenario analysed herein, we consider the power expenditures of the irradiated DL data signal (with the amplifier efficiency), $P_{\mathrm{TX}}^{\mathrm{DL}}$, the UL training, $P_{\mathrm{TX}}^{\mathrm{tr}}$, the channel estimation, P_{CE}, the coding/decoding, $P_{\mathrm{C/D}}$, the backhaul, P_{BH}, the linear processing computation, P_{PR}, the transceiver chains, P_{TC}, and a fixed quantity regarding the circuitry power consumption required for site-cooling, control signaling, and load-independent power of backhaul infrastructure and baseband processors, P_{FIX}. Thus, the overall power consumption results

$$
\mathcal{P} = P_{\mathrm{TX}}^{\mathrm{DL}} + P_{\mathrm{TX}}^{\mathrm{tr}} + P_{\mathrm{CE}} + P_{\mathrm{C/D}} + P_{\mathrm{BH}} + P_{\mathrm{PR}} + P_{\mathrm{TC}} + P_{\mathrm{FIX}}. \quad (31)
$$

Our objective here is to investigate the dependence of the selected subset of antennas, $\mathcal{A}$, with the total energy efficiency of the system. Note that the total energy efficiency of the system depends on $\mathcal{A}$ in different ways. First, the sum rate of the system depends on the SE of the users, which is a function of their SINRs dependent of $\mathcal{A}$. Moreover, the sum rate impacts on the power expenditures of the coding/decoding, and the backhaul. Besides, the power consumption of the

transceiver chains is modeled as

$$P_{\text{TC}} = P_{\text{SYN}} + |\mathcal{A}|P_{\text{BS}} + KP_{\text{MT}}, \qquad (32)$$

in which P_{SYN} is the power of the local oscillator, P_{BS} is the power required to each active BS antenna operate, while P_{MT} is the power required to each single-antenna mobile terminal (MT) operate. Note that M is usually very high in an XL-MIMO system[2], while P_{BS} accounting for the power-hungry RF chains is considered in [16] as 1 W per antenna. Thus, activating the RF chains of all BS antennas would result in a very large power expenditure, in such a way that it is very important to perform a suitable antenna selection procedure.

The power consumed with processing, P_{PR}, corresponds to the power required to obtain the transmit signal in (4), to obtain the precoding matrix, and to obtain the AS set. Note that this power is also dependent on the number of active antennas $|\mathcal{A}|$. Following the model in [16], but including the term of power related to the AS processing, we have

$$P_{\text{PR}} = B\left(1 - \frac{\tau}{S}\right)\frac{\mathcal{C}_{\text{ts}}}{\mathcal{L}_{\text{BS}}} + \frac{B}{S}\frac{\mathcal{C}_{\text{prec}}}{\mathcal{L}_{\text{BS}}} + \frac{1}{T_{\text{LT}}}\frac{\mathcal{C}_{\text{as}}}{\mathcal{L}_{\text{BS}}}, \qquad (33)$$

in which τ is the length of the uplink pilot signals, S is the coherence block size, $\mathcal{C}_{\text{ts}}$ is the computational complexity for evaluating eq. (4). Besides, $\mathcal{L}_{\text{BS}}$ is the computational efficiency of the BS (in W/$flop$), $\mathcal{C}_{\text{prec}}$ is the complexity of obtaining the precoding vectors for all users, T_{LT} is the long-term fading coherence time, and $\mathcal{C}_{\text{as}}$ is the complexity of obtaining the antenna selection set. The obtained AS set remains valid for a long-term coherence interval, since our analysis is based only in long-term fading parameters. One can see from (33) that this approach results in a lower influence of the AS set computation in P_{PR}, since it is multiplied by the factor $1/T_{\text{LT}}$, which is much lower than B/S and $B\left(1 - \frac{\tau}{S}\right)$.

Following the analysis in [16], [17], we consider 1 $flop$ as an arithmethic operation between two complex numbers. Thus, the multiplication between a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ and a matrix $\mathbf{B} \in \mathbb{C}^{n \times p}$ spends $2mnp$ flops. Therefore, we have $\mathcal{C}_{\text{ts}} = 2|\mathcal{A}|K$ $flops$ from [17]. Besides, if using the CB precoder, $\mathcal{C}_{\text{prec}} = \mathcal{C}_{\text{CB}} = 3|\mathcal{A}|K$ $flops$ from [17], against $\mathcal{C}_{\text{prec}} = \mathcal{C}_{\text{ZF}} = K^3/3 + 3|\mathcal{A}|K^2 + |\mathcal{A}|K$ $flops$ if adopting ZF. The complexity $\mathcal{C}_{\text{as}}$ is discussed in the next Section. Besides, the terms in (31) not discussed in this Section can be computed in the same way as in [16].

Finally, we can rewrite (31) as

$$\mathcal{P} = \mathcal{P}^{\dagger} + P_{\text{CE}} + P_{\text{C/D}} + P_{\text{BH}} + P_{\text{PR}} + |\mathcal{A}|P_{\text{BS}}, \qquad (34)$$

in which we have gathered the power components that do not depend of $\mathcal{A}$ in the term:

$$\mathcal{P}^{\dagger} = P_{\text{TX}}^{\text{DL}} + P_{\text{TX}}^{\text{tr}} + P_{\text{SYN}} + KP_{\text{MT}} + P_{\text{FIX}}. \qquad (35)$$

The dependence of the terms in (34) with $\mathcal{A}$ can be justified as follows: P_{CE} depends on $\mathcal{A}$ since the short-term channel estimates are obtained only for the active antennas, $P_{\text{C/D}}$ and P_{BH} because they depend on the system sum-rate, which depends on $\mathcal{A}$, and P_{PR} because the processing complexity is dependent on the number of active antennas.

[2]Typically hundreds or even thousands of antennas.

III. ANTENNA SELECTION SCHEMES

In this section we propose different AS schemes for XL-MIMO aiming to obtain a suitable subset of antennas $\mathcal{A}$ selected to transmit the DL signal to the mobile users subject to channel non-stationarities. First we propose a simple, deterministic, greedy scheme based on the *highest received normalized power* (HRNP) criterion. Then, three heuristic schemes are proposed using the HRNP active antennas set as initial solution: *local search* (LS), *genetic algorithm* (GA), and *particle swarm optimization* (PSO).

A. HRNP criterion

A first and greedy approach is to select just the M_s antennas responsible for the major part of the power received by the users. However, since closer users receive more power, this should be performed in a normalized fashion in order to achieve a fair result for all users. In this case, we first compute the metric:

$$\varphi_m = \sum_{k=1}^{K} \frac{\beta_{m,k}}{\sum_{j=1}^{M} \beta_{j,k}}, \quad m = 1, 2, \ldots, M. \qquad (36)$$

Then, the selected subset of antennas $\mathcal{A}^{\text{HRNP}}$ will be composed by the M_s antennas with the highest values of φ_m. A pseudo-code for the HRNP-AS procedure is presented in Algorithm 1, in which $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \ldots \varphi_M]$.

The complexity[3] of the HRNP AS scheme is described by

$$\mathcal{C}_{\text{as}}^{\text{HRNP}} = 3MK + M\log(M) \ [flops], \qquad (37)$$

corresponding to the computation of (36) for all antennas, and a sorting algorithm to select the M_s antennas with highest φ_m. It is noteworthy, however, that the HRNP EE performance is highly dependent on the M_s choice, since the system would provide low sum-rates with few active antennas, or it would consume a high power with many active antennas. Thus, we propose in Section IV an approximated closed-form analytical expression for the EE of the XL-MIMO system employing ZF and HRNP-AS as a function of M_s. Then, we propose an iterative method for obtaining the M_s value which maximizes this expression. We do not consider the complexity of this method in eq. (37) since it is not dependent on the channel parameters, but only controlled by the system parameters, such as the number of users, transmit power, dimensions of XL-MIMO array and coverage area. Therefore, its computation can be performed over larger time periods. We discuss in Section V-A the complexity of the proposed method for obtaining the optimal M_s value.

B. LS-based Antenna Selection

A simple strategy for seeking a better active antennas set is to perform a local search (LS) in the neighborhood of the HRNP solution. For this purpose, we first represent the set $\mathcal{A}$ as a binary vector $\mathbf{a}$ of length M, in which if $m \in \mathcal{A}$, $a_m = 1$; otherwise $a_m = 0$. Then, we compute the total

[3]We evaluate the computational complexities of the investigated schemes in terms of *floating point operations (flops)*, defined as an addition, subtraction, multiplication or division between two floating point numbers [22].

Algorithm 1 Proposed HRNP AS Scheme

Input: M_s, $\beta_{m,k}$, $\forall m, k$.

1: Initialize $\mathcal{A}$ as an empty set;
2: **for** $m = 1, 2, \ldots, M$ **do**
3:    Evaluate φ_m as in (36);
4: **end for**
5: **for** $n = 1, 2, \ldots, M_s$ **do**
6:    Evaluate $a = \arg\max_m \varphi_m$;
7:    Update $\mathcal{A}$ as $\mathcal{A} = \mathcal{A} \cup a$;
8:    Remove φ_a from $\boldsymbol{\varphi}$;
9: **end for**

Output: $\mathcal{A}^{\mathrm{HRNP}}$.

---

energy efficiency (14) of every candidate within a certain Hamming distance d_{Ham} from it. If a better candidate is found, the solution is updated, and the procedure is repeated on its neighborhood. This iterative procedure is repeated for a predefined number of iterations or until the convergence. A pseudo-code representation of the LS-based AS scheme is provided in Algorithm 2, in which N_{el} as defined in step 3 is the number of elements within the Hamming distance d_{Ham} from the current solution. For simplicity, we have limited our search with a unitary Hamming distance.

The complexity of the LS-AS scheme is

$$\mathcal{C}_{\mathrm{as}}^{\mathrm{LS}} = \mathcal{C}_{\mathrm{as}}^{\mathrm{HRNP}} + \overline{N}_{it} M \mathcal{C}_{\mathrm{EE}} \quad [flops], \qquad (38)$$

in which $\overline{N}_{it}$ is the average number of iterations until convergence, and $\mathcal{C}_{\mathrm{EE}} = 2MK^2 [flops]$ is the complexity of computing the total energy efficiency cost function. An interesting point to observe in the LS algorithm is that if a new solution is not found into an iteration, the search can be interrupted, since the algorithm has converged. This contributes to decrease the complexity of the algorithm, and, therefore, improve EE.

Algorithm 2 Proposed LS-based AS Scheme

Input: d_{Ham}, $N_{it}^{\max}$, $\mathcal{A}^{\mathrm{HRNP}}$, $\beta_{m,k}$, $\forall m, k$.

1: Initialize $\mathbf{a}^0$ as the binary vector representation of $\mathcal{A}^{\mathrm{HRNP}}$;
2: Initialize η_e^{best} as the total energy efficiency of $\mathbf{a}^0$;
3: Evaluate $N_{el} = \binom{M}{d_{\mathrm{Ham}}}$;
4: **for** $n = 1, 2, \ldots, N_{it}^{\max}$ **do**
5:    Generate the search space matrix $\mathbf{S}$ of size $M \times N_{el}$ with all vectors within the distance d_{Ham} from $\mathbf{a}^{n-1}$;
6:    **for** $\ell = 1, 2, \ldots, N_{el}$ **do**
7:       Evaluate η_e as the total energy efficiency of $\mathbf{s}_\ell$;
8:       **if** $\eta_e > \eta_e^{\mathrm{best}}$ **then**
9:          Update $\eta_e^{\mathrm{best}} = \eta_e$, and $\mathbf{a}^n = \mathbf{s}_\ell$;
10:       **else**
11:          **Break;**
12:       **end if**
13:    **end for**
14: **end for**

Output: $\mathcal{A}^{\mathrm{LS}}$ as the set representation of $\mathbf{a}^n$.

C. GA-based Antenna Selection

The genetic algorithm is a widely-known bio-inspired heuristic optimization algorithm, which has been used to solve optimization problems in different areas. In the context of massive MIMO antenna selection, GA has been employed in the conventional stationary case in [18]. Herein, we employ a similar algorithm from [18], but adjusted to the non-stationary XL-MIMO configurations. The GA-AS uses the HRNP output as initial solution, also, other random candidates forming an initial population of size p_{GA}, which is evaluated in terms of the cost function in (14). A given number ϕ of the best candidates in this population is selected as parents, which will generate descendants in a new population. For this purpose, two parents are selected at random for each descendant, and the crossover operator is applied with a random crossover point. Then, the mutation operator is also applied, which inverts the entries of each candidate with certain probability p_{mut}. After a predefined number of iterations or until the convergence of the algorithm, it returns the best solution found so far. A pseudo-code representation for the GA-based AS scheme is provided in Algorithm 3.

The complexity of our proposed GA-AS procedure is

$$\mathcal{C}_{\mathrm{as}}^{\mathrm{GA}} = \mathcal{C}_{\mathrm{as}}^{\mathrm{HRNP}} + \overline{N}_{it}[p_{\mathrm{GA}} \mathcal{C}_{\mathrm{EE}} + p_{\mathrm{GA}} \log(p_{\mathrm{GA}})] \quad [flops], \quad (39)$$

due to the cost function evaluation of each candidate in the population, and a sorting algorithm for selecting the best candidates.

Algorithm 3 Proposed GA-based AS Scheme

Input: p_{GA}, ϕ, p_{mut}, $\mathcal{A}^{\mathrm{HRNP}}$, $N_{it}^{\max}$, $\beta_{m,k}$, $\forall m, k$.

1: Initialize the population $\boldsymbol{\Theta}^{\mathrm{GA}}$ with the binary vector representation of $\mathcal{A}^{\mathrm{HRNP}}$ and other p_{GA}-1 random binary vectors;
2: Evaluate the total energy efficiency of each candidate in $\boldsymbol{\Theta}^{\mathrm{GA}}$, forming the vector $\boldsymbol{\eta}_e^{\mathrm{GA}}$;
3: Sort $\boldsymbol{\eta}_e^{\mathrm{GA}}$ in descending order, reorganizing the columns of $\boldsymbol{\Theta}^{\mathrm{GA}}$ accordingly;
4: Initialize $\eta_e^{\mathrm{best}} = \eta_{e,1}^{\mathrm{GA}}$, and $\mathbf{a}^{\mathrm{GA}} = \boldsymbol{\theta}_1^{\mathrm{GA}}$;
5: **for** $n = 2, 3, \ldots, N_{it}^{\max}$ **do**
6:    **for** $\ell = 1, 2, \ldots, p_{\mathrm{GA}}$ **do**
7:       Generate two different random integers $\in [1, \phi]$ to be the parents of $\boldsymbol{\theta}_\ell^{\mathrm{GA}}$, applying the crossover operator in a random crossover point $\in [2, M]$;
8:       Apply the mutation operator in $\boldsymbol{\theta}_\ell^{\mathrm{GA}}$ with probability p_{mut};
9:       Evaluate the total energy efficiency of $\boldsymbol{\theta}_\ell^{\mathrm{GA}}$, and assign it to $\eta_{e,\ell}^{\mathrm{GA}}$;
10:    **end for**
11:    Sort $\boldsymbol{\eta}_e^{\mathrm{GA}}$ in descending order, reorganizing the columns of $\boldsymbol{\Theta}^{\mathrm{GA}}$ accordingly;
12:    **if** $\eta_{e,1}^{\mathrm{GA}} > \eta_e^{\mathrm{best}}$ **then**
13:       Update $\eta_e^{\mathrm{best}} = \eta_{e,1}^{\mathrm{GA}}$, and $\mathbf{a}^{\mathrm{GA}} = \boldsymbol{\theta}_1^{\mathrm{GA}}$;
14:    **end if**
15: **end for**

Output: $\mathcal{A}^{\mathrm{GA}}$ as the set representation of $\mathbf{a}^{\mathrm{GA}}$.

D. PSO-based Antenna Selection

The particle swarm optimization algorithm is another bio-inspired optimization algorithm, similarly as GA. However, it is commonly recognized as a simpler algorithm, in terms of fewer mechanisms to escape from local maxima, and reduced computational complexity per iteration. Therefore, we also suggest the use of a PSO-based AS scheme for the non-stationary XL-MIMO case, similarly as proposed in [19] for conventional stationary massive MIMO scenario.

The PSO-AS algorithm uses the HRNP output as initial solution, as well as other random candidates to form an initial swarm of p_{PSO} particles. At each iteration, each particle updates its position in terms of its previous velocity (inertial effect, with inertia weight ν), its individual best solution found (cognitive information, with cognitive factor μ_c), and the best solution found by all particles (social information, with social factor μ_s). After a predefined number of iterations or the convergence of the algorithm, it returns the best solution found. A pseudo-code representation for the PSO-based AS scheme is provided in Algorithm 4, in which $\mathbf{\Gamma} \in \mathbb{R}^{M \times p_{\text{PSO}}}$ is a random matrix generated each time it is called with each element uniformly distributed in $[0, 1]$ interval, and $\texttt{binround}(x)$ is the binary round operator, which returns 1 if $x > 0.5$, and 0 otherwise.

The complexity of the proposed PSO-AS algorithm is

$$\mathcal{C}_{\text{as}}^{\text{PSO}} = \mathcal{C}_{\text{as}}^{\text{HRNP}} + \overline{N}_{it}(p_{\text{PSO}}\mathcal{C}_{\text{EE}} + p_{\text{PSO}}) \quad [flops], \quad (40)$$

due to the cost function evaluation (14) for all particles and finding the maximum EE particle, at each iteration.

Algorithm 4 Proposed PSO-based AS Scheme

Input: p_{PSO}, ν, μ_c, μ_s, $\mathcal{A}^{\text{HRNP}}$, $N_{it}^{\max}$, $\beta_{m,k}$, $\forall m, k$.

1: Initialize the positions $\mathbf{\Theta}^{\text{PSO}}$ with the binary vector representation of $\mathcal{A}^{\text{HRNP}}$ and other p_{PSO}-1 random binary vectors;
2: Evaluate the total energy efficiency of each candidate in $\mathbf{\Theta}^{\text{PSO}}$, forming the vector $\boldsymbol{\eta}_e^{\text{PSO}}$;
3: Initialize the social information $\eta_e^{\text{best}} = \eta_{e,\phi}^{\text{PSO}}$, and $\mathbf{a}^{\text{PSO}} = \boldsymbol{\theta}_\phi^{\text{PSO}}$, in which $\phi = \arg\max_n \eta_{e,n}^{\text{PSO}}$;
4: Initialize the cognitive information $\boldsymbol{\eta}_e^c = \boldsymbol{\eta}_e^{\text{PSO}}$, and $\mathbf{\Theta}_c^{\text{PSO}} = \mathbf{\Theta}^{\text{PSO}}$;
5: Initialize the velocity matrix $\mathbf{V} \in \mathbb{R}^{M \times p_{\text{PSO}}}$ with random elements uniformly distributed in $[-1, 1]$;
6: **for** $n = 2, 3, \ldots, N_{it}^{\max}$ **do**
7:     Update the velocity matrix $\mathbf{V} = \nu\mathbf{V} + \mu_c\mathbf{\Gamma}\left[\mathbf{\Theta}_c^{\text{PSO}} - \mathbf{\Theta}^{\text{PSO}}\right] + \mu_s\mathbf{\Gamma}\left[\mathbf{a}^{\text{PSO}} - \mathbf{\Theta}^{\text{PSO}}\right]$;
8:     Update the positions $\mathbf{\Theta}^{\text{PSO}} = \texttt{binround}\left(\mathbf{\Theta}^{\text{PSO}} + \mathbf{V}\right)$;
9:     **for** $\ell = 1, 2, \ldots p_{\text{PSO}}$ **do**
10:         Evaluate the total energy efficiency of $\boldsymbol{\theta}_\ell^{\text{PSO}}$, and assign it to $\eta_{e,\ell}^{\text{PSO}}$;
11:         **if** $\eta_{e,\ell}^{\text{PSO}} > \eta_{e,\ell}^c$ **then**
12:             Update $\eta_{e,\ell}^c = \eta_{e,\ell}^{\text{PSO}}$, and $\boldsymbol{\theta}_{c,\ell}^{\text{PSO}} = \boldsymbol{\theta}_\ell^{\text{PSO}}$;
13:             **if** $\eta_{e,\ell}^{\text{PSO}} > \eta_e^{\text{best}}$ **then**
14:                 Update $\eta_e^{\text{best}} = \eta_{e,\ell}^{\text{PSO}}$, and $\mathbf{a}^{\text{PSO}} = \boldsymbol{\theta}_\ell^{\text{PSO}}$;
15:         **end if**
16:     **end if**
17:     **end for**
18: **end for**

Output: $\mathcal{A}^{\text{PSO}}$ as the set representation of $\mathbf{a}^{\text{PSO}}$.

IV. OPTIMAL NUMBER OF SELECTED ANTENNAS: AN ITERATIVE-ANALYTICAL METHOD

In this Section we derive approximated performance analytical expressions for XL-MIMO systems employing the ZF precoder and the HRNP-based AS method. Such expressions are compared with numerical results obtained via Monte-Carlo simulation method in Section V, confirming the tightness of the derivations proposed herein. Then, based on these analytical expressions, we devise an analytical iterative algorithm based on Newton-Raphson (NR) method to determine the optimal number of activated antennas for XL-MIMO systems, which maximizes the approximated EE expression.

In order to compute the average ZF SINR expression, one can directly evaluate from eq. (30):

$$\mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right] = \mathbb{E}\left[\frac{P_{\max}}{K\sigma^2}\left(\sum_{m \in \mathcal{A}}\beta_{m,k} - \sum_{j=1,j\neq k}^{K}\frac{\sum_{m \in \mathcal{A}}\beta_{m,k}\beta_{m,j}}{\sum_{m \in \mathcal{A}}\beta_{m,j}}\right)\right]$$

$$= \mathbb{E}\left[\frac{P_{\max}}{K\sigma^2}\left(\sum_{m \in \mathcal{A}}\beta_{m,k} - \sum_{m \in \mathcal{A}}\beta_{m,k}\sum_{j=1,j\neq k}^{K}\frac{\beta_{m,j}}{\sum_{n \in \mathcal{A}}\beta_{n,j}}\right)\right]$$

$$= \mathbb{E}\left[\frac{P_{\max}}{K\sigma^2}\sum_{m \in \mathcal{A}}\beta_{m,k}\left(1 - \sum_{j=1,j\neq k}^{K}\frac{\beta_{m,j}}{\sum_{n \in \mathcal{A}}\beta_{n,j}}\right)\right]$$

$$= \frac{P_{\max}}{K\sigma^2}\sum_{m \in \mathcal{A}}\mathbb{E}\left[\beta_{m,k}\right]\left(1 - \sum_{j=1,j\neq k}^{K}\mathbb{E}\left[\frac{\beta_{m,j}}{\sum_{n \in \mathcal{A}}\beta_{n,j}}\right]\right) \quad (41)$$

in which the expectation is taken with respect to the random users' positions.

Instead of advancing with (41) seeking an exact solution, we approximate the average SINR by the SINR of a *user in the most expected position* (UMEP). Given the uniform distribution of the users as illustrated in Fig. 1, this most expected position would be as depicted in Fig. 2.
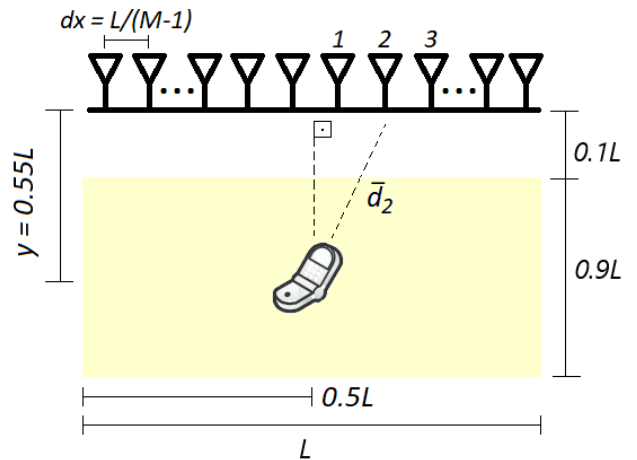

Fig. 2. Illustration of the most expected user's position (UMEP).

Then, considering this position for the users, and noting that the HRNP AS activate in this case the M_s closest antennas,

the ZF SINR expression becomes

$$\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right] \approx \frac{P_{\max}}{K\sigma^2}\left(\sum_{m\in\mathcal{A}}\overline{\beta}_m - (K-1)\frac{\sum_{m\in\mathcal{A}}\overline{\beta}_m^2}{\sum_{m\in\mathcal{A}}\overline{\beta}_m}\right),$$

$$\approx \frac{P_{\max}}{K\sigma^2}\left(2\sum_{m=1}^{M_s/2}\overline{\beta}_m - (K-1)\frac{2\sum_{m=1}^{M_s/2}\overline{\beta}_m^2}{2\sum_{m=1}^{M_s/2}\overline{\beta}_m}\right), \quad (42)$$

with $\overline{\beta}_m = q\cdot(\overline{d}_m)^{-\kappa}$, and $\overline{d}_m = \sqrt{y^2+[(m-\frac{1}{2})dx]^2} = y\sqrt{1+[(m-\frac{1}{2})\frac{dx}{y}]^2} \approx y\sqrt{1+(m\frac{dx}{y})^2}$ is represented in Fig. 2 for $m=2$. Eq. (42) can thus be simplified as in the next page, in which from (48) to (49) we have used the binomial approximation $(1+x)^\alpha \approx 1+\alpha x$ for $|\alpha x| \ll 1$. In our scenario, this condition becomes

$$\frac{\kappa M_s^2 dx^2}{8 y^2} \ll 1, \quad (43)$$

which usually holds for typical XL-MIMO systems. For example, the binomial approximation results in relative errors lower than 5% for $|\alpha x| < 0.25$, which in our XL-MIMO scenario corresponds to $M_s < 225$. Besides, with this approximated ZF HRNP-AS SINR expression, we can also approximate the EE expression as in eq. (44). Moreover, by expanding all the power terms in the denominator of (44), as discussed in Section II-C, and grouping them according to their dependence with M_s, we arrive at eq. (45), in which $\mathcal{P}_{BC} = \mathcal{P}_{COD} + \mathcal{P}_{DEC} + \mathcal{P}_{BT}$, and $\mathcal{T}_0$, $\mathcal{T}_1$ are defined in eq. (46) and (47), respectively.

$$\eta_e \approx \frac{BK\log_2\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]\right)}{P_{TX}^{DL}+P_{TX}^{tr}+P_{CE}+P_{C/D}+P_{BH}+P_{PR}+P_{TC}+P_{FIX}}, \quad (44)$$

$$\approx \frac{BK\log_2\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]\right)}{\mathcal{P}_{BC}BK\log_2\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]\right)+\mathcal{T}_0+\mathcal{T}_1 M_s}. \quad (45)$$

$$\mathcal{T}_0 = \mathcal{P}^\dagger + \frac{3MK+M\log(M)}{T_{LT}\mathcal{L}_{BS}} + \frac{BK^3}{3\mathcal{S}\mathcal{L}_{BS}}. \quad (46)$$

$$\mathcal{T}_1 = \mathcal{P}_{BS} + \frac{5BK^2}{\mathcal{S}\mathcal{L}_{BS}} + \left(1-\frac{\tau}{\mathcal{S}}\right)\frac{2BK}{\mathcal{L}_{BS}} + \frac{BK}{\mathcal{S}\mathcal{L}_{BS}}. \quad (47)$$

A. Optimal Number of Activated Antennas

Considering our previous analytical results, we propose in this Section a method for obtaining the optimal M_s value when employing ZF with HRNP AS, by taking the derivative of eq. (45), with the SINR given in eq. (50), with respect to M_s, and equaling it to 0 when $M_s = M_s^*$. Following this procedure, and after some simplifications, we arrive at $f(M_s^*) = 0$, with $f(M_s)$ defined as

$$f(M_s) = \frac{\partial\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]}{\partial M_s} - \frac{\mathcal{T}_1\ln(2)\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]\right)\log_2\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]\right)}{\mathcal{T}_0+\mathcal{T}_1 M_s}, \quad (54)$$

where $\frac{\partial\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]}{\partial M_s}$ is given in (52).

Since $\mathbb{E}\left[\overline{\gamma}_k^{(ZF)}\right]$ and its derivative are dependent of M_s, we cannot arrive at a closed-form expression for M_s^*. However, we can find the root of $f(M_s)$ by applying some iterative numerical method, like Newton-Raphson (NR) method, which obtains a sequence of M_s values $M_{s,0}, M_{s,1}, M_{s,2}, \ldots M_{s,n}$

converging to M_s^* if the starting point $M_{s,0}$ is not too far from it. The values in the sequence obey

$$M_{s,n} = M_{s,n-1} - \frac{f(M_{s,n-1})}{\frac{\partial f(M_s)}{\partial M_s}\Big|_{M_{s,n-1}}}, \quad (55)$$

in which the derivative of $f(M_s)$ is given in (51).

V. NUMERICAL RESULTS AND DISCUSSION

Our adopted simulation parameters are indicated in Table I. While we have chosen very similar power consumption parameters than that of [16], [17], the XL-MIMO system parameters are chosen similarly as [3]–[5], as well as in accordance with common XL-MIMO scenario applications. Considering $M = 512$ antennas at the XL-MIMO BS, Fig. 3 depicts the SINR, sum SE and the energy efficiency as a function of number of users K (from 1 to $M/2$), for both CB and ZF precoders. The sum SE is presented in units of bits per channel use (bpcu). One can note that ZF precoding always achieve a higher total energy efficiency than CB in the scenario investigated. The presented results were averaged among 1000 random realizations of the users' positions. It is also shown in the Figure the equivalence between the results of performance expressions from [5], eq. (12) and (13), and the expressions with our proposed simplifications, eq. (21) and (28).

TABLE I
SIMULATION PARAMETERS.

Parameter	Value
Carrier frequency: f	2.6 GHz
Number of BS antennas M	$[500; 512]$
XL-MIMO array length: L	30 m
Distance of users to BS:	$[0.1\cdot L, L]$
Path loss decay exponent: κ	3
Path loss at the reference distance: q	$10^{-3.53}$
Transmission bandwidth: B	20 MHz
Channel coherence bandwidth: B_C	100 kHz
Channel coherence time: T_C	2 ms
Long-term fading coherence time: T_{LT}	2 s
Total noise power: σ^2	-96 dBm
UL pilot transmit power: ρ_p	20 mW
DL radiated power: $P_{\max} = \frac{\rho\sigma^2}{qL^{-\kappa}}$	0.23 mW
Coherence block: $\mathcal{S}$	200 symbols
Length of the uplink pilot signals: τ	K
Computational efficiency at BSs: $\mathcal{L}_{BS}$	12.8 $\left[\frac{\text{Gflops}}{\text{W}}\right]$
Fraction of DL transmission: ξ^d	1
Fraction of UL transmission: ξ^u	0
PA efficiency at the BS: η^d	0.39
PA efficiency at the MTs: η^{uT}	0.50
Fixed power consumption: P_{FIX}	18 W
Power for local oscillators at BSs: P_{SYN}	2 W
Power for circuit components BSs: P_{BS}	1 W
Power for circuit components MTs: P_{MT}	0.10 W
Power density for coding data: $\mathcal{P}_{COD}$	0.10 $\left[\frac{\text{W}}{\text{Gbit/s}}\right]$
Power density for decoding data: $\mathcal{P}_{DEC}$	0.80 $\left[\frac{\text{W}}{\text{Gbit/s}}\right]$
Power density for backhaul traffic: $\mathcal{P}_{BT}$	0.25 $\left[\frac{\text{W}}{\text{Gbit/s}}\right]$

$$\mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right] \approx \frac{P_{\max}q}{K\sigma^2}\left(2\sum_{m=1}^{M_s/2}(\overline{d}_m)^{-\kappa} - (K-1)\frac{2\sum_{m=1}^{M_s/2}(\overline{d}_m)^{-2\kappa}}{2\sum_{m=1}^{M_s/2}(\overline{d}_m)^{-\kappa}}\right),$$

$$\approx \frac{P_{\max}q}{K\sigma^2}\left\{2\sum_{m=1}^{M_s/2}y^{-\kappa}\left[1+\left(\frac{m\,dx}{y}\right)^2\right]^{-\frac{\kappa}{2}} - (K-1)\frac{2\sum_{m=1}^{M_s/2}y^{-2\kappa}\left[1+\left(\frac{m\,dx}{y}\right)^2\right]^{-\kappa}}{2\sum_{m=1}^{M_s/2}y^{-\kappa}\left[1+\left(\frac{m\,dx}{y}\right)^2\right]^{-\frac{\kappa}{2}}}\right\},$$

$$\approx \frac{P_{\max}q\,y^{-\kappa}}{K\sigma^2}\left\{2\sum_{m=1}^{M_s/2}\left[1+\left(\frac{m\,dx}{y}\right)^2\right]^{-\frac{\kappa}{2}} - (K-1)\frac{2\sum_{m=1}^{M_s/2}\left[1+\left(\frac{m\,dx}{y}\right)^2\right]^{-\kappa}}{2\sum_{m=1}^{M_s/2}\left[1+\left(\frac{m\,dx}{y}\right)^2\right]^{-\frac{\kappa}{2}}}\right\} \quad (ZF_{ME}), \tag{48}$$

$$\approx \frac{P_{\max}q\,y^{-\kappa}}{K\sigma^2}\left\{2\sum_{m=1}^{M_s/2}\left[1-\frac{\kappa}{2}\left(\frac{m\,dx}{y}\right)^2\right] - (K-1)\frac{2\sum_{m=1}^{M_s/2}\left[1-\kappa\left(\frac{m\,dx}{y}\right)^2\right]}{2\sum_{m=1}^{M_s/2}\left[1-\frac{\kappa}{2}\left(\frac{m\,dx}{y}\right)^2\right]}\right\}, \tag{49}$$

$$\approx \frac{P_{\max}q\,y^{-\kappa}}{K\sigma^2}\left[\left(\mathcal{T}_{1,1}M_s - \mathcal{T}_{1,2}M_s^2 - \mathcal{T}_{1,3}M_s^3\right) - (K-1)\frac{\left(\mathcal{T}_{2,1}M_s - \mathcal{T}_{2,2}M_s^2 - \mathcal{T}_{2,3}M_s^3\right)}{\left(\mathcal{T}_{1,1}M_s - \mathcal{T}_{1,2}M_s^2 - \mathcal{T}_{1,3}M_s^3\right)}\right] \quad (ZF_{BA}), \tag{50}$$

$$\text{with} \quad \mathcal{T}_{1,1} = 1 - \frac{K\,dx^2}{12\,y^2}, \quad \mathcal{T}_{1,2} = \frac{K\,dx^2}{8\,y^2}, \quad \mathcal{T}_{1,3} = \frac{K\,dx^2}{24\,y^2}, \quad \mathcal{T}_{2,1} = 1 - \frac{K\,dx^2}{6\,y^2}, \quad \mathcal{T}_{2,2} = \frac{K\,dx^2}{4\,y^2}, \quad \mathcal{T}_{2,3} = \frac{K\,dx^2}{12\,y^2}.$$

$$\frac{\partial f(M_s)}{\partial M_s} = \frac{\partial^2 \mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]}{\partial M_s^2} - \frac{\mathcal{T}_1^2 \ln(2)\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]\right)\log_2\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]\right) - \mathcal{T}_1\left(\mathcal{T}_0+\mathcal{T}_1 M_s\right)\frac{\partial \mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]}{\partial M_s}\left(1+\ln(2)\log_2\left(1+\mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]\right)\right)}{\left(\mathcal{T}_0+\mathcal{T}_1 M_s\right)^2}, \tag{51}$$

$$\text{with} \quad \frac{\partial \mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]}{\partial M_s} = \frac{P_{\max}q\,y^{-\kappa}}{K\sigma^2}\left[\mathcal{F}_1' - (K-1)\frac{\mathcal{F}_1\mathcal{F}_2' - \mathcal{F}_2\mathcal{F}_1'}{\mathcal{F}_1^2}\right], \tag{52}$$

$$\text{and} \quad \frac{\partial^2 \mathbb{E}\left[\overline{\gamma}_k^{(\text{ZF})}\right]}{\partial M_s^2} = \frac{P_{\max}q\,y^{-\kappa}}{K\sigma^2}\left[\mathcal{F}_1'' - (K-1)\frac{\mathcal{F}_1^2\left(\mathcal{F}_1\mathcal{F}_2''-\mathcal{F}_2\mathcal{F}_1''\right) - 2\left(\mathcal{F}_1\mathcal{F}_2'-\mathcal{F}_2\mathcal{F}_1'\right)\mathcal{F}_1\mathcal{F}_1'}{\mathcal{F}_1^4}\right], \tag{53}$$

in which $\mathcal{F}_1 = \mathcal{T}_{1,1}M_s - \mathcal{T}_{1,2}M_s^2 - \mathcal{T}_{1,3}M_s^3$, $\quad \mathcal{F}_2 = \mathcal{T}_{2,1}M_s - \mathcal{T}_{2,2}M_s^2 - \mathcal{T}_{2,3}M_s^3$,

$\mathcal{F}_1' = \mathcal{T}_{1,1} - 2\mathcal{T}_{1,2}M_s - 3\mathcal{T}_{1,3}M_s^2$, $\quad \mathcal{F}_2' = \mathcal{T}_{2,1} - 2\mathcal{T}_{2,2}M_s - 3\mathcal{T}_{2,3}M_s^2$, $\quad \mathcal{F}_1'' = -2\mathcal{T}_{1,2} - 6\mathcal{T}_{1,3}M_s$, $\quad \mathcal{F}_2'' = -2\mathcal{T}_{2,2} - 6\mathcal{T}_{2,3}M_s$.
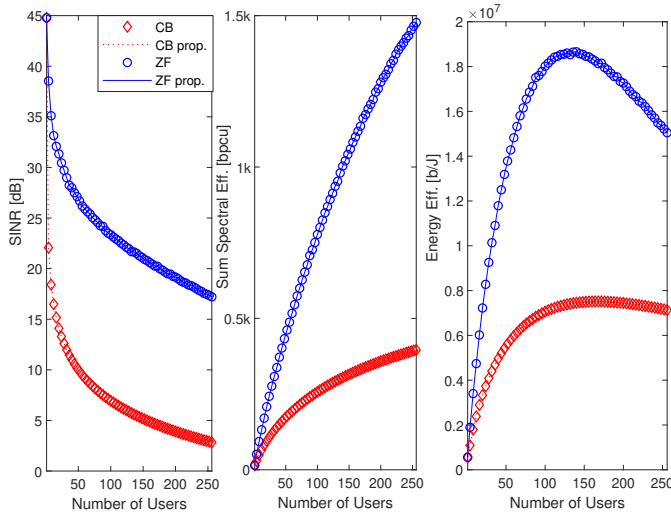

Fig. 3. (a) SINR, (b) sum-SE, and (c) EE *vs.* K for $M = 512$ antennas, selecting all available antennas. Proposed eq. (21) and (28), are represented by dotted and solid line curves, respectively, while the performances from [5], eq. (12) and (13), are indicated by the curves with '$\Diamond$' and 'o' markers.

Now, considering $M = 500$ antennas at the XL-MIMO BS, and the same power consumption parameters, Fig. 4 shows the SINR, sum SE and the EE as a function of $M_s \in \{100; M\}$, with $K = 100$ users, for both CB and ZF precoders when employing the HRNP AS scheme. Notice that ZF precoding achieves a higher total energy efficiency than CB in the scenario investigated. Besides, by activating a number of $M_s = 146$ BS antennas, one can attain the maximum total energy efficiency for ZF precoder with $K = 100$ users ("M_s^* by NR" point in Fig. 4.c), as found by our proposed NR method of Section IV-A. Fig. 4 also compares the performance obtained by averaging eq. (30) with several random realizations for the users' positions (denoted as ZF), with the approximated deterministic result from eq. (48), denoted as ZF_{ME}, and with the binomial approximation in eq. (50), denoted as ZF_{BA}. It also shows the results in terms of sum SE and EE of the XL-MIMO system. One can conclude that both proposed approximations are tight, and that the M_s values that maximize them are nearly the same.

Next, in order to obtain the performance results of GA, LS, and PSO-based AS schemes, we have set the maximum number of iterations $N_{it}^{\max} = 60$ for such schemes, and analysed their convergence for $K = 100$ users, as depicted in Fig. 5.a. One can see from the Figure that the LS-AS convergence presents a non-decreasing behavior, since when a new solution is not found in certain iteration, the algorithm interrupts its search, and does not spend more processing power. On the other hand, for GA and PSO-based AS for XL-MIMO systems, if the algorithms do not find new solutions
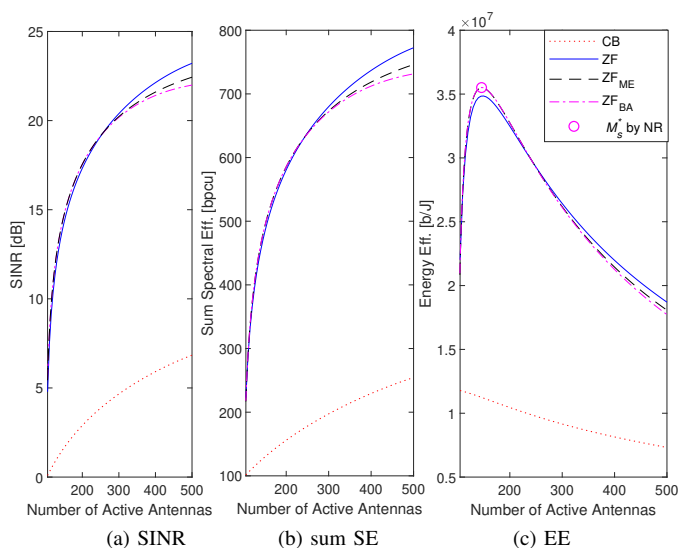
Fig. 4. HRNP-AS scheme under ZF and CB precoders: (a) SINR, (b) sum SE, and (c) EE as a function of M_s for $M = 500$ antennas and $K = 100$.

and keep searching during additional iterations, the EE of that solution decreases due to the progressive processing power consumed in the subsequent iterations. Therefore, it is not efficient to predefine the number of iterations for these two schemes in the XL-MIMO antenna selection problem, since in this optimization problem it would be very difficult do adjust the number of iterations in such a way to obtain a suitable EE solution for the algorithms. To circumvent while taking advantage of this feature, we implement an *early-interruption* criterion, in which if the GA or the PSO-based AS schemes do not find a new solution within 5 iterations, the search is interrupted, obtaining the convergences depicted in Figure 5.b. Besides, for the GA-based AS scheme, we have considered a population size of $M/2$, of which 10% are selected as parents at each iteration, and a mutation probability of 2%. For the PSO-based one, we have considered a swarm of $M/5$ particles, and an inertia weight, cognitive factor and social factor of 0.5.
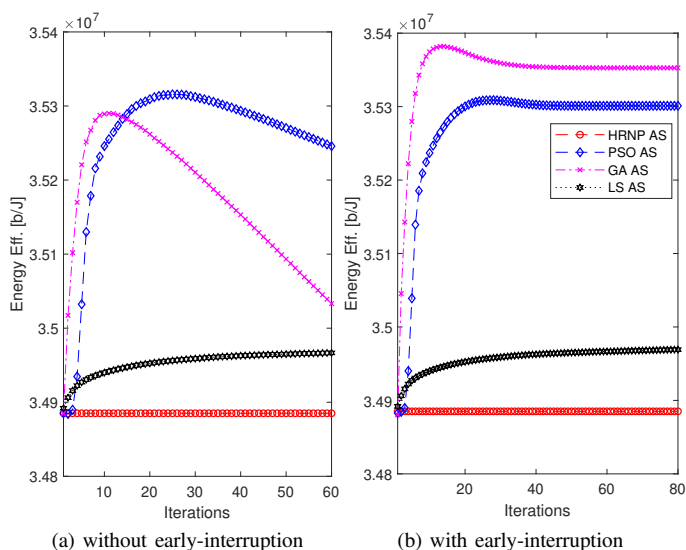

Fig. 5. Convergence of the AS schemes: (a) without, and (b) with early-interruption stopping search criterion. $K = 100$ and $M = 500$ antennas.

Fig. 6 depicts the SINR, sum SE, and EE as a function of K for the HRNP, GA, LS, and PSO-based AS schemes employing ZF precoding, with $M = 500$ antennas at the XL-MIMO array. While the achieved sum SE performance is nearly the same for all investigated schemes, the graphs reveal that SINR and EE gains can be achieved in comparison with HRNP. The Figure also shows that, in terms of SINR and EE, the GA, LS, and PSO-based AS schemes achieve a similar performance, and their gains in comparison with HRNP AS are small, since the processing required for finding a suitable antennas subset in the XL-MIMO system increases the energy consumption; thus, the EE gains become marginal. Except for small number of users, the GA AS scheme achieves one of the best EEs in most part of the investigated scenario, although for high number of users, its performance becomes very similar to HRNP AS scheme. Besides, due to its simplicity and celerity to return the results, one can point out that the HRNP criterion coupled to the NR procedure for M_s^* selection represents a very promising XL-MIMO AS scheme.
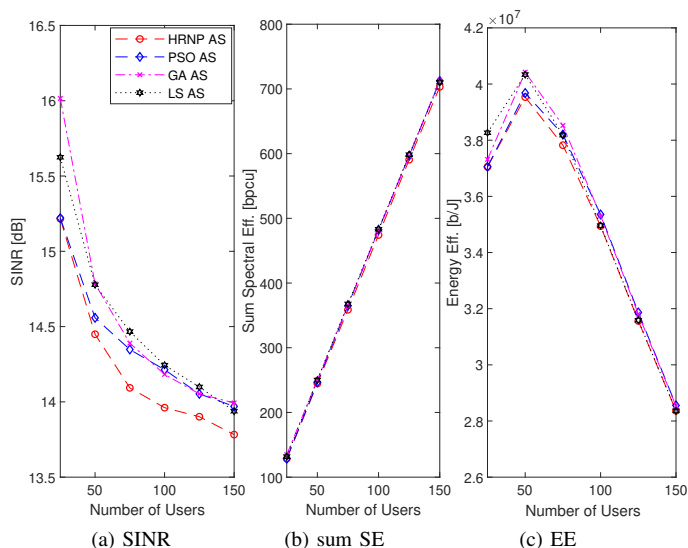

Fig. 6. AS schemes for ZF precoding: (a) SINR, (b) sum SE, and (c) EE as a function of K for $M = 500$ antennas.

A. Complexity of XL-MIMO AS Methods

Fig. 7.a depicts the average number of active antennas as a function of K for the investigated AS methods. One can see that the M_s^* value obtained by our proposed NR method usually matches the number of antennas selected by LS, PSO, and GA-based AS schemes, corroborating the tightness of the approximations made and the effectiveness of the method. The major advantage of our proposed NR method for obtaining M_s^* is that it can be evaluated for any system configuration satisfying eq. (43). In our numerical simulations, the method has converged in at most 3 iterations from the starting point $M_{s,0} = 1.5K$. Besides, the M_s^* value is not dependent on the channel coefficients, but only on the system parameters, like number of users, transmit power, dimensions of XL-MIMO array and coverage area. Therefore, once found M_s^*, the NR method just has to be evaluated again when one of these parameters change. The fixed complexity of evaluating M_s^*

under 3 NR iterations is about 380 *flops*, which is negligible in comparison with that of selecting the antennas subset, eq. (37), (38), (39), and (40), besides of remaining valid for larger time periods.
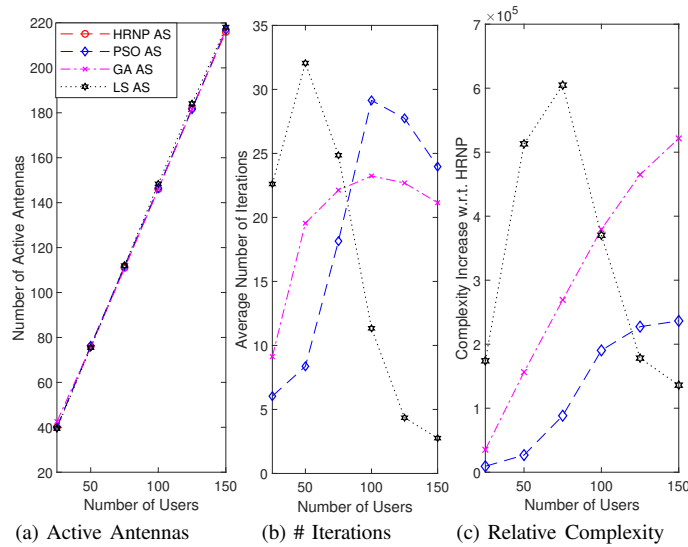

Fig. 7. (a) Average number of active antennas, (b) average number of iterations, and (c) complexity increase of the AS schemes w.r.t. HRNP AS approach as a function of K for $M = 500$ antennas.

Fig. 7.b depicts the average number of iterations required by each investigated AS scheme, recalling that the number of iterations are not fixed, since the LS interrupts when a new solution is not find in an iteration, and GA and PSO implement the early-interruption criterion. Besides, due to the non-decreasing behavior of the LS convergence depicted in Fig. 5, the average number of iterations for this scheme in Fig. 7.b does not correspond to the point in which the LS convergence curve becomes horizontal. Besides, the advantage of HRNP criterion in selecting antennas within the XL-MIMO array can also be confirmed by the extra computational complexity required for the other analysed methods. Hence, considering the average number of iterations from Fig. 7.b, the *relative complexity increment* of the LS, GA and PSO AS schemes w.r.t. the HRNP AS method are depicted in Fig. 7.c. The relative complexity increment metric is defined as:

$$\Delta_{\mathcal{C}} = \frac{\mathcal{C}_{\text{as}}^{\text{LS, GA,PSO}} - \mathcal{C}_{\text{as}}^{\text{HRNP}}}{\mathcal{C}_{\text{as}}^{\text{HRNP}}}$$

considering typical XL-MIMO network configurations for K users and M BS antennas. One can confirm the very large relative complexity increase of the AS methods for XL-MIMO, *i.e.*, this complexity increment is in the order of 10^5, which make the benefits they would bring less significant in terms of energy efficiency.

It is noteworthy that the computational complexity spent with the AS methods is included in the EE values, in terms of the processing power. In summary, the performance improvement of the AS scheme comes at the expense of high complexity, which results in marginal EE gains. On the other hand, the HRNP-AS procedure is able to achieve an improved EE of 34.85 Mbit/J for $K = 100$ users, in comparison with 18.71 Mbit/J of selecting all antennas, *i.e.*, not applying any

AS procedure, corresponding in a 86.3% of EE increasing, as one can infer from Fig. 4.

Elaborating further regarding the dependence of the optimal number of selected antennas M_s^* on the system parameters, such as number of users, total transmit power available, dimensions of XL-MIMO array, and coverage area, one can argue that such system parameters vary quite slowly with respect to the data symbol period. Therefore, it could be possible to evaluate the proposed AS scheme, and turning-on the optimal number of RF chains M_s^*, which are then switched to the best antenna subset according to our proposed HRNP criterion. Notice that only when the number of users changes significantly that it would be necessary to re-evaluate the (54)-(55), and then turning-on or turning-off some RF chains. Besides, the proposed method for finding the optimal number of selected antennas can provide very useful information for XL-MIMO system designers.

VI. CONCLUSION

In this paper, we have investigated the XL-MIMO systems subject to channel non-stationarities. First, we have revisited the performance expressions from [5], and proposed to incorporate the power constraint at the SINR expressions of CB and ZF to arrive at more lean and comprehensive results. Then, based on such obtained expressions, we have proposed four XL-MIMO AS schemes aiming at maximizing the EE based on the following criteria: HRNP, LS, GA, and PSO. Some simplifying assumptions allowed us to derive closed-form EE expressions, based on which we proposed a NR iterative method to obtain the optimal number of active antennas. Numerical results have shown that GA usually achieves one of the best EEs, although the gains were marginal in comparison with HRNP, since the processing required for achieving a suitable antennas subset increases the consumed energy, limiting the achieved EE gains. Thus, due to its simplicity and celerity in returning results, the proposed HRNP-AS scheme, with the NR method providing the optimal subarray size value M_s^*, can be seen as a very promising solution for AS XL-MIMO systems, achieving an EE gain of 86.3% in comparison with selecting all antennas strategy.

REFERENCES

[1] T. Marzetta, E. Larsson, H. Yang, and H. Ngo, *Fundamentals of Massive MIMO*. New York, NY, USA: Cambridge University Press, 2016.

[2] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[3] A. Amiri, C. N. Manchón, and E. de Carvalho, "Deep learning based spatial user mapping on extra large MIMO arrays," *arXiv. 2002.00474*, 2020.

[4] X. Yang, F. Cao, M. Matthaiou, and S. Jin, "On the Uplink Transmission of Multi-user Extra-large Scale Massive MIMO Systems," *arXiv. 1909.06760*, 2019.

[5] A. Ali, E. D. Carvalho, and R. W. Heath, "Linear Receivers in Non-Stationary Massive MIMO Channels With Visibility Regions," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 885–888, June 2019.

[6] S. Chen, S. Sun, G. Xu, X. Su and Y. Cai, "Beam-Space Multiplexing: Practice, Theory, and Trends, From 4G TD-LTE, 5G, to 6G and Beyond," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 162–172, April 2020.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2020.3022708, IEEE Transactions on Vehicular Technology

13

[7] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality – What is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, pp. 3–20, 2019.

[8] E. de Carvalho, A. Ali, A. Amiri, M. Angjelichinoski, and R. W. Heath, "Non-Stationarities in Extra-Large Scale Massive MIMO," *arXiv. 1903.03085v2*, Oct 2019.

[9] A. Amiri, S. Rezaie, C. N. Manchón, and E. de Carvalho, "Distributed Receivers for Extra-Large Scale MIMO Arrays: A Message Passing Approach," *arXiv. 2007.06930*, July 2020.

[10] V. C. Rodrigues, A. Amiri, T. Abrão, E. de Carvalho, and P. Popovski, "Low-Complexity Distributed XL-MIMO for Multiuser Detection," 2020 IEEE International Conference on Communications Workshops (ICC Workshops), Dublin, Ireland, 2020.

[11] H. Wang, A. Kosasih, C. Wen, S. Jin and W. Hardjawana, "Expectation Propagation Detector for Extra-Large Scale Massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2036–2051, March 2020.

[12] Y. Han, S. Jin, C.-K. Wen, and X. Ma, "Channel Estimation for Extremely Large-Scale Massive MIMO Systems," *IEEE Wireless Communications Letters*, pp. 1–5, 2020.

[13] J. Flordelis, X. Li, O. Edfors and F. Tufvesson, "Massive MIMO Extensions to the COST 2100 Channel Model: Modeling and Validation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 380–394, Jan. 2020.

[14] Y. Fu, C. Wang, X. Fang, L. Yan and S. Mclaughlin, "BER Performance of Spatial Modulation Systems Under a Non-Stationary Massive MIMO Channel Model," *IEEE Access*, vol. 8, pp. 44547–44558, Feb. 2020.

[15] O. S. Nishimura, J. C. Marinello, and T. Abrão, "A Grant-based Random Access Protocol in Extra-Large Massive MIMO System," *IEEE Communications Letters*, Early Access, July 2020.

[16] J. C. Marinello, C. Panazio, T. Abrão, and S. Tomasin, "Total Energy Efficiency of TR-MRC and FD-MRC Receivers for Massive MIMO Uplink," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2285–2296, Sep. 2019.

[17] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal Design of Energy-Efficient Multi-User MIMO Systems: Is Massive MIMO the Answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, June 2015.

[18] B. Makki, A. Ide, T. Svensson, T. Eriksson, and M. Alouini, "A Genetic Algorithm-Based Antenna Selection Approach for Large-but-Finite MIMO Networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6591–6595, July 2017.

[19] Z. Liu, W. Du, and D. Sun, "Energy and Spectral Efficiency Tradeoff for Massive MIMO Systems With Transmit Antenna Selection," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4453–4457, May 2017.

[20] 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, February 2013.

[21] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[22] G. H. Golub and C. F. V. Loan, "Matrix Computations," Maryland, USA: Johns Hopkins University Press, 1996.

Taufik Abrão (SM'12, SM-SBrT, Pq-1D-CNPq) received the B.S., M.Sc., and Ph.D. degrees in electrical engineering from the Polytechnic School of the University of São Paulo, São Paulo, Brazil, in 1992, 1996, and 2001, respectively. Since March 1997, he has been with the Communications Group, Department of Electrical Engineering, Londrina State University, Paraná, Brazil, where he is currently an Associate Professor in Telecommunications and the Head of the Telecomm. & Signal Processing Lab. From July-October 2018 he was with the Connectivity section, Aalborg University as a Guest Researcher. In 2012, he was an Academic Visitor with the Southampton Wireless Research Group, University of Southampton, U.K. He has also served as an Associate Editor for the IEEE Systems Journal, ETT-Wiley journal AEUe-Elsevier and IET Signal Processing. Previously, he served as AE of the IEEE Communication Surveys & Tutorials (2013-2017) and IEEE Access 2016-2019. His research interests include massive and XL-MIMO, URLLC, mMTC, detection and estimation, multicarrier systems, cooperative communication, resource allocation, convex optimization for 5G and 6G wireless systems. He has supervised +30 M.Sc. and +10 Ph.D. students, co-authored +10 book chapters on mobile radio communications and +280 research papers published in international journals and conferences.

Abolfazl Amiri currently is PhD Fellow at the Department of Electronic Systems, The Technical Faculty of IT and Design, Connectivity section, at Aalborg University, Denmark. His main expertise is in signal processing for MIMO communications, with recent focus on massive MIMO, and XL-MIMO, including channel measurements, channel modeling, beamforming, and protocol aspects. Email: aba@es.aau.dk

Elisabeth de Carvalho received the Ph.D. degree in electrical engineering from Telecom ParisTech, France. She was a Post-Doctoral Fellow with Stanford University, Standford, CA, USA, and then with industry in the field of DSL and wireless LAN. Since 2005, she has been an Associate Professor with Aalborg University, where she has led several research projects in wireless communications. She co-authored the text book A Practical Guide to the MIMO Radio Channel. Her main expertise is in signal processing for MIMO communications, with recent focus on massive MIMO, including channel measurements, channel modeling, beamforming, and protocol aspects.

José Carlos Marinello received his B.S. and M.S. degree in Electrical Engineering (the first with Summa Cum Laude) from Londrina State University, PR, Brazil, in December 2012 and Sept 2014, respectively. He hold his Ph.D. in 2018 at Polytechnic School of São Paulo University, São Paulo, Brazil. His research interests include signal processing and wireless communications, especially massive multiple-antenna precoding/decoding techniques, acquisition of channel-state information, multicarrier modulation, cross-layer optimization of MIMO/OFDM systems, interference management and 5G

Petar Popovski . (S'97–M'04–SM'10–F'16) received the Dipl.-Ing. degree in electrical engineering and the Magister Ing. degree in communication engineering from Sts. Cyril and Methodius University, Skopje, Macedonia, in 1997 and 2000, respectively, and the Ph.D. degree from Aalborg University, Denmark, in 2004. He is currently a Professor in wireless communications with Aalborg University. He has over 270 publications in journals, conference proceedings, and books. He holds over 30 patents and patent applications. His research interests are in the areas of wireless communication and networking, and communication/information theory. He received the Consolidator Grant from the European Research Council. He was a recipient of the Elite Researcher Award in Denmark in 2016. From 2012 to 2014, he served as the Chair of the IEEE ComSoc Emerging Technology Committee on Smart Grid Communications. He is a Steering Committee Member of the IEEE Internet of Things Journal and a Steering Committee Member of the IEEE SmartGridComm. He is currently an Editor of the IEEE Transactions on Communications and an Area Editor of the IEEE Transactions Wireless Communications.