Reduced Switching Connectivity for Power-Efficient Large Scale Antenna Selection

Adrian Garcia-Rodriguez, Student Member, IEEE, Christos Masouros, Senior Member, IEEE, and Pawel Rulikowski, Member, IEEE

Abstract—In this paper, we explore reduced-connectivity radio frequency (RF) switching networks for reducing the analog hardware complexity and power losses in antenna selection (AS) systems. In particular, we analyze different hardware architectures for implementing the RF switching matrices required in AS designs with a reduced number of RF chains. We explicitly show that fully-flexible (FF) switching matrices, which facilitate the selection of any plausible subset of antennas and attain the maximum theoretical performance of AS, present numerous drawbacks such as the introduction of significant insertion losses (ILs), particularly pronounced in large scale antenna systems (LSAS). Since these disadvantages make FF switching suboptimal in the energy efficiency sense, we further consider partially-connected (PC) switching networks as an alternative switching architecture with reduced hardware complexity, which we characterize in this work. In this context, we also analyze the impact of reduced switching connectivity on the analog hardware and digital signal processing of AS schemes that rely on channel power information. Overall, the analytical and simulation results shown in this paper demonstrate that PC switching maximizes the energy efficiency of massive MIMO systems for a reduced number of RF chains, while FF switching offers sub-optimal energy efficiency benefits due to its significant power losses.

Index Terms—Antenna selection, large scale antenna systems, RF switching matrices, insertion losses, energy efficiency.

I. INTRODUCTION

T
HE constant growth in the number of mobile devices as well as the development of data-hungry applications have driven the design of novel wireless communications solutions. In this context, systems incorporating a large number of antennas at the base stations (BS) and occupying larger bandwidths are leading candidates for 5G [1], [2]. In particular, LSAS for conventional microwave frequencies and millimeter waves implement a large number of antennas for achieving unprecedented spatial resolutions and for compensating for the larger signal attenuations introduced at high frequencies [3]–[7]. Simultaneously, these solutions motivate the design of novel digital and analog signal processing strategies, since the economic cost, power consumption, and hardware and digital signal processing complexities of systems with a dedicated RF chain per antenna might become burdensome [7]–[9].

In order to overcome the above challenges, a number of solutions have been recently the focus of intense study. In particular, both discrete lens arrays and hybrid analog-digital precoding and combining schemes have shown to be a promising solution for reducing the number of RF chains at millimeter wave frequencies [7], [10], [11]. Spatial modulation, which relies on activating a subset of antennas and employing the antenna indices as an additional source of information, also aims at reducing the RF complexity [12]–[14]. AS has been posed as a feasible alternative for reducing the complexity in both small and large scale MIMO systems [15]–[17], and comprises the focus of this work.

The characterization and development of AS systems have been the focus of intensive research. For instance, a large number of sub-optimal antenna selection algorithms for reducing the complexity of the optimal selection have been proposed in the related literature (see, e.g. [16], [18]–[22] and references therein). Numerous performance analyses in terms of attainable rates and diversity of these algorithms for transmit and receive AS are also available [15]–[17], [23]–[27]. Moreover, there are several works studying different aspects of AS systems such as the acquisition of channel state information (CSI) [28]–[30], their energy efficiency improvements [31]–[33], or their practical implementation aspects [34], [35]. The implementation of AS as a means of reducing the excessive number of RF chains in LSAS has also attracted considerable attention [36]–[46]. Particularly relevant in this context are the results presented in [38], [46], [47], which illustrate that a large portion of the full-RF massive MIMO rates can be attained via AS in realistic propagation scenarios, hence motivating the practical application of AS in LSAS. However, none of the above-mentioned works presents a comprehensive analysis of a critical aspect: the design of the RF switching matrix that implements the AS.

The RF switching matrix represents the hardware components required in AS for interconnecting the RF chains with their selected antennas in implementations with a reduced number of RF chains [16], [34]. Indeed, the colossal growth of antenna numbers in LSAS also increases the complexity of RF switching, hence making these components a significant performance factor. In spite of acknowledging the crucial importance of this component [16], the impact of the switching network on the performance and energy efficiency of AS has been commonly ignored in the related literature. Only recently, the large ILs and complexity of switching matrices in LSAS have been considered in [46] and [48]. Specifically, [48] considers a switching network that only connects each RF chain to predefined subsets of antennas for reducing the complexity of switching between the large
number of antennas required at millimeter wave frequencies. Similarly, [46] proposes to implement the switching matrix via binary switches to alleviate the ILs. However, this comes at the expense of reducing the input-output connectivity of the resultant switching architecture, which is PC. Still, these works do not perform a thorough and detailed analysis of the hardware implications behind their designs and only focus on specific implementations such as binary switching matrices.

Considering the above, in this contribution we generalize the above-mentioned works to arbitrary switching architectures and concentrate on providing a detailed analysis of switching networks to characterize their influence in AS systems. In particular, we present a number of specific hardware implementations of switching matrices that are optimized under different criteria such as the number of internal connections or the power losses. In this context, we accurately characterize their ILs, which are shown to be critical for LSAS due to the large number of inputs (RF chains) and outputs (antennas) required. Moreover, we consider architectures with limited connectivity as a means for reducing the complexity of the fully-flexible (FF) switching networks conventionally considered. In this line, we also determine the performance loss introduced due to the limited connectivity for power-based (PB) AS systems, since PB-AS can offer a superior performance when practical CSI acquisition procedures are considered. Altogether, the results and generalized designs considered in this work provide a comprehensive view of the impact of RF switching matrices on the performance, hardware architecture and energy efficiency of AS systems. For clarity, we summarize the main contributions of this paper as follows:

- We propose a number of architectures for FF switching matrices with various optimization criteria such as the number of hardware components or the ILs.
- We explore and characterize the performance of PC switching matrices with an arbitrary number of input RF chains and output antenna ports as a means for reducing the switching complexity of AS in MIMO systems.
- We reveal that the PB-AS procedure should be redesigned when PC switching matrices are implemented, due to the need of imposing additional constraints and the permuted connectivity between RF chains and antenna ports.

This paper is organized as follows: Sec. II describes the system model employed in this work. Sec. III describes FF hardware architectures for RF switching matrices, whereas PC switching is presented in Sec. [V]. The performance loss due to considering switching matrices with limited input-output connectivity is analyzed in Sec. V. Subsequently, Sec. VI introduces the energy efficiency model considered in this work and Sec. VII presents numerical results. Lastly, the conclusions are drawn in Sec. VIII.

II. PRELIMINARIES

A. Downlink System Model

Let us consider a time-division duplex (TDD) multi-user MIMO system comprised of a BS with $N$ transmit antennas and $K$ single-antenna mobile stations (MSs). The BS incorporates $K \leq M \leq N$ RF chains to convey $K$ independent data symbols to the MSs via AS as illustrated in Fig. 1. The composite signal $y \in \mathbb{C}^{K \times 1}$ received by the MSs can be expressed as

$$y = \sqrt{\rho} H_{[M]} x + w,$$

where $H \in \mathbb{C}^{K \times N}$ denotes the channel response between the BS antennas and the users, $x \in \mathbb{C}^{M \times 1}$ is the transmit signal and $w \in \mathbb{C}^{K \times 1} \sim \mathcal{CN}(0, I_K)$. In the previous expressions, $\sim$ means “distributed as”, $I_K$ is the $K \times K$ identity matrix and $\mathcal{CN}(a, A)$ represents a complex normally distributed vector with means $a$ and covariance matrix $A$. Moreover, $H_{[M]} \in \mathbb{C}^{K \times M}$ is a submatrix of $H$ built by selecting the columns specified by the set $M \subseteq \{1, \ldots, N\}$ with cardinality $|M| = M$. Here, $A \subseteq B$ denotes that $A$ is a subset of $B$. We let $\mathbb{E}[x^H x] = K$, where $\mathbb{E}[]$ is the expectation operator. Based on the above, $\rho$ represents the average transmission power per mobile terminal.

B. Antenna Selection Benchmarks

The capacity $C$ of the downlink system described in (1) is given by [49]

$$C = \max_{\mathbf{P}} \left( \frac{\eta_{dl}}{\eta_{coh}} \right) \log_2 \det \left( I_K + \rho \mathbf{P} H_{[M]} (H_{[M]}^H \rho \mathbf{P} H_{[M]} = \mathbf{P} \rho (\mathbf{H}_{[M]} H_{[M]}^H) \right),

where $(\cdot)^H$ denotes the Hermitian transpose, $\eta_{dl}$ is the number of symbols dedicated to downlink transmission and $\eta_{coh}$ denotes the number of symbols where the channel remains constant, as shown in Fig. 2. Moreover, $\mathbf{P} \in \mathbb{R}^{K \times K}$ is a diagonal power allocation matrix satisfying $\sum_{i=1}^{K} P_{i,i} = K$, where $A_{i,j}$ denotes the $i,j$-th entry of $A$. A conventional criterion for AS consists in maximizing the system’s capacity $C$, which results in the optimization problem [25], [47]

$$P_1 : \text{maximize } \mathbb{E}[\log_2(\det(\mathbf{I}_K + \rho \mathbf{P} \mathbf{H} \mathbf{S})) \rho]$$

subject to $\sum_{i=1}^{N} S_{i,i} = M$,

$$S_{i,i} \in \{0, 1\}, \forall i \in \{1, \ldots, N\},$$

where $S \in \mathbb{B}^{N \times N}$ is a diagonal binary matrix satisfying

$$S_{i,i} = \begin{cases} 1, & \text{if } i \in M, \\ 0, & \text{otherwise}. \end{cases}$$

The optimization problem $P_1$ is not convex due to the binary constraints $S_{i,i} \in \{0, 1\}, \forall i \in \{1, \ldots, N\}$ [50]. However, these constraints can be relaxed to the weaker $0 \leq S_{i,i} \leq 1$,
hence turning $P_1$ into the convex optimization problem \[ P_2 : \max_{S,P} \log_2 \det (I_K + \rho P H S H^H) \] with
\[
\begin{align*}
\text{subject to} \quad & \sum_{i=1}^{N} S_{i,i} = M, \\
& 0 \leq S_{i,i} \leq 1, \forall i \in \{1, \ldots, N\}.
\end{align*}
\]
The final solution is attained by selecting the antennas corresponding to the $M$ largest diagonal entries in $S$. At this point we note that $P_2$ can also have additional constraints depending on the specific architecture of the RF switching matrix, since there might be specific antenna combinations that cannot be simultaneously selected due to having a limited input-output connectivity [46]. For ease of exposition, in this section we shall assume that $P_2$ is solved while satisfying them and leave their detailed definition for Sec. [IV].

In resemblance with [47], one can compute the system’s capacity for AS by first optimizing (5) over $S$ while setting $P = I_K$, which produces $S^*$. Subsequently, the user power allocation matrix $P^*$ is computed by solving $P_2$ with $S^*$ fixed. This is a convex optimization problem with a waterfiling-type solution [51]. We note that this procedure is particularly efficient for large $P$’s, since the optimal power allocation matrix $P^* \approx I_K$ in $P_2$ [49]. Moreover, inspired by [47], in this work we also consider the practical linear zero-forcing (ZF) precoder, whose sum rates can be obtained as the resultant value of the objective function in the optimization problem
\[
\begin{align*}
P_3 : \max_{D} & \frac{\eta_{dl}}{\eta_{coh}} \sum_{i=1}^{K} \log_2 (1 + \rho D_{i,i}) \\
\text{subject to} & \sum_{i=1}^{K} D_{i,i} \left( (H S^* H^H)^{-1} \right)_{i,i} = K,
\end{align*}
\]
where $(\cdot)^{-1}$ denotes the inverse matrix and $D \in \mathbb{R}^{K \times K}$ is the diagonal power allocation matrix for ZF. The solution to $P_3$ can also be obtained via conventional water-filling [52]. We note that while $S^*$ obtained from $P_2$ does not necessarily maximize $P_3$, this choice adopted in [46], [47] is supported by the results in [20], where it is shown that employing $S^*$ has a slight impact on the resultant system’s performance when compared with other AS methods specifically designed for linear precoders.

C. Channel State Information Acquisition for AS Systems

The optimization problems $P_2$ and $P_3$ defined in Sec. [II-B] can only be solved provided that full knowledge of the channel response $H$ is available at the BS. However, CSI acquisition poses a major challenge in AS systems with reduced RF chains since, as opposed to full-RF systems, the CSI acquisition requires multiplexed training slots to obtain full CSI for all transmit antennas.

1) Instantaneous CSI Acquisition for full-RF systems ($M = N$): The conventional training stage in TDD systems where $M = N$ consists of the transmission of orthogonal training sequences from the users to the BS throughout $K \leq \eta_{tr} \leq \eta_{coh}$ training symbols [9], [53]. Considering channel reciprocity, the signal received at the BS $Q \in \mathbb{C}^{N \times n_{tr}}$ can be expressed as [9], [53]
\[
Q = H^H \Phi + N,
\]
where $\Phi \in \mathbb{C}^{K \times n_{tr}}$ denotes the orthogonal training signals conveyed by the users, and $N \in \mathbb{C}^{N \times n_{tr}}$ is the standard noise matrix with independent and identically distributed (i.i.d.) entries satisfying $n_{i,j} \sim \mathcal{CN}(0, \sigma_{n}^2)$. The communication channel $H$ can be directly estimated from (7) via least-squares (LS) estimation by correlating the received signal with the known unitary pilot sequences [53], [55].

\[
\hat{H}^H = (H^H \Phi + N)^H H^H + \hat{N},
\]
where $\hat{N}$ is identically distributed to $N$ [54]. Minimum mean square error (MMSE) estimation is also commonly applied provided that the channel statistics are known at the BS [55].

2) Instantaneous CSI Acquisition for reduced-RF systems ($M < N$): In contrast with the full-RF system considered in (7), the AS system considered in this paper only implements $M \leq N$ RF chains. This constraint entails that only $M$ signals from the antenna ports can be processed simultaneously, i.e., the training signal received after $K$ training symbols reads as
\[
\tilde{Q} = (H_{|M})^H \Phi + \tilde{N},
\]
where $\tilde{Q} \in \mathbb{C}^{M \times n_{tr}}$, and $\tilde{N} \in \mathbb{C}^{M \times n_{tr}}$ is comprised of i.i.d. entries following $\tilde{n}_{i,j} \sim \mathcal{CN}(0, \sigma_{n}^2)$. As a result, only partial CSI $H_{|M}$ of the overall channel matrix $H$ can be derived from (9). This entails that a multiplexed training stage is required to estimate the channels of all $N$ antennas [28]. Indeed, the minimum number of training symbols required to estimate the channel in AS systems with a reduced number of RF chains is given by
\[
\eta_{tr} = K \times \left\lceil \frac{N}{M} \right\rceil,
\]
where $\lceil \cdot \rceil$ rounds to the highest closer integer. While this extended training might have a negligible impact on the attainable sum rates of slowly varying channels with large $\eta_{coh}$, their influence can instead be significant for fast varying channels, a trade-off that we explicitly study in Sec. [VII] for LSAS. The resultant number of symbol periods dedicated to downlink data transmission is given by
\[
\eta_{dl} = \eta_{coh} - \eta_{ul} = \eta_{coh} - \eta_{dl} = \left( K \times \left\lceil \frac{N}{M} \right\rceil \right), \quad (11)
\]
where $\eta_{ul}$ refers to the number of symbols employed for uplink data transmission as illustrated in Fig. [2].
3) Power-Based AS (PB-AS) for reduced-RF systems ($M < N$): An elegant solution to the CSI acquisition problem in AS systems consists in adopting a selection decision hinging on the norm of the channel entries. With this purpose, let \( \mathbf{h}_i \) denote the channel power measured per antenna element, where \( \mathbf{h}_i \) denotes the \( i \)-th column of \( 
abla \mathbf{H} \). The diagonal selection matrix \( \mathbf{S} \) defined in (4) can be obtained as

\[
\mathcal{P}_4 : \text{maximize } \sum_{i=1}^{N} \mathbf{S}_{i} = M, \\
\text{subject to } \mathbf{S}_{i,i} \in \{0,1\}, \forall i \in \{1, \ldots, N\}.
\]

This strategy is commonly referred to as power-based (PB) or norm-based antenna selection. The solution to the PB-AS of \( \mathcal{P}_4 \) is straightforward for the case of FF architectures, i.e.

\[
\tilde{\mathbf{S}}^* = \max_{\mathcal{M}} ||\mathbf{h}_i||^2,
\]

where \( \max_{\mathcal{M}} (\cdot) \) selects the largest \( M \) entries. While sub-optimal, PB-AS is capable of reducing the amount of time resources spent for CSI acquisition in systems where RF power meters instead of full RF chains are attached to each antenna port [23]. This is because a) the channel power information can be acquired from the prior uplink stage and b) this information can be subsequently employed for AS as per (12).

Therefore, the number of pilot symbols employed for downlink transmission \( \eta_{dl} \) for the case of PB-AS can be expressed as

\[
\eta_{dl} = \eta_{coh} - \eta_{ul} - \eta_{tr} = \eta_{coh} - \eta_{ul} - K,
\]

since instantaneous CSI is only required for the \( M \) antennas chosen for data transmission and we consider \( \eta_{tr} = K [9] \).

D. Sources of Losses in the RF Switching Matrices of AS

The design of the RF switching matrix in AS plays a fundamental role in the overall system performance [16]. Among the multiple technical aspects that should be considered from a system-level design perspective, the most relevant ones are:

- **Insertion losses (ILs).** RF switching matrices introduce ILs that generally grow with the number of input and output ports [56], [57]. This is a critical parameter for LSAS, where both the number of RF chains and antennas are significantly large [3]–[5], [46].

- **Coupling between ports.** The coupling of the switching matrix determines the fraction of the signals that appear at a specific port, but were intended for other ports. This parameter depends on the network of connections inside the switching matrix and the power leakage of the internal switching devices (e.g., FET transistors or electromechanical switches) [58].

- **Transfer function balance.** The transfer function of each input-output combination between RF chain and antenna element should be ideally identical to ensure that the baseband model [1] accurately characterizes the system’s operation [16].

From the above three sources of losses, here we focus on the ILs. This is because a) transfer function imbalances can be compensated via calibration prior to the normal system operation [16] and b) coupling between internal switching imbalances can be in the order of \(-20 \) to \(-30 \) dB, hence effectively making unintentional power transfers between nearby ports negligible [59]. Instead, the ILs introduced by switching matrices with a large number of input and output ports can be in the order of \(-2 \)–\(-3 \) dB [57], hence dominating the overall performance loss. For this reason, Sec. III explores a number of implementations of conventional FF switching matrices, with emphasis on the number of switching components required and the associated ILs.

III. FULLY-Flexible (FF) Switching for Antenna Selection

RF switching matrices in conventional AS have two essential requirements: a) connecting each RF chain to the antenna ports (full flexibility) and b) allowing bidirectional switching for uplink-downlink operation [16], [56]. These characteristics promote the implementation of the so-called blocking switching matrices, where the interconnection between the input and the output ports is performed by concatenated a number of switches of smaller size [56], [57]. An illustrative example of a blocking switching matrix is shown in Fig. 3(a), where the block diagram of a 4 \times 8 switching matrix can be observed. This figure shows that a large switching matrix is comprised of two main switching stages (represented by the dashed boxes in the figure) with multiple switches of smaller size: one stage at the RF-chain ports, referred to as RF-chain switching stage in the sequel, and a subsequent stage at the antenna ports, referred to as antenna switching stage.

Each of the above two switching stages is comprised by a number of smaller switches with a smaller number of ports, which will be referred to as basic switches in the following. Indeed, the RF-chain switching stage of Fig. 3(a) illustrates that several of these basic switches can be concatenated to produce the desired number of ports [57]. The basic switches considered in this paper conventionally follow the nomenclature SP\( X \)T (single-pole X-throw) which refers to the number of separate ports with independent signals that the basic switch can control (poles) and the number of different signal paths that the switch allows for each pole (throws) [59], [60]. For instance, a SP\( 3 \)T switch is capable of routing one signal from or towards three different ports (throws). The range of basic switches is usually SP\( 2 \)T-SP\( 10 \)T and their cost and ILs generally grow with the number of output ports [59], [60]. Without loss of generality, we consider the SP\( 2 \)T-SP\( 4 \)T basic switches with their IL detailed in Table I for the illustrative architectures shown in this paper [59].
Conventional AS systems consider FF switching matrices between the $M$ RF chains and the $N$ antenna ports, which allows any possible combination of $M$ antennas to be simultaneously selected. However, a number of implementations for the design of switching matrices with varying complexity and IL can be implemented. Importantly, different architectures may be implemented. Importantly, different architectures may result in RF-chain and antenna switching stages with different number of ports (threws). An accurate characterization of the maximum number of throws per switching stage is crucial, since they determine overall IL of the critical signal path, i.e. the signal path with largest power losses.

With the above purpose, let us define $T_{RF}$ and $T_{AN}$ as the maximum number of throws per RF chain ($T_{RF}$) in the RF-chain switching stage or per antenna ($T_{AN}$) in the antenna switching stage, as represented in Fig. 3(a). Moreover, let $\mathcal{T}$ be the set with elements in decreasing order given by the number of throws in the basic switches, i.e., $\mathcal{T} = \{4, 3, 2\}$ for the basic switches considered in Table I, $T_j$ corresponds to the $j$-th entry in $\mathcal{T}$ and the cardinality of $\mathcal{T}$ is $|\mathcal{T}| = N_t$. Intuitively, $N_t$ refers to the number of different basic switches considered, i.e. $N_t = 3$ for the basic switches of Table I. The total IL measured in dB of the critical signal path $L$ for a given switching architecture can be computed as

$$L = \sum_{j=1}^{N_t} (S_{T_j}^{\text{RF}} + S_{T_j}^{\text{AN}}) \times L_{T_j} = \sum_{j=1}^{N_t} S_{T_j} \times L_{T_j},$$

where $S_{T_j}^{\text{RF}}$ and $S_{T_j}^{\text{AN}}$ represent the number of consecutive basic switches with $T_j$ throws that signals cross in the RF-chain and in the antenna switching stages, respectively, and $L_{T_j}$ denotes the IL in dB introduced by a basic switch with $T_j$ throws. For instance, we have $L_2 = 0.25$ dB, $L_3 = 0.45$ dB and $L_4 = 0.45$ dB for the basic switches considered in Table I. Moreover, $S_{T_j} = S_{T_j}^{\text{RF}} + S_{T_j}^{\text{AN}}$ refers to the total number of switches with $T_j$ throws in the critical signal path of the overall switching matrix.

The number of basic switches with $T_j$ throws crossed by the transmit signals in the RF-chain and antenna switching stages can be iteratively computed as

$$S_{T_j}^{\text{RF, AN}} = \text{fact} \left( Q_{T_j}^{\text{RF, AN}}, T_j \right), j \in \{1, \ldots, N_t\},$$

where $\text{fact} \left( a, b \right)$, $b < a$ denotes the number of times $b$ appears in the integer factorization of $a$ and $Q_{T_j}^{\text{RF, AN}}$ is given by

$$Q_{T_j}^{\text{RF, AN}} = \begin{cases} T_j & \text{if } j = 1, \\ \left( Q_{T_{j-1}}^{\text{RF, AN}} \right) / \max \left( T_j \left( S_{T_{j-1}}^{\text{RF, AN}} \right), 1 \right), & \text{otherwise}. \end{cases}$$

In plain words, $T_j S_{T_j}$ represents the number of throws obtained by $S_{T_j}$ subsequent switching stages comprised of basic switches with $T_j$ throws, as described in the illustrative example below.

**Illustrative example:** Let us illustrate the procedure for computing the IL in (15) for a system with $T_{RF} = 8$ and $T_{AN} = 4$. This can correspond to a system with $M = 4$ and $N = 8$, as shown in Fig. 3(a). In this case, multiple basic switches from Table I are required at the RF-chain switching stage for connecting the RF chains with the $N = 8$ antennas. Specifically, since $Q_4^{\text{RF}} = 8$ as per (17), the number of consecutive SP4T switching stages is $S_{RF}^{\text{RF}} = \text{fact} (8, 4) = \text{fact} (4 \times 2, 4) = 1$ as per (16). By substituting the above values in (17) for $j = 2$ we have $Q_2^{RF} = Q_4^{RF} / T_4^{s_{RF}} = 8/4 = 2$. This results in $S_3^{RF} = \text{fact} (2, 3) = 0$. Since no SP3T switches are required, we finally have $Q_2^{RF} = 2$ as per (16) for $j = 3$, which entails that an additional stage of SP2T ($S_2^{RF} = 1$) switches is required for implementing the $T_{RF} = 8$ throws required for transmission per RF chain. Overall, the transmit signal of each RF chain must always cross one SP4T ($S_4^{RF} = 1$) and one SP2T ($S_2^{RF} = 1$) basic switches, as explicitly illustrated in Fig. 3(a) by the coloured signal path. This figure also shows that only one SP4T ($S_4^{AN} = \text{fact} (4, 4) = 1$) switch is required per antenna element. More intricate examples are also provided in Table I for the $N = 32$, $M = 6$ case.

Considering the above, in this work we explore three implementations for the design of FF switching matrices:

1. **Architecture 1. Conventional FF architecture with full connectivity (FF-FC).** This architecture is illustrated in Fig. 3(a), where it can be seen that each RF chain is connected to every antenna port. In this particular case

   $$T_{RF} = N, \text{ and } T_{AN} = M.$$  

2. **Architecture 2. FF architecture with minimum connectivity (FF-MC).** This architecture minimizes the maximum number of ports at the RF-chain and antenna switching stages. The block diagram of this architecture is shown in Fig. 3(b) for an illustrative $3 \times 5$ RF switching matrix. Here, it can be seen that there are additional constraints regarding the connectivity of each antenna. For instance, Fig. 3(b) shows that the first RF chain does not connect to antenna ports A-4 and A-5. In spite of this, a full

---

1For simplicity, it has been considered that $T_{RF}$ and $T_{AN}$ can be factorized into the integers contained in $\mathcal{T}$. 

---

Fig. 3. Block diagrams of FF architectures: (a) a conventional $4 \times 8$ switching matrix with each input port connected to every output port (FF-FC) and (b) a switching matrix minimizing the number of internal connections (FF-MC).
flexibility is guaranteed provided that

\[ T_{RF} = N - M + 1, \text{ and } T_{AN} = \min (M, T_{RF}). \tag{19} \]

Overall, the FF-MC architecture selects the basic switches such that \( T_{RF} \) and \( T_{AN} \) are minimized, while ensuring that any combination of antennas can be simultaneously activated. For instance, Table II shows that, for the 6 × 32 switching matrix, the \( T_{RF} = 32 \) ports from the FF-FC switching architecture can be reduced to \( T_{RF} = 27 \) ports when the FF-MC architecture is employed. Hence, this architecture aims to minimize the number of connections to simplify the hardware design and possibly the hardware monetary cost.

3) Architecture 3. FF architecture with minimum losses (FF-ML). The FF-MC architecture does not guarantee that the ILs are minimized, since designing networks with larger \( T_{RF} \) or \( T_{AN} \) might actually reduce the total IL. This counter-intuitive behaviour is illustrated in Table II for an AS system with \( N = 32 \) and \( M = 6 \), where it can be observed that minimizing the \( T_{RF} \) and \( T_{AN} \) as per the FF-MC architecture does not minimize the total IL. This is because the ILs introduced by the basic switches vary depending on their number of ports as per Table II, an effect particularly noticeable for large \( T_{RF} \) and \( T_{AN} \). In contrast, the FF-ML architecture selects the \( T_{RF} \) and \( T_{AN} \) that minimize the IL by relaxing (19) into

\[ T_{RF} \geq N - M + 1, \text{ and } T_{AN} \geq \min (M, T_{RF}). \tag{20} \]

At this point we note that while the different design criteria considered above guarantee full flexibility, their IL can still be as significant as 2-3 dB for LSAS. For instance, it can be shown that a system with \( N = 128 \) and \( M = 76 \) introduces ILs of \( L = 2.95 \), \( L = 3.2 \) and \( L = 3.4 \) dB for the FF-ML, FF-MC and FF-FC switching architectures, respectively. For this reason, in the following we concentrate on reducing the complexity and losses of RF switching matrices at the expense of reducing their flexibility.

### TABLE II

| Parameter | FF-FC | FF-MC | FF-ML | PC |
|-----------|-------|-------|-------|----|
| \( T_{RF} \) | 32    | 27    | 32    | 6  |
| \( T_{AN} \) | 6     | 6     | 6     | 1  |
| \( Q_{RF}^{\{4,3,2\}} \) as per (17) | \( Q_{RF}^{\{4,3,2\}} = \{32, 2, 2\} \) | \( Q_{RF}^{\{4,3,2\}} = \{27, 27, 0\} \) | \( Q_{RF}^{\{4,3,2\}} = \{32, 2, 2\} \) | \( Q_{RF}^{\{4,3,2\}} = \{6, 6, 2\} \) |
| \( S_{RF}^{\{4,3,2\}} \) basic switches required per RF chain | \( S_{RF}^{\{4,3,2\}} = \{2, 0, 1\} \) | \( S_{RF}^{\{4,3,2\}} = \{0, 3, 0\} \) | \( S_{RF}^{\{4,3,2\}} = \{2, 0, 1\} \) | \( S_{RF}^{\{4,3,2\}} = \{0, 1, 1\} \) |
| \( Q_{AN}^{\{4,3,2\}} \) as per (17) | \( Q_{AN}^{\{4,3,2\}} = \{6, 6, 2\} \) | \( Q_{AN}^{\{4,3,2\}} = \{6, 6, 2\} \) | \( Q_{AN}^{\{4,3,2\}} = \{6, 6, 2\} \) | \( Q_{AN}^{\{4,3,2\}} = \{1, 1, 1\} \) |
| \( S_{AN}^{\{4,3,2\}} \) basic switches required per antenna | \( S_{AN}^{\{4,3,2\}} = \{0, 1, 1\} \) | \( S_{AN}^{\{4,3,2\}} = \{0, 1, 1\} \) | \( S_{AN}^{\{4,3,2\}} = \{0, 1, 1\} \) | \( S_{AN}^{\{4,3,2\}} = \{0, 0, 0\} \) |
| Total number of basic switches \( (S_{AN}^{\{4,3,2\}}) \) | \( S_{4,3,2} = \{2, 1, 2\} \) | \( S_{4,3,2} = \{0, 4, 1\} \) | \( S_{4,3,2} = \{2, 1, 2\} \) | \( S_{4,3,2} = \{0, 1, 1\} \) |
| Total IL in dB \( (L) \) as per (15), where \( L_{T_{RF}} \) are given by Table II | \( L = 2 \times 0.45 + 1 \times 0.45 + 2 \times 0.25 = 1.85 \) dB | \( L = 4 \times 0.45 + 1 \times 0.25 = 2.05 \) dB | \( L = 2 \times 0.45 + 1 \times 0.45 + 2 \times 0.25 = 1.85 \) dB | \( L = 1 \times 0.45 + 1 \times 0.25 = 0.7 \) dB |

| ![Illustrative block diagrams for the PC switching architectures of (a) and (b)](image-url) |

IV. PARTIALLY-CONNECTED SWITCHING ARCHITECTURES AND RESULTING AS CONSTRAINTS

The significant IL and complexity of the FF matrices introduced for LSAS motivate the design of alternative low-complexity switching architectures [46]. In this section we consider a switching architecture with partial connectivity, referred to as partially-connected (PC) architecture, with arbitrary \( N \) and \( M \) for alleviating the above-mentioned concerns. Specifically, the PC architecture is designed to reduce the number of internal connections and basic switches, i.e.

\[ T_{RF} = \left\lceil \frac{N}{M} \right\rceil, \text{ and } T_{AN} = \begin{cases} 1, & \text{if } \left\lceil \frac{N}{M} \right\rceil \geq 2, \\ 2, & \text{otherwise,} \end{cases} \tag{21} \]

where the maximum number of throws per RF chain in the RF-chain switching stage, \( T_{RF} \), is defined to guarantee the essential constraint of connecting every antenna to at least one RF chain. Moreover, (21) shows that there are two specific cases depending on the ratio \( \frac{N}{M} \) and whose practical implications are considered in the following:

1) Case for \( \left\lceil \frac{N}{M} \right\rceil \geq 2 \). PC switching schemes satisfying this condition are those where each antenna is connected.
to a single RF chain \((T_{AN} = 1)\). Accordingly, basic switches SP3T or with a larger number of outputs have to be employed at the RF-chain switching stage to ensure that all antennas can be employed for transmission. An illustrative example is shown in Fig. 4(a), where \(N = 5\) and \(M = 2\). In this architecture, it can be observed that one SP3T must be implemented at the RF-chain switching stage for guaranteeing that every antenna is connected to one RF chain. We note that this architecture also comprehends the particular case \(N/M = 2\) considered in [46], where only binary (SP2T) switches are required to ensure antenna connectivity.

2) Case for \(\lfloor N/M \rfloor < 2\). In these architectures each antenna might be connected to more than one RF chain \((T_{AN} = 2)\). Accordingly, employing basic switches SP2T suffices to guarantee connectivity for all \(N\) antennas. A specific switching architecture satisfying \(N/M < 2\) is shown in Fig. 4(b) for \(N = 5\) and \(M = 3\). This figure shows that at least one antenna is connected to more than one RF chain, hence introducing additional ILs in the output switching stage when compared with the case \(N/M \geq 2\). However, incorporating additional RF chains simultaneously enhances the rates, which results in a relevant trade-off that will be studied in Sec. VII.

### A. Optimization constraints for AS with Instantaneous CSI

Overall, the above PC architectures impose a limited flexibility in the antenna selection procedure, since the simultaneous selection of certain antenna combinations is not implementable. In this context, the specific restrictions introduced by the partial connectivity should be considered when defining the optimization problem \(P_1\) for selecting the antennas as per Sec. II-B. With this purpose, let \(\mathcal{N} = \{1, \ldots, N\}\) and \(\mathcal{M} = \{1, \ldots, M\}\) be sets indexing the antennas and the RF chains, respectively. We define the antenna subsets \(\mathcal{N}^i, i \in \{1, \ldots, S_{\text{cons}}\}\) as disjoint sets of antennas sharing at least one common RF chain, i.e. \(\mathcal{N}^1 \cup \mathcal{N}^2 \cup \ldots \cup \mathcal{N}^{S_{\text{cons}}} = \mathcal{N}\) and \(\mathcal{N}^i \cap \mathcal{N}^j = \emptyset, \forall \ i, j \in \{1, \ldots, S_{\text{cons}}\}\). Here, \(\emptyset\) denotes the empty set and \(S_{\text{cons}}\) denotes the number of constraints detailed below. Similarly, \(\mathcal{M}^i \subset \mathcal{M}, i \in \{1, \ldots, S_{\text{cons}}\}\) represent disjoint sets of RF chains sharing at least one common antenna. For the illustrative example of Fig. 4(b), the sets \(\mathcal{M}^i\) are defined as \(\mathcal{M}^1 = \{1, 3\}\), since these RF chains inter-connect antennas \(\mathcal{N}^1 = \{1, 3, 5\}\), and \(\mathcal{M}^2 = \{2\}\), which inter-connects antennas \(\mathcal{N}^2 = \{2, 4\}\). Intuitively, the cardinality of the antenna groups sharing at least one RF chain \(S_{\text{cons}}\) also represents the additional number of constraints required in \(P_1\) to account for partial connectivity, which is given by

\[
S_{\text{cons}} = M - N_{ov},
\]  

where \(N_{ov}\) represents the number of antennas with the possibility of connecting to more than one RF chain. This is because the partial connectivity architecture imposes an additional constraint for every subset of antennas \(\mathcal{N}^i \subset \mathcal{N}, i \in \{1, \ldots, S_{\text{cons}}\}\), interconnected with their corresponding subset of RF chains \(\mathcal{M}^i \subset \mathcal{M}, i \in \{1, \ldots, S_{\text{cons}}\}\). For instance, Fig. 4(b) shows that only two out of antennas \(\mathcal{N}^1 = \{1, 3, 5\}\)

\(\text{can be simultaneously active, since these antennas are only connected to two RF chains} (|\mathcal{M}^1| = 2)\). Similarly, only one antenna out of antennas \(\mathcal{N}^2 = \{2, 4\}\) can be active because they share a single RF chain \((|\mathcal{M}^2| = 1)\).

Without loss of generality, we consider in the following that each RF chain is connected to antennas physically separated as shown in Fig. 4(a) and Fig. 4(b). Based on the above, let us define

\[
N_{\text{dist}} = \begin{cases} M, & \text{if } \lfloor N/M \rfloor \geq 2, \\ N - M, & \text{otherwise.} \end{cases}
\]  

(23)

In plain words, \(N_{\text{dist}}\) represents the distance between antennas inter-connected to a given RF chain. For instance, Fig. 4(a) and (b) show that each RF chain connects to antennas separated by \(N_{\text{dist}} = M = 2\) and \(N_{\text{dist}} = N - M = 2\) indices, respectively. Similarly, the distance between the RF chains connected to a given subset of antennas \(M_{\text{dist}}\) is given by

\[
M_{\text{dist}} = \begin{cases} 1, & \text{if } \lfloor N/M \rfloor \geq 2, \\ N - M, & \text{otherwise.} \end{cases}
\]  

(24)

For example, Fig. 4(b) shows that RF chains \(\mathcal{M}^1 = \{1, 3\}\), which inter-connect antennas \(\mathcal{N}^1 = \{1, 3, 5\}\), are separated by \(M_{\text{dist}} = N - M = 2\) because RF chain \(\mathcal{M}^2 = \{2\}\) is situated in-between. Accordingly, we can express \(\mathcal{N}^i\) and \(\mathcal{M}^i\) for the non-adjacent antenna connectivity considered in this paper as

\[
\mathcal{N}^i = \{i, i + N_{\text{dist}}, \ldots, i + \left(\frac{N - i + 1}{N_{\text{dist}}} - 1\right)N_{\text{dist}}\},
\]  

(25)

where each RF chain connects to antennas with indices separated by \(N_{\text{dist}}\) and

\[
\mathcal{M}^i = \{i, i + M_{\text{dist}}, \ldots, i + \left(\frac{M - i + 1}{M_{\text{dist}}} - 1\right)M_{\text{dist}}\},
\]  

(26)

where each antenna connects to RF chains with indices separated by \(M_{\text{dist}}\). Considering the above, the convex AS optimization problem for general PC architectures can be formulated as

\[
P_5 : \max_{S, \rho} \log_2 \det (\mathbf{I}_K + \rho \mathbf{P} \mathbf{H}^H)
\]  

(27a)

subject to

\[
\sum_{i=1}^{N} S_{i,i} = M, \quad (27b)
\]

\[0 \leq S_{i,i} \leq 1, \forall i \in \{1, \ldots, N\}, \quad (27c)
\]

\[\sum_{j \in N^i} S_{j,j} = |\mathcal{M}^i|, \forall i \in \{1, \ldots, S_{\text{cons}}\}, \quad (27d)
\]

where w.r.t. \(P_2\), it can be seen that additional constraints due to PC have been added as per (27d). Similarly to \(P_2\), the final solution to the AS can be obtained by choosing the \(M\) largest entries in the solution of \(P_5\) compliant with the partial connectivity constraints.

\[2\text{This consideration is motivated by the results obtained in [46] for real propagation environments, where it was shown that inter-connecting a given RF chain to non-adjacent antennas provides a better performance than the connection to adjacent antennas.}\]
B. Optimization constraints for AS with Power-Based CSI

The specific RF switching architecture selected also determines the procedure for performing AS based on the channel power at each of the antenna ports. Indeed, the solution to select the antennas is not straightforward as per the case of FF RF switching in (13), since additional constraints due to the partial connectivity must be incorporated to \( P_4 \). As a result, the convex optimization problem for PB-AS under limited connectivity can be expressed as

\[
P_6 : \text{maximize } \sum_{S} \mathbf{h}^T \mathbf{S},
\]

subject to \( \sum_{i=1}^{N} \mathbf{S}_{i,i} = M, \)

\( 0 \leq \mathbf{S}_{i,i} \leq 1, \forall i \in \{1, \ldots, N\}, \)

\( \sum_{j \in \mathcal{M}_i} \mathbf{S}_{j,j} = |\mathcal{M}_i|, \forall i \in \{1, \ldots, S_{\text{cons}}\}, \)

where the sets \( \mathcal{N}_i \) and \( \mathcal{M}_i \) are defined in (25) and (26), respectively, and it can be seen that new constraints are added w.r.t. \( P_4 \) as per (28d). The final solution is obtained by selecting the \( M \) largest entries satisfying the optimization constraints.

C. Implications of Reducing Connectivity

The additional constraints introduced to the optimization problems \( P_1 \) in \( P_4 \) have a number of practical system-level implications that we detail as follows:

- **Insertion losses.** The ILS introduced will be smaller in the design with partial switching connectivity due to the smaller number of basic switches required to implement the switching matrix. For instance, in the example of Table II, the ILS are reduced to \( L = 0.7 \) dB, since \( T_{RF} = 6 \) and \( T_{AN} = 1 \), i.e. at most two basic switches (SP2T and SP3T) are required per RF chain.
- **Power-based AS.** The need of solving \( P_6 \) entails that the signal processing complexity of the PB-AS under restricted RF switching connectivity is increased w.r.t. its FF counterpart, where only the largest entries have to be selected as per (13). Therefore, there exists a trade-off between hardware complexity reduction and signal processing increase when considering PC switching networks. However, while not as simple as solving (13), \( P_6 \) is convex and can be solved using standard algorithms [50]. Moreover, note that this optimization problem only has to be solved at the start of the channel coherence block. Therefore, the substantial hardware complexity savings demonstrated in Sec. VII motivate the employment of PC RF switching architectures.
- **System performance.** Limiting the connectivity affects the number of possible antenna combinations that can be selected. Due to its importance, this aspect is studied in Sec. VIII where the performance loss w.r.t. FF architectures is characterized for the relevant case of PB-AS.
- **Baseband signal processing.** The digital signal processor (DSP) must account for the reduced connectivity in order to perform the precoding / detection operations. This is because the order of the antenna channels might be different from the conventional order at the antenna ports. Note that this might also happen for FF architectures minimizing the number of connections (FF-MC) and the power losses (FF-ML). For instance, consider the architecture shown in Fig. 5(a) and assume that antennas \( N^2 = 2 \) and \( N^1 = 5 \) are those that maximize capacity. For correct symbol-to-antenna mapping, the DSP should be aware that the channel employed for precoding/detection might have a different order w.r.t. the conventional antenna port order, since now the first RF chain will be connected to antenna port 5 whereas the second RF chain will be connected to antenna port 2.
- **Analog hardware complexity.** The employment of a reduced number of basic switches and connections can reduce cross-coupling between hardware components that are physically close, the time required for calibrating the input-output transfer function and the overall economic cost for implementing the switching matrix.

D. Practical Hardware Implementation for PB-AS in LSAS

The solution advocated in [23] for reducing the CSI acquisition time in PB-AS is based on implementing analog power estimators at each antenna as shown in Fig. 5(a), i.e. \( N_{PM} = N \), where \( N_{PM} \) refers to the number of power meters. Note that here we consider that the speed and resolution requirements of the analog-to-digital converters (ADC) required in the parallel power-meter chains can be relaxed due to the non-sensitive nature of the power measurements. Nevertheless, additional data ports are required at the digital signal processor (DSP) to acquire this data.

The excessive number of antennas implemented in LSAS also motivates the implementation of a considerable number of RF chains, even when AS is implemented [37]. Since each RF chain also captures the channel power information during the uplink data stage, this entails that the solution illustrated in Fig. 5(a) acquires a significant amount of redundant data. This redundancy is eliminated in the architecture considered...
in Fig. 5(b), where only \( N_{PM} = N - M \) power meters are required. The particular example shown in Fig. 5(b) corresponds to an architecture with \( N = 2M \). When compared with the scheme of Fig. 5(a), it can be seen that, while additional RF switches are required, the number of parallel power-meter chains and data ports at the DSP can be substantially reduced. Since the technical specifications of the additional RF switches can be relaxed due to their non-critical operation, this solution constitutes an attractive alternative for implementing AS in LSAS.

V. PERFORMANCE ANALYSIS: DEGRADATION DUE TO LIMITED SWITCHING CONNECTIVITY

The partial connectivity architecture presented in Sec. VI reduces the IL introduced by the switching matrix at the expense of limiting the number of number of antenna combinations that can be simultaneously selected, which entails a loss in performance. In this section, we characterize this loss by following the intuitive approach adopted in [27] for PB-AS. The reasons for focusing on the simpler PB-AS are twofold: a) its reduced CSI acquisition time as detailed in Sec. II-C by use of RF power meters, and b) the negligible performance loss introduced by this scheme when compared with the selection based on instantaneous CSI for the channels considered in this paper. Indeed, it will be shown in Sec. VII that, due to the above, PB-AS outperforms AS based on instantaneous CSI in terms of sum rates when CSI acquisition overheads are considered.

A. Ergodic Capacity Approximation for FF-AS

The analysis developed in [27] essentially relies on approximating \( \mathbf{H}_{[\mathbf{A}]} \) in (2) by a matrix \( \mathbf{V} \in \mathbb{C}^{K \times M} \) with entries following an identical distribution but with each column of the having different variances, i.e.

\[
\mathbf{H}_{[\mathbf{A}]} \approx \mathbf{V}_{f(\mathbf{C})} = \mathbf{G} \Theta_{f(\mathbf{C})},
\]

(29)

where \( \mathbf{G} \in \mathbb{C}^{K \times M} \) is a matrix whose entries follow the same distribution of those from \( \mathbf{H} \), \( \mathbf{A}_{f(\mathbf{C})} \) indicates that \( \mathbf{A} \) is a function of the specific switching connectivity and \( \Theta \in \mathbb{R}^{M \times M} \) is a diagonal matrix whose definition is considered in the following. Specifically, let \( B_{t} = \sum_{i=1}^{K} |h_{i,t}|^2 \) denote the norm of the \( t \)-th column of \( \mathbf{H} \) and define \( B_{t;N} \) as the \( t \)-th smallest column norm of \( \mathbf{H} \) as per \( \mathcal{W} = \{ B_{1;N}, B_{2;N}, \ldots, B_{N;N} \} \). Then, the diagonal entries of \( \Theta \) are given by [27]

\[
\theta_{i,t} = \mathbb{E}[B_{t;N}] / \sqrt{K}, \quad t_i \in \{1, \ldots, N\}.
\]

(30)

The definition of the indices \( t_i \) is straightforward and deterministic for the case of FF switching networks, since the PB-AS will always selects the antennas corresponding to the largest column norms of \( \mathbf{H} \), i.e.

\[
\Theta = \text{diag} \left( \sqrt{B_{N;N}}, \sqrt{B_{N-1;N}}, \ldots, \sqrt{B_{N-M+1;N}} \right) / \sqrt{K}.
\]

(31)

A simpler approximation of the resultant channel in (29) can be obtained by averaging the power scaling (PS) factors of the selected ordered statistics in (30), which for FF switching networks yields

\[
\bar{P}_{FF} = \frac{1}{KM} \sum_{i=1}^{M} \mathbb{E}_H [B_{N-i+1;N}],
\]

(32)

where \( \mathbb{E}_H \) denotes that the expectation is taken over the small-scale fading parameters of the random channel \( \mathbf{H} \). For the case of uncorrelated Rayleigh flat-fading channels, the first moment of the \( t \)-th ordered random variable \( B_{t;N} \) is given by [61]; [62]

\[
\mathbb{E}_H [B_{t;N}] = \frac{N!}{(M-1)! (t-1)! (N-t)!} \sum_{r=0}^{t-1} (-1)^r \binom{t-1}{r} \frac{(K+s)!}{(N-t+r+1)^{K+s+1}},
\]

(33)

where \( a_s (N - t + r) \) represent the polynomial coefficient of \( x^s \) in the expansion

\[
\left( \sum_{s=0}^{K-1} \frac{x^s}{s!} \right)^N = a_0 x + a_1 x + a_2 x^2 + \cdots + a_N x^N.
\]

(34)

Note that \( a_s (N - t + r) \) can be recursively computed from \( a_1 (\cdot) \) \([61]; [62]\). The resultant ergodic capacity with the PS approximation can be subsequently expressed as [27]

\[
C_{PS-FF} = \mathbb{E}_G \left[ \log_2 \det \left( I_K + \rho \bar{P}_{FF} \mathbf{G} \mathbf{G}^H \right) \right],
\]

(35)

where the expectation over \( G \) has been analytically derived in [63]; [65] for different correlated and uncorrelated communication channels.

B. Ergodic Capacity Approximation for PC-AS

While the approximation in (35) accurately characterizes the ergodic capacity attainable by FF schemes, an optimal PB-AS cannot be guaranteed under limited connectivity in general, since selecting the antenna combination with the largest column norms of \( \mathbf{H} \) might not be feasible due to the additional constraints imposed in the optimization problem. Moreover, note that selecting some combinations of ordered column norms is more probable than selecting others, since the probability that the \( t \)-th ordered coefficient \( B_{t;N} \) is not chosen is different for every \( t \in \{1, \ldots, N\} \) as per the specific input-output connectivity in the implemented switching matrix. This entails that \( \mathbf{1}^\text{T} \) and \( \mathbf{2}^\text{T} \) are no longer valid for PC networks. For this reason we propose to further approximate [27] by taking into consideration the probability of selecting a given subset of the ordered set \( \mathcal{W} = \{ B_{1;N}, B_{2;N}, \ldots, B_{N;N} \} \).

With the above purpose, let us define \( B_j, j \in \{1, \ldots, N! \} \) of cardinality \( \mathcal{B} = M \) as discrete sets containing a given combination of ordered column norms of \( \mathbf{H} \). For instance, in the architecture shown in Fig. 4(a), the relevant sets \( \mathcal{B} \) with non-zero probability of being selected following a PB criterion are \( \mathcal{B}^1 = \{1, 2\} \), \( \mathcal{B}^2 = \{1, 3\} \) and \( \mathcal{B}^3 = \{1, 4\} \). Note that the antenna with the largest power can always be selected even under partial connectivity restrictions. Moreover, let \( T_j \in \{0, 1\} \) be a binary random variable that determines
whether the specific combination of columns of $\mathbf{H}$ determined by $\mathbf{B}^j$ is selected or not. Intuitively, the limited connectivity restricts a given combination of antennas $\mathbf{B}^j$ to be selected with a given probability. This entails that, in contrast with (32), the expectation in (30) must also be taken with respect to the discrete random variables $T_j$. Therefore, we propose to theoretically approximate the performance of PB-AS for PC switching architectures by employing

$$\tilde{P}_{PC} = \frac{1}{KM} \sum_{j=1}^{\left(\frac{N}{M}\right)} \left( \sum_{i=1}^{M} \mathbb{E}_{\mathbf{H}} \left[ \mathbf{B}_{i;N}^j \right] \right) \times P(T_j)$$

where $P(T_j)$ refers to the $i$-th entry of a set $A$ and the outer expectation is taken over the set of random discrete variables $T_j$. Closed-form expressions for the inner expectation in (36) are already available for multiple channels such as those with Rayleigh flat-fading as per (33) [27]. Instead, the outer expectation corresponds to that of a discrete random variable, which is given by

$$\tilde{P}_{PC} = \frac{1}{KM} \sum_{j=1}^{\left(\frac{N}{M}\right)} \left( \sum_{i=1}^{M} \mathbb{E}_{\mathbf{H}} \left[ \mathbf{B}_{i;N}^j \right] \right) \times P(T_j)$$

$$= \sum_{j=1}^{\left(\frac{N}{M}\right)} \left( \tilde{P}_j \times P(T_j) \right),$$

where $P(T_j)$ refers to the probability of jointly selecting the ordered channel columns determined by $\mathbf{B}^j$. The probabilities associated with selecting $\mathbf{B}^j$ can be computed depending on the specific switching connectivity and the particular channel characteristics. Specifically, by following the chain rule of probability we have

$$P(T_j) = P \left( \bigcap_{i=1}^{M} \text{selecting } \mathbf{B}_{i}^j \right)$$

$$= \prod_{i=1}^{M} P \left( \text{selecting } \mathbf{B}_{i}^j \right)$$

where $\cap$ denotes the intersection of events. To simplify the derivation of the joint probabilities in (38), in the following we concentrate on channels $\mathbf{H}$ adopting the form (3)

$$\mathbf{H} = \mathbf{R} \mathbf{F},$$

where $\mathbf{R} \in \mathbb{C}^{K \times K}$ is the deterministic channel covariance matrix and $\mathbf{F} \in \mathbb{C}^{K \times N}$ is a matrix of independent and identically distributed (i.i.d.) random variables following $\mathbf{f}_{i,j} \sim \mathcal{CN}(0,1)$. Note that $\mathbf{R}$ is a diagonal matrix with diagonal entries characterizing the large-scale channel fading if the users are placed sufficiently apart (3).

The above assumption entails that the probability of finding the $t$-th ordered statistic $B_{t;N}, t \in \{1, \ldots, N\}$ at a given antenna is equal for all antenna elements $\mathcal{N}$, which makes computing the probabilities in (38) straightforward for a given switching connectivity. An example of this derivation for $N = 5$ and $M = 2$ is provided in Appendix A for completeness. Once the probabilities of selecting a given ordered channel column combination $P(T_j)$ in (37) are determined, an approximation to the analytical ergodic capacity of AS systems with PC switching matrices can be expressed as

$$C_{PS-PC} \approx \mathbb{E}_{\mathbf{G}} \left[ \log_2 \det \left( \mathbf{I}_K + \rho \tilde{P}_{PC} \mathbf{GG}^H \right) \right],$$

An alternative approximation to the ergodic capacity of PC switching matrices can be derived by directly computing the expectation of the capacity over both the discrete random variables $T_j$ and $\mathbf{G}$. In this particular case, the approximation of the ergodic capacity is given by

$$C_{PS-PC} \approx \sum_{j=1}^{\left(\frac{N}{M}\right)} \mathbb{E}_{\mathbf{G}} \left[ \log_2 \det \left( \mathbf{I}_K + \rho \tilde{P}_j \mathbf{GG}^H \right) \right] \times P(T_j),$$

where $\tilde{P}_j$ is defined in (37).

The accuracy of the derived approximations can be observed in Fig. 6, which represents the ergodic capacity attained by the schemes considered in this paper against increasing $M$ for a system with $N = 8$, $K = 2$ and $\rho = 10$ dB. Here we consider a small-scale MIMO setup for illustrative reasons since, as shown in Sec. VII, the differences between the switching architectures vanish for large $N$. For the results of this figure, we consider an uncorrelated Rayleigh flat-fading channel and ignore the overheads associated with CSI acquisition. In spite of this unfavourable assumption for the PB criterion, it can be seen that the PB-AS approaches the performance attained of the instantaneous CSI-based AS for different values of $M$, an observation consistent with previous works [17], [27]. Fig. 6 also shows that only a slight performance loss is experienced for the switching architecture with partial connectivity w.r.t. the FF scheme, which is also coherent with the results obtained for $N = 2M$ in the real propagation channels of [46]. The slight performance loss can be explained by noting that, in general, there exists a large probability of selecting antennas with significant channel powers, even if these are not strictly the largest ones as shown in Appendix A. This result combined with those obtained in Sec. VII motivate the employment of PB-AS with partial connectivity, especially for LSAS where the ILs introduced by the switching matrices can be significant.
Moreover, Fig. 6 shows that the proposed approximations in (40) and (41) are able to capture the performance loss produced by the partial connectivity architecture, hence validating the approach adopted in this section.

VI. ENERGY EFFICIENCY

The enhancement of the communication’s energy efficiency is a key driver for considering AS implementations in LSAS due to the possibility of reducing the number of RF chains simultaneously active [46], [47], [66]. This comes, however, at the expense of introducing additional ILs in the RF switching stage, hence posing an energy efficiency trade-off that we aim to analyze in this paper. The energy efficiency can be expressed as

\[ \eta = \frac{R_{\text{sum}}}{P_{\text{tot}}} = \frac{R_{\text{sum}}}{P_{\text{PA}} + P_{\text{RF}}}, \]  

(42)

where \( R_{\text{sum}} \) refers to the sum rates in bits/s/Hz, \( P_{\text{tot}} \) is the total system power consumption, and \( P_{\text{PA}} = \frac{P_t}{M} \) is the power consumed by power amplifiers (PAs) with efficiency \( \kappa \) to produce an output signal of \( P_t \) Watts. Moreover, \( P_{\text{RF}} = MP_{\text{CIR}} + P_{\text{LO}} \) represents the power consumed by the analog hardware components excluding the power amplifiers. Here, \( P_{\text{CIR}} \) is the power consumed by the analog circuitry required in each of the RF chains, and \( P_{\text{LO}} \) denotes the power consumed by the local oscillator.

To explicitly account for the impact of the ILs on the system’s energy efficiency, in this work we incorporate \( L_{\text{IL}} \) directly in \( P_t \) above, that is, we consider that the RF switching matrix is placed after the power amplifiers in the transmission chain [15], [16]. This entails that power amplifiers in systems with AS will be required to produce output signals with larger power to compensate for the ILs of the switching stage, i.e. \( P_t^{\text{AS}} = P_t^{\text{no-AS}} \times L_{\text{IL}} \) [15], [16]. While other hardware solutions such as placing the switching matrices before the power amplifiers are certainly feasible, considering the impact of the ILs on the resultant system’s energy efficiency also becomes more intricate. This is because precise knowledge of the power amplifiers’ response, which is non-linear and component-dependent, would be required to quantify the additional power required per RF chain. Instead, the above consideration allows us to provide meaningful insights on the trade-offs that arise when employing AS with the switching architectures considered in this paper, as shown in the following.

VII. SIMULATION RESULTS

In this section we present numerical results for characterizing the performance of the switching architectures considered in this paper. In particular, we consider AS schemes based on both PB and instantaneous CSI decisions under uncorrelated Rayleigh flat-fading channels. Moreover, we show the sum rates attained via dirty paper coding (DPC) and for the more practical ZF precoder. We concentrate on characterizing the impact of varying the number of RF chains to provide guidelines for determining their optimal number in AS systems implementing the realistic switching architectures considered in this work. The results are obtained for LSAS, since the ILs and complexity of the FF switching networks are critical in these systems due to both the large number of RF chains and antennas deployed. For simplicity and without loss of generality, in the following we consider that the power losses introduced by the switching matrices correspond to those of the input-output combination with largest IL, i.e. the critical signal path.

Fig. 7 represents the sum rates of the schemes considered in this paper against increasing number of RF chains \( M \) in a system with \( N = 128 \), \( K = 16 \) and \( \rho = 15 \) dB. The instantaneous CSI is assumed to be acquired at no cost, i.e. an infinite channel coherence block is considered, while the impact of CSI acquisition is explained in subsequent results. It can be observed that ZF approaches the performance of DPC as \( M \) grows, due to the favourable propagation conditions experienced by LSAS [3]–[5]. The results of figure also show that there exists a very slight penalty when employing PB-AS against the AS based on instantaneous CSI in LSAS, a phenomenon already observed for small-scale MIMO systems and for LSAS in realistic propagation environments [17], [47]. Interestingly, it can also be observed that, independently of \( M \), the performance of PC-AS approaches that of the FF selection. This is occurs because, while not being able to select the optimal antenna combination due to a limited switching connectivity, there are numerous antennas combinations with a similar performance due to the large number of antennas (channels) available to perform the selection.

The results of Fig. 8 illustrate the impact of considering a limited channel coherence time and the need for training in AS for the same simulation of parameters of Fig. 7 with \( \eta_{\text{coh}} = 200 \). This corresponds to a fast-varying communication channel [8]. Without loss of generality, we consider that 70% of the remaining time resources after channel estimation are allocated to downlink transmission, i.e. \( \eta_{\text{dl}} = 0.7 \times (\eta_{\text{coh}} - \eta_{\text{tr}}) \) in (2), (6), (27a) and (28a), where \( \eta_{\text{tr}} \) is given by (11) and (14) for full-CSI and PB-AS, respectively. Fig. 8 critically shows that full CSI acquisition becomes suboptimal and PB-AS becomes a more attractive approach when a realistic channel coherence
time is considered. This is because acquiring the instantaneous CSI to perform the selection requires a larger training period, which results in a performance loss particularly significant for reduced \( M \). Instead, PB-AS can employ the \( N_{PM} = N - M \) power meters integrated in the analog stage for estimating the channel powers and, exploiting channel reciprocity, for performing the selection without requiring accurate CSI for all antennas. The results of Fig. 8 also indicate the most efficient ranges for selecting \( M \) in the instantaneous CSI-based decision without power meters, since sudden rate variations can be observed for increasing \( M \) due to the different training times required as per (10). Specifically, the stepwise behaviour is a direct consequence of \( \eta_t \), being an integer multiple of \( K \) in (10), as per the multiplexed training operation detailed in Sec. II-C.

The results of Fig. 9 illustrate the sum rates against increasing values of \( N \) for a system with \( M = 32, K = 16, \rho = 15 \text{ dB} \) and \( \eta_{coh} = 200 \). It can be observed that increasing \( N \) does not always provide higher attainable rates for the case of instantaneous CSI acquisition due to the larger training time required when \( M \) is fixed. Instead, it can be observed that PB-AS strictly enhances the sum rates as \( N \) grows. Moreover, Fig. 9 also shows that the performance differences between FC and PC switching are approximately preserved independently of \( N \). Similarly to Fig. 8 the stepwise trend is produced by the multiplexed training operation in (10).

Fig. 10 represents the ILs of (15) introduced by the FF and PC switching architectures considered in this paper against increasing \( M \). We set \( N = 128 \) and consider the basic switches described in Table I for implementing the switching network. The results of Fig. 10 clearly show the benefits of considering PC architectures when compared with FF schemes, which motivates their employment when simultaneously considering their small performance loss depicted in Fig. 7. For instance, the ILs can be reduced by up to 2.5 dB for \( M = N/2 \), which is the point with minimal IL for the PC network and the specific case considered in [46]. This can be explained by noting that \( T_{RF} = 2 \) and \( T_{AN} = 1 \) for \( M = N/2 \). Instead, implementing a larger \( M \) requires additional RF switches at the output stage of the switching network as illustrated in Fig. 3b, hence introducing additional losses in the critical signal path. However, we remark that considering a larger number of RF chains might also be required in realistic systems for satisfying specific sum rate requirements.

Fig. 10 also shows that there are non-desirable areas where the power losses can be substantially increased, if a FF architecture is preferred. Let us concentrate on understanding the behaviour these FF architectures. As expected, the architecture designed to minimize the power losses (FF-ML) introduces smaller ILs than those with different criteria. Indeed, it can be observed that there are points where minimizing the number of connections as per FF-MC also entails larger ILs than both FF-ML and the architecture with complete input-output connectivity (FF-FC). This counter-intuitive behaviour can be explained by detailing the number of basic RF switches required at the input switching stage, which is explicitly detailed in Table II for the illustrative case of \( N = 32 \) and \( M = 6 \). Overall, the results of Fig. 10 provide meaningful insights for the design of AS systems, since it can be observed...
that the ILs of different switching architectures do not follow a monotonic trend with $M$. This behaviour, not considered in the related literature, arises due to the differences in the IL of the basic switches and the specific number of ports $T_{RF}$ and $T_{AN}$ required at the switching matrices, as analyzed in Sec. III.

The total system power consumption $P_{\text{tot}}$ in (42) is shown in Fig. 11 against increasing values of $M$. Here, we have set $N = 128$, $P_t = 30$ dBm and $P_{RF} = \gamma M (P_t/100)$, i.e. $\gamma$ indicates the relative power consumption of the RF circuitry when compared with the total transmission power. We adopt this approach to make our conclusions independent of the specific transmission power, since the relative power consumption of the RF chains when compared with the power consumed by the power amplifiers strongly depends on the transmission power ($70$, $71$). Fig. 11 considers $\gamma = 0.5$, which corresponds to a setup where the PA power consumption dominates the total power consumption, and $\gamma = 5$, where the circuitry power consumption dominates instead. Note that the circuitry power consumption can dominate in realistic systems such as those of macro BSs with $P_t = 46$ dBm ($70$, $71$), $P_{\text{CH}} = 1$ W and $P_{\text{LO}} = 2$ W ($67$), which results in $\gamma \approx 2.5$.

The results of Fig. 11 illustrate the importance of reducing the ILs in the switching stage, since substantial power savings over FF-AS can be attained when employing the PC architecture. Moreover, it can be observed that employing AS might be unfavourable in systems where the power consumed by the PAs dominates ($\gamma = 0.5$), since the additional PA power required to compensate for the additional ILs is not counterbalanced by the savings in the power consumption of the RF circuitry. In other words, the most energy efficient option for FF-AS as per (42) may well be a MIMO system with $N = M$, since the power consumption is reduced for $N = M$ and the attainable rates grow with $M$. Instead, significant enhancements in the total system power consumption can be attained via PC-AS in systems where the RF power consumption dominates, hence motivating their implementation in these setups.

The energy efficiency $\xi$ in (42) is shown vs. $M$ in Fig. 12 for a ZF precoding system with $N = 128$, $K = 16$, $\rho = 15$ dB, $\eta_{\text{coh}} = 200$, $P_t = 46$ dBm, $\kappa = 0.38$, $P_{\text{CH}} = 1$ W and $P_{\text{LO}} = 2$ W ($67$). Fig. 12 illustrates the benefits of employing both a switching network with limited connectivity and relying on power estimates for performing the AS decision. Specifically, it can be observed that the energy efficiency of the power-based PC-AS is maximized for $M \approx 32$, which can be explained by noting a) the significantly reduced IL when compared with smaller $M$ in Fig. $10$ and b) that a large portion of the maximum attainable sum rates for $M = 128$ is already obtained for $M \approx 32$ as shown in Fig. $8$.

We remark that, in spite of their reduced energy efficiency, the range $M > 40$ is still of interest, since some systems may pose specific sum rate constraints that should be met. Overall, Fig. 12 demonstrates the importance of considering the ILs in the switching stage for energy-efficient system design.

VIII. CONCLUSION

In this paper, we analyze the impact of implementing different RF switching architectures in AS systems. The hardware features of a number of conventional fully-flexible switching designs are characterized, which motivates the implementation of switching matrices with partial connectivity. We characterize the performance loss introduced by these reduced-complexity architectures as well as relevant differences that should be considered when implemented in combination with power-based AS. The results obtained in this paper promote the employment of PC switching networks, particularly for LSAS, and provide practical insights for the designs of AS systems. Future work includes studying the impact of relying on the uplink data for performing PB-AS and incorporating hardware non-idealities in the parallel power meter chains.

APPENDIX A. COMPUTATION OF THE JOINT PROBABILITIES IN (38)

In this Appendix we outline the procedure for computing the probabilities $P(T_i)$ in (38) for completeness. In particular, we do this for the PC architecture in Fig. 4(a) for reasons...
of illustration, where a scheme with $M = 2$ and $N = 5$ is considered. As detailed in Sec. [V] the sets containing the combinations of the ordered column norms with non-zero probability of being selected are $B^1 = \{1, 2\}$, $B^{2} = \{1, 3\}$ and $B^3 = \{1, 4\}$. Let us start by computing $P(T_1)$, which is given by

$$P(T_1) = P(\text{selecting } B^1_1 \cap \text{selecting } B^1_2) =$$

$$= (\text{a}) P(\text{selecting } B^1_1) P(\text{selecting } B^1_2|\text{selected } B^1_1)$$

$$= (\text{a}) P(\text{selecting } B^1_1) P(\text{selecting } B^1_2|\text{selected } B^1_1)$$

$$= \left(\frac{5}{3}\right) P(\text{not selecting } B^1_2) \cap \text{selecting } B^1_2 = \left\{1, 3, 5\right\}$$

$$= \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{1}{10}$$

where (a) and (b) hold because the only possibility of selecting the antenna with the $B^3$ is 4-th largest norm is that the channels with the first, second and third largest norms are in antennas $\{1, 3, 5\}$. Finally, $P(T_2) = 1 - P(T_1) = P(T_3) = \frac{3}{10}$.

REFERENCES

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” IEEE Commun. Magazine, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[2] J. G. Andrews et al., “What will 5G be?” IEEE Journal on Sel. Areas in Commun., vol. 32, no. 6, pp. 1065–1082, June 2014.

[3] F. Rusek et al., “Scaling up MIMO: Opportunities and challenges with very large arrays,” IEEE Signal Process. Magazine, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[4] K. Zheng et al., “Survey of large-scale MIMO systems,” IEEE Commun. Surveys Tutorials, vol. 17, no. 3, pp. 1738–1760, thirdquarter 2015.

[5] L. Lu et al., “An overview of massive MIMO: Benefits and challenges,” IEEE Journal of Sel. Topics in Signal Process., vol. 8, no. 5, pp. 742–758, Oct. 2014.

[6] T. S. Rappaport et al., “Millimeter wave mobile communications for 5G cellular: It will work!” IEEE Access, vol. 1, pp. 335–349, 2013.

[7] Z. Pi and F. Khan, “An introduction to millimeter-wave mobile broadband systems,” IEEE Commun. Magazine, vol. 49, no. 6, pp. 101–107, June 2011.

[8] E. Björnson, E. G. Larsson, and T. L. Marzetta, “Massive MIMO: Ten myths and one critical question,” IEEE Commun. Magazine, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[9] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. on Wireless Commun., vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[10] E. Ayach et al., “Spatially sparse precoding in millimeter wave MIMO systems,” IEEE Trans. on Wireless Commun., vol. 13, no. 3, pp. 1499–1513, March 2014.

[11] P. V. Amadori and C. Masouros, “Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection,” IEEE Trans. on Commun., vol. 63, no. 6, pp. 2212–2223, June 2015.

[12] M. D. Renzo, H. Haas, A. Ghayeb, S. Sagiuara, and L. Hanzo, “Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation,” Proceedings of the IEEE, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[13] P. Yang et al., “Single-carrier spatial modulation: A promising design for large-scale broadband antenna systems,” IEEE Commun. Surveys Tutorials, vol. PP, 2016.

[14] A. García-Rodríguez and C. Masouros, “Low-complexity compressive sensing detection for spatial modulation in large-scale multiple access channels,” IEEE Trans. on Commun., vol. 63, no. 7, pp. 2565–2579, July 2015.

[15] A. F. Molisch, “MIMO systems with antenna selection – an overview,” in Radio and Wireless Conf., Aug. 2003, pp. 167–170.

[16] A. F. Molisch and M. Z. Win, “MIMO systems with antenna selection,” IEEE Microwave Magazine, vol. 5, no. 1, pp. 46–56, March 2004.

[17] S. Sanayei and A. Nosratinia, “Antenna selection in MIMO systems,” IEEE Commun. Magazine, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[18] A. Gorokhov, D. A. Gore, and A. J. Paulraj, “Receive antenna selection for MIMO spatial multiplexing: Theory and algorithms,” IEEE Trans. on Signal Process., vol. 51, no. 11, pp. 2796–2807, Nov. 2003.

[19] R. W. Heath and A. Paulraj, “Antenna selection for spatial multiplexing systems based on minimum error rate,” in IEEE Int. Conf. on Commun. (ICC), vol. 7, 2001, pp. 2276–2280.

[20] R. W. Heath, S. Sandhu, and A. Paulraj, “Antenna selection for spatial multiplexing systems with linear receivers,” IEEE Commun. Letters, vol. 5, no. 4, pp. 142–144, April 2001.

[21] Y.-S. Choi, A. F. Molisch, M. Z. Win, and J. H. Winters, “Fast algorithms for antenna selection in MIMO systems,” in IEEE Vehicular Technology Conf. (VTC Fall), vol. 3, Oct. 2003, pp. 1723–1737.

[22] A. Gorokhov, D. Gore, and A. Paulraj, “Receive antenna selection for MIMO flat-fading channels: Theory and algorithms,” IEEE Trans. on Inf. Theory, vol. 49, no. 10, pp. 2687–2696, Oct. 2003.

[23] S. Sanayei and A. Nosratinia, “Capacity of MIMO channels with antenna selection,” IEEE Trans. on Inf. Theory, vol. 53, no. 11, pp. 4356–4362, Nov. 2007.

[24] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes, “On the diversity order of spatial multiplexing systems with transmit antenna selection: A geometrical approach,” IEEE Trans. on Inf. Theory, vol. 52, no. 12, pp. 5297–5311, Dec. 2006.

[25] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, “Capacity of MIMO systems without antenna selection,” IEEE Trans. on Wireless Commun., vol. 4, no. 4, pp. 1759–1772, July 2005.

[26] A. Ghayeb and T. M. Duman, “Performance analysis of MIMO systems with antenna selection over quasi-static fading channels,” IEEE Trans. on Vehicular Technology, vol. 52, no. 2, pp. 281–288, March 2003.

[27] P. J. Smith, T. W. King, L. M. Garth, and M. Mohler, “A power scaling analysis of norm-based antenna selection techniques,” IEEE Trans. on Wireless Commun., vol. 7, no. 8, pp. 3140–3149, Aug. 2008.

[28] A. B. Narasimhamurthy and C. Tepedelenlioglu, “Antenna selection for MIMO-OFDM systems with channel estimation error,” IEEE Trans. on Vehicular Technology, vol. 58, no. 5, pp. 2269–2278, June 2009.

[29] V. Krishnamurthy and C. Tepedelenlioglu, “Sensing detection for spatial modulation in large-scale multiple access channels,” IEEE Trans. on Info. Theory, vol. 53, no. 1, pp. 435–437, Jan. 2007.

[30] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, “Capacity of MIMO systems with antenna selection,” IEEE Trans. on Wireless Commun., vol. 5, no. 4, pp. 1759–1772, July 2005.

[31] A. Ghayeb and T. M. Duman, “Performance analysis of MIMO systems with antenna selection over quasi-static fading channels,” IEEE Trans. on Vehicular Technology, vol. 52, no. 2, pp. 281–288, March 2003.

[32] P. J. Smith, T. W. King, L. M. Garth, and M. Mohler, “A power scaling analysis of norm-based antenna selection techniques,” IEEE Trans. on Wireless Commun., vol. 7, no. 8, pp. 3140–3149, Aug. 2008.

[33] A. B. Narasimhamurthy and C. Tepedelenlioglu, “Antenna selection for MIMO-OFDM systems with channel estimation error,” IEEE Trans. on Vehicular Technology, vol. 58, no. 5, pp. 2269–2278, June 2009.

[34] V. Krishnamurthy and C. Tepedelenlioglu, “Sensing detection for spatial modulation in large-scale multiple access channels,” IEEE Trans. on Info. Theory, vol. 53, no. 1, pp. 435–437, Jan. 2007.
[33] A. Dua, K. Medepalli, and A. J. Paulraj, “Receive antenna selection in massive MIMO systems using convex optimization,” *IEEE Trans. on Wireless Commun.*, vol. 5, no. 9, pp. 2353–2357, Sept. 2006.

[34] A. F. Molisch et al., “Implementation aspects of antenna selection for MIMO systems,” in *Int. Conf. on Commun. and Networking in China*, Oct. 2006, pp. 1–7.

[35] J. Ahmadi-Shokouh, S. H. Jamali, S. Safavi-Naeini, and G. Z. Rafi, “Switch loss and antenna directivity effects on MIMO antenna selection,” in *Canadian Conf. on Electrical and Computer Engineering (CCECE)*, May 2008, pp. 641–646.

[36] P. V. Amadori and C. Masouros, “Interference driven antenna selection for massive multi-user MIMO,” *IEEE Trans. on Vehicular Technology*, vol. PP, 2015.

[37] K. Dong, N. Prasad, X. Wang, and S. Zhu, “Adaptive antenna selection and Tx/Rx beamforming for large-scale MIMO systems in 60 GHz channels,” *EURASIP Journal on Wireless Commun. and Networking*, vol. 11, no. 11, pp. 1–15, 2011. [Online]. Available: http://dx.doi.org/10.1186/1687-1499-2011-39

[38] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, “ Massive MIMO performance evaluation based on measured propagation data,” *IEEE Trans. on Wireless Commun.*, vol. 14, no. 3, pp. 3899–3911, July 2015.

[39] H. Li, L. Song, and M. Debbah, “Energy efficiency of large-scale multiple antenna systems with transmit antenna selection,” *IEEE Trans. on Commun.*, vol. 62, no. 2, pp. 638–647, Feb. 2014.

[40] B. M. Lee, J. Choi, J. Bang, and B. C. Kang, “An energy efficient antenna selection for large scale green MIMO systems,” in *IEEE Int. Symp. on Circuits and Systems (ISCAS)*, May 2013, pp. 950–953.

[41] S. Mahboob, R. Ruby, and V. C. M. Leung, “Transmit antenna selection for downlink transmission in a massively distributed antenna system using convex optimization,” in *Int. Conf. on Broadband, Wireless Computing, Commun. and Applications (BWCCA)*, Nov. 2012, pp. 228–233.

[42] M. Gkizeli and G. N. Karystinos, “Maximum-SNR antenna selection among a large number of transmit antennas,” *IEEE Journal of Sel. Topics in Signal Process.*, vol. 8, no. 5, pp. 891–901, Oct. 2014.

[43] R. Kataoka, K. Nishimori, N. Tran, and T. Imai, “Performance evaluation by antenna selection using real propagation channel on massive MIMO,” in *IEEE Int. Workshop on Electromagnetics (iWEM)*, Aug. 2014, pp. 227–228.

[44] M. Benmimoune, E. Driouch, W. Ajib, and D. Massicotte, “Feedback reduction and efficient antenna selection for massive MIMO system,” in *IEEE Vehicular Technology Conf. (VTC Fall)*, Sept. 2015, pp. 1–6.

[45] B. Fang, Z. Qian, W. Shao, and W. Zhong, “Raise: A new fast transmit antenna selection algorithm for massive MIMO systems,” *Wireless Personal Commun.*, vol. 80, no. 3, pp. 1147–1157, 2014. [Online]. Available: http://dx.doi.org/10.1007/s11277-014-2077-4

[46] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, “Multi-switch for antenna selection in massive MIMO,” in *IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[47] ——, “Massive MIMO in real propagation environments: Do all antennas contribute equally?” *IEEE Trans. on Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.

[48] R. Ménéndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, “Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?” *IEEE Access*, vol. 4, pp. 247–267, 2016.

[49] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE Journal on Sel. Areas in Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[50] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[51] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates, and sum-rate capacity of gaussian MIMO broadcast channels,” *IEEE Trans. on Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[52] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[53] T. L. Marzetta, “How much training is required for multiuser MIMO?” in *Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, Oct. 2006, pp. 359–363.

[54] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel,” *IEEE Trans. on Commun.*, vol. 61, no. 6, pp. 2350–2361, June 2013.

[55] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, “Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics,” *IEEE Journal of Sel. Topics in Signal Process.*, vol. 8, no. 5, pp. 815–830, Oct. 2014.