Exploiting the Increasing Correlation of Space Constrained Massive MIMO for CSI Relaxation

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

Abstract—In this paper, we explore low-complexity transmission in physically-constrained massive multiple-input multiple-output (MIMO) systems by means of channel state information (CSI) relaxation. In particular, we propose a strategy to take advantage of the correlation experienced by the channels of neighbour antennas when deployed in tightly packed antenna arrays. The proposed scheme is based on collecting CSI for only a subset of antennas during the pilot training stage and, subsequently, using averages of the acquired CSI for the remaining closely-spaced antennas. By doing this, the total number of radio frequency (RF) chains, for both CSI acquisition and data transmission, and the baseband signal processing are reduced, hence simplifying the overall system operation. At the same time, this impacts the quality of the channel estimation produced after the CSI acquisition process. To characterize this tradeoff, we explore the impact that the number of antennas with instantaneous CSI has on the performance, signal processing complexity, and energy efficiency of time-division duplex (TDD) systems. The analytical and simulation results presented in this paper show that the application of the proposed strategy in size-constrained antenna arrays is able to significantly enhance the energy efficiency against systems with full CSI availability, while approximately preserving their average performance.

Index Terms—Massive MIMO, incomplete CSI, antenna correlation, energy efficiency.

I. INTRODUCTION

THE DESIGN of future wireless communication systems centers upon rigorous spectral and energy efficiency requirements [1]–[3]. Specifically, novel communication strategies aim at meeting the exponentially increasing data rate demands while simultaneously offering a reduced power consumption [1]. A feasible solution to satisfy these requirements has been provided by massive or large-scale MIMO systems, which are based on incorporating a large number of antennas at the base stations (BSs) [2], [3]. These systems are able to significantly outperform their small-scale counterparts by using conventional linear detection and precoding schemes, which have been shown to offer a close-to-optimal performance for a large but finite number of antennas [4], [5]. This is because

the impact of harmful communication effects such as the environmental noise or the imperfect CSI vanish as more antennas are allocated at the BSs [2], [4], [5]. Nevertheless, the use of these systems has also posed new challenges that need to be solved before their practical implementation [2].

One of the main concerns is related to the deployment of a large number of antennas in constrained physical spaces [2], [6], [7]. This leads to an increased inter-antenna correlation due to insufficient antenna separation, which may affect the system capacity unless conveniently approached [2], [8]–[16]. Indeed, both spatial correlation and mutual coupling effects have been extensively studied in the MIMO literature [9]–[16]. In this context, the excessive number of antennas deployed in massive MIMO makes the consideration of these effects critical. The significant power consumption introduced by the additional RF chains implemented in massive MIMO systems can also have a negative impact on the energy efficiency [17]. Moreover, the collection of accurate CSI complicates the application of massive MIMO to frequency-division duplex (FDD) scenarios. This is because the CSI acquisition time of the conventional FDD approaches is proportional to the number of antennas allocated at the BSs [2], [3]. As a consequence, a significant part of the literature has focused on the operation of TDD systems and the pilot contamination problem [3], [18]. Additionally, even for TDD the large number of antennas entails a notable increase in the number of signal processing operations performed by channel estimation processes and linear precoding and detection strategies [19], [20]. In fact, this constitutes one of the bottlenecks in the practical implementation of these systems, since the complexity of linear precoding schemes such as zero-forcing (ZF) is still considered burdensome [19].

The above-mentioned challenges have been explored in a number of related works. For instance, the effects of fitting a massive number of antennas in size-constrained structures have been studied in [6]–[8], [21]. Specifically, [6] explores the impact of physical space constraints on the sum rate performance of massive MIMO systems whereas [8] demonstrates that inter-user interference does not vanish in the large-scale limit when the physical size of the antenna array is constrained. While focused on massive MIMO setups, these works lean on the results previously obtained in [9]–[16]. Indeed, these results have also been leveraged in [12], [22]–[24] for reconfiguring the physical location of the antenna elements with the purpose of maximizing performance. The influence of the power consumption of the RF circuitry on the energy efficiency and strategies to reduce the number of electronic components have been considered in [17], [25]–[28]. Additionally, several alternatives have been proposed to improve the acquisition of the

CSI for both TDD and FDD systems (see, e.g., [18], [28]–[33] and references therein). Some of these works exploit the appearance of low-rank channel matrices due to the inter-antenna correlation of massive antenna arrays or finite scattering to enhance the channel estimation process. Particularly related to this work is [34], where the large correlation between adjacent antennas is exploited to optimize the design of the feedback information during the CSI acquisition stage of FDD systems. A number of solutions to reduce the computational complexity involved in the channel estimation [20] and signal precoding or detection [19], [35] have been also proposed. In line with these contributions, here we study a solution particularly tailored for size-constrained antenna arrays that addresses several of the above-mentioned challenges.

Specifically, in this paper we propose a strategy to optimize the trade-off between the quality of the acquired CSI and the energy efficiency in tightly packed antenna arrays with a large number of antennas. Here, we do not rely on reconfiguring the antenna array in contrast with [12], [22]–[24], or on exploiting channel sparsity due to having directional links with a limited number of scattering clusters as per the millimeter wave designs of [36]–[40]. Instead, the proposed scheme is based on exploiting inter-antenna correlation to relax CSI acquisition by completely disposing with the CSI of a subset of antennas. This is done by deactivating a number of antennas during the pilot training stage, hence acquiring incomplete information of the global communication channel. After this process, the CSI of the antennas that were inactive during the CSI acquisition stage is obtained by simply averaging the instantaneous CSI of the neighbour antennas that collected this information. As a result, the signal processing load and the power consumption of the RF circuitry during the CSI acquisition process is reduced at the cost of decreasing the accuracy of the channel estimation. We later show that this benefit holds true for the data transmission stage as well. This motivates us to study the effects that the number of antennas without instantaneous CSI has on the performance, computational load and energy efficiency of size-constrained massive BSs.

In particular, although the collection and use of incomplete CSI has notable implications in the implementation of both TDD and FDD scenarios, in the following we concentrate on TDD systems for both reasons of simplicity and for their practical importance in massive MIMO systems [2], [3], [41]. The relevance of TDD systems for future massive MIMO deployments is justified by the independence of their CSI pilot training length on the excessive number of antennas implemented at the BS. The latter dependence arises in FDD systems operating in non-sparse channels [2], [3], [41]. In this setting, we observe that the transmission performance of massive MIMO BSs deployed in constrained physical spaces is highly tolerant to the acquisition of incomplete CSI. This translates to significant gains in the energy efficiency due to the simultaneous reduction in the number of RF chains and signal processing load offered by the proposed scheme. Additionally, the acquisition of incomplete CSI can be combined with other channel estimation and pilot design techniques to further enhance performance [29], [30]. For clarity, the contributions of this paper can be summarized as follows:

1) We propose a low-complexity scheme based on the acquisition of incomplete CSI to exploit the inter-antenna correlation that arises in massive MIMO systems deployed in constrained physical spaces.
2) We perform a mathematical analysis of the error introduced by the proposed strategy in the channel estimation process to characterize the impact of the incomplete CSI acquisition.
3) We carry out a) complexity and b) energy efficiency analyses and use these to examine the trade-off between CSI relaxation and performance against systems with complete CSI availability.

II. SYSTEM MODEL

A. TDD System Model and CSI Acquisition

A generic multi-user TDD system comprised of K single-antenna users and a large-scale BS with $N \gg K$ antennas is considered in this paper. The communication channel follows a block-fading propagation model, hence remaining constant throughout a number of η_{coh} symbols in which a frame is conveyed, and varying independently between frames [4]. Initially, the CSI is acquired at the BS via the predefined orthogonal training signals transmitted from the mobile stations (MSs) during the first $\eta_{\text{tr}} \geq K$ symbol slots [42]–[45]. Specifically, let $\mathbf{H} \in \mathbb{C}^{N \times K}$ denote the uplink communication channel matrix. The signal received by the BS during the training stage is given by [42], [46]

$$\mathbf{P} = \sqrt{\rho_{\text{ul}}} \cdot \mathbf{H}\boldsymbol{\Phi} + \mathbf{N}, \qquad (1)$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{K \times \eta_{\text{tr}}}$ represents the orthogonal training signals transmitted by the users, ρ_{ul} is the signal-to-noise ratio (SNR) of the reverse channel, and $\mathbf{N} \in \mathbb{C}^{N \times \eta_{\text{tr}}}$ is an additive white Gaussian noise matrix with independent and identically distributed (i.i.d.) entries $n_{i,j} \sim \mathcal{CN}(0, 1)$. Here, $\sim$ can be read as "distributed as", and $\mathcal{CN}(\mu, \sigma^2)$ represents a circularly symmetric complex Gaussian random variable with mean μ and variance σ^2. The training signal matrix is given by $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_K]^T$, where $\boldsymbol{\phi}_k \in \mathbb{C}^{\eta_{\text{tr}} \times 1}$ denotes the orthogonal training signal assigned to the k-th user and $(\cdot)^T$ represents the transpose of a matrix or vector. The correlation of the received signals during the training stage with the pilot sequences yields the decision metric [5], [42], [46]

$$\widehat{\mathbf{P}} = \left(\sqrt{\rho_{\text{tr}}} \cdot \mathbf{H}\boldsymbol{\Phi} + \mathbf{N}\right) \boldsymbol{\Phi}^H = \sqrt{\rho_{\text{tr}}} \cdot \mathbf{H} + \mathbf{W}, \qquad (2)$$

where ρ_{tr} represents the effective training SNR [5], $(\cdot)^H$ denotes the Hermitian transpose and $\mathbf{W} = \mathbf{N}\boldsymbol{\Phi}^H$ is comprised of i.i.d. entries given by $w_{i,j} \sim \mathcal{CN}(0, 1)$. The estimated uplink channel can be directly obtained from (2) by applying estimation strategies such as those based on minimizing the mean square error (MMSE) [20], [46]. Moreover, provided that channel reciprocity holds, the downlink channel can be straightforwardly obtained as the conjugate transpose of the uplink estimate [5].

Once the pilot transmission stage is completed, the remaining symbol slots $\eta_{\text{coh}} - \eta_{\text{tr}}$ of the channel coherence time are used

for downlink and uplink data transmission [42]–[45], [47]. In particular, the signal received by the k-th MS during a downlink symbol slot can be expressed as

$$y_k = \sqrt{\rho_f} \mathbf{h}_k^H \mathbf{x} + z_k, \tag{3}$$

where $\mathbf{x} \in \mathbb{C}^{N \times 1}$ denotes the symbols transmitted by the BS, $\mathbf{h}_k^H \in \mathbb{C}^{1 \times N}$ represents the downlink communication channel from the BS to the k-th user, ρ_f denotes the SNR of the forward link, and $z_k \sim \mathcal{CN}(0, 1)$ is the standard additive white Gaussian noise. In this massive MIMO setting, linear precoding strategies approach the performance of the more computationally complex non-linear schemes, hence making their use convenient [2], [5]. For this reason, we decompose the transmit signal as

$$\mathbf{x} = \nu \cdot \mathbf{Fs} = \frac{\mathbf{Fs}}{\sqrt{\mathbb{E}\left\{\text{Tr}\left[\mathbf{F}^H \mathbf{F}\right]\right\}}}, \tag{4}$$

where $\mathbf{F} \in \mathbb{C}^{N \times K}$ represents an arbitrary linear precoding matrix, $\mathbf{s} \in \mathbb{C}^{K \times 1}$ denotes the user constellation symbols to be conveyed, and $\nu = 1/\mathbb{E}\left\{\sqrt{\text{Tr}\left[\mathbf{F}^H \mathbf{F}\right]}\right\}$ [5]. Here, $\text{Tr}(\cdot)$ is the trace of a matrix and the normalization constant ν guarantees that the average transmission power is constrained to $P_t = \mathbb{E}\left\{\mathbf{x}^H \mathbf{x}\right\} = 1$. In the previous expressions, $\mathbb{E}\{\cdot\}$ stands for the expectation operator. To maintain the focus of the paper on the proposed concept, in the following we concentrate on zero-forcing (ZF) precoding, although it is clear that the benefits of the proposed strategy also apply to other precoding schemes. The signal transmitted by the ZF precoder is given by

$$\mathbf{x} = \nu \cdot \left(\widetilde{\mathbf{H}}^H\right)^\dagger \mathbf{s} = \frac{\widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}\right)^{-1} \mathbf{s}}{\sqrt{\mathbb{E}\left\{\text{Tr}\left[\left(\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}\right)^{-1}\right]\right\}}}, \tag{5}$$

where $\widetilde{\mathbf{H}}^H$ represents the estimate of the downlink channel available at the BS and $(\cdot)^\dagger$ denotes the matrix pseudoinverse.

B. Physically-Constrained Channel Model for 2D Antenna Arrays

In this section we describe the physical channel model that will be used in the following to characterize the particular properties of size-constrained massive MIMO BSs [6], [8], [30], [46], [48]. It is anticipated that future massive MIMO transceivers will be deployed within the limited spaces of today's BSs. Accordingly, the MIMO arrays will experience transmit correlation and mutual coupling that will be dependent on the increasing number of antenna elements. In particular, since large-scale BSs are expected to be implemented in 2D and 3D antenna structures that allow packing more elements in a fixed physical space, we concentrate on describing the physical channel model of a planar antenna array [2]. In this study we neglect the effects of mutual coupling since they can be compensated by impedance matching techniques [49]. Therefore, the propagation channel from the k-th user to the BS can be expressed as [46], [48]

$$\mathbf{h}_k = \mathbf{A}_k^H \mathbf{g}_k, \tag{6}$$

where $\mathbf{g}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{D_k})$, D_k corresponds to the number of directions in which the angular domain is divided, and $\mathbf{A}_k \in \mathbb{C}^{D_k \times N}$ is the steering matrix containing the D_k transmit steering vectors of the antenna array. The transmit steering matrix of the k-th user is given by [6], [48]

$$\mathbf{A}_k = \frac{1}{\sqrt{D_k}} \left[\mathbf{a}^T(\phi_{k,1}, \theta_{k,1}), \dots, \mathbf{a}^T(\phi_{k,D_k}, \theta_{k,D_k})\right]^T. \tag{7}$$

Here, the directions of departure (DoD) of the k-th user are characterized by $\phi_{k,n} \in [\phi_k^{\min}, \phi_k^{\max}]$ and $\theta_{k,n} \in [\theta_k^{\min}, \theta_k^{\max}]$, which specify the azimuth and elevation angles of departure respectively. The transmit steering vectors of an arbitrary rectangular array can be expressed as [50]

$$\mathbf{a}(\phi_{k,n}, \theta_{k,n}) = \left[1, e^{j2\pi[d_h \sin(\theta_{k,n}) \sin(\phi_{k,n})]}, \dots, \right. \tag{8}$$
$$\left. e^{j2\pi[(M_h-1)d_h \sin(\theta_{k,n}) \sin(\phi_{k,n}) + (M_v-1)d_v \sin(\theta_{k,n}) \cos(\phi_{k,n})]}\right].$$

Here, d_h and d_v denote the inter-antenna spacing in the horizontal and vertical axes normalized by the carrier wavelength λ, and M_h and M_v represent the number of antennas allocated in the horizontal and vertical axes respectively. Clearly, $d_h = d_v$ for the particular case of uniform rectangular arrays (URAs). The total number of antennas is therefore given by $N = M_h \times M_v$. Moreover, to provide a fair comparison, in this work we consider that the horizontal and vertical inter-antenna distances at the BS are given by

$$d_{\{h,v\}} = \frac{L_{\{h,v\}}}{\left(M_{\{h,v\}} - 1\right)}, \tag{9}$$

where $L_{\{h,v\}}$ denotes the fixed horizontal and vertical lengths of the antenna array. The total surface in which the antennas are deployed is denoted as $L = L_h \times L_v$. Note that the above model assumes that there is no spatial correlation between users, a reasonable consideration since the inter-user distance is usually larger than λ [6].

Remark 1: Since the physical space dimensions of the antenna array are predetermined, the above entails that the inter-antenna correlation depends on the total number of antennas installed. This model allows us to accurately capture the inter-antenna correlation that arises due to insufficient antenna separation, which constitutes our focus in the following.

Remark 2: Linear antenna arrays can be regarded as a particular case of the described rectangular array by letting $M_v = 1$ and setting $\theta_{k,n} = \pi/2 \, \forall k \in [1, K], n \in [1, D_k]$ [6].

III. RELAXING THE CSI ACQUISITION: INCOMPLETE CSI

The dense deployment of antenna arrays leads to an increased correlation between the antenna elements, which also translates to a larger similarity between the communication channels of closely spaced antennas [2], [6]. This circumstance motivates us to consider the practicability of collecting the CSI for a subset of antennas during the CSI acquisition stage and analyzing its effects in the operation of the communication system. Formally, let the set that indexes all the antenna elements allocated at the BS be defined as
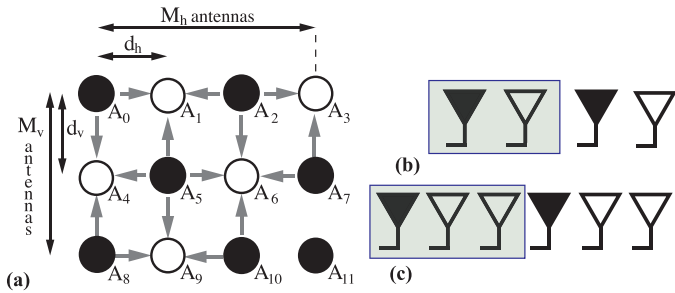
$$\mathbb{N} = \{1, \dots, N\}. \tag{10}$$

Fig. 1. (a) Example of CSI distribution for $M_h = 4$, $M_v = 3$ and $N_c = 7$. (b), (c) Basic CSI distribution patterns. Black and white elements represent antennas with and without acquired CSI respectively.

Each entry of the set indexes the n-th antenna of the planar array A_n as shown in Fig. 1(a). Furthermore, let us define the sets $\mathcal{B} \subset \mathcal{N}$ and $\mathcal{C} \subset \mathcal{N}$ as two subsets of $\mathcal{N}$ satisfying

$$\mathcal{B} \cap \mathcal{C} = \varnothing, \mathcal{B} \cup \mathcal{C} = \mathcal{N}. \tag{11}$$

In the previous expressions, $\subset$, $\cup$ and $\cap$ denote the subset, union and intersection of sets respectively. Specifically, the subset $\mathcal{B}$ indexes the antennas whose CSI has been acquired during the pilot training stage, and $\mathcal{C}$ determines the antennas without instantaneous CSI availability. Subsequently, $|\mathcal{B}| = N_c \leq N$ and $|\mathcal{C}| = N - N_c$ determine the number of antennas with and without instantaneous CSI knowledge respectively. Here $|\cdot|$ denotes the cardinality of a set.

The proposed scheme aims at limiting the collection of instantaneous CSI to a number of antennas $N_c \leq N$. Subsequently, the large inter-antenna correlation that arises in tightly packed antenna arrays is exploited to determine the CSI for the rest of antennas. In other words, only the antennas indexed by the subset $\mathcal{B}$ will be active during pilot signal reception, and the averages of the information acquired will be subsequently used as CSI for the remaining antennas for the purposes of precoding and detection. This brings the following critical benefits that will be studied in the following:

- The signal processing load of the channel estimation and signal transmission processes can be dramatically reduced. This can be attained by limiting the number of antennas simultaneously active during pilot reception and by performing an insightful design of the transmit signals.
- The number of receive RF chains activated throughout the pilot training and downlink transmission stages is constrained to N_c, hence reducing the circuit power consumption of the related circuitry.

At this point we remark that the incomplete acquisition process affects the decision metric obtained after pilot correlation in (2), which for $\eta_{tr} = K$ can be expressed as

$$\widehat{\mathbf{P}}\Big|_{[\mathcal{B},\mathcal{K}]} = \sqrt{\rho_{tr}} \cdot \mathbf{H}\Big|_{[\mathcal{B},\mathcal{K}]} + \mathbf{W}\Big|_{[\mathcal{B},\mathcal{K}]}, \tag{12}$$

where $\mathbf{C}\Big|_{[\mathcal{T}_1,\mathcal{T}_2]}$ represents the sub-matrix obtained by selecting the rows given by $\mathcal{T}_1$ and the columns specified by $\mathcal{T}_2$. In this notation, the sets $\mathcal{K} = \{1, \ldots, K\}$ and $\mathcal{N} = \{1, \ldots, N\}$ are commonly employed to select all the columns or rows depending on the context. Similarly, we use $\mathbf{c}\Big|_{[\mathcal{T}]}$ to denote

the sub-vector obtained by selecting the entries determined by $\mathcal{T}$. In the following we also let $c_{k,n}$ denote the n-th entry of the vector $\mathbf{c}_k$ for ease of notation.

After the incomplete channel estimation process, the channel of the k-th user can be expressed as

$$\widetilde{\mathbf{h}}_k\Big|_{[\mathcal{B}]} = \widehat{\mathbf{h}}_k, \tag{13}$$

$$\widetilde{\mathbf{h}}_k\Big|_{[\mathcal{C}]} = f\left(\widetilde{\mathbf{h}}_k\Big|_{[\mathcal{B}]}\right), \tag{14}$$

where $\widehat{\mathbf{h}}_k \in \mathbb{C}^{N_c \times 1}$ denotes the uplink channel of the k-th user for the antennas with CSI acquired during the training stage (subject to CSI acquisition errors), and $f(\cdot) : \mathbb{C}^{N_c} \rightarrow \mathbb{C}^{(N-N_c)}$ is a linear function whose definition will be considered in the following.

A. Distribution of the CSI on the Antenna Array

The definition of $\mathcal{B}$, $\mathcal{C}$ and $f(\cdot)$ impacts the performance and complexity of the proposed scheme. Ideally, an optimization problem could be formulated for determining the antennas that collect instantaneous CSI during the pilot training stage. The objective of this problem would be minimizing the effect of the incomplete CSI acquisition on the accuracy of the resulting channel estimate. This however entails significant complexity and is therefore out of the scope of our study, which focuses on a low-complexity CSI distribution and averaging approach. For this reason, in this paper we propose a CSI distribution strategy that leverages on the use of predefined CSI distribution patterns. The considered scheme aims at obtaining a low-complexity solution by exploiting the following intuition: the larger the correlation between the channels without instantaneous CSI and the channels employed to derive their CSI, the smaller the CSI error introduced. In other words, for any antenna deactivated during the training stage, there must be others with their instantaneous CSI available in its proximity. Specifically, the proposed CSI distribution procedure is based on the combination of the two basic CSI distribution patterns shown in Fig. 1(b) and (c) with colored subsets. In these figures, the black and white antennas represent the antennas with and without instantaneous CSI acquired during the training stage respectively. These patterns allow us to efficiently distribute the CSI for the cases in which $(N_c/N) \geq (1/3)$ that are explored in the following[1].

The pseudocode of the proposed CSI distribution procedure is shown in Algorithm 1, where $\varnothing$ represents the empty set, $(\cdot)^C$ denotes the complementary of a set, and $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling functions, respectively. An illustrative example of the application of this strategy is shown in Fig. 1(a) for the case of $M_h = 4$, $M_v = 3$ and $N_c = 7$. The application of Algorithm 1 generalizes the CSI distribution process to more complicated cases. Overall, this representative example shows that the proposed CSI distribution aims at maximizing the number of adjacent antennas with acquired CSI for each antenna

[1]Note that this is adequate since it will be shown that $(N_c/N) \geq (1/3)$ is required to obtain a satisfactory performance in realistic massive MIMO systems. This is because the double polarized antennas to be employed in future deployments occupy a physical space that might impede placing them arbitrarily close.

Algorithm 1. Pseudocode of the CSI distribution algorithm

Inputs: N, N_c, M_h, M_v

1: **Outputs:** $\mathcal{B}$, $\mathcal{C}$

2: $\mathcal{B}$, $\mathcal{C}$ ← ∅　　　　　　　　　　　　　　{Initialization}

　{3 − 4: Maximum and minimum number of antennas with instantaneous CSI per row}

3: $N_{\max}$ ← $\lceil N_c/M_v \rceil$

4: $N_{\min}$ ← $\lfloor N_c/M_v \rfloor$

　{5 − 6: Number of rows with the maximum and minimum number of antennas with CSI per row}

5: $M_{\max}$ ← $(N_c - M_v N_{\min})$

6: $M_{\min}$ ← $M_v - M_{\max}$

7: {$\mathcal{B}_{\min}$, $\mathcal{B}_{\max}$ ← Create reference CSI distribution patterns per row with $N_{\min}$ ($\mathcal{B}_{\min}$) and $N_{\max}$ ($\mathcal{B}_{\max}$) antennas with CSI by combining the basic patterns shown in Fig. 1(b) and (c) and adding additional antennas with CSI where required}

8: **for** $j = 0 \rightarrow (M_v - 1)$ **do**

9: 　**if** $j < M_{\min}$ **then**

10: 　　$\mathcal{B}$ ← $\mathcal{B} \cup \mathcal{B}_{\min}$

11: 　**else**

12: 　　$\mathcal{B}$ ← $\mathcal{B} \cup \mathcal{B}_{\max}$

13: 　**end if**

14: 　$\mathcal{B}_{\min}$ ← {Circular shift of $\mathcal{B}_{\min}$ to determine the antennas with CSI in the j-th row for $j < M_{\min}$}

15: 　$\mathcal{B}_{\max}$ ← {Circular shift of $\mathcal{B}_{\max}$ to determine the antennas with CSI in the j-th row for $j \geq M_{\min}$}

16: **end for**

17: $\mathcal{C}$ ← $\mathcal{B}^C$

---

without this information to efficiently exploit the inter-antenna correlation.

The operation of the proposed CSI distribution algorithm can be described as follows: Initially, the number of antennas per row with instantaneous CSI is determined by assigning a similar number of antennas with CSI per row to evenly distribute the CSI. For instance, Fig. 1(a) shows that the first $M_{\min} = 2$ rows of antennas have $N_{\min} = 2$ antennas with CSI, whereas the last one has $N_{\max} = 3$ antennas with CSI. After this, the antennas with CSI are distributed for each row by combining the antenna patterns depicted in Fig. 1(b) and (c), and adding antennas with CSI to these to ensure that the total number of antennas with CSI per dimension is the one previously specified. For instance, Fig. 1(a) shows that the CSI in the first row of antennas is distributed by consecutively combining two of the antenna patterns described in Fig. 1(b), whereas for the CSI distribution in the third row an additional antenna with CSI has been added. We note that the additional antennas with CSI are solely required for the cases $(N_c/N) > 1/2$, and that their specific positions have shown to have a negligible performance impact. Subsequently, we shift circularly the patterns to determine the distribution of the CSI in the following rows. This is illustrated in Fig. 1(a), where it can be seen that the CSI distribution of the second row is obtained by shifting the distribution of the previous one. This procedure ensures that the CSI of the antennas is evenly distributed, thus reducing the distance of the antennas with and without instantaneous CSI.

Remark 3: The employment of the basic CSI distribution patterns does not restrict the proposed scheme to a solution with fixed N_c. Instead, they allow us to analyze the complexity-performance trade-off that arises for varying N_c, since different system requirements might motivate the employment of distinct solutions as studied in the following.

After defining the sets $\mathcal{B}$ and $\mathcal{C}$, the remaining step consists on determining the CSI of the antennas that were inactive during the training stage. Let $\mathcal{T}_n$ denote the n-th entry of a set. Then, the CSI of the antennas without instantaneous CSI is obtained by averaging the CSI of the closest antennas with this information. Subsequently, the $\mathcal{C}_n$-th entry of the channel vector of the k-th user can be expressed as

$$\widetilde{\mathbf{h}}_{k,\mathcal{C}_n} = \frac{1}{M_{\mathcal{C}_n}} \sum_{i=1}^{M_{\mathcal{C}_n}} \widehat{\mathbf{h}}_{k,\mathcal{B}_i^{\mathcal{C}_n}}. \tag{15}$$

Here, $M_{\mathcal{C}_n}$ denotes the number of neighbour antennas with indexes $\mathcal{B}^{\mathcal{C}_n} \subset \mathcal{B}$ used to average the CSI for the $\mathcal{C}_n$-th antenna. In other words, the CSI of a given antenna without instantaneous CSI knowledge is obtained by averaging the CSI of the antennas belonging to $\mathcal{B}^{\mathcal{C}_n}$. The global estimated communication channel with incomplete CSI after the averaging operations can be therefore expressed as

$$\widetilde{\mathbf{H}} = \mathbf{R}^H \widehat{\mathbf{H}}, \tag{16}$$

where $\widehat{\mathbf{H}} \in \mathbb{C}^{N \times K}$ is the estimated channel matrix formed by the entries that correspond to the antennas active during the training stage, which are given by $\widehat{\mathbf{h}}_k \big|_{[\mathcal{B}]}$, and zeros elsewhere. Moreover, $\mathbf{R} \in \mathbb{R}^{N \times N}$ describes how the available CSI is combined to obtain the CSI for the rest of the antennas. This matrix is referred to as the CSI averaging matrix and its n-th column is given by

$$\mathbf{R} \big|_{[\mathcal{N},n]} = \begin{cases} \mathbf{I} \big|_{[\mathcal{N},n]}, & \forall n \in \mathcal{B}, \\ \frac{1}{M_n} \mathbf{r}_{\mathcal{B}^n}, & \forall n \in \mathcal{C}, \end{cases} \tag{17}$$

where $\mathbf{I} \big|_{[\mathcal{N},n]} \in \mathbb{Z}^{N \times 1}$ denotes the n-th column of an $N \times N$ identity matrix, and $\mathbf{r}_{\mathcal{B}^n} \in \mathbb{Z}^{N \times 1}$ is a vector with all its entries set to zero except for those found in the positions given by the set $\mathcal{B}^n$. Similarly to (15), M_n corresponds to the number $|\mathcal{B}^n|$ of antennas employed for interpolating the CSI of the n-th antenna. Note that $\mathbf{R} = \mathbf{I}_N$ when complete CSI is acquired.

Remark 4: Although more intricate strategies can be implemented, in this paper we focus on the simple averaging operation for reasons of illustration and for not detracting the attention from the basic principle introduced in this work. Moreover, this operation is especially practical due to both the hardware and the signal processing advantages detailed hereafter. Future work will analyze the consequences of using different strategies on the resulting performance.

B. Implications of the Acquisition of Incomplete CSI

The averaging of the CSI has several implications on a range of communication aspects such as the signal processing load or the hardware complexity that are considered in this section. In particular, we concentrate on the training and downlink stages, which constitute the main focus of this paper.
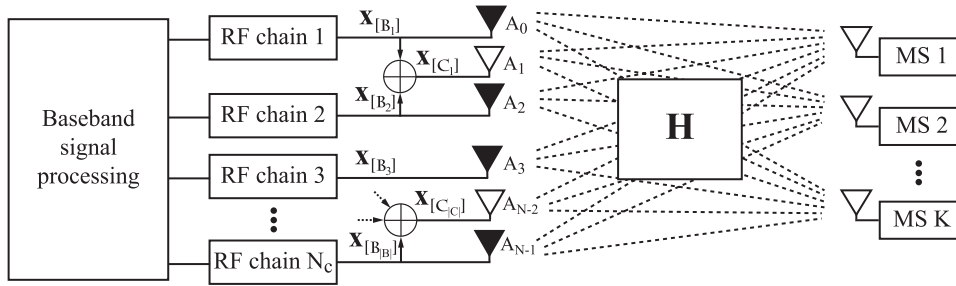
Fig. 2. Block diagram of the proposed transmission scheme. Black and white antennas represent elements with and without acquired CSI respectively.

1) Dimensionality Reduction of the Received Training Signals: Inherently, the dimensions of the received training signals are reduced when incomplete CSI is acquired. An immediate consequence of this is that only $N_c \leq N$ RF chains remain active during the pilot training stage, which allows us to reduce the total power consumption of standard channel estimation approaches [42]. Simultaneously, the computational complexity of the CSI estimation process is reduced due to the smaller dimensions of the acquired pilot signals. Particularly, apart from a decrease in the complexity of the conventional pilot correlation process shown in (2), the dimensionality reduction can be particularly beneficial when complex channel estimation strategies are employed due to the polynomial dependence of their complexity on N_c [20], [30]. This comes at the cost of introducing an error in the channel estimate, which clearly depends on the inter-antenna distance, the quality of the acquired CSI, and the number of antennas with instantaneous CSI as studied in Sec. IV.

2) Data Transmission Stage: The differences in the estimated channel with full and incomplete CSI have a direct effect on the achievable rates of the system that will be explored in Sec. VII. However, the incomplete acquisition of CSI also allows us to derive an interesting relationship between the signals transmitted from the different antennas, which can be re-expressed as

$$\mathbf{x} = \nu \cdot \mathbf{F}\mathbf{s} \Longleftrightarrow \begin{cases} \mathbf{x}\big|_{[\mathcal{B}]} = \dfrac{\mathbf{F}\big|_{[\mathcal{B},\mathcal{K}]} \cdot \mathbf{s}}{\mathbb{E}\left\{\sqrt{\mathrm{Tr}\left[\mathbf{F}^H\mathbf{F}\right]}\right\}}, \\ \mathbf{x}\big|_{[\mathcal{C}]} = \left(\mathbf{R}\big|_{[\mathcal{B},\mathcal{C}]}\right)^H \cdot \mathbf{x}\big|_{[\mathcal{B}]}. \end{cases} \quad (18)$$

for the conventional matched filter (MF), ZF and regularized ZF precoding schemes. The above expression entails that the signals of the antennas that were deactivated during the training stage can be obtained by combining the signals generated for the antennas whose CSI was acquired. This has a direct consequence on the signal processing load per transmitted symbol since we can exploit that, similarly to (15), the signal generation for the set $\mathcal{C}$ can be done in the analog domain by signal splitting and combining to avoid the computation of $\mathbf{x}\big|_{[\mathcal{C}]}$ in the baseband processing stage. This dispenses with the need of $N - N_c$ RF chains as shown in Fig. 2 w.r.t. conventional designs, hence dramatically reducing the circuit power consumption accordingly.

IV. CSI ERROR UNDER INCOMPLETE CSI

First, let us clarify that in the following we refer to imperfect CSI when solely describing CSI acquisition errors, whereas the term incomplete CSI is reserved for the cases in which the CSI of a subset of antennas is not acquired. The acquisition of incomplete and imperfect CSI impacts on the quality of the channel estimated and the resulting performance of precoding and detection schemes. For this reason, in this section we concentrate on characterizing the error introduced by the proposed strategy in the channel estimation process. The channel estimate generated after incomplete CSI acquisition for the k-th user (16) can be re-expressed as [5]

$$\widetilde{\mathbf{h}}_k = \mathbf{h}_k - \overline{\mathbf{h}}_k, \quad (19)$$

where $\overline{\mathbf{h}}_k \in \mathbb{C}^{N \times 1}$ represents the error introduced by the acquisition of incomplete and imperfect CSI. The imperfect estimate of the communication channel for the antennas with CSI can be characterized as [4], [48]

$$\widehat{\mathbf{h}}_k\Big|_{[\mathcal{B}]} = \left(\mathbf{A}_k^H \left(\sqrt{1 - \tau_k^2}\mathbf{g}_k + \tau_k\mathbf{q}_k\right)\right)\Big|_{[\mathcal{B}]}, \quad (20)$$

where $\mathbf{q}_k \sim \mathcal{CN}(0, \mathbf{I}_{D_k})$ is uncorrelated with $\mathbf{g}_k$ [4], [48] and τ_k is a variable that models the quality of the acquired CSI. The stochastic imperfect CSI model employed in (20) facilitates the derivation of intuitive expressions for characterizing the behaviour of the proposed scheme. In particular, $\tau_k = 0$ indicates that perfect CSI is collected, whereas $\tau_k = 1$ corresponds to the case where only statistical information is available at the transmitter. Moreover, the error of the channel after the averaging operation for a relevant antenna without CSI is clearly conditioned by its position on the antenna array, the inter-element spacing, and the number of antennas selected for averaging. In the following we model the impact of the incomplete and imperfect CSI acquisition on the accuracy of the channel estimation by using the channel error factor [51]

$$\Delta_k = \frac{\mathbb{E}\left\{\left\|\overline{\mathbf{h}}_k\right\|_2^2\right\}}{\mathbb{E}\left\{\|\mathbf{h}_k\|_2^2\right\}} = \frac{\mathbb{E}\left\{\left\|\mathbf{h}_k - \widetilde{\mathbf{h}}_k\right\|_2^2\right\}}{\mathbb{E}\left\{\|\mathbf{h}_k\|_2^2\right\}}. \quad (21)$$

Specifically, since the focus of this paper is placed on physically-constrained systems, the following theorem details the behaviour of the above metric for the channel model introduced in Sec. II-B.

Theorem 1: Let the communication channel be described as $\mathbf{h}_k = \mathbf{A}_k^H \mathbf{g}_k$ with $\mathbf{g}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{D_k})$ for a given $\mathbf{A}_k$. Also let $\widehat{\mathbf{h}}_k \big|_{[\mathcal{B}]} \in \mathbb{C}^{N_c \times 1}$ be the noisy channel estimate of the channel modeled by following (20). Moreover, consider the channel of the antennas without CSI to be computed by averaging the imperfect CSI from the nearby antennas as described in Sec. III. Then, the channel error factor is given by

$$\Delta_k = \frac{1}{N} \sum_{n \in \mathcal{B}} \tau_k^2 \mathbf{\Theta}_k \big|_{[n,n]} + \frac{1}{N} \sum_{n \in \mathcal{C}} \left[\mathbf{\Theta}_k \big|_{[n,n]} \right.$$
$$+ \frac{1}{(M_n)^2} \sum_{i \in \mathcal{B}^n} \mathbf{\Theta}_k \big|_{[i,i]} - \frac{2\sqrt{1-\tau_k^2}}{M_n} \mathrm{Re}\left(\sum_{i \in \mathcal{B}^n} \mathbf{\Theta}_k \big|_{[n,i]} \right)$$
$$\left. + \frac{2}{(M_n)^2} \mathrm{Re}\left(\sum_{i,j \in \mathcal{B}^n, i > j} \mathbf{\Theta}_k \big|_{[j,i]} \right) \right], \qquad (22)$$

where $\mathbf{\Theta}_k$ is the correlation matrix of the channel of the k-th user given by $\mathbf{\Theta}_k = \mathbb{E}\left\{ \mathbf{h}_k \mathbf{h}_k^H \right\} = \mathbf{A}_k^H \mathbf{A}_k$. Moreover, $\mathcal{B}^n$ and M_n represent the indices and total number of antennas used for computing the CSI of the n-th antenna as defined in Sec. III respectively.

Proof: The proof of Theorem 1 is given in Appendix A. ∎

The above theorem allows us to determine the impact of applying the proposed strategy based on the entries of the channel correlation matrix of the true channel, the number of antennas with and without instantaneous CSI, the number of antennas employed for the averaging operation, and the quality of the acquired CSI. In the following we provide a set of relevant corollaries based on the results of Theorem 1.

Corollary 1: Dependence of Δ_k on τ_k. The channel error factor Δ_k is an increasing function of the imperfect CSI parameter τ_k for transmitters in which the channels of the neighbour antennas are strongly correlated, i.e., those channels satisfying $\mathrm{Re}\left(\sum_{i \in \mathcal{B}^n} \mathbf{\Theta}_k \big|_{[n,i]} \right) > 0$.

Proof: The proof follows from the analysis of the terms in (22) influenced by the factor τ_k

$$t_1 = \frac{1}{N} \sum_{n \in \mathcal{B}} \tau_k^2 \mathbf{\Theta}_k \big|_{[n,n]}, \, n \in \mathcal{B}, \qquad (23)$$

and

$$t_2 = -\frac{2\sqrt{1-\tau_k^2}}{M_n} \mathrm{Re}\left(\sum_{i \in \mathcal{B}^n} \mathbf{\Theta}_k \big|_{[n,i]} \right), \, n \in \mathcal{C}. \qquad (24)$$

Clearly, both terms are increasing functions of τ_k provided that $\mathrm{Re}\left(\sum_{i \in \mathcal{B}^n} \mathbf{\Theta}_k \big|_{[n,i]} \right) > 0$, which holds true when the channels of the adjacent antennas are severely correlated. ∎

Note that the above corollary corroborates the intuition that the collection of imperfect CSI during the training stage influences the quality of the channel approximation generated for the antennas deactivated throughout this process. This can be explained by the dependence of t_2 on τ_k.

Additionally, it is also intuitive that the channel error factor Δ_k is an increasing function of the inter-antenna distance d for
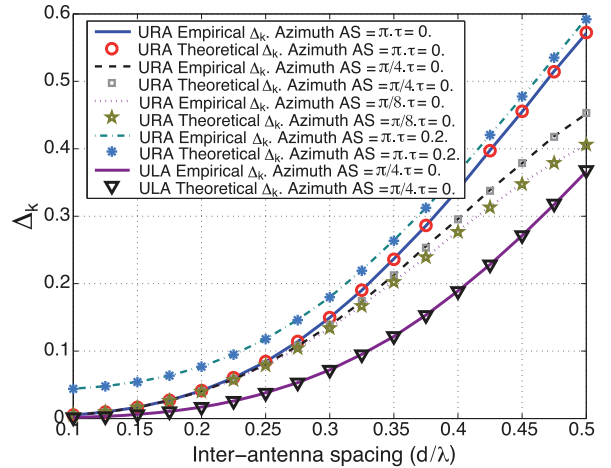

Fig. 3. Δ_k vs. inter-antenna spacing for azimuth sector angles of π, $\pi/4$ and $\pi/8$ radians (URA) and $\pi/4$ (ULA). $N = 144$, $N_c = N/2$ and $K = 16$.
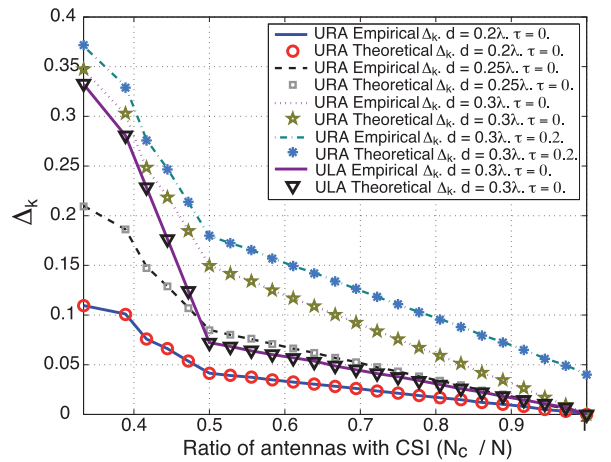

Fig. 4. Δ_k vs. ratio of antennas with CSI (N_c/N) for different inter-antenna spacings d. $N = 144$ and $K = 16$.

the ranges typically considered, since the inter-antenna correlation depends on this parameter. However, explicitly showing this for the physical channel model described in Sec. II-B and a general case is infeasible due to the complexity of the expressions involved. For this reason, the following corollary explores the particular case of a uniform linear array (ULA) with reduced angle spreads, which can arise naturally as manifested by the one-ring correlation model [52].

Corollary 2: Dependence of Δ_k on d. Let the antennas at the BS be placed in a ULA configuration. Moreover, consider that perfect CSI is acquired for the active antennas during the training stage, $N/N_c \geq 2$ and small angle spreads such that $\sin(\phi) \approx \phi$ and $\cos(\phi) \approx (1 - \phi^2/2)$. Then, the channel error factor Δ_k is a monotonically increasing function of the inter-antenna spacing d.

Proof: The proof is given in Appendix B. ∎

The above observation can be explicitly seen in Fig. 3, which represents the variation of the theoretical and empirical metric Δ_k in (21) for increasing values of inter-antenna spacing when the channel model of Sec. II-B is employed. The theoretical

results in Figs. 3 and 4 have been obtained by direct application of (22). We have considered a transmitter with either $M_h = M_v = 12$ (URA) or $M_v = 1$ and $M_h = 144$ (ULA) antennas, $N_c = N/2$, and $K = 16$ users. The number of angular directions is set to $D_k = 50$ in (7) [46], and the elevation angles of departure are equidistributed within sector angles of $\pi/3$ radians. We note that, while other arbitrary choices for D_k could be adopted, the modification of this parameter does not modify the trends and conclusions derived in the following. The different azimuth angle spreads, π, $\pi/4$ and $\pi/8$ radians, characterize the effects of having a limited angle spread on the proposed scheme. In this figure and throughout this paper the number of antennas whose CSI is averaged are selected depending on the number of antennas with CSI as follows. For simplicity, we also consider $\tau_k = \tau$, $\forall k \in [1, K]$. On the one hand, for the cases in which $(N_c/N) \geq (1/2)$, only the CSI of the vertically and horizontally adjacent antennas with CSI is averaged. On the other hand, when the number of antennas with CSI is further reduced so that $(N_c/N) < (1/2)$, the CSI of the diagonally adjacent antennas is used to guarantee that the CSI of a sufficient number of antennas is averaged. Overall, it can be seen that the error of the channel generated after the averaging operation decreases as the antenna inter-spacing is reduced. This can be explained by the larger correlation experienced between neighbour antennas, which also increases when the angle spread is constrained.

The evolution of Δ_k for a varying number of antennas with CSI is shown in Fig. 4 for the same setup of Fig. 3. We consider an azimuth angle spread of π radians for the URA BS and $\pi/4$ for the ULA BS, and different inter-antenna spacings d. From the results of this figure it can be concluded that small variations in the number of antennas with CSI strongly influence the error for $N_c/N < (1/2)$, whereas this impact is reduced for $N_c/N \geq (1/2)$. Intuitively, the reason for this behaviour is related to the number of antennas whose CSI is averaged for a given antenna without CSI, and it is especially pronounced for the ULA BS. Furthermore, it can be seen that the theoretical derivation closely resembles the empirical results in both Fig. 3 and Fig. 4 with perfect ($\tau_k = 0$) and imperfect CSI ($\tau_k = 0.2$).

V. COMPLEXITY ANALYSIS

In this section we study the improvements offered by the proposed strategy regarding the signal processing load. Specifically, we show that the smaller number of antennas simultaneously active during the CSI acquisition procedure reduces the number of required signal processing operations for channel estimation and precoding. In this section we follow the analysis of [47] for reasons of reference.

The signal processing operations performed in a TDD communication system can be divided into three different stages, namely CSI acquisition stage, downlink signal transmission, and uplink signal reception [43]–[45], [47]. These stages correspond to the communication phases in which the channel coherence time is divided and their respective lengths determine the global complexity as shown in the following.

A. CSI Acquisition Stage

Throughout this stage, the BS receives the pilot signals transmitted by the MSs and uses these to estimate the communication channel [42]–[45], [47]. For simplicity, we consider that the coherence time of the long-term statistics is significantly larger than the channel coherence time. The above entails that a significant portion of the operations to perform the channel estimation after the pilot correlation process remain valid for a large number of coherence blocks. This makes the pilot correlation process dominate the complexity per coherence block [20]. The total number of real floating-point operations (flops) of this procedure as performed in (2), C^{tr}, depends on the training length in number of symbols η_{tr} and is given by [47], [53], [54]

$$C^{tr} = 8N_c K \eta_{tr} \overset{(a)}{=} 8N_c K^2, \tag{25}$$

where (a) holds because it is considered that orthogonality between the pilot training signals of the system users is preserved by letting $\eta_{tr} = K$ [47]. In this expression and throughout this paper, the common assumption that real multiplications (divisions) have the same complexity of real additions (subtractions) has been followed [54]. Moreover, we also consider that complex additions and multiplications involve two and six real floating-point {operations respectively [53], [54].

B. Data Transmission (Downlink) Stage

The complexity of the downlink stage is determined by the operations required to generate the precoding matrix at the beginning of each coherence period, and the operations to generate the transmitted signal in every channel use [47]. For the sake of brevity, in the following we consider that the ZF precoding matrix is computed via the conventional singular value decomposition (SVD) approach. Other strategies such as the Cholesky decomposition have been shown to offer similar complexity results for the scenarios considered in this paper [51]. In particular, let $\widetilde{\mathbf{H}}^H = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$ be the SVD of the estimated downlink channel matrix. Here, $\mathbf{U} \in \mathbb{C}^{K \times K}$ and $\mathbf{V} \in \mathbb{C}^{N \times N}$ are unitary matrices comprised of the left and right singular vectors of $\widetilde{\mathbf{H}}^H$ respectively, and $\boldsymbol{\Sigma} \in \mathbb{C}^{K \times N}$ is a diagonal matrix containing the ordered singular values of $\widetilde{\mathbf{H}}^H$. The pseudo-inverse precoding matrix can be subsequently obtained from the SVD as $\left(\widetilde{\mathbf{H}}^H\right)^\dagger = \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^H$. The complexity of this process is given by

$$C_{inv}^d = C_{SVD}^d + C_{div}^d + C_{mult}^d = \left(24K^3 + 16K^2N\right)$$
$$+ K + \left(2KN + 8K^2N\right), \tag{26}$$

where C_{SVD}^d characterizes the complexity of performing a complex SVD, C_{div}^d refers to the complexity of dividing the K real diagonal elements of $\boldsymbol{\Sigma}$, and C_{mult}^d accounts for the operations required to multiply a complex full matrix with a real diagonal matrix and two full complex matrices [54]. Note that this computationally intensive operation only has to be performed at the beginning of each channel coherence period, since the result stays valid as long as the channel remains approximately constant.

In conventional MIMO systems, the pseudo-inverse matrix must be multiplied with the symbols to be transmitted to generate the final precoded signal $\mathbf{x} \in \mathbb{C}^{N \times 1}$ [47]. However, the proposed strategy obtains the signal output of $N - N_c$ antennas by splitting and combining the signal components of $\mathbf{x}\big|_{[\mathcal{B}]} \in \mathbb{C}^{N_c \times 1}$ in the analog domain, hence reducing the dimensions of the required matrix-vector multiplication[2]. Based on this observation, the complexity of computing the precoding signal can be expressed as

$$C_d^d = 8 N_c K \eta_{\text{dl}}, \tag{27}$$

Here, η_{dl} denotes the number of symbol periods assigned to downlink transmission. The complexity benefits offered by the proposed in the precoding stage can be clearly observed in (27), since the resulting signal processing load is a linear function of the number of antennas with acquired CSI, N_c. Combining (26) and (27), the global complexity of the precoding process during a channel coherence period can be expressed as

$$C^d = C_{\text{inv}}^d + C_d^d \approx 24 K^3 + 24 K^2 N + 8 N_c K \eta_{\text{dl}}. \tag{28}$$

C. Data Reception (Uplink) Stage

The operations to be performed during the data reception stage mirror the ones to be performed during data transmission. In the case of TDD systems, the detection matrix is readily obtained from the precoding one by exploiting channel reciprocity and the relationship $\left(\widetilde{\mathbf{H}}^H\right)^\dagger = \left(\widetilde{\mathbf{H}}^\dagger\right)^H$. Therefore, the complexity of the uplink stage can be computed as

$$C^u = 8 N K \eta_{\text{ul}}, \tag{29}$$

where η_{ul} stands for the number of symbol periods assigned to the uplink stage.

D. Total Complexity

The total complexity is given by

$$C_{\text{tot}} \simeq C^u + C^d + C^{\text{tr}}, \tag{30}$$

where, in general, $C^u \ll C^d$ since more time resources are conventionally allocated to the downlink transmission, i.e., $\eta_{\text{ul}} \ll \eta_{\text{dl}}$ [56]. This entails that the operations performed during the downlink stage dominate the global complexity, which maximizes the benefits of the proposed due to the reduced complexity required for precoding as summarized in Table I. Clearly, the complexity with full CSI corresponds to $N_c = N$ in all expressions above.

VI. ACHIEVABLE RATES AND ENERGY EFFICIENCY

A. Downlink Achievable Rates With Incomplete CSI

In this section we concentrate on the effects that the use of incomplete CSI has on the spectral and energy efficiency of the communication system. With this purpose and following [4], let

[2]The computational load of the signal generation process can also be reduced when each antenna has its dedicated RF chain [55].

TABLE I
COMPLEXITY IN NUMBER OF REAL FLOATING-POINT OPERATIONS (FLOPS) OF A BASE STATION WITH INCOMPLETE CSI ACQUISITION

Communication phase	Complexity in flops
CSI acquisition stage	$C^{\text{tr}} = 8 N_c K \eta_{\text{tr}}$
Downlink stage • Signal generation • Precoding matrix	• $C_d^d = 8 N_c K \eta_{\text{dl}}$ • $C_{\text{inv}}^d = \left(24K^3 + 16K^2N\right) + K + \left(2KN + 8K^2N\right)$
Uplink stage	$C^u = 8 N K \eta_{\text{ul}}$
Total complexity	• $C_{\text{tot}} \approx 24K^3 + 24K^2\left(N + N_c/3\right)$ $+ 8K \left(N_c \eta_{\text{dl}} + N \eta_{\text{ul}}\right)$

us define the ergodic downlink sum rates of the communication system as

$$R_{\text{sum}} = \frac{\eta_{\text{dl}}}{\eta_{\text{coh}}} \cdot \mathbb{E}\left\{\sum_{k=1}^{K} R_k\right\}, \tag{31}$$

where $\varphi = \eta_{\text{dl}}/\eta_{\text{coh}}$ accounts for the loss in the time dedicated to downlink transmission due to the required CSI acquisition and uplink stages. Moreover, a set of ergodic achievable rates $\mathbb{E}\{R_k\}$ can be expressed as [5]

$$\mathbb{E}\{R_k\} = B \log_2\left(1 + \gamma_k\right) = B S_k, \tag{32}$$

where B stands for the system bandwidth, S_k represents the achievable spectral efficiency of the k-th user, and γ_k refers to the associated SINR for the k-th user given by [5], [42]

$$\gamma_k = \frac{\nu^2 |\mathbb{E}\{\mathbf{h}_k^H \mathbf{f}_k\}|^2}{\frac{1}{\rho_f} + \nu^2 \text{var}\left(\mathbf{h}_k^H \mathbf{f}_k\right) + \nu^2 \sum_{i \neq k} \mathbb{E}\{|\mathbf{h}_k^H \mathbf{f}_i|^2\}}, \tag{33}$$

where $\mathbf{f}_k \in \mathbb{C}^{N \times 1}$ corresponds to the k-th column of the precoding matrix $\mathbf{F}$. The availability of incomplete and imperfect CSI impacts on the resultant SINR γ_k via the precoding vectors $\mathbf{f}_k$. Specifically, for the case of ZF precoding and incomplete CSI, $\mathbf{f}_k$ can be expressed as

$$\mathbf{f}_k = \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H\right)^{-1} \widetilde{\mathbf{h}}_k, \tag{34}$$

where $\widetilde{\mathbf{h}}_k \in \mathbb{C}^{N \times 1}$ is the k-th column of the incomplete CSI matrix given by $\widetilde{\mathbf{H}} = \mathbf{R}^H \widehat{\mathbf{H}}$ as per (16). To constrain the focus of the paper on the proposed CSI reduction solution for massive MIMO systems, we designate the mathematical study of the above sum rates as the focus of future work. Here, we employ the above sum rate expression to study the improved energy efficiency of our approach in the following.

B. Energy Efficiency Model

The optimization of the energy efficiency has attracted much interest recently due to the need of controlling the power consumption required by future wireless communication systems [1], [57]. Since the proposed technique presents reduced hardware and signal processing complexities, in this section we present an energy efficiency model to characterize these

improvements. The energy efficiency in the downlink of a communication system is measured in bits per Joule and it can be expressed as [17], [25], [57]–[60]

$$\epsilon = \frac{R_{\text{sum}}}{P_{\text{tot}}}. \tag{35}$$

Here, R_{sum} refers to the system achievable sum rates in bits per second defined in (31) and P_{tot} denotes the total power in Watts consumed during transmission, which can be further decomposed as [17], [61]

$$P_{\text{tot}} = \frac{P_{\text{PA}} + P_{\text{RF}} + P_{\text{BB}}}{(1 - \sigma_{\text{DC}})(1 - \sigma_{\text{MS}})(1 - \sigma_{\text{cool}})}, \tag{36}$$

where σ_{DC} and σ_{MS} characterize the losses of the DC-DC and main power supplies whereas σ_{cool} refers to the active cooling losses [60]. These losses scale linearly with the power consumption of the electronic components and an accurate characterization of their values has been provided in [60], [61] for different BSs. Moreover, P_{PA} denotes the average power required by the power amplifiers (PAs) and it is given by

$$P_{\text{PA}} = \frac{P_{\text{t}}}{\varsigma}, \tag{37}$$

where ς corresponds to the PA efficiency and P_{t} refers to the power required at the output of the power amplifiers. Here, we remark that only N_{c} PAs are required to generate the output signals as shown in Fig. 2. Moreover, P_{RF} denotes the active power consumption of the electronic components in the RF chains without accounting for that of the power amplifiers. Therefore, P_{RF} can be further expressed as

$$P_{\text{RF}} = N_{\text{c}}(P_{\text{DAC}} + P_{\text{mix}} + P_{\text{filt}}) + P_{\text{syn}}, \tag{38}$$

where P_{syn} represents the power consumed by the frequency synthesizer, and P_{DAC}, P_{mix} and P_{filt} denote the power consumption of the digital-to-analog converters (DACs), signal mixers and filters included in each RF chain respectively [17], [62]. Additionally, $P_{\text{BB}} = p_{\text{c}}C$ corresponds to the power consumption of the digital signal processor (DSP). Here, p_{c} determines the power consumption per real flop and C refers to the average number of real flops per second determined in Sec. V. Since P_{RF} depends on the transmission power, it should be noted that the power consumption of the RF circuitry components for the different BSs has been approximated by scaling the data available in [62], which is considered to correspond to a macro BS, by a proportionality factor that relates to the BSs currently deployed [60], [61].

VII. SIMULATION RESULTS

To demonstrate the performance and complexity of the proposed strategy, in this section we present numerical results obtained via Monte Carlo simulations. The simulation setup consists of a BS with a planar array comprised of N antennas, a transmission channel following the physical channel model described in Sec. II-B, and K single-antenna receivers. For a fair comparison, all transmitters have the same physical size
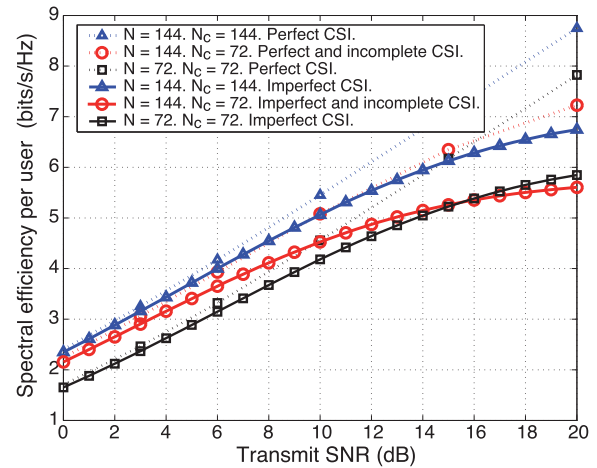

Fig. 5. Spectral efficiency per user (S_k) vs. SNR for $L = 10.9\lambda^2$. $N = 144$ and $N = 72$, $K = 12$ and $N_{\text{c}} = N/2$ for incomplete CSI acquisition.

independently of N. The channel model considers the practical case of imperfect CSI acquisition unless otherwise stated, an equal number of angular directions $D_k = 50$ in (7), and that the angle spread of the azimuth and elevation angles of departure is fixed to π and $\pi/6$ radians respectively [50]. The imperfect CSI scenario is modeled considering MMSE estimation for the active antennas after the pilot training stage as per (12)–(14) with $\rho_{\text{tr}} = 15$ dB [5]. The azimuth and elevation angles of departure of each user are considered to be independent and to follow a random uniform distribution within the relevant angle spread [6], [48]. In the following, we also consider a standard LTE frame with a duration of 10 milliseconds comprised of 10 subframes with 14 OFDM symbols each, which results in a coherence block of $\eta_{\text{coh}} = 140$ symbols [56]. Moreover, we follow the uplink-downlink configuration 4 of LTE TDD, in which a total of 7 subframes are used for downlink transmission [56]. This determines the time spent for CSI acquisition and data transmission, which in turn influences the signal processing complexity and the achievable sum rates as per (30) and (31) respectively.

The performance of the communication systems considered in this paper can be observed in Fig. 5, where a comparison of the spectral efficiency per user in (32) is shown for increasing levels of SNR. In this figure, the antennas are deployed in an area of $L = 10.9\lambda^2$ and we consider both systems with complete and incomplete CSI. In the following, the number of users is fixed to $K = 12$, and the performance for $N = 144$ ($M_{\text{h}} = M_{\text{v}} = 12$) and $N = 72$ ($M_{\text{h}} = 9$, $M_{\text{v}} = 8$) antennas with full CSI is depicted as a reference. This establishes an inter-antenna distance $d = 0.3\lambda$ in Fig. 5 for the cases of $N = 144$. Moreover, it is considered that $N_{\text{c}}/N = 1/2$ for the case of incomplete CSI acquisition. At this point we remark that, although related to the considered scheme, antenna selection strategies are not directly comparable. This is because they require a larger CSI acquisition time due to the limitation of having a limited number of RF chains and the need of acquiring CSI for all the antennas to perform the selection [63], [64]. Moreover, antenna selection techniques might require solving complex optimization problems prior to the application of precoding schemes
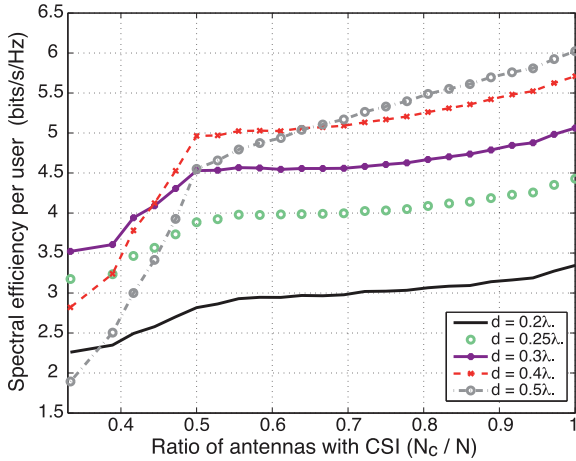
Fig. 6. Spectral efficiency per user (S_k) vs. ratio of antennas with CSI (N_c/N). $N = 144$, $K = 12$ and SNR = 10 dB.

Fig. 7. Spectral efficiency per user (S_k) vs. inter-antenna distance (d/λ). $N = 144$, $K = 12$ and SNR = 10 dB.

in the correlated channels considered in this paper, hence not complying with our objective of reducing the signal processing complexity to a minimum [64].

Overall, the results of Fig. 5 show that the proposed strategy outperforms a system with N_c antennas, full CSI and the same number of RF chains for the low-to-intermediate SNR region. We would like to note that the large number of antennas implemented massive MIMO systems also motivate their operation in this region due to their improved spectral efficiency and the need of limiting the total power consumption [5], [17]. This validates the benefits of using incomplete CSI in this range as well as the employment of the proposed scheme based on the basic CSI distribution patterns shown in Fig. 1. It can also be observed that the performance of the incomplete CSI scheme is limited by the errors introduced by the averaging procedure at high SNRs when perfect CSI is considered. Moreover, it can be seen that the consideration of imperfect CSI also bounds the performance of conventional transmission strategies at high SNRs, hence minimizing the differences between the schemes with full and incomplete CSI.

The effect of N_c on the performance of the proposed strategy can be observed in Fig. 6, where the spectral efficiency vs. the percentage of antennas with CSI is shown for a fixed SNR of 10 dB. From the results of this figure it can be concluded that a larger number of antennas with CSI is required as the inter-antenna spacing increases to reach a given percentage of the ultimate performance obtained when (N_c/N) = 1. This result is intuitive since the communication channels of the different antennas exhibit a lower correlation for larger antenna separations, hence degrading the performance of the proposed scheme due to the errors introduced by the averaging procedure. In spite of this, it can be seen that activating half of the antennas during the training stage for $d = 0.3\lambda$ suffices to achieve 75 % of the maximum system throughput with full CSI and $d = 0.5\lambda$, where the occupied area is larger.

The impact of varying the inter-antenna distance on the spectral efficiency per user can be clearly seen in Fig. 7, where an identical setup to that of Fig. 6 has been considered. Fig. 7
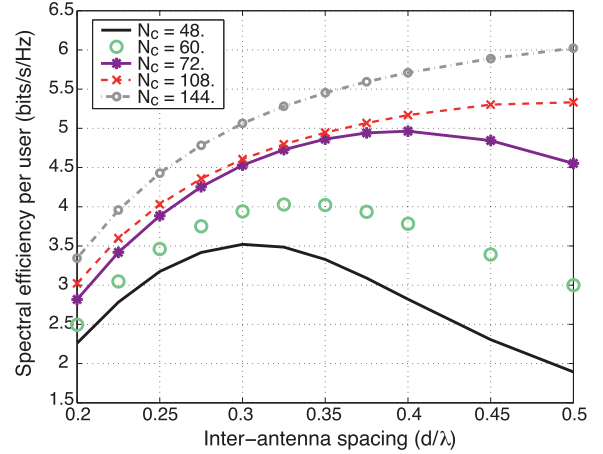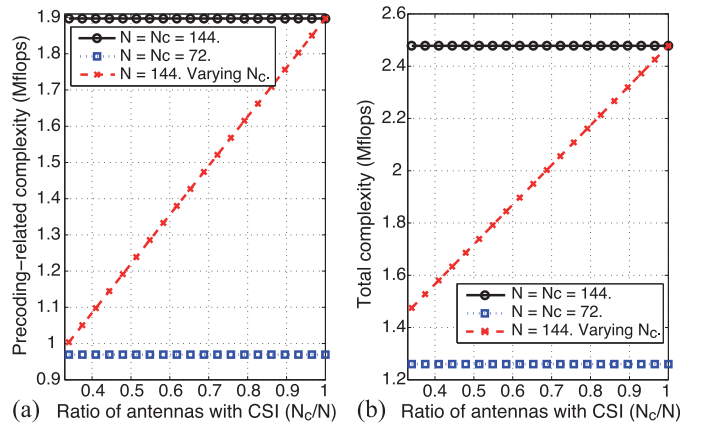

Fig. 8. (a) Precoding-related complexity and (b) total complexity per frame for increasing N_c. $K = 12$ users, $\eta_{coh} = 140$, $\eta_{tr} = K = 12$ and $\eta_{dl} = 7 \times 14$.

characterizes the trade-off that arises by varying the inter-antenna distances: while the proposed scheme reduces the CSI approximation errors for small inter-element spacings, the loss produced by a larger spatial correlation can harm the spectral efficiency. In this line, the results of Fig. 7 allow determining the optimal inter-antenna distance for a given ratio of antennas with CSI (N_c/N). Fig. 7 also indicates the possibility of intentionally reducing the inter-antenna spacing in massive antenna arrays due to the computational and energy efficiency benefits offered by the proposed strategy as shown in the following.

The number of real flops during the precoding stage and the global communication per frame are depicted in Fig. 8(a) and (b) respectively for increasing levels of N_c. The results of this figure show the notable complexity savings that can be experienced when the proposed incomplete CSI scheme is employed. Specifically, it can be seen that the number of precoding-related operations of the proposed scheme can be dramatically reduced w.r.t. a system with the same number of antennas but full CSI acquisition. This translates to a significant reduction of the global complexity as depicted in Fig. 8(b). The global complexity reduction can be explained by noting that the complexity of the precoding process accounts for a significant part of the
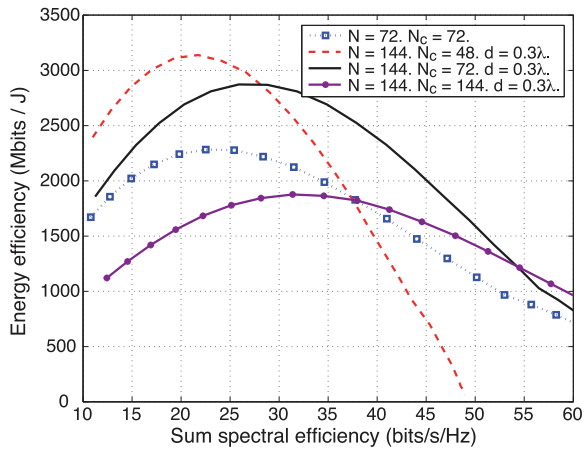
Fig. 9. Energy efficiency vs. sum spectral efficiency for $L = 10.9\lambda^2$. $N = 144$ and $N = 72$, $K = 12$.

TABLE II
POWER CONSUMPTION OF MICRO AND FEMTO BASE STATIONS

Parameter		Micro	Femto
Total transmission power [61],	P_t	38 dBm	17 dBm
Power amplifier efficiency [59],	ς	0.38	
Digital-to-analog converter [61], [62],	P_{DAC}	7.8 mW	0.46 mW
Signal mixer [61], [62],	P_{mix}	15.2 mW	0.9 mW
Signal filtering [61], [62],	P_{filt}	10 mW	0.6 mW
Frequency synthesizer [61], [62],	P_{syn}	25 mW	1.5 mW
Computational efficiency [58],	p_c^{-1}	12.8 Mflops / mW	
DC-DC losses [61],	σ_{DC}	0.075	0.09
Main supply losses [61],	σ_{MS}	0.09	0.11
Cooling losses [61],	σ_{cool}	0	0

global signal processing load, hence maximizing the benefits of the incomplete CSI scheme.

Fig. 9 depicts the energy efficiency metric (35) for increasing levels of spectral efficiency and $L_h = L_v = 3.3\lambda$. The results of this figure have been obtained by fixing the total noise power $B\sigma^2$ to 10 dBm for reference, by setting the total system bandwidth to $B = 20$ MHz, and by computing the transmission power required to achieve a given spectral efficiency [56]. Moreover, since the circuit power consumption varies depending on the total transmission power [61], in this figure we have performed a linear interpolation of the data shown in Table II to estimate the approximate power consumed by the RF chains. For simplicity, it has been considered that the analog signal combining and splitting processes do not introduce any additional losses, i.e., the considered schemes have the same transmission power. We remark that, although dynamic power losses appear in the power combining process [65], determining the additional power consumption required by the transmission chains is intricate due to the possibility of designing different solutions for their compensation.

The results of Fig. 9 show that, under a moderate antenna correlation, deactivating half of the antennas during the pilot training stage constitutes the most energy-efficient alternative for a wide range of spectral efficiencies. In particular, it can be seen that using the proposed scheme is beneficial for low and intermediate spectral efficiencies. Instead, conventional massive MIMO strategies become more energy-efficient for large
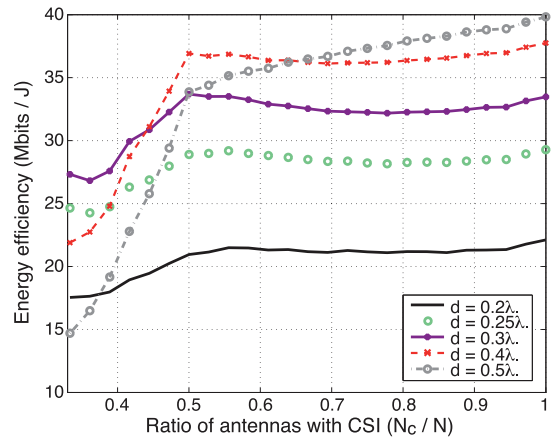

Fig. 10. Energy efficiency vs. ratio of antennas with CSI (N_c/N) for a microcell scenario and varying inter-antenna distance d. $N = 144$, $K = 12$ and SNR = 10 dB.
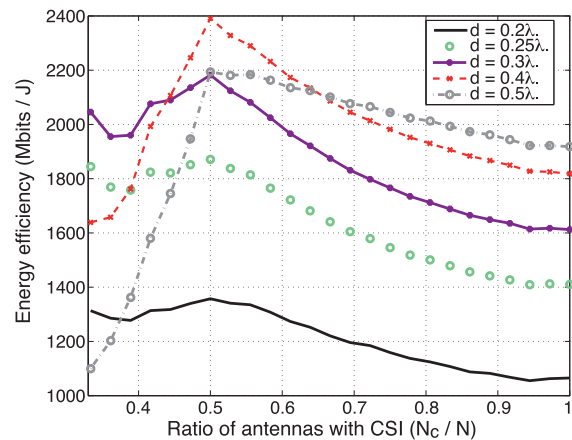

Fig. 11. Energy efficiency vs. ratio of antennas with CSI (N_c/N) for a femtocell scenario and varying inter-antenna distance d. $N = 144$, $K = 12$ and SNR = 10 dB.

spectral efficiencies. This behaviour can be explained by noting that the energy efficiency benefits of the proposed scheme are especially noticeable when the circuit power consumption strongly influences the total power consumption. We note that this can potentially occur in massive BSs due to the significant increase in the number of RF components required when compared with traditional BSs [17], [60]. In these cases, reducing the number of RF chains is beneficial for the global energy efficiency as shown in Fig. 9. As higher spectral efficiencies are required, the dominant factor in the total power consumption is the power consumed by the power amplifiers following (36), which is not affected by the reduced number of RF chains offered by the proposed strategy. Nevertheless, it should be noted that the incomplete CSI schemes are capable of offering significant flexibility benefits as well as reductions of the signal processing complexity, hence extending their range of application to high spectral efficiencies where the computational load is expected to be cumbersome [19].

The effect of the percentage of active antennas during the training stage on the energy efficiency metric is shown in Fig. 10 and Fig. 11 for microcell and femtocell BSs respectively. The power consumption values from Table II have been

employed for each of the scenarios. The results are shown for SNR = 10 dB and a varying inter-antenna distance d. From the results of these figures it can be concluded that reducing the number of active antennas during the training stage and the inter-antenna spacing can be beneficial from an energy efficiency perspective. Specifically, Fig. 10 shows that a large percentage of the maximum energy efficiency attained for $d = 0.5\lambda$ and $N_c = N = 144$ can be obtained by acquiring CSI for half of the antennas and $d = 0.3\lambda$. We remark that in this second case the size of the deployed massive antenna array is reduced w.r.t. the case of $d = 0.5\lambda$. Moreover, Fig. 11 shows that the energy efficiency is maximized for the femtocell scenario when $N_c = N/2 = 72$ antennas are active during the CSI acquisition process and the distance between antennas is $d = 0.4\lambda$. Intuitively, this occurs because the small loss in the achievable rates is compensated by the substantial reduction experienced in the power consumption of the RF components and the digital signal processor.

VIII. CONCLUSION

In this paper, a low-complexity scheme to exploit the inter-antenna correlation that arises in space-constrained massive MIMO BSs has been presented. The proposed strategy reduces the required number of RF chains by deactivating a number of antennas during the CSI acquisition stage. Subsequently, the CSI for the remaining neighbour antennas is obtained by averaging the information acquired. This is also extended to the transmission stage, where a reduced RF chain approach is illustrated. Overall, the results presented in this paper confirm that the proposed scheme is able to approximately preserve the performance of tightly-packed BSs, while simultaneously reducing the computational complexity and enhancing the energy efficiency. The study of the implications of the use of incomplete CSI in FDD massive MIMO systems constitutes the subject of subsequent work. The consideration of alternative channel models is also of interest [66]. Moreover, more intricate strategies for determining the antennas with instantaneous CSI and for interpolating this information will also be explored in the future.

APPENDIX A
PROOF OF THEOREM 1

The proof of Theorem 1 relies on determining the error factor for the antennas with imperfect CSI and the antennas whose CSI is derived following the CSI averaging procedure described in Sec. III. First, note that Δ_k in (21) can be further decomposed as

$$
\Delta_k = \frac{\mathbb{E}\left\{\sum_{n\in\mathcal{B}}|\bar{h}_{k,n}|^2 + \sum_{n\in\mathcal{C}}|\bar{h}_{k,n}|^2\right\}}{\mathbb{E}\left\{\sum_{n=1}^{N}|h_{k,n}|^2\right\}}
$$

$$
\stackrel{(a)}{=} \frac{\sum_{n\in\mathcal{B}}\mathbb{E}\left\{|\bar{h}_{k,n}|^2\right\} + \sum_{n\in\mathcal{C}}\mathrm{var}\left(\bar{h}_{k,n}\right)}{N}, \qquad (39)
$$

where $\bar{h}_{k,n}$ denotes the error of the channel from the n-th antenna of the BS to the k-th user and $\mathrm{var}(\cdot)$ denotes the

variance of a random variable. In the above expression, $\stackrel{(a)}{=}$ holds because it is considered that the channel for the n-th antenna without CSI is generated by following the averaging operation (15) and $\mathbb{E}\{\widehat{\mathbf{h}}_k\} = \mathbb{E}\{\mathbf{h}_k\} = \mathbf{0}$, which entails that $\mathrm{var}\left(\bar{h}_{k,n}\right) = \mathbb{E}\{|\bar{h}_{k,n}|^2\}$. Moreover, the channel coefficients of the physically-constrained model defined in (6) clearly satisfy $\mathbb{E}\{|h_{k,n}|^2\} = 1$.

Now we concentrate on the two terms found in the numerator of (39). In particular, it can be seen that the first term only involves the antennas with instantaneous CSI availability and accounts for the errors due to imperfect channel estimation. The factors of this term can be expressed as

$$
\mathbb{E}\left\{|\bar{h}_{k,n}|^2\right\} = \tau_k^2\mathbb{E}\left\{\left(\mathbf{A}_k\Big|_{[\mathcal{D},n]}\right)^H \mathbf{q}_k\mathbf{q}_k^H \mathbf{A}_k\Big|_{[\mathcal{D},n]}\right\}
$$

$$
= \tau_k^2\mathbf{\Theta}_k\big|_{[n,n]}, n\in\mathcal{B}, \qquad (40)
$$

where $\{\mathbf{A}_k\}_{[\mathcal{D},n]}$ denotes the n-th column of the transmit steering matrix $\mathbf{A}_k$ defined in (7). Instead, the second term of the numerator in (39) characterizes the combination of the errors introduced by having an inaccurate channel estimate and the averaging operation for the antennas without CSI. By combining (15) and (19) this factor can be re-expressed as

$$
\mathrm{var}\left(\bar{h}_{k,n}\right) = \mathrm{var}\left(h_{k,n} - \frac{1}{M_n}\sum_{i=1}^{M_n}\widehat{\mathbf{h}}_{k,\mathcal{B}_i^n}\right), n\in\mathcal{C}, \qquad (41)
$$

which corresponds to the variance of a subtraction of complex correlated random variables due to the compactness of the antenna array. This expression can be rewritten by iterative application of the statistical identities [67]

$$
\mathrm{var}\left(X\pm Y\right) = \mathrm{var}\left(X\right) + \mathrm{var}\left(Y\right) \pm 2\mathrm{Re}\left(\mathrm{cov}\left(X,Y\right)\right),
$$

$$
\mathrm{cov}\left(X\pm Y,Z\right) = \mathrm{cov}\left(X,Z\right)\pm\mathrm{cov}\left(Y,X\right), \qquad (42)
$$

where $\mathrm{cov}\left(\cdot,\cdot\right)$ denotes the covariance between two random variables and $\mathrm{Re}\left(\cdot\right)$ represents the real part of the argument. As a result, we obtain

$$
\mathrm{var}\left(h_{k,n} - \frac{1}{M_n}\sum_{i=1}^{M_n}\widehat{\mathbf{h}}_{k,\mathcal{B}_i^n}\right) = \mathrm{var}\left(h_{k,n}\right)
$$

$$
+ \frac{1}{(M_n)^2}\sum_{i\in\mathcal{B}^n}\mathrm{var}\left(\widehat{h}_{k,i}\right) - \frac{2}{M_n}\mathrm{Re}\left(\sum_{i\in\mathcal{B}^n}\mathrm{cov}\left(h_{k,n},\widehat{h}_{k,i}\right)\right)
$$

$$
+ \frac{2}{(M_n)^2}\mathrm{Re}\left(\sum_{i,j\in\mathcal{B}^n,i>j}\mathrm{cov}\left(\widehat{h}_{k,j},\widehat{h}_{k,i}\right)\right), n\in\mathcal{C}. \qquad (43)
$$

The terms involved in the above formulation can be further decomposed and given as a function of the true channel correlation matrix $\mathbf{\Theta}_k$. With this purpose, $\mathrm{cov}\left(h_{k,n},\widehat{h}_{k,i}\right)$ can be obtained as

$$
\mathrm{cov}\left(h_{k,n},\widehat{h}_{k,i}\right) = \mathrm{cov}\left(h_{k,n},\left(\sqrt{1-\tau_k^2}h_{k,i} + \tau_k\bar{h}_{k,i}\right)\right)
$$

$$
= \sqrt{1-\tau_k^2}\mathbf{\Theta}_k\big|_{[n,i]}, \qquad (44)
$$

where it has been considered that $\mathbf{q}_k$ is uncorrelated with $\mathbf{g}_k$ in (20) [4]. The rest of the terms involved in (43) can be obtained similarly, and are omitted here for brevity. Finally, (22) is obtained by substituting (40) and (43) into (39), which completes the proof.

APPENDIX B
PROOF OF COROLLARY 2

The objective of the proof is to show that the channel error factor Δ_k is a monotonically increasing function of d for the physical channel model described in Sec. II-B, i.e.,

$$\frac{\partial \Delta_k}{\partial d} = \frac{\partial}{\partial d} \left\{ \sum_{n \in \mathcal{C}} \left[\underbrace{-\frac{2}{M_n} \mathrm{Re}\left(\sum_{i \in \mathcal{B}^n} \Theta_k \big|_{[n,i]} \right)}_{S_1} \right. \right.$$

$$\left. \left. + \underbrace{\frac{2}{(M_n)^2} \mathrm{Re}\left(\sum_{i,j \in \mathcal{B}^n, i>j} \Theta_k \big|_{[j,i]} \right)}_{S_2} \right] \right\} > 0. \quad (45)$$

Here, we have removed the terms independent of d and the factor $\sqrt{1 - \tau_k^2}$ from (22), since $\tau_k = 0$ when perfect CSI is considered as per (20). In the following we also drop the user index k for notational convenience. The i,j-th entry of the channel correlation matrix Θ for a ULA following the physical channel model described in Sec. II-B can be expressed as

$$\Theta \big|_{[i,j]} = \frac{1}{D} \sum_{m=1}^{D} e^{j 2\pi d(j-i)\sin(\phi_m)}, \quad (46)$$

where D denote the total number of angles of arrival (AoA) and ϕ_m is the m-th azimuth AoA. In the following we focus on showing that

$$Q_n = \frac{\partial S_1}{\partial d} + \frac{\partial S_2}{\partial d} > 0, \forall n \in \mathcal{C}, \quad (47)$$

which in turn ensures that (45) holds. Here, since the real part of a function is not analytic, it is convenient to express S_1 and S_2 in (45) as

$$S_1 = -\frac{1}{M_n} \left(\sum_{i \in \mathcal{B}^n} \Theta \big|_{[n,i]} + \Theta \big|_{[i,n]} \right), \quad (48)$$

and

$$S_2 = \frac{1}{(M_n)^2} \left(\sum_{i,j \in \mathcal{B}^n, i>j} \Theta \big|_{[j,i]} + \Theta \big|_{[i,j]} \right). \quad (49)$$

Let us concentrate on S_1. For the case of ULAs and $N/N_c \geq 2$, the CSI for each of the $|\mathcal{C}|$ antennas without CSI is obtained by averaging the CSI of the $M_n = 2$ adjacent antennas. Note that this is a consequence of the application of the basic CSI

distribution pattern shown in Fig. 1(b). Therefore, S_1 can be re-expressed as

$$S_1 = -\left(\Theta \big|_{[n,i]} + \Theta \big|_{[i,n]} \right). \quad (50)$$

Substituting (46) into (50) and differentiating w.r.t. d results in

$$\frac{\partial S_1}{\partial d} = \frac{4\pi}{D} \sum_{m=1}^{D} \sin(\phi_m) \sin(2\pi d \sin(\phi_m)), \quad (51)$$

where we have considered that $(n - i) = 1$ because the CSI of the adjacent antennas is averaged, and $\sin(\phi) = \left(e^{j\phi} - e^{-j\phi} \right)/2j$. Following a similar process while noting that the distance between the antennas with CSI used for averaging satisfies $(i - j) = 2$ in (49), we can express $\frac{\partial S_2}{\partial d}$ as

$$\frac{\partial S_2}{\partial d} = \frac{2\pi}{D} \sum_{m=1}^{D} \sin(\phi_m) \sin(4\pi d \sin(\phi_m)) \quad (52)$$

$$\overset{(a)}{=} \frac{4\pi}{D} \sum_{m=1}^{D} \sin(\phi_m) \sin(2\pi d \sin(\phi_m)) \cos(2\pi d \sin(\phi_m)),$$

where the trigonometric identity $\sin(2\phi) = 2\sin(\phi)\cos(\phi)$ has been employed in $\overset{(a)}{=}$. Subsequently, substituting (51) and (52) into (47) we can write Q_n as

$$Q_n = \frac{4\pi}{D} \sum_{m=1}^{D} \sin(\phi_m)\sin(2\pi d \sin(\phi_m)) \left(1 - \cos(2\pi d \sin(\phi_m))\right). \quad (53)$$

To conclude, we incorporate the small-angle approximations $\sin(\phi) \approx \phi$ and $\cos(\phi) \approx \left(1 - \phi^2/2\right)$ into (53), which yields

$$Q_n = \frac{4\pi (2\pi d)}{D} \sum_{m=1}^{D} \phi_m^2 \left(\frac{(2\pi d \phi_m)^2}{2} \right) > 0, \quad (54)$$

and completes the proof.

REFERENCES

[1] G. Li et al., "Energy-efficient wireless communications: Tutorial, survey, and open issues," IEEE Wireless Commun., vol. 18, no. 6, pp. 28–35, Dec. 2011.
[2] F. Rusek et al., "Scaling up MIMO: Opportunities and challenges with very large arrays," IEEE Signal Process. Mag., vol. 30, no. 1, pp. 40–60, Jan. 2013.
[3] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," IEEE J. Sel. Topics Signal Process., vol. 8, no. 5, pp. 742–758, Oct. 2014.
[4] S. Wagner et al., "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," IEEE Trans. Inf. Theory, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.
[5] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" IEEE J. Sel. Areas Commun., vol. 31, no. 2, pp. 160–171, Feb. 2013.
[6] C. Masouros et al., "Large-scale MIMO transmitters in fixed physical spaces: The effect of transmit correlation and mutual coupling," IEEE Trans. Commun., vol. 61, no. 7, pp. 2794–2804, Jul. 2013.
[7] S. Biswas, C. Masouros, and T. Ratnarajah, "Performance analysis of large multi-user MIMO systems with space-constrained 2D antenna arrays," IEEE Trans. Wireless Commun., 2016, to be published.
[8] C. Masouros and M. Matthaiou, "Space-constrained massive MIMO: Hitting the wall of favorable propagation," IEEE Commun. Lett., vol. 19, no. 5, pp. 771–774, May 2015.

[9] A. Tulino, A. Lozano, and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[10] V. Raghavan and A. Sayeed, "Sublinear capacity scaling laws for sparse MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 345–364, Jan. 2011.

[11] W. Hachem *et al.*, "A new approach for mutual information analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, Sep. 2008.

[12] A. Sayeed and V. Raghavan, "Maximizing MIMO capacity in sparse multipath with reconfigurable antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 156–166, Jun. 2007.

[13] S. Simon *et al.*, "Capacity and character expansions: Moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.

[14] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of 'Kronecker' MIMO radio channel model," *Electron. Lett.*, vol. 39, no. 16, pp. 1209–1210, Aug. 2003.

[15] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[16] P. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[17] D. Ha, K. Lee, and J. Kang, "Energy efficiency analysis with circuit power consumption in massive MIMO systems," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, Sep. 2013, pp. 938–942.

[18] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[19] A. Kammoun *et al.*, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 861–875, Oct. 2014.

[20] N. Shariati *et al.*, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 815–830, Oct. 2014.

[21] J. Zhang, L. Dai, M. Matthaiou, C. Masouros, and S. Jin, "On the spectral efficiency of space-constrained massive MIMO with linear receivers," in *Proc. IEEE Int. Conf. Commun. (ICC)*, to be published.

[22] B. Cetiner, H. Jafarkhani, J.-Y. Qian, H. J. Yoo, A. Grau, and F. De Flaviis, "Multifunctional reconfigurable MEMS integrated antennas for adaptive MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 62–70, Dec. 2004.

[23] D. Piazza *et al.*, "Design and evaluation of a reconfigurable antenna array for MIMO systems," *IEEE Trans. Antennas Propag.*, vol. 56, no. 3, pp. 869–881, Mar. 2008.

[24] T. Muharemovic, A. Sabharwal, and B. Aazhang, "Antenna packing in low-power systems: Communication limits and array design," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 429–440, Jan. 2008.

[25] E. Björnson *et al.*, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[26] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

[27] A. Liu and V. Lau, "Phase only RF precoding for massive MIMO systems with limited RF chains," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4505–4515, Sep. 2014.

[28] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[29] P.-H. Kuo *et al.*, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2012, pp. 492–497.

[30] S. L. H. Nguyen and A. Ghrayeb, "Compressive sensing-based channel estimation for massive multiuser MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2013, pp. 2890–2895.

[31] Z. Jiang, A. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2868–2882, May 2015.

[32] S. Noh, M. Zoltowski, Y. Sung, and D. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 787–801, Oct. 2014.

[33] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Access*, vol. 2, pp. 947–959, Aug. 2014.

[34] B. Lee et al., "Antenna grouping based feedback reduction for FDD-based massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 4477–4482.

[35] H. Prabhu et al., "Approximate matrix inverse computations for very-large MIMO and applications to linear pre-coding systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2013, pp. 2710–2715.

[36] A. Sayeed and N. Behdad, "Continuous aperture phased MIMO: A new architecture for optimum line-of-sight links," in *Proc. IEEE Int. Symp. Antennas Propag. (APSURSI)*, Jul. 2011, pp. 293–296.

[37] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[38] L. Dai et al., "Near-optimal hybrid analog and digital precoding for downlink mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1334–1339.

[39] X. Gao *et al.*, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *CoRR*, vol. abs/1507.04592, 2015 [Online]. Available: http://arxiv.org/abs/1507.04592.

[40] P. V. Amadori and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212–2223, Jun. 2015.

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

[42] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[43] S. Ray, M. Medard, and L. Zheng, "On noncoherent MIMO channels in the wideband regime: Capacity and reliability," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 1983–2009, Jun. 2007.

[44] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[45] L. Zheng, D. Tse, and M. Medard, "Channel coherence in the low-SNR regime," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 976–997, Mar. 2007.

[46] 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. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[47] H. Yang and T. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[48] C. Wang and R. Murch, "Adaptive downlink multi-user MIMO wireless systems for correlated channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2435–2446, Sep. 2006.

[49] K. F. Warnick and M. Jensen, "Optimal noise matching for mutually coupled arrays," *IEEE Trans. Antennas Propag.*, vol. 55, no. 6, pp. 1726–1731, Jun. 2007.

[50] C. A. Balanis, *Antenna Theory: Analysis and Design*. Hoboken, NJ, USA: Wiley, 2012.

[51] L. Vandenberghe and S. Boyd, "Applied numerical computing," University Lecture, Dept. Elect. Eng., UCLA, Los Angeles, CA, USA, 2012.

[52] C. Oestges and B. Clerckx, *MIMO Wireless Communications: From Real-World Propagation to Space-Time Code Design*. Amsterdam, The Netherlands: Elsevier, 2010.

[53] R. Hunger. (2006). *Floating Point Operations in Matrix-Vector Calculus* [Online]. Available: https://mediatum.ub.tum.de/doc/625604/625604.pdf.

[54] M. Arakawa, "Computational workloads for commonly used signal processing kernels," DTIC Document, Project Rep. SPR-9, Mass. Inst. Technol., Cambridge, MA, USA, 2006.

[55] A. Garcia-Rodriguez and C. Masouros, "Exploiting the tolerance of massive MIMO to incomplete CSI for low-complexity transmission," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[56] M. Rumney *et al.*, *LTE and the Evolution to 4G Wireless: Design and Measurement Challenges*. Hoboken, NJ, USA: Wiley, 2013.

[57] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[58] H. Yang and T. Marzetta, "Total energy efficiency of cellular large scale antenna system multiple access mobile networks," in *Proc. IEEE Online Conf. Green Commun. (GreenCom)*, Oct. 2013, pp. 27–32.

[59] J. Xu and L. Qiu, "Energy efficiency optimization for MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 690–701, Feb. 2013.

[60] G. Auer *et al.*, "Cellular energy efficiency evaluation framework," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, 2011, pp. 1–6.
[61] G. Auer *et al.*, "D2. 3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," Energy Aware Radio and NeTwork TecHnologies, Tech. Rep. INFSOICT-247733 EARTH, 2010.
[62] H. Kim *et al.*, "A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4264–4275, Aug. 2009.
[63] P. V. Amadori and C. Masouros, "Interference driven antenna selection for massive multi-user MIMO," *IEEE Trans. Veh. Technol.*, 2015, to be published.
[64] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization," in *Proc. IEEE Globecom Workshops*, Dec. 2013, pp. 129–134.
[65] D. M. Pozar, *Microwave Engineering*. Hoboken, NJ, USA: Wiley, 2009.
[66] A. S. Y. Poon, R. W. Brodersen, and D. N. C. Tse, "Degrees of freedom in multiple-antenna channels: A signal space approach," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 523–536, Feb. 2005.
[67] R. J. Pruim, *Foundations and Applications of Statistics: An Introduction Using R. Providence*, RI, USA: AMS, 2011, vol. 13.

Adrian Garcia-Rodriguez (S'13) received the M.S. degree in telecommunications engineering from the Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain, in 2012. He is currently pursuing the Ph.D. degree at the Department of Electrical and Electronic Engineering, University College London, London, U.K. His industry experience includes an internship at Nokia Bell Labs, Dublin, Ireland, in 2015. His research interests include signal processing and wireless communications with emphasis on energy-efficient and multiantenna communications.

Christos Masouros (M'06–SM'14) received the diploma in electrical and computer engineering from the University of Patras, Patras, Greece, in 2004, the M.Sc. by research and the Ph.D. degree in electrical and electronic engineering from the University of Manchester, Manchester, U.K., in 2006 and 2009, respectively. He has previously held a Research Associate position with the University of Manchester, and a Research Fellow position with Queen's University Belfast, Belfast, U.K. He is currently a Lecturer with the Department of Electrical and Electronic Engineering, University College London, London, U.K.

He holds a Royal Academy of Engineering Research Fellowship (2011–2016) and is the Principal Investigator of the EPSRC project EP/M014150/1 on large scale antenna systems. His research interests include wireless communications and signal processing with particular focus on green communications, large scale antenna systems, cognitive radio, interference mitigation techniques for MIMO, and multicarrier communications. He is an Associate Editor for IEEE COMMUNICATIONS LETTERS, and was named the Exemplary Reviewer for the IEEE TRANSACTIONS ON COMMUNICATIONS in 2015. He was the recipient of the Best Paper Award in the IEEE GLOBECOM 2015 for his work on interference exploitation.