Low-Complexity Compressive Sensing Detection for Spatial Modulation in Large-Scale Multiple Access Channels

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

***Abstract*—In this paper, we propose a detector, based on the compressive sensing (CS) principles, for multiple-access spatial modulation (SM) channels with a large-scale antenna base station (BS). Particularly, we exploit the use of a large number of antennas at the BSs and the structure and sparsity of the SM transmitted signals to improve the performance of conventional detection algorithms. Based on the above, we design a CS-based detector that allows the reduction of the signal processing load at the BSs particularly pronounced for SM in large-scale multiple-input–multiple-output (MIMO) systems. We further carry out analytical performance and complexity studies of the proposed scheme to evaluate its usefulness. The theoretical and simulation results presented in this paper show that the proposed strategy constitutes a low-complexity alternative to significantly improve the system's energy efficiency against conventional MIMO detection in the multiple-access channel.**

***Index Terms*—Spatial modulation, large-scale MIMO, multiple access, compressive sensing, energy efficiency.**

I. Introduction

THE exponential growth of the data rates in wireless communications has caused a significant increase in the total energy consumption required to establish the communication links [1]. This is because novel transceiver structures with a higher number of antennas, transmission power or complexity in their signal processing algorithms have been designed to accommodate this growth [1], [2]. For this reason, the energy efficiency (EE) of the multi-user wireless transmission constitutes one of the main areas of research interest at present [1], [2]. Technologies such as large-scale MIMO and SM have been developed with the main objective of satisfying the EE requirements of future wireless communication systems [3], [4].

Massive MIMO technologies increase the EE by incorporating a large number of antennas at the BSs [4]–[7]. This leads to communication systems in which the use of conventional linear detection and precoding techniques becomes optimal in the large-scale limit [4], [5]. However, the increased number of radio frequency (RF) chains have a considerable influence on the EE, hence severely affecting the large-scale benefits from this perspective [8]. To alleviate this impact, SM poses as a reduced RF-complexity scheme by exploiting the transmit antenna indices as an additional source of information [3], [9]. Intuitively, instead of activating all the antennas simultaneously as in conventional MIMO transmission, SM proposes to switch on a subset of them and modify the receiver's operation to detect both the active antenna indices and the amplitude-phase symbols. This reduces the number of RF chains when compared to conventional MIMO systems at the cost of decreasing the maximum achievable rates [3], [9], [10].

So far, the literature of SM has mostly focused on developing strategies for point-to-point links [3], [9]–[15]. For instance, several low-complexity detectors that approach the performance of the optimal maximum likelihood (ML) estimation have been proposed [15]–[18]. In this context, the use of a normalized CS detection algorithm as a low-complexity solution for space shift keying (SSK) and generalized space shift keying (GSSK) peer-to-peer (P2P) systems was introduced in [18]. In this work, the authors apply a normalization to the channel matrix before the application of the greedy compressive detector to improve performance. However, the authors restrict its application to single-user SSK and GSSK systems, which constrains its use to low data rate transmission. The performance of the normalized CS detection for SSK and GSSK has been recently enhanced in [19], where the improvement is obtained by pre-equalizing the received signal.

More recently, the use of SM has been extended to the multiple access channel (MAC) as a way of enhancing the achievable rates of the conventional single-antenna devices considered in this setting [20]–[22]. The literature related to the study of multiple access SM builds upon [20], [21], where a characterization of the bit error rate (BER) probabilities of the optimal ML detector for both spatial and SSK modulations is performed. However, the signal processing load of the ML detector makes it impractical in the MAC as it grows exponentially with this parameter [20]. In this setting, several detection schemes to reduce the complexity of the ML detector in the MAC have been proposed and studied in [22], where the focus is on Q-ary SSK modulation.

A number of related works have concentrated on the design of detection schemes to account for the particularities of the large-scale MAC. The development of detection schemes in

Manuscript received July 14, 2014; revised March 4, 2015; accepted May 8, 2015. Date of publication May 18, 2015; date of current version July 13, 2015. This work was supported in part by the Royal Academy of Engineering, U.K., and in part by EPSRC under Grant EP/M014150/1. The material in this paper was presented in part at the IEEE International Conference on Acoustics, Speech and Signal Processing, Brisbane, Australia, April 2015, and the IEEE International Conference on Communications, London, U.K., June 2015. The associate editor coordinating the review of this paper and approving it for publication was A. Ghrayeb.

The authors are with the Department of Electronic and Electrical Engineering, University College London, London WC1E 6BT, U.K. (e-mail: adrian.rodriguez.12@ucl.ac.uk; chris.masouros@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2015.2434817

this setting is motivated by the intractable complexity of non-linear detectors such as the sphere decoder when a high number of antennas and users are considered [4], [15]. The use of a message passing detection (MPD) algorithm is proposed in [23] for the MAC with a high number of antennas at the BS. This algorithm offers a performance improvement with respect to conventional MIMO systems with the same spectral efficiency. However, both the storage requirements and the total number of operations are conditioned by the high number of messages transmitted between all the nodes, which must be updated in every iteration [24]. A more complex local search detection algorithm based on finding the local optimum in terms of the maximum likelihood cost is also introduced in [23]. An iterative detector for large-scale MACs is developed in [25]. Here, the authors decouple the antenna and symbol estimation processes to reduce the global detection complexity. The algorithm introduced in [26] accounts for the sparsity and signal prior probability of SM transmission in the MAC. In this work, the use of stage-wised linear detection is discarded due to its high complexity and the authors propose a generalized approximate message passing detector. A related approach has been developed in [27] to deal with quantized measurements and spatial correlation. Still, the above algorithms do not fully account for the particularities of iterative detection processes and the complexity benefits that can be obtained by leveraging the principles behind CS algorithms.

In this paper we propose a low-complexity detector based on CS for the MAC of SM systems with large-scale BSs. In particular, we show that the signal structure of SM in the MAC can be exploited to provide additional information and improve the performance of the CS algorithms [28], [29]. Moreover, the use of a high number of antennas in the MAC allows us to eliminate the error floor that greedy CS techniques show in noisy scenarios for practical uncoded BERs [18]. Indeed, contrary to the common CS knowledge, in this paper we show by means of a thorough complexity analysis that the trade-off between complexity and performance is especially favorable for CS-based detection schemes in scenarios with a high number of receive antennas.

Furthermore, in this paper we compare the EE and signal processing (SP) complexity of the proposed strategy with the conventional zero forcing (ZF) and minimum mean square error (MMSE) linear detectors. In particular, we show that the improvements offered by the proposed technique allow enhancing the EE achieved with these detectors while reducing their computational complexity. In fact, the detailed complexity analysis leads us to derive interesting conclusions regarding the algorithms that must be used to solve the ZF and MMSE detection problems in large-scale SM-MIMO systems. To summarize, the contributions of this paper can be stated as follows:

- We propose and validate the use of CS-based algorithms as an efficient alternative to recover SM signals under large-scale MIMO conditions. Moreover, we design a CS-based detector specifically tailored for the MAC of SM systems to improve the performance and the EE of the detectors conventionally used for large-scale MIMO.
- As opposed to the previous literature, we perform an analytical characterization of the computational complexity of the iterative CS algorithm to determine the conditions in which the use of CS-based detection in SM is especially beneficial.
- We carry out a mathematical analysis of the performance and convergence of the CS greedy algorithms when applied to the proposed large-scale MAC scenario.
- We use the above complexity and performance analyses to characterize the EE of the proposed, which is shown to be superior to that of the existing linear detectors.

II. Preliminaries

A. System Model of the Multiple Access Channel (MAC)

The model considered throughout this paper characterizes the MAC of a multi-user MIMO system comprised of K mobile stations (MSs) with n_t antennas each, and a single BS with N receive antennas. The total number of antennas allocated at the MSs is denoted as $M = K \times n_t$. The behavior of the multiple access system can be described by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{N\times 1}$ is the signal received at the BS, and $\mathbf{x} \in \mathbb{C}^{M\times 1}$ denotes the signal transmitted by the MSs. Moreover, $\mathbf{w} \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I}_N)$ denotes the standard additive white-Gaussian noise vector with variance σ_n^2, and $\mathbf{H} \in \mathbb{C}^{N\times M} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{I}_M)$ is a matrix whose m, n-th complex coefficient, $h_{m,n}$, represents the frequency flat fading channel gain between the n-th transmit antenna and the m-th receive antenna. In the previous expressions, $\sim$ indicates "distributed as", $\mathbf{I}_A$ is an $A \times A$ identity matrix, and $\otimes$ denotes the Kronecker product [30]. As typically assumed in the SM detection literature, the BS is expected to have a perfect knowledge of the communication channel $\mathbf{H}$ [3], [4].

Throughout this paper we assume that the data symbols transmitted by the active antennas belong to a normalized Q-QAM constellation satisfying $\mathbb{E}[E_s] = 1$, where E_s refers to the symbol energy and the operator $\mathbb{E}$ denotes the expected value. Based on this, the total average signal-to-noise ratio (SNR) of the MAC ρ can be expressed as

$$\rho = \frac{\mathbb{E}[\mathbf{x}^H\mathbf{x}]}{\sigma_n^2} = \frac{S \cdot \mathbb{E}[E_s]}{\sigma_n^2}, \tag{2}$$

where $(\cdot)^H$ denotes the Hermitian transpose and $S \leq M$ is the total number of antennas simultaneously active amongst the MSs.

B. Multiple Access Spatial Modulation

SM and its generalized version reduce the hardware complexity of multiple-antenna devices by limiting the number of active antennas per user and conveying additional information onto their spatial position [3], [9]. In this section we focus on describing the operation of generalized SM transmission with a single RF chain since particularization to conventional SM is straightforward by letting $S = K$ and forcing the number of

antennas per user to be a power of two [9], [31]. Throughout this paper we use the term SM when referring to both conventional and generalized SM for ease of description. In SM, each transmitter conveys the same constellation symbol by activating a given number of antennas n_a according to the input bit sequence [3], [9], [31]. Without loss of generality, in the following we assume that the users activate the same number of antennas, i.e. we have $S = n_a \times K$. Mathematically, the transmit signal $\mathbf{x}_u \in \mathbb{C}^{n_t \times 1}$ of the u-th SM transmitter can be expressed as [9]

$$\mathbf{x}_u = \left[0 \ \cdots \ s_{l_1}^q \ \cdots \ s_{l_k}^q \cdots \ 0\right]^T, \tag{3}$$

where $l_k \in [1, n_t]$ denotes the active antenna index and s^q represents the q-th symbol of the transmit constellation $\mathcal{B}$. The number of bits that can be encoded on the antenna indices is $\lfloor \log_2 \binom{n_t}{n_a} \rfloor$ [8], [31]. In the previous expression, $\binom{\cdot}{\cdot}$ denotes the binomial operation and $\lfloor \cdot \rfloor$ is the floor function. Therefore, the number of possible antenna combinations at the transmitter is given by $r = 2^b$, where $b \triangleq \lfloor \log_2 \binom{n_t}{n_a} \rfloor$. This determines the cardinality of $\mathcal{A}$, the set comprised of the possible antenna groups with $\mathcal{A}_l$ being the l-th element. Note that there may be invalid antenna groups to preserve an integer length of the bit stream. The composite transmit vector $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is obtained by concatenating the transmit signals as $\mathbf{x} = [\mathbf{x}_1^T \mathbf{x}_2^T \cdots \mathbf{x}_K^T]^T$.

When compared with conventional MIMO technologies, SM reduces the inter-user interference and the circuit power consumption of the MSs for the same number of antennas. However, this causes a reduction of the achievable rates with respect to conventional MIMO for the same number of antennas [3]. Particularly, while a conventional MIMO transmitter is able to convey $B_{\text{MIMO}} = n_t \cdot \log_2(Q)$ bits, a single SM transmitter encodes $B_{\text{SM}} = b + \log_2(Q)$ bits in every channel use. At the receiver, detection schemes exploit the channel knowledge to determine the active antennas and the conveyed constellation symbols [3]. Among these, the optimum detector follows the ML criterion and its output reads as

$$\hat{\mathbf{x}} = \arg\min_{\tilde{\mathbf{s}}_p} \left\| \mathbf{y} - \mathbf{H}\tilde{\mathbf{s}}_p \right\|_2^2. \tag{4}$$

Here, the signal $\tilde{\mathbf{s}}_p \in \mathbb{C}^{M \times 1}$ belongs to the set that includes all the possible transmit signals Γ and $\| \cdot \|_p$ denotes the ℓ_p norm. The cardinality of Γ, $|\Gamma| = (Q \times r)^K$, exponential with the number of users K, establishes an upper bound on the complexity of SM detection.

C. Large-Scale MIMO and Low-Complexity Detection

The large-scale MIMO theory focuses on analyzing the benefits of communication systems with a high number of antennas at the BSs [4], [5]. One of the fundamental results in this field states that, provided that $N \gg M$, the received signal after linear detection $\mathbf{g} \in \mathbb{C}^{M \times 1}$ satisfies

$$\mathbf{g} = \mathbf{D}(\mathbf{H}\mathbf{x} + \mathbf{w}) \xrightarrow[N \to \infty,\ M = \text{const.}]{\text{a.s.}} \mathbf{x}, \tag{5}$$

where $\mathbf{D} \in \mathbb{C}^{M \times N}$ is a linear detection matrix that for the matched filter (MF), ZF and MMSE detectors read as [32]

$$\mathbf{D}^{\text{MF}} = \mathbf{H}^H, \tag{6}$$

$$\mathbf{D}^{\text{ZF}} = \mathbf{H}^\dagger = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H, \tag{7}$$

$$\mathbf{D}^{\text{MMSE}} = (\mathbf{H}^H\mathbf{H} + \varsigma\mathbf{I})^{-1}\mathbf{H}^H. \tag{8}$$

In the above, $\varsigma = M/\rho$ [32], $(\cdot)^\dagger$ denotes the pseudoinverse operation and $(\cdot)^{-1}$ refers to the inverse matrix. Let $\mathbf{g}^{\{u\}}$ be the decision vector corresponding to the u-th user. In the following we adopt the sub-optimal but low-complexity approach of decoupling the estimation of the spatial and amplitude-phase modulated symbols [22], [33]. Specifically, the estimated active antenna indices $\hat{\mathcal{A}}_l$ and the transmitted constellation symbol $\hat{q}$ for the u-th user are obtained from (5) as

$$\hat{\mathcal{A}}_l = \arg\max_l \left\| \mathbf{g}_{A_l}^{\{u\}} \right\|_2, \tag{9}$$

$$\hat{q} = \mathcal{D}\left(\mathbf{g}_{\hat{A}_l}^{\{u\}}\right), \tag{10}$$

where $g_{\{\mathcal{A}_l, \hat{\mathcal{A}}_l\}}^{\{u\}}$ represent the entries of the decision vector of the u-th user $\mathbf{g}^{\{u\}}$ determined by the sets $\mathcal{A}_l$ and $\hat{\mathcal{A}}_l$ respectively, and $\mathcal{D}$ denotes the demodulation function.

Note that the necessary increase in the number of transmit antennas when SM is used degrades the performance of the above-mentioned detectors due to the worse conditioning of the channel matrix. For this reason, in this paper we propose a low-complexity solution inspired by CS to take advantage of the large-scale MIMO benefits while, at the same time, reducing the detection complexity. In other words, in the following we look at scenarios where $N \gg M$ does not necessarily hold but $N \gg K$ brings the massive MIMO effect.

III. The Trivial Approach: Direct Application of the CS Algorithms for SM Detection

The main issue with the conventional ZF and MMSE linear detectors when applied to SM and generalized SM detection is that the entire channel matrix $\mathbf{H} \in \mathbb{C}^{N \times M}$ must be used for detection even though only S columns contribute to the acquisition of the amplitude-phase signal information. We circumvent this by exploiting the sparsity of SM signals to reduce the complexity of linear detectors. The signals conveyed by SM are defined as S-sparse because they only contain $S \ll M$ non-zero entries equal to the number of antennas simultaneously active S [18]. This property has been exploited by CS to improve signal estimation from compressive measurements. Specifically, CS capitalizes on signal sparsity to guarantee a reliable signal recovery with efficient algorithms [34], [35]. The CS measurements $\mathbf{y} \in \mathbb{R}^{N \times 1}$ of a sparse signal $\mathbf{x}$ can be expressed as [35]–[37]

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{e}, \tag{11}$$

where $\mathbf{x} \in \mathbb{R}^{M \times 1}$ represents the original sparse signal, $\mathbf{\Phi} \in \mathbb{R}^{N \times M}$ is the measurement matrix, and $\mathbf{e} \in \mathbb{R}^{N \times 1}$ is a

measurement error term. Note that the complex-valued system in (1) can be straightforwardly re-expressed to resemble the real-valued one in (11) [18]. In this paper, we exploit the similarity between (1) and (11) to improve the detection performance of the conventional linear MIMO detectors.

In CS, the restricted isometry property (RIP) determines whether signal recovery guarantees are fulfilled or not for any communication channel $\mathbf{\Phi} = \mathbf{H}$ [35], [36]. For the case of the MIMO channel, the RIP of order S is satisfied for a channel matrix $\mathbf{H}$ if, for any S-sparse signal $\mathbf{x}$, the relationships

$$(1 - \delta_S)\|\mathbf{x}\|_2^2 \leq \|\mathbf{H}\mathbf{x}\|_2^2 \leq (1 + \delta_S)\|\mathbf{x}\|_2^2 \tag{12}$$

hold for a constant $\delta_S \in (0, 1)$. For instance, a matrix comprised of independent and identically distributed Gaussian random variables is known to satisfy $\delta_S \leq 0.1$ provided that $N \geq cS\log(M/S)$, with c being a fixed constant [36]. Note that this kind of channel communication matrix conventionally arises in rich scattering environments with Rayleigh fading [4].

Once the signal measurements are acquired and contrary to the ML detector given in (4), the detection of the SM signals in CS relies on the sparsity of SM transmission to generate an estimate. For the case with $\delta_{2S} < \sqrt{2} - 1$ [37] we solve (4) in a low-complexity CS-inspired fashion as

$$\begin{aligned} &\text{minimize } \|\mathbf{x}\|_1 \\ &\text{subject to } \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2 \leq \mu, \end{aligned} \tag{13}$$

In the above expression the constant μ limits the noise power $\|\mathbf{w}\|_2 \leq \mu$. Although the above optimization problem can be solved with well-known convex approaches, these alternatives are often computationally intensive, so faster techniques that offer a trade-off between performance and complexity such as greedy algorithms are commonly used instead [36].

From the vast variety of CS greedy algorithms, in this section we choose one of the most efficient schemes to approximate the solution of (13): the Compressive Sampling Matching Pursuit (CoSaMP) [38]. The CoSaMP is a low-complexity algorithm that goes through an iterative reconstruction process to recover both the active antenna indices and the amplitude-phase information of the transmitted signals. Moreover, this algorithm provides optimal error guarantees for the detection of sparse signals since, similarly to the more complex convex algorithms, a stable signal recovery is guaranteed under noisy conditions with a comparable number of receive antennas [38]. In our results we show that the large number of antennas at the BS benefits the use of this algorithm for SM detection by performing a thorough complexity analysis as opposed to [18], [38]. Moreover, as the structure of the transmitted signals in the MAC is not accounted for in the generic CS detection, in the following we consider an approach specifically tailored for the considered scenarios to improve the detection performance.

IV. Proposed Enhanced CS Technique: Spatial Modulation Matching Pursuit (SMMP)

One of the key characteristics of the conventional greedy CS algorithms is that no prior knowledge of the sparse signal other than the number of non-zero entries is assumed. However, when applied to the proposed scenario, this condition can generate situations in which the output of the detector does not have physical sense. For instance, the detected signal could have more than one active antenna per user, which is not possible when conventional SM modulation is used [3]. This undesired operating condition is caused by the noise and inter-user interference effects that arise in the MAC. To mitigate these, in this sub-scheme we incorporate the additional prior knowledge about the distribution of the non-zero entries in the transmitted signal to further enhance performance [28].

Algorithm 1 Spatial Modulation Matching Pursuit

Inputs: $\mathbf{H}$, $\mathbf{y}$, S, n_a, i_{max}.
Output: $\tilde{\mathbf{x}}^{i_{end}} \triangleq S$-sparse approximation
$\tilde{\mathbf{x}}^0 \leftarrow \mathbf{0}$, $i \leftarrow 0$ {Initialization}
while halting criterion *false* **do**
  $\mathbf{r} \leftarrow \mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}^i$ {Update residual}
  $i \leftarrow i + 1$
  $\mathbf{c} \leftarrow \mathbf{H}^H\mathbf{r}$ {MF to estimate active antenna indices}
  {7–11: Detect n_a indices with highest energy per user}
  $\Omega \leftarrow \varnothing$
  for $j = 1 \rightarrow K$ **do**
    $\mathcal{M} \leftarrow \{(j-1)\cdot n_t, \ldots, j\cdot n_t - 1\}$
    $\Omega \leftarrow \mathcal{M}(\arg\max\{|\mathbf{c}|_{\mathcal{M}}|\}_{n_a}) \cup \Omega$
  end for
  {12–13: Detect remaining $k - S$ highest-energy indices}
  $\mathbf{c}(\Omega) \leftarrow \mathbf{0}$
  $\Omega \leftarrow \arg\max\{|\mathbf{c}|\}_{(k-S)} \cup \Omega$
  $\mathcal{T} \leftarrow \Omega \cup \text{supp}(\tilde{\mathbf{x}}^{i-1})$ {Merge supports}
  $\mathbf{b}|_{\mathcal{T}} \leftarrow \mathbf{H}_{\mathcal{T}}^{\dagger}\mathbf{y}$ {Least squares problem}
  $\mathbf{b}|_{\mathcal{T}^C} \leftarrow \mathbf{0}$
  {17–21: Obtain next signal approximation}
  $\tilde{\mathbf{x}}^i \leftarrow \mathbf{0}$
  for $j = 1 \rightarrow K$ **do**
    $\mathcal{M} \leftarrow \{(j-1)\cdot n_t, \ldots, j\cdot n_t - 1\}$
    $\tilde{\mathbf{x}}^i|_{\mathcal{M}(\arg\max\{|\mathbf{b}|_{\mathcal{M}}|\}_{n_a})} \leftarrow \max\{|\mathbf{b}|_{\mathcal{M}}|\}_{n_a}$
  end for
end while

The detection algorithm considered in this paper is referred to as spatial modulation matching pursuit (SMMP) to explicitly indicate that it corresponds to a particularization to SM operation of the structured CoSaMP iteration developed in [29]. In particular, SMMP reduces the errors in the identification of the active antennas by exploiting the known distribution of the non-zero entries [28], [29]. This also improves the convergence speed of the algorithm as less iterations are required to determine the active antenna indices. While it is intuitive that the concept behind this strategy could be incorporated to other CS detection algorithms, in the following we focus on the CoSaMP algorithm for clarity.

The pseudocode of the proposed strategy is shown in Algorithm 1 for convenience and its operation can be described as follows [29], [38]. The algorithm starts by producing an estimate of the largest components of the transmitted signal to

identify the active antennas [38]. For this, the algorithm employs the residual signal $\mathbf{r} \in \mathbb{C}^{N\times 1}$ given by

$$\mathbf{r} \triangleq \mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}^i = \mathbf{H}(\mathbf{x} - \tilde{\mathbf{x}}^i) + \mathbf{w}, \tag{14}$$

where $\tilde{\mathbf{x}}^i \in \mathbb{C}^{M\times 1}$ is the approximation of the transmit signal at the i-th iteration. The residual signal concentrates the energy on the components with a largest error in the estimated received signal $\tilde{\mathbf{y}} = \mathbf{H}\tilde{\mathbf{x}}^i$ [25], [39]. The decision metric used to determine the plausible active antennas $\mathbf{c} \in \mathbb{C}^{M\times 1}$ is obtained as the output of a MF and can be expressed as

$$\mathbf{c} = \mathbf{H}^H\mathbf{r}. \tag{15}$$

From this decision metric, the active antenna estimation process forms a set Ω of decision variables with cardinality $|\Omega| = k \geq S$. The set Ω provides an estimate of the plausible active antennas. Note that these entries do not have to correspond to the S coefficients with highest estimated energy in the transmitted signal as in the CoSaMP algorithm [38]. For instance, let $\mathbf{x} = [0, 0, -1, 0|0, 1, 0, 0]$ be the signal transmitted in a MAC with $K = S = 2$ users with $n_t = 4$ antennas each and BPSK modulation. The total number of transmit antennas is $M = n_t \times K = 8$. Then, the set Ω in the conventional CoSaMP algorithm could be formed by any $S = 2$ entries of the set of integers $\Upsilon_{\text{CoSaMP}} = \{1, 2, \ldots, 8\}$ without any restriction. This contrasts with the considered SMMP algorithm, in which the set Ω is formed by at least $n_a = 1$ entry of the set $\Upsilon^1_{\text{SMMP}} = \{1, 2, 3, 4\}$ and another one from $\Upsilon^2_{\text{SMMP}} = \{5, 6, 7, 8\}$. In simple terms, the proposed algorithm forces that at least n_a entries per user are selected as we exploit the knowledge that

$$\|\mathbf{x}_k\|_0 = n_a, \quad k \in [1, K], \tag{16}$$

where $\|\cdot\|_0$ determines the number of non-zero entries [34], [36]. This is represented in lines 7–11 of Algorithm 1, where the $\arg\max\{\cdot\}_p$ and $\max\{\cdot\}_p$ functions return the indices and the entries of the p components with largest absolute value in the argument vector, and $\varnothing$ denotes the empty set. The remaining $k - S$ entries are instead selected as the entries with highest energy independently of the user distribution. These additional entries aim at improving the support detection process by involving the LS problem, which offers an enhanced performance in the antenna identification with respect to the estimate provided by the MF [4].

Once the entries with highest error energy in the current residual have been estimated, the set $\mathcal{T}$ is obtained as

$$\mathcal{T} \triangleq \Omega \cup \text{supp}(\tilde{\mathbf{x}}^{i-1}), \tag{17}$$

where $\text{supp}(\cdot)$ identifies the indices of the non-zero entries. This set provides a final estimate of the plausible active antennas used for transmission by incorporating the ones considered in the previous iteration [39]. Therefore, the set $\mathcal{T}$ determines the columns of the matrix $\mathbf{H}$ used to solve the unconstrained least squares (LS) problem given by

$$\underset{\mathbf{b}|_{\mathcal{T}}}{\text{minimize}} \ \|\mathbf{H}_{\mathcal{T}}\mathbf{b}|_{\mathcal{T}} - \mathbf{y}\|_2^2 \rightarrow \mathbf{b}|_{\mathcal{T}} = \mathbf{H}^{\dagger}_{\mathcal{T}}(\mathbf{H}\mathbf{x} + \mathbf{w}), \tag{18}$$

where $\mathbf{b}|_{\mathcal{T}}$ denotes the entries of $\mathbf{b} \in \mathbb{C}^{M\times 1}$ supported in $\mathcal{T}$. This notation differs from $\mathbf{H}_{\mathcal{T}}$, which refers to the submatrix obtained by selecting the columns of $\mathbf{H}$ determined by $\mathcal{T}$.

The LS approximation is a crucial step as the complexity reduction and the performance improvement w.r.t. the conventional linear alternatives depend on the efficiency of this process [38]. This procedure can also be seen as implementing a ZF detector in which, instead of inverting all the columns of the channel matrix, only the columns that have been previously included in the support are inverted. This allows exploiting the large-scale MIMO detection benefits as the equivalent ZF detector generally satisfies $N \gg 2k$ [5]. Due to the possibility of using different algorithms to solve this problem, a detailed analysis of their complexity is developed in Section V-A.

After solving the LS problem, the signal approximation of the i-th iteration $\tilde{\mathbf{x}}^i$ is built by selecting the entries with highest energy at the output of the LS problem based on a user-by-user criterion following (16). Finally, the sparse output of the algorithm $\tilde{\mathbf{x}}^{i_{end}}$ is obtained after the algorithm reaches the maximum pre-defined number of iterations i_{max} or a halting criterion is satisfied [38]. Overall, although sub-optimal for a large but finite number of antennas, the proposed scheme exploits the high performance offered by linear detection schemes together with the structured sparsity inherent to SM transmission to reduce complexity. We also note that, although shown via iterative structures in Algorithm 1, the additional operations required in the considered algorithm can be implemented via vector operations with reduced computational time.

Regarding the compromise between complexity and performance of the considered technique, it should be noted that several parameters can be modified to adjust this trade-off [38]:

- *Maximum number of iterations of SMMP* (i_{max}): The total number of iterations determines the complexity and detection accuracy of the algorithm. This parameter can be used to adjust the performance depending on the computational capability of the BS as shown hereafter.
- *Number of entries detected at the output of the MF* (k): The parameter k determines the dimensions of the LS problem, hence severely affecting the SP complexity. It also influences the detection performance since the solution is conditioned by the LS matrix $\mathbf{H}_{\mathcal{T}}$.
- *Maximum number of iterations of the iterative LS* (i^{ls}_{max}): The accuracy of the LS solution is improved in every iteration when iterative algorithms are used [41]. Hence, there exists a trade-off between complexity and performance that can be optimized at the BS depending on the communication requirements. As this parameter greatly affects the global complexity, a detailed study of the required number of iterations is developed in Section VI-A.

At this point we also point out that the main difference of the proposed approach with respect to the typical CS is that we also consider over-determined scenarios, i.e. communication systems where $N \geq M$. As shown hereafter, this entails that SMMP is not only able to reduce the complexity of the conventional linear detectors in scenarios where $N \geq M$, but it also provides performance guarantees when this condition does not hold.

TABLE I
COMPLEXITY IN NUMBER OF REAL FLOATING-POINT OPERATIONS (FLOPS) TO SOLVE A $m \times n$ LEAST SQUARES PROBLEM

Operation	*Complexity in flops*
$\mathbf{H}^{\dagger}\mathbf{y}$ via QR decomposition [38], [39] • QR decomposition • Intermediate computations and back substitution	$C_T \simeq 8n^2m - \left(\frac{8}{3}\right)n^3 + 8mn + 4n^2$ • $C_1 = 8n^2m - \left(\frac{8}{3}\right)n^3$ • $C_2 = 4n^2 + 8mn$
$\mathbf{H}^{\dagger}\mathbf{y}$ via Cholesky decomposition [39], [40] • $C = \mathbf{A}^H\mathbf{A}$ • Cholesky decomposition • Intermediate computations, forward and backward substitutions	$C_T \simeq 4n^2m + \left(\frac{4}{3}\right)n^3 + 8mn + 11n^2$ • $C_1 = 4n^2m$ • $C_2 = \left(\frac{4}{3}\right)n^3 + 3n^2$ • $C_3 = 8n^2 + 8mn$
$\mathbf{H}^{\dagger}\mathbf{y}$ via Richardson Iteration [36], [38], [39] • SVD to compute optimal α • Richardson iterations	$C_T \simeq \left(\frac{16}{3}\right)n^3 + i^{ls}_{max} \times (24mn + 6n)$ • $C_1 = \left(\frac{16}{3}\right)n^3$ • $C_2 = i^{ls}_{max} \times (24mn + 6n)$
$\mathbf{H}^{\dagger}\mathbf{y}$ via Conjugate Gradient [36], [39] • Algorithm without initial approximation • Algorithm with initial approximation • Conjugate Gradient iterations	$C_T \simeq C_{1,\{1,2\}} + i^{ls}_{max} \times (16mn + 27n)$ • $C_{1,1} = 8mn + 7n$ • $C_{1,2} = 16mn + 18n$ • $C_2 = i^{ls}_{max} \times (16mn + 27n)$

V. COMPLEXITY ANALYSIS

The analysis of the computational complexity is commonly performed by determining the complexity order. This is the approach adopted in [18], [23], [38] to determine the complexity of the proposed algorithms. Instead, in this paper we adopt a more precise approach since, as shown in the following, the complexity order does not provide an accurate characterization of the total number of operations due to the iterative nature of the greedy CS-based algorithms. In fact, as opposed to the results obtained in [18] and [38], here we conclude that the LS problem can dominate the global complexity.

A. Complexity of the Least Squares Problem

An efficient implementation of the LS algorithm is required to reduce the complexity of the proposed approach since otherwise it can dominate the total number of floating point operations (flops) [38]. In general, the methods to solve the LS problem can be classified according to their approach to obtain the solution into direct and iterative procedures [41]:

- Direct methods include the QR and the Cholesky decompositions and they are based on producing a system of equations that can be easily solved via backward and forward substitutions. The total number of flops is conditioned by the costly decompositions that must be performed at the beginning of each coherence period [43].
- Iterative methods solve the LS problem by refining an initial solution based on the instantaneous residual [41]. These approaches prevent the storage intensive decompositions required by direct methods, an aspect especially beneficial in the proposed scenario due to the large dimensions of the channel matrix $\mathbf{H}$.

The number of flops of the QR decomposition, Cholesky decomposition, Richardson iteration and conjugate gradient (CG) LS algorithms are detailed in Table I. Note that the fact that each complex calculus involves the computation of several real operations have been taken into account in Table I. Moreover, for simplicity, it has been considered that a real product (division) has the same complexity of a real addition (subtraction), a usual assumption in the related literature [40].

Regarding the complexity of the Richardson iteration, note that it depends on a parameter $0 < \alpha < 2/\lambda^2_{\max}(\mathbf{A})$ that determines the convergence rate of the algorithm [41]. In the last expression, $\lambda_{\max}(\mathbf{A})$ denotes the maximum singular value of an arbitrary matrix $\mathbf{A}$. The determination of this parameter becomes necessary if a high convergence speed is required and these additional operations are included in Table I. Concerning the CG algorithm, Table I shows that the complexity differs depending on the availability of an initial approximation to the LS solution [41]. This difference must be considered because, as opposed to the other detectors, the complexity reduction of the greedy CS-based approach is based on the increasing accuracy of the approximations as the algorithm evolves, hence improving the convergence speed as explained in Section IV.

The comparison between the direct LS methods shows that, even though the QR decomposition is more numerically accurate, it also is more complex than the Cholesky decomposition [43]. Therefore, the Cholesky decomposition is preferred for the considered scenario due to the high dimensions of the matrices involved in the LS problem [43]. Regarding the complexity of the iterative methods, the results in Table I describe the reduction of the complexity order when compared to the direct methods pointed out in [38]. However, this improvement does not guarantee a reduction in the number of operations as the total complexity is highly dependent on the number of iterations required until convergence as shown in Section VIII.

B. Overall Complexity of the Proposed Algorithm

Based on the above analyses, a tight upper bound on the total number of real flops of the ZF and SMMP detectors is shown in Table II. The complexity of the ZF detector is solely determined by the operations required to solve the LS problem. Therefore, the increase in the number of transmit antennas of SM may have a significant impact on the total complexity. This is because, while $M = K$ for the conventional large-scale MIMO scenarios with single-antenna users, the relationship $M = K \times n_t$ is satisfied for SM transmitters in the MAC. However, the performance improvements offered by SM justify the complexity increase at the BSs, where computational resources are expected to be available [20].

TABLE II
COMPLEXITY IN NUMBER OF REAL FLOATING-POINT OPERATIONS (FLOPS) OF DIFFERENT LARGE-SCALE MIMO DETECTORS

Detector	*Complexity in flops*
Zero Forcing detector • QR factorization • Cholesky factorization • Richardson iteration • Conjugate gradient	• $C_{\text{QR}} \simeq 8M^2N - \left(\frac{8}{3}\right) M^3 + 8MN + 4N^2$ • $C_{\text{Chol}} \simeq 4M^2N + \left(\frac{4}{3}\right) M^3 + 8MN + 11M^2$ • $C_{\text{Rich}} \simeq \left(\frac{16}{3}\right) M^3 + \left[i_{max}^{ls} \times (24MN + 6M)\right]$ • $C_{\text{CG}} \simeq (8MN + 7M) + \left[i_{max}^{ls} \times (16MN + 27M)\right]$
Spatial Modulation Matching Pursuit • Matched filter • Least Squares problem – First SMMP iteration – Rest of SMMP iterations • Compute residual	$C_{\text{SMMP}} \simeq i_{max} \times (8MN + 8SN) + C_{ls}^{1} + \left[(i_{max} - 1) \times C_{ls}^{i>1}\right]$ • $C_{\text{MF}} = 8MN$ • Least Squares problem (C_{ls}) – C_{ls}^{1}. C_T from Table I with $m = N$, $n = k$ and $C_{1,1}$ for CG. – $C_{ls}^{i>1}$. C_T from Table I with $m = N$, $n = (k + S)$ and $C_{1,2}$ for CG. • $C_{\text{res}} \simeq 8SN$

TABLE III
COMPLEXITY AND BER OF A BASE STATION WITH $N = 128$, $K = S = 12$, $n_t = 8$, $i_{max} = 2$ AND SNR = 5 dB

Detection scheme	*Complexity*	*BER*
ZF - Cholesky decomposition	6098 kflops	2.1×10^{-2}
ZF - Iterative CG detection	696.5 kflops	1×10^{-3}
SMMP - Cholesky decomposition	655.3 kflops	1.12×10^{-5}
SMMP - Iterative CG detection	507 kflops	2.08×10^{-5}

Additionally, we point out that the flop calculation for the proposed algorithm has been performed for the worst-case scenario in which none of the S entries from the previous SMMP iteration coincide with the k coefficients selected at the output of the MF. This means that the average number of operations is generally smaller than the one shown in Table II. In spite of this, this tight upper bound allows us to determine the conditions under which CS-based detection is convenient.

The total number of real flops and the BERs for a specific large-scale MIMO scenario are represented in Table III as an illustrative example. The relevant system parameters are $N = 128$, $K = S = 12$, $n_t = 8$, $i_{max} = 2$ and SNR = 5 dB. The number of CG iterations i_{max}^{ls} is set to ensure the maximum attainable performance. Overall, it can be seen that the SMMP algorithm offers a considerable performance improvement with reduced complexity. The results of Table III, which are further developed in Section VIII, also lead us to conclude that the use of iterative LS algorithms for conventional linear detection may be able to reduce the detection complexity over a channel coherence time. Note that the solution of the LS problem is not exact in these cases, which influences the resulting BERs as shown in Table III. This results in a performance improvement for the conventional ZF detector due to the energy concentration on the components that correspond to the active antennas. Indeed, the significant complexity benefits attained by the use of iterative LS algorithms motivate the study of their convergence speed in the following section.

VI. CONVERGENCE AND PERFORMANCE ANALYSES

A. Convergence Rate of the Iterative LS in Large-Scale MIMO

The complexity of the proposed algorithm highly depends on the number of iterations required to obtain a solution to the LS problem. Although the required number of iterations to achieve convergence cannot be known in advance due to the random nature of the communication channel, in this section we derive two expressions to determine the number of iterations depending on the required output accuracy: a straightforward but less accurate one based on an asymptotic analysis, and a more complex and precise one that resorts to the cumulative distribution function (CDF) of the condition number. This offers a more refined and intuitive approach than the one adopted in [38] as the resultant expressions directly depend on the number of antennas of the communication system. Furthermore, the following study can be used to determine whether it is convenient from a numerical perspective to use iterative LS algorithms to solve the ZF problem when applied to large-scale MIMO scenarios without SM.

The convergence rate of the iterative LS methods is determined by the condition number of the LS matrix $\mathbf{H}_{\mathcal{L}} \in \mathbb{C}^{N \times |\mathcal{L}|}$ [38], [41]. Throughout this section, $\mathcal{L}$ is a set defined as $\mathcal{L} \triangleq \mathcal{T}$ when the SMMP algorithm is used whereas it is given by $\mathcal{L} \triangleq \{1, 2, \ldots, M\}$ for ZF. The standard condition number $\Theta \in [1, \infty]$ of the matrix $\mathbf{H}_{\mathcal{L}}$ is defined as [41], [44]

$$\Theta(\mathbf{H}_{\mathcal{L}}) = \frac{\lambda_{\max}(\mathbf{H}_{\mathcal{L}})}{\lambda_{\min}(\mathbf{H}_{\mathcal{L}})} = \sqrt{\frac{\sigma_{\max}\left(\mathbf{H}_{\mathcal{L}}^{H}\mathbf{H}_{\mathcal{L}}\right)}{\sigma_{\min}\left(\mathbf{H}_{\mathcal{L}}^{H}\mathbf{H}_{\mathcal{L}}\right)}} = \sqrt{\Xi(\mathbf{W})}, \quad (19)$$

where $\lambda_{\max}$ and $\lambda_{\min}$ denote the maximum and minimum singular values of the argument matrix respectively, $\sigma_{\max}$ and $\sigma_{\min}$ denote the maximum and minimum eigenvalues, and $\Xi(\mathbf{W})$ refers to the modified condition number defined in [44]. In the last equality, $\mathbf{W} \in \mathbb{C}^{s \times s}$ is a Wishart matrix with t degrees of freedom $\mathbf{W} \sim \mathcal{CW}_s(t, \mathbf{I}_s)$ defined as [44], [45]

$$\begin{cases} \mathbf{W} \triangleq \mathbf{H}_{\mathcal{L}}\mathbf{H}_{\mathcal{L}}^{H}, & \text{if } N \leq |\mathcal{L}|, \\ \mathbf{W} \triangleq \mathbf{H}_{\mathcal{L}}^{H}\mathbf{H}_{\mathcal{L}}, & \text{otherwise.} \end{cases} \quad (20)$$

In the above, $s = \min(N, |\mathcal{L}|)$ and $t = \max(N, |\mathcal{L}|)$.

The conditions of large-scale MIMO greatly favor the use of the iterative algorithms thanks to the reduced difference between the maximum and minimum singular values of the communication channel, which in turn ensures a fast convergence [4], [41]. This effect is enhanced by the proposed algorithm due to the operation with a better conditioned matrix than conventional ZF or MMSE detectors. In other words, even though several LS problems are solved when the SMMP algorithm is used, the overall complexity can be reduced w.r.t. the

conventional ZF or MMSE detectors because the dimensions of the LS matrices are smaller and the convergence speed of the algorithm is faster. In fact, we argue that the effectiveness of the CS greedy algorithms in terms of complexity is maximized when a high number of antennas at the BS is considered due to the faster convergence when compared to conventional linear detectors.

To characterize the above-mentioned convergence rate, in this paper we use the results of conventional and large-scale MIMO theory related to the distribution of the condition number of complex Gaussian and Wishart matrices [44], [45]. Throughout, we focus on the CG algorithm for brevity, but similar conclusions can be derived for other iterative LS methods [38], [41]. In particular, the residual generated by each iteration of the CG algorithm satisfies [41]

$$\left\|\mathbf{b}_i - \mathbf{H}_{\mathcal{L}}^{\dagger}\mathbf{y}\right\|_2 \leq 2 \cdot \varrho^i \cdot \left\|\mathbf{b}_0 - \mathbf{H}_{\mathcal{L}}^{\dagger}\mathbf{y}\right\|_2, \tag{21}$$

where the index i denotes the iteration number, $\mathbf{b}_i$ is the LS approximation at the i-th iteration and ϱ is defined as

$$\varrho = \frac{\Theta(\mathbf{H}_{\mathcal{L}}) - 1}{\Theta(\mathbf{H}_{\mathcal{L}}) + 1}. \tag{22}$$

Based on the above result, an upper bound in the number of CG iterations can be expressed as [41]

$$i^{ls} < \frac{1}{2}\Theta(\mathbf{H}_{\mathcal{L}}) \cdot \log\left(\frac{2}{\varepsilon}\right), \tag{23}$$

where ε is the relative error defined as

$$\varepsilon = \frac{\left\|\mathbf{b}_i - \mathbf{H}_{\mathcal{L}}^{\dagger}\mathbf{y}\right\|_2}{\left\|\mathbf{b}_0 - \mathbf{H}_{\mathcal{L}}^{\dagger}\mathbf{y}\right\|_2}. \tag{24}$$

From the above expressions it can be concluded that the convergence speed of the iterative LS methods increases when a smaller number of entries are selected at the output of the MF in the proposed algorithm. This is because the high number of antennas available at the BS greatly favors the convergence of the CG due to the reduced condition number. This is characterized by the following proposition on the convergence of the CG algorithm in the large antenna number limit.

Proposition 1: An upper bound in the number of iterations required by the LS CG algorithm to achieve a relative error reduction of ε under large-scale MIMO conditions is given by

$$i^{ls} < \frac{1}{2} \cdot \left|\frac{1 + \sqrt{\beta\left(|\mathcal{L}|\right)}}{1 - \sqrt{\beta\left(|\mathcal{L}|\right)}}\right| \cdot \log\left(\frac{2}{\varepsilon}\right), \tag{25}$$

where the function $\beta(V)$, $V \in \mathbb{Z}^+$ is defined as

$$\begin{cases} \beta(V) = N/V & \text{if } N \geq V, \\ \beta(V) = V/N & \text{otherwise.} \end{cases} \tag{26}$$

Proof: The large-scale limit theory establishes that the condition number of a channel matrix $\mathbf{H} \in \mathbb{C}^{N \times k}$ with entries $h_{m,n} \sim \mathcal{CN}(0, 1)$ independent and identically distributed (i.i.d.) converges almost surely in the asymptotic limit of transmit and receive antennas to [45, Theorem 7.3]

$$\Theta\left(\mathbf{H}_{\mathcal{T}}\right) \xrightarrow[N,|\mathcal{L}|\to\infty]{} \frac{1 + \sqrt{1/\beta\left(|\mathcal{L}|\right)}}{1 - \sqrt{1/\beta\left(|\mathcal{L}|\right)}} = \left|\frac{1 + \sqrt{\beta\left(|\mathcal{L}|\right)}}{1 - \sqrt{\beta\left(|\mathcal{L}|\right)}}\right|. \tag{27}$$

Equation (27) provides a useful approximation to determine the maximum condition number of a Rayleigh channel with a high number of receive antennas [4]. This can be seen in Fig. 1, where the CDF of the condition number of a Rayleigh fading channel matrix with $N = 128$ receive antennas is depicted. The number of columns is $|\mathcal{L}| = 8$ and $|\mathcal{L}| = 32$ for Fig. 1(a) and (b) respectively. From the results of this figure it can be concluded that the condition number of the channel matrix is below the bound shown in (27) with a high probability. To conclude the argument, (27) is substituted into (23). □

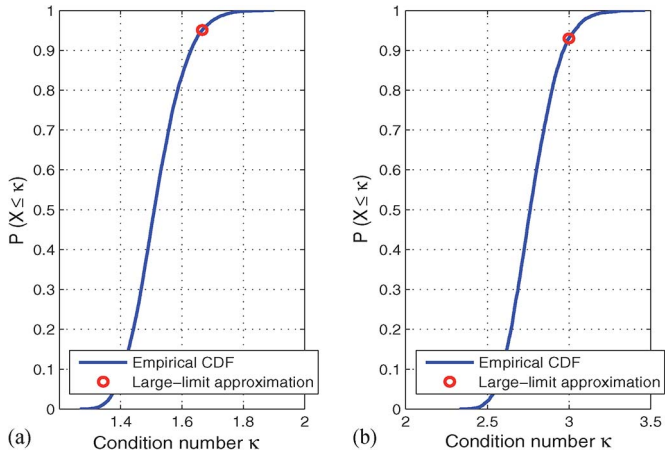

Fig. 1. Empirical CDF and limit condition number for (a) $N = 128$, $|\mathcal{L}| = 8$ and (b) $N = 128$, $|\mathcal{L}| = 32$.

In spite of being valid in a high number of cases, the probability of requiring a higher number of iterations for certain badly conditioned channels cannot be quantified with the above approximation. For this reason, in the following we resort to the analysis of the CDF of the modified condition number $F_{\Xi}(\xi) = P(\Xi \leq \xi)$ with $\xi \geq 1$ developed in [44]. To do so, we first define $\boldsymbol{\sigma} \triangleq [\sigma_1, \sigma_2, \ldots, \sigma_k]$ as the ordered eigenvalues of the Wishart matrix defined in (20) so that $0 < \sigma_1 \leq \ldots \leq \sigma_k$. Moreover, let $\mathbf{V}(\sigma)$ be the Vandermonde matrix with m, n-th entry $v_{m,n} = \sigma_n^{m-1}$ [30, Section 4.6]. The CDF of the modified condition number Ξ of an uncorrelated central Wishart matrix is given by [44, Equation (9)]

$$F_{\Xi}(\xi) = \left[\prod_{m=1}^{k}(s-m)!\prod_{n=1}^{k}(t-n)!\right]^{-1} \sum_{l=1}^{k}\int_0^{\infty} |\Upsilon| d\sigma_k. \tag{28}$$

Here, $|\boldsymbol{\Upsilon}|$ denotes the determinant of the matrix $\boldsymbol{\Upsilon}$, whose m, n-th entry $\upsilon_{m,n}$ is defined as

$$\upsilon_{m,n} = \begin{cases} \gamma(t-k+m+n-1, \xi\sigma_k) & \\ -\gamma(t-k+m+n-1, \sigma_k) & , \quad m \neq l \\ v_{m,n}^2 \sigma_l^{t-k} e^{-\sigma_l} & , \quad m = l \end{cases} \tag{29}$$

where $\gamma(a, b)$ is the lower incomplete gamma function given by $\gamma(a, b) = \int_0^b e^{-t} t^{a-1} dt$. Equations (28) and (29) allows us to

estimate the CDF of the condition number of a communication channel, which in turn is necessary to determine the number of LS iterations as shown in the following. Using the above results, the following theorem can be stated:

Theorem 1: The probability that a given number of LS iterations i^{ls}_{max} suffices to achieve a relative error reduction of ε in the LS CG algorithm is given by

$$P\left(i^{ls} \leq i^{ls}_{max}\right) = F_{\Xi}\left(\left[\frac{2 \cdot i^{ls}_{max}}{\log\left(\frac{2}{\varepsilon}\right)}\right]^2\right). \tag{30}$$

Proof: The first step to derive (30) is to note that the definition of the condition number Ξ used in [44] varies with respect to the one employed in [41], [45]. Attending to their relationship, which is shown in (19), the CDF of the standard condition number Θ can be expressed as

$$F_{\Theta}(\theta) = P(\Theta \leq \theta) = F_{\Xi}(\theta^2). \tag{31}$$

In plain words, $F_{\Xi}(\theta^2)$ gives the probability that for a given constant θ, the condition number of the LS matrix is below that value. The solution of the above expression can be immediately obtained via numerical integration [44]. Once the CDF of the standard condition number has been characterized, the number of iterations of the CG algorithm in the considered scheme can be determined. Particularly, by using (23) the probability that a given number of LS iterations i^{ls}_{max} achieves a relative error reduction of ε can be expressed as

$$P\left(i^{ls} \leq i^{ls}_{max}\right) = P\left(\Theta \leq \frac{2 \cdot i^{ls}_{max}}{\log\left(\frac{2}{\varepsilon}\right)}\right). \tag{32}$$

The proof is completed by substituting (32) into (31). □

The result of Theorem 1 can be used to show the trade-off between accuracy and complexity of the CG when the number of LS iterations is varied. In other words, (30) characterizes the impact of varying the number of iterations of the CG algorithm in the performance of the ZF and greedy CS detectors.

B. Convergence Rate and Error Analysis of CS-Based Detection in Large-Scale MIMO

The evolution of the error at the i-th iteration between the original signal and the sparse approximation is studied in [38] provided that the assumption $\delta_{4S} < 0.1$ on the restricted isometry constant is satisfied. In this section we derive a more intuitive metric to characterize the error reduction based on the number of antennas used for signal transmission and reception. In particular, we apply the results on the maximum and minimum singular values of a random Gaussian matrix to the analysis developed in [38] to derive the following bound.

Theorem 2: The Euclidean norm of the error between the sparse signal at the i-th iteration of the generic and SMMP CS algorithms and the transmitted signal is upper bounded by

$$\|\mathbf{x} - \tilde{\mathbf{x}}^i\|_2 \leq c_1(N,S)\|\mathbf{x} - \tilde{\mathbf{x}}^{i-1}\|_2 + c_2(N,S)\|\mathbf{w}\|_2. \tag{33}$$

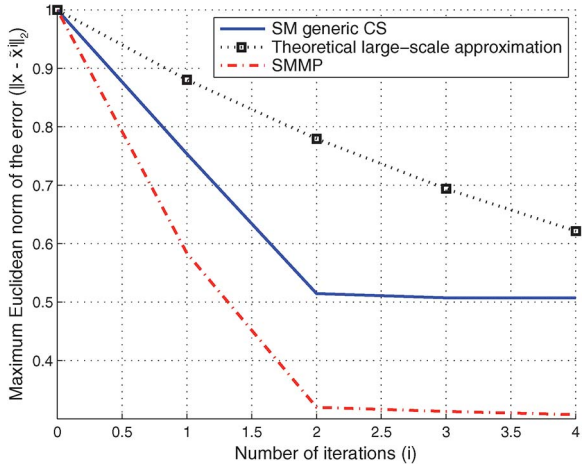

Fig. 2. Theoretical and empirical evolution of the maximum Euclidean norm of the error vs. number of iterations for $N = 128$, $K = 16$, $n_t = 4$, $n_a = 1$, $i_{max} = 4$, $k = 2K$, SNR = 4 dB, 4-QAM and 10^5 channel realizations.

Here, $c_1(N,S)$ and $c_2(N,S)$ depend on the number of active antennas for transmission and reception and are given by

$$c_1 = \left(2 + \frac{T_{4S}\left(\frac{2+4\sqrt{\beta(4S)}}{\beta(4S)}\right)}{T_{3S}\left(1 - \frac{1}{\sqrt{\beta(3S)}}\right)^2}\right) \times \frac{\frac{T_{4S}}{T_{2S}}\left(\frac{1+2\sqrt{\beta(4S)}}{\beta(4S)}\right) + \left(\frac{1+2\sqrt{\beta(2S)}}{\beta(2S)}\right)}{\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2}, \tag{34}$$

$$c_2 = \left(2 + \frac{T_{4S}\left(\frac{2+4\sqrt{\beta(4S)}}{\beta(4S)}\right)}{T_{3S}\left(1 - \frac{1}{\sqrt{\beta(3S)}}\right)^2}\right)\left(\frac{\left(2 + \frac{2}{\sqrt{\beta(2S)}}\right)}{\sqrt{T_{2S}}\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2}\right) + \frac{2}{\sqrt{T_{3S}} - \frac{\sqrt{T_{3S}}}{\sqrt{\beta(3S)}}}, \tag{35}$$

where $T_C \triangleq \max(N, C)$, $C \in \mathbb{R}^+$.

Proof: The proof is shown in Appendix A. □

Note that, as opposed to [38], (40) and (41) show the direct relationship between the norm bounds and the dimensions of the measurement matrix. In other words, the RIP hypothesis is not longer necessary thanks to the use of the random matrix theory results. The bound derived in (33) allows characterizing the error reduction per iteration as a function of the number of receive antennas. In other words, the convergence speed and performance of the algorithm is upper bounded by (33) as a smaller value of the $c_1(N,S)$ will allow obtaining a faster convergence whereas $c_2(N,S)$ characterizes the error floor. This is shown in Fig. 2, where the maximum empirical Euclidean norm of the error for the proposed algorithms is obtained over 10^5 channel realizations and compared with the analytical bound. The theoretical result is obtained by using (33) with a number of receive antennas that guarantees algorithmic convergence [38]. In this figure, it can be seen that the use of a theoretical large-scale approximation allows us to upper

bound the evolution of the empirical maximum ℓ_2 norm of the error for a practical number of iterations. Moreover, Fig. 2 also depicts the faster convergence and reduction in the maximum Euclidean norm of the error offered by the SMMP algorithm w.r.t. the straightforward CS approach.

VII. Energy Efficiency

The study of the EE becomes especially important in the uplink of multi-user scenarios due to the necessity of finding energy-efficient schemes that allow increasing the battery lifetime [1]. In this section, we define the EE model that will be used hereafter to characterize the EE improvements offered by the proposed technique in the MAC when compared to other detection schemes. Towards this end, we express the EE as the rate per milliwatt of total consumed power by using the metric [26], [46]–[52]

$$\epsilon = \frac{S_e}{\sum_{u=1}^{K} P_u} \text{ subject to BER} \leq \text{BER}_{obj}. \tag{36}$$

Here, BER_{obj} is the objective average BER, S_e refers to the spectral efficiency in bits per channel use (bpcu), and P_u is the total power consumption of the u-th MS in milliwatts required to achieve a given BER_{obj}. The total power consumption per MS can be expressed as [47]

$$P_u = P_{Cu} + P_{Tu} = [P_\psi + P_\Psi] + \left[\left(\sum_{j=1}^{n_t} p_{u,j}\right) \cdot \zeta\right]. \tag{37}$$

In the previous expression, $P_{Cu} = P_\psi + P_\Psi$ denotes the total circuit power consumption excluding the power amplifier (PA) and it is divided into two components: P_ψ that represents the circuit power consumption that depends on the number of active antennas, and P_Ψ that corresponds to the static power consumption and it is fixed to a reference value of 5 mW per MS [47]. In particular, P_ψ comprises the additional power consumption required to activate the circuitry of the RF chains and the digital signal processors for transmission. Moreover, $P_{Tu} = \sum_{j=1}^{n_t} p_{u,j} \cdot \zeta$ refers to the power consumption of the PAs and it depends on the factor $\zeta = \frac{\nu}{\eta}$ and the power of the signal that is required to be transmitted by each of the antennas $p_{u,j}$. In the last expression, the factor ν is the modulation-dependent peak to average power ratio (PAPR) and η corresponds to the PA efficiency [46]. Based on the above, the global EE can be expressed as

$$\epsilon = \frac{S_e}{\sum_{u=1}^{K} \left\{[P_\psi + P_\Psi] + \left[\left(\sum_{j=1}^{n_t} \frac{\nu}{\eta} \cdot p_{u,j}\right)\right]\right\}},$$

$$\text{subject to BER} \leq \text{BER}_{obj}. \tag{38}$$

For the simulations of this work, the efficiency factor corresponds to the one of a class-A amplifier, $\eta = 0.35$, which is commonly used in this setting due to the linearity required to transmit QAM signals [46].

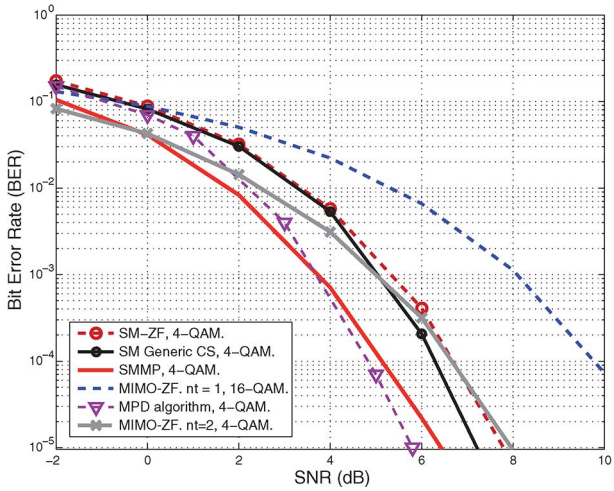

Fig. 3. BER vs. SNR for $N = 128$, $K = S = 16$, $n_t = 4$, $n_a = 1$, $i_{max} = 2$, $k = K$ and $S_e = 64$ bpcu.

VIII. Simulation Results

Monte Carlo simulations have been performed to characterize the performance and complexity improvements of the proposed technique. To maintain the coherence with the SM literature, we compare conventional massive MIMO systems and SM systems with the same spectral efficiency. The performance and energy efficiency results of spatial multiplexing systems with the same number of transmit antennas have not been represented in these figures to avoid congestion and because they exhibit a worse behavior when compared to the spatial multiplexing systems considered here. In the following, SM-ZF and SM-MMSE refer to the linear ZF and MMSE detectors introduced in Section II-C, SM Generic CS denotes the CoSaMP algorithm without the improvement described in Section III, and the techniques with MIMO on their description correspond to those of a conventional MIMO scenario without SM. For reasons of brevity, in this section we select $k = S$ for the generic and SMMP CS algorithms. Note that this decision entails a minimization of the computational complexity for the CS-based algorithms following the results of Table II. Additionally, we also consider the MPD algorithm in spite of its increased computational complexity [23]. The performance results, unless stated otherwise, have been obtained with an exact solution of the LS problem via Cholesky decomposition for clarity.

Fig. 3 characterizes the performance of the above-mentioned detectors in a scenario with $N = 128$, $K = 16$, $n_t = 4$, $n_a = 1$ corresponding to the single active antenna SM, and a resulting spectral efficiency of $S_e = 64$ bpcu. The results of this figure have been obtained with $i_{max} = 2$ iterations of Algorithm 1, which ensures a reduced complexity when compared to the algorithms developed in [23]. Note that an error floor is expected due to the non-feasible solutions produced by the generic CS algorithm. This, however, cannot be appreciated in these results since the high number of antennas considered at the BS moves the error floor to very low BER values [18]. Fig. 3 shows that the proposed algorithm is able to reduce the required transmission power by more than 4 dB w.r.t.

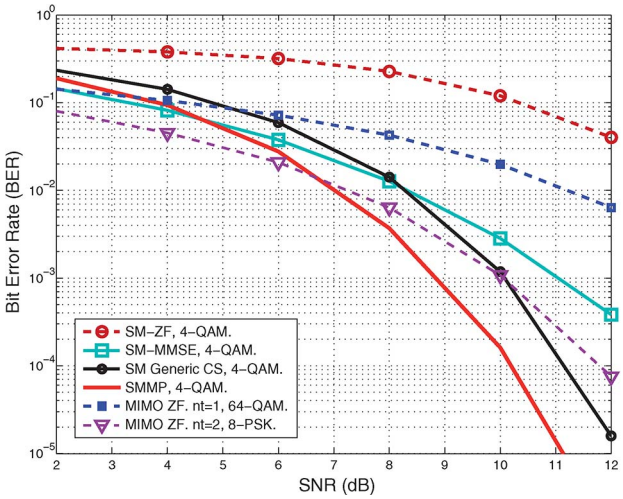

Fig. 4. BER vs. SNR for $N = 128$, $K = 16$, $n_t = 7$, $n_a = 2$, $i_{max} = 3$, $k = S = 2K$ and $S_e = 96$ bpcu.

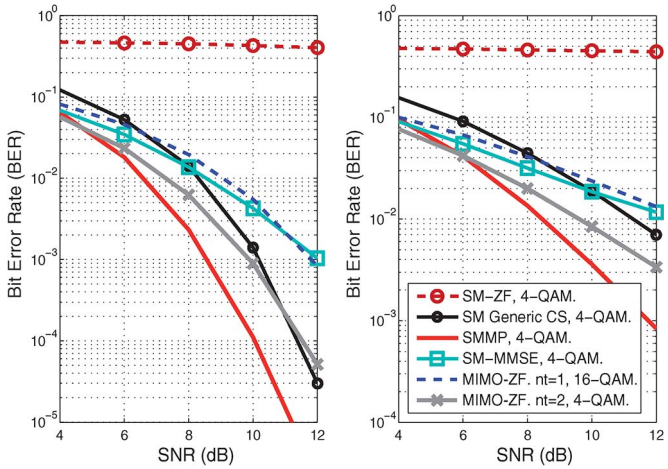

Fig. 5. BER vs. SNR for $N = 128$, $K = S = 32$, $n_t = 4$, $n_a = 1$, $i_{max} = 3$, $k = K$ and $S_e = 128$ bpcu with (a) perfect and (b) imperfect CSI ($\tau = 0.25$).

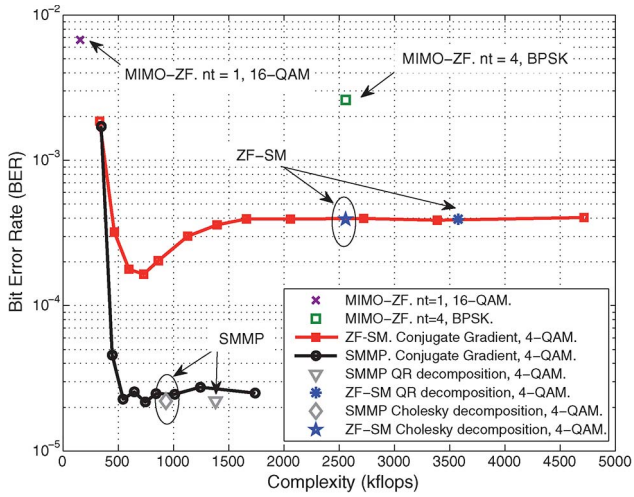

Fig. 6. BER vs. complexity for $N = 128$, $K = S = 16$, $n_t = 4$, $n_a = 1$, SNR = 6 dB, $S_e = 64$ bpcu, $i_{max} = 2$, $k = K$ with different LS methods.

conventional MIMO with $n_t = 1$ and 16-QAM to achieve a target BER of 10^{-4}. Moreover, it is portrayed that SMMP approaches the performance of the more complex MPD and outperforms the SM-ZF detector. This is because the proposed strategy is able to iteratively identify the active antenna indices and then perform a selective channel inversion with a matrix of reduced dimensions. Instead, the performance of conventional MIMO with two active antennas per user approaches the performance of SMMP. However, we remark that in this case the power consumption and complexity of the MSs is increased due to the additional RF chains implemented.

Fig. 4 shows the performance of the considered detectors in a scenario with $N = 128$, $K = 16$, $n_t = 7$, generalized SM with $n_a = 2$ active antennas per user, and a resulting spectral efficiency of $S_e = 96$ bpcu. The results of this figure show that generalized SM systems with SMMP detection are capable of improving the performance of conventional MIMO systems employing the same number of RF chains ('MIMO-ZF. $n_t = 1$'). The use of SMMP also allows outperforming conventional MIMO systems with a pair of RF chains ('MIMO-ZF. $n_t = 2$') for the range of practical BERs. Moreover, it can be seen that the SMMP algorithm clearly improves the performance of other linear detectors such as the SM-MMSE detector. In the following we focus our attention on single active antenna SM for reasons of brevity, although it is clear that the resultant conclusions also extend to generalized SM transmission.

The number of users is increased from $K = 16$ of Fig. 3 to $K = 32$ in Fig. 5, which shows a scenario especially favorable for the proposed technique. Fig. 5 shows that the proposed strategy is able to provide an enhancement of up to three orders of magnitude in the BER at high SNR values w.r.t. conventional large-scale MIMO, and both ZF and MMSE detectors in SM. The performance improvement offered by SM when compared to conventional MIMO transmission is coherent with the behavior described in [20], [53]. The effect of acquiring inaccurate channel state information (CSI) on the performance is also shown in Fig. 5(b). The imperfect CSI is modeled following $\widehat{\mathbf{H}} = \sqrt{1-\tau^2}\mathbf{H} + \tau\mathbf{B}$, where $\tau \in [0, 1]$ regulates CSI quality and $\mathbf{B} \in \mathbb{C}^{N\times M} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{I}_M)$ characterizes the channel estimation error [54]. The results of this figure for $\tau = 0.25$ highlight the robustness of SMMP when compared to traditional CS-based detection, which makes it the best alternative in terms of performance under imperfect CSI.

Fig. 6 shows the trade-off between performance and complexity in real flops for a MAC with $N = 128$, $K = 16$, $n_a = 1$, $S_e = 64$ bpcu, $i_{max} = 2$, $k = K$ and SNR = 6 dB. The number of antennas per user is $n_t = 4$ when SM is used, whereas it is $n_t = 1$ and $n_t = 4$ for conventional MIMO. Note that the complexity and performance of the CG algorithm varies depending on the number of iterations as explained in Section V-A. From the results of this figure it can be concluded that the SMMP approach offers the best performance with a restrained complexity. When compared to large-scale MIMO systems with the same spectral efficiency and alike number of RF chains ('MIMO-ZF, $n_t = 1$'), the use of SM increases the complexity due to the higher number of antennas at the MSs, but in turn offers a performance improvement of more than two orders of

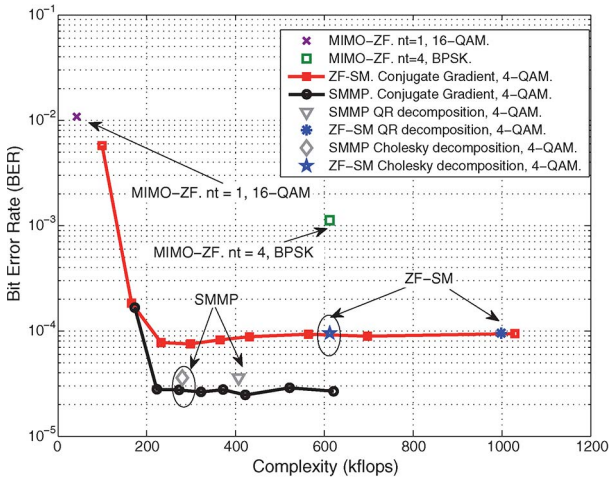

Fig. 7. BER vs. complexity for $N = 128$, $K = S = 8$, $n_t = 4$, $n_a = 1$, SNR = 2 dB, $S_e = 32$ bpcu, $i_{max} = 2$, $k = K$ with different LS methods.

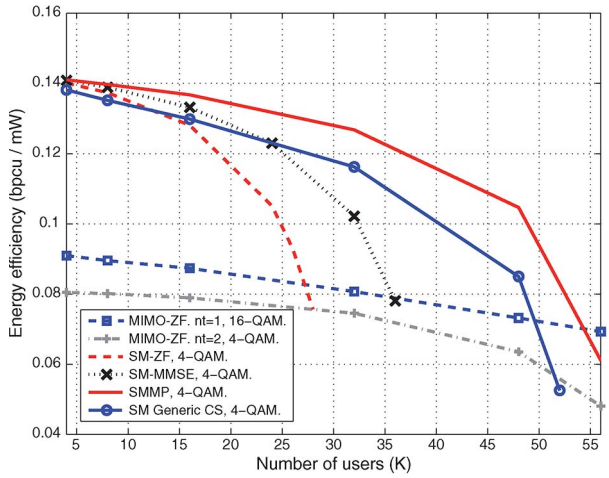

Fig. 8. EE vs. number of users K to achieve BER $= 10^{-3}$. $N = 128$, $n_t = 4$, $n_a = 1$. $P_\psi = 20$ mW.

magnitude. This improvement is not as significant with respect to the spatial multiplexing system with the same number of antennas ('MIMO-ZF, $n_t = 4$'), where instead the complexity is dramatically reduced. Fig. 6 also shows the complexity improvements that can be obtained when the iterative CG algorithm termed 'ZF-SM. Conjugate Gradient' considered in this paper is used to solve the LS problems against traditional direct methods such as 'ZF-SM. QR decomposition' and 'ZF-SM. Cholesky decomposition' in Fig. 6. The number of iterations and, inherently, the complexity can be adjusted depending on the BER required at the BS. This leads us to conclude that the use of the CS-based detection is especially convenient in fast fading scenarios with a reduced channel coherence period. This is because whereas CS-based detection methods perform the same computations in every channel use, the linear detectors solved with direct methods focus the intensive computations at the beginning of the coherence period and reduce the complexity afterwards [2].

A similar conclusion can be obtained from the results of Fig. 7, where we focus on controlling the attainable performance by varying the number of LS iterations in the CG algorithm. We also note that, as opposed to the conclusions achieved in [18], [38], the iterative LS algorithm accounts for 60% of the global detection complexity, hence justifying the need of an accurate complexity characterization. Overall, it can be concluded that the proposed strategy offers significant performance and complexity improvements with respect to conventional detection schemes in systems with the same number of antennas.

Regarding the EE of the CS-based detection, Fig. 8 shows this metric for a MAC with $N = 128$ and a varying number of users. The transmission power is varied depending on the number of users so that $\text{BER}_{obj} = 10^{-3}$ in (36). The circuit power consumption depending on the number of active antennas is set to a realistic value of $P_\psi = 20$ mW [46], and the noise variance is fixed to $\sigma^2 = 0.01$. We note that the PAPR factor of the single-antenna users is increased w.r.t. SM due to the use of a higher modulation order Q. From the results of this figure it can be concluded that the use of SM allows to significantly

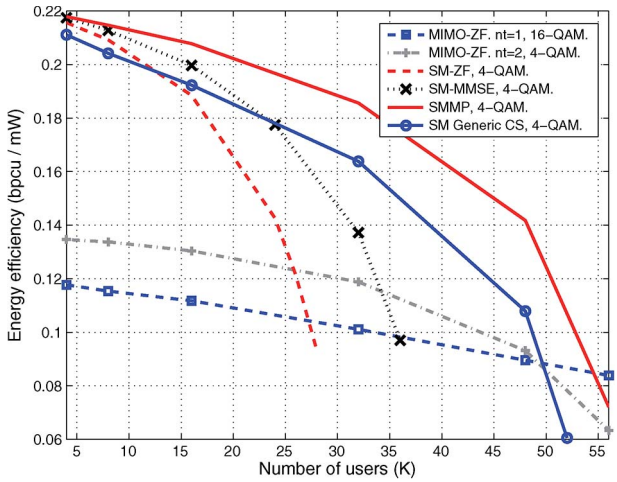

Fig. 9. EE vs. number of users K to achieve BER $= 10^{-3}$. $N = 128$, $n_t = 4$, $n_a = 1$. $P_\psi = 10$ mW.

increase the EE of the conventional large-scale MIMO for low and intermediate system loading factors. This improvement due to the reduced circuit power consumption and PAPR was already noticed in [55] for the downlink of small-scale P2P systems. Moreover, the proposed technique outperforms the rest of conventional detectors, hence constituting an energy-efficient alternative in the MAC.

Fig. 8 also shows that, in spite of the transmission energy savings that can be obtained when $n_t = 2$ and 4-QAM are used in MIMO systems, the increased circuit power consumption caused by the higher number of RF chains penalizes the EE. To show this effect, in Fig. 9 we reduce the power consumed by the RF chains P_ψ to half, i.e., $P_\psi = 10$ mW. By doing so, now it can be seen that the use of a MIMO system with $n_t = 2$ outperforms the option of having single-antenna devices. Still, the EE of the SM alternatives is significantly higher than the one of MIMO strategies for a low and intermediate number of users due to the reduced transmission power required to compensate the inter-user interference.

IX. Conclusion

In this paper, a low-complexity detection algorithm for SM has been presented. The proposed strategy is based upon CS by incorporating the additional structure and sparsity of the transmitted signals in the MAC. Our complexity and performance analyses show that the benefits of the proposed are maximized when a high number of receive antennas at the BS are used due to its faster convergence and improved performance. Overall, the results derived in this paper confirm that the CS-based detection for SM constitutes a low-complexity alternative to increase the EE in the MAC. Possible future work can be carried out in the analytic characterization of the bit error rate performance of the proposed scheme in the large-scale regime.

Appendix A
Proof of Theorem 2

The proof of Theorem 2 commences by recalling the results regarding the maximum and minimum singular values of Wishart matrices with large dimensions [4]. Let $\mathbf{W}$ be the Wishart matrix defined in (20) with the entries of the matrix $\mathbf{H}_{\mathcal{L}} \in \mathcal{C}^{N\times|\mathcal{L}|}$ satisfying $h_{m,n} \sim \mathcal{CN}(0,1)$. For these expressions, no assumptions on the definition and cardinality of $\mathcal{L}$ and $\mathcal{R}$ are adopted. For large N and $|\mathcal{L}|$, the maximum and minimum eigenvalues of $\mathbf{W}$ converge to

$$\begin{aligned}\sigma_{min}(\mathbf{W}) &\underset{N,|\mathcal{L}|\to\infty}{\longrightarrow} T_{|\mathcal{L}|}\times\left(1-\frac{1}{\sqrt{\beta\left(|\mathcal{L}|\right)}}\right)^2\\ \sigma_{max}(\mathbf{W}) &\underset{N,|\mathcal{L}|\to\infty}{\longrightarrow} T_{|\mathcal{L}|}\times\left(1+\frac{1}{\sqrt{\beta\left(|\mathcal{L}|\right)}}\right)^2\end{aligned} \tag{39}$$

where $\beta(V)$ was defined in (26) and $T_C \triangleq \max(N, C), C \in \mathbb{R}^+$. Based on the above expressions, we can redefine some relationships that will be useful in the sequel

$$\begin{aligned}\left\|\mathbf{H}_{\mathcal{L}}^H\mathbf{y}\right\|_2 &\leq \sqrt{T_{|\mathcal{L}|}}\times\left(1+\frac{1}{\sqrt{\beta\left(|\mathcal{L}|\right)}}\right)\|\mathbf{y}\|_2,\\ \left\|\mathbf{H}_{\mathcal{L}}^{\dagger}\mathbf{y}\right\|_2 &\leq \left(\frac{1}{\sqrt{T_{|\mathcal{L}|}}\times\left(1-\frac{1}{\sqrt{\beta(|\mathcal{L}|)}}\right)}\right)\|\mathbf{y}\|_2,\\ \left\|\mathbf{H}_{\mathcal{L}}^H\mathbf{H}_{\mathcal{L}}\mathbf{x}\right\|_2 &\underset{\geq}{\leq} T_{|\mathcal{L}|}\times\left(1\pm\frac{1}{\sqrt{\beta\left(|\mathcal{L}|\right)}}\right)^2\|\mathbf{x}\|_2,\\ \left\|\left(\mathbf{H}_{\mathcal{L}}^H\mathbf{H}_{\mathcal{L}}\right)^{-1}\mathbf{x}\right\|_2 &\underset{\geq}{\leq} \left(\frac{1}{\sqrt{T_{|\mathcal{L}|}}\times\left(1\pm\frac{1}{\sqrt{\beta(|\mathcal{L}|)}}\right)}\right)^2\|\mathbf{x}\|_2.\end{aligned} \tag{40}$$

Note that the last two inequalities can determine upper and lower bounds [38]. Moreover, the following relationship is also satisfied

$$\left\|\mathbf{H}_{\mathcal{R}}^H\mathbf{H}_{\mathcal{L}}\right\| \leq \frac{T_B}{\beta(B)}+\frac{2T_B}{\sqrt{\beta(B)}}, \tag{41}$$

where the set $\mathcal{G}$ with cardinality B is given by $\mathcal{G} \triangleq \mathcal{R}\cup\mathcal{L}$. As the derivation of the previous expression is based upon the ideas developed in [38], but considering instead that the spectral norm of $\mathbf{H}_{\mathcal{R}}^H\mathbf{H}_{\mathcal{L}}$ is always smaller than the one of $\mathbf{H}_{\mathcal{G}}^H\mathbf{H}_{\mathcal{G}}-T$ and that the function $\beta:\mathbb{Z}^+\to\mathbb{R}^+$, here we avoid a detailed description for brevity.

The objective of the proof is to establish an upper bound on the ℓ_2 norm of the error between the transmitted signal and the approximation obtained by the generic CS algorithm at the i-th iteration [38], i.e.,

$$\|\mathbf{x}-\tilde{\mathbf{x}}^i\|_2 \leq R_{max}, \tag{42}$$

where R_{max} is the upper bound to be derived and $\tilde{\mathbf{x}}^i$ is the estimated signal at the i-th iteration. The upper bound is obtained by following the arguments developed in [38] and applying the results obtained in (40) and (41) where convenient. To preserve the coherence with [38], in the following it has been assumed that $|\mathcal{T}| \leq 3S$ or, in other words, $|\Omega| = 2S$ entries are selected at the output of the MF per iteration.

To derive the desired result, first note that the difference between the S-sparse transmitted signal $\mathbf{x}$ and the output of the LS problem $\mathbf{b}$ is bounded when selecting the best S-sparse approximation $\tilde{\mathbf{x}}^i$ [38]. Formally,

$$\|\mathbf{x}-\tilde{\mathbf{x}}^i\|_2 \leq \|\mathbf{x}-\mathbf{b}\|_2+\|\mathbf{b}-\tilde{\mathbf{x}}^i\|_2 \leq 2\|\mathbf{x}-\mathbf{b}\|_2, \tag{43}$$

where the basic norm property $\|\mathbf{a}+\mathbf{z}\| \leq \|\mathbf{a}\|+\|\mathbf{z}\|$ has been applied. With the purpose of deriving an upper bound to the right-hand side of (43), we first express $\|\mathbf{x}-\mathbf{b}\|_2$ as

$$\|\mathbf{x}-\mathbf{b}\|_2 \leq \|\mathbf{x}|_{\mathcal{T}^C}\|_2+\|\mathbf{x}|_{\mathcal{T}}-\mathbf{b}|_{\mathcal{T}}\|_2, \tag{44}$$

where the inequality holds because $\mathbf{b}$ is supported on $\mathcal{T}$ following (18). We now focus on deriving an upper bound for $\|\mathbf{x}|_{\mathcal{T}}-\mathbf{b}|_{\mathcal{T}}\|_2$, which is given by

$$\begin{aligned}&\|\mathbf{x}|_{\mathcal{T}}-\mathbf{b}|_{\mathcal{T}}\|_2\\ &\quad\overset{(a)}{=}\left\|\mathbf{x}|_{\mathcal{T}}-\mathbf{H}_{\mathcal{T}}^{\dagger}(\mathbf{H}\mathbf{x}|_{\mathcal{T}}+\mathbf{H}\mathbf{x}|_{\mathcal{T}^C}+\mathbf{w})\right\|_2\\ &\quad\overset{(b)}{\leq}\left\|\left(\mathbf{H}_{\mathcal{T}}^H\mathbf{H}_{\mathcal{T}}\right)^{-1}\mathbf{H}_{\mathcal{T}}^H\mathbf{H}\mathbf{x}|_{\mathcal{T}^C}\right\|_2+\left\|\mathbf{H}_{\mathcal{T}}^{\dagger}\mathbf{w}\right\|_2\\ &\quad\overset{(c)}{\leq}\left(\frac{\left(\frac{T_{4S}}{\beta(4S)}+\frac{2T_{4S}}{\sqrt{\beta(4S)}}\right)}{T_{3S}\left(1-\frac{1}{\sqrt{\beta(3S)}}\right)^2}\right)\|\mathbf{x}|_{\mathcal{T}^C}\|_2+\frac{\|\mathbf{w}\|_2}{\sqrt{T_{3S}}-\frac{\sqrt{T_{3S}}}{\sqrt{\beta(3S)}}}.\end{aligned} \tag{45}$$

In the above expressions, $\overset{(a)}{=}$ holds by definition (18), and $\overset{(b)}{\leq}$ follows from the definition of the pseudoinverse matrix and by noting that $\mathbf{H}_{\mathcal{T}}^{\dagger}\mathbf{H}\mathbf{x}|_{\mathcal{T}}=\mathbf{x}|_{\mathcal{T}}$. Moreover, $\overset{(c)}{\leq}$ is obtained by using the relationships derived in (40) and (41) along with the fact that, by definition, $|\mathcal{T}| \leq 3S$ and $\mathbf{x}$ is S-sparse. Substituting the last inequality of (45) into (44) we obtain the desired upper bound

$$\|\mathbf{x}-\mathbf{b}\|_2 \leq \left(1+\frac{\left(\frac{T_{4S}}{\beta(4S)}+\frac{2T_{4S}}{\sqrt{\beta(4S)}}\right)}{T_{3S}\left(1-\frac{1}{\sqrt{\beta(3S)}}\right)^2}\right)\|\mathbf{x}|_{\mathcal{T}^C}\|_2+\frac{\|\mathbf{w}\|_2}{\sqrt{T_{3S}}-\frac{\sqrt{T_{3S}}}{\sqrt{\beta(3S)}}}. \tag{46}$$

To derive an upper bound on $\|\mathbf{x}|_{\mathcal{T}^C}\|_2$, first define the signal $\mathbf{m} = \mathbf{x} - \tilde{\mathbf{x}}^{i-1}$. Considering this, the following relationship holds [38]

$$\|\mathbf{x}|_{\mathcal{T}^C}\|_2 \overset{(a)}{=} \left\|(\mathbf{x} - \tilde{\mathbf{x}}^{i-1})|_{\mathcal{T}^C}\right\|_2 = \|\mathbf{m}|_{\mathcal{T}^C}\|_2 \overset{(b)}{\leq} \|\mathbf{m}|_{\Omega^C}\|_2, \quad (47)$$

where $\overset{(a)}{=}$ is satisfied because the signal $\tilde{\mathbf{x}}^{i-1}$ is supported in $\mathcal{T}$ whereas $\overset{(b)}{\leq}$ holds because $\Omega \subset \mathcal{T}$.

Based on (47), the next step consists on upper bounding $\|\mathbf{m}|_{\Omega^C}\|_2$. Let the set $\mathcal{P}$ be the support of the signal $\mathbf{m}$ with cardinality $|\mathcal{P}| \leq 2S$. Since Ω is formed by the $2S$ entries with highest energy from the output of the MF $\mathbf{c}$, the relationship

$$\|\mathbf{c}|_{\mathcal{P}-\Omega}\|_2 \leq \|\mathbf{c}|_{\Omega-\mathcal{P}}\|_2 \quad (48)$$

is satisfied. Recall from (14) that the residual signal at the start of the i-th iteration satisfies $\mathbf{r} = \mathbf{H}\mathbf{m} + \mathbf{w}$. By making use of (40) and (41), it can be seen that

$$\begin{aligned}
&\|\mathbf{c}|_{\mathcal{P}-\Omega}\|_2 \\
&\quad = \left\|\mathbf{H}^H_{\mathcal{P}-\Omega}\mathbf{r}\right\|_2 \\
&\quad = \left\|\mathbf{H}^H_{\mathcal{P}-\Omega}(\mathbf{H}\mathbf{m} + \mathbf{w})\right\|_2 \\
&\quad \geq \left\|\mathbf{H}^H_{\mathcal{P}-\Omega}\mathbf{H}\mathbf{m}|_{\mathcal{P}-\Omega}\right\|_2 - \left\|\mathbf{H}^H_{\mathcal{P}-\Omega}\mathbf{H}\mathbf{m}|_{\Omega}\right\|_2 - \left\|\mathbf{H}^H_{\mathcal{P}-\Omega}\mathbf{w}\right\|_2 \\
&\quad \overset{(a)}{\geq} T_{2S}\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2 \|\mathbf{m}|_{\mathcal{P}-\Omega}\|_2 - \left(\sqrt{T_{2S}} + \frac{\sqrt{T_{2S}}}{\sqrt{\beta(2S)}}\right) \\
&\qquad \times \|\mathbf{w}\|_2 - \left(\frac{T_{2S}}{\beta(2S)} + \frac{2T_{2S}}{\sqrt{\beta(2S)}}\right)\|\mathbf{m}\|_2, \quad (49)
\end{aligned}$$

where $\overset{(a)}{\geq}$ follows from $|\mathcal{P} - \Omega| \leq 2S$. The same expressions can be used to derive an upper bound to $\|\mathbf{c}|_{\Omega-\mathcal{P}}\|_2$, yielding

$$\begin{aligned}
&\|\mathbf{c}|_{\Omega-\mathcal{P}}\|_2 \\
&\quad = \left\|\mathbf{H}^H_{\Omega-\mathcal{P}}\mathbf{r}\right\|_2 = \left\|\mathbf{H}^H_{\Omega-\mathcal{P}}(\mathbf{H}\mathbf{m} + \mathbf{w})\right\|_2 \\
&\quad \leq \left\|\mathbf{H}^H_{\Omega-\mathcal{P}}\mathbf{H}\mathbf{m}\right\|_2 + \left\|\mathbf{H}^H_{\Omega-\mathcal{P}}\mathbf{w}\right\|_2 \\
&\quad \overset{(a)}{\leq} \left(\frac{T_{4S}}{\beta(4S)} + \frac{2T_{4S}}{\sqrt{\beta(4S)}}\right)\|\mathbf{m}\|_2 + \left(\sqrt{T_{2S}} + \frac{\sqrt{T_{2S}}}{\sqrt{\beta(2S)}}\right)\|\mathbf{w}\|_2. \quad (50)
\end{aligned}$$

Here, $\overset{(a)}{\leq}$ is obtained by noting that $|(\Omega - \mathcal{P}) \cup \mathcal{P}| \leq 4S$. By using (48), (49) and (50) the following threshold is derived

$$\begin{aligned}
\|\mathbf{m}|_{\mathcal{P}-\Omega}\|_2 = \|\mathbf{m}|_{\Omega^C}\|_2 &\leq \frac{\left(2 + \frac{2}{\sqrt{\beta(2S)}}\right)}{\sqrt{T_{2S}}\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2}\|\mathbf{w}\|_2 \\
&+ \left(\frac{\left(\frac{T_{4S}}{\beta(4S)} + \frac{2T_{4S}}{\sqrt{\beta(4S)}} + \frac{T_{2S}}{\beta(2S)} + \frac{2T_{2S}}{\sqrt{\beta(2S)}}\right)}{T_{2S}\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2}\right)\|\mathbf{m}\|_2. \quad (51)
\end{aligned}$$

The proof of (33) is completed by combining the results from (43)–(51). Formally,

$$\begin{aligned}
&\|\mathbf{x} - \tilde{\mathbf{x}}^i\|_2 \\
&\quad \overset{(43)}{\leq} 2\|\mathbf{x} - \mathbf{b}\|_2 \\
&\quad \overset{(46)-(47)}{\leq} \left(\frac{2\|\mathbf{w}\|_2}{\sqrt{T_{3S}} - \frac{\sqrt{T_{3S}}}{\sqrt{\beta(3S)}}}\right) + \left(2 + \frac{\left(\frac{2T_{4S}}{\beta(4S)} + \frac{4T_{4S}}{\sqrt{\beta(4S)}}\right)}{T_{3S}\left(1 - \frac{1}{\sqrt{\beta(3S)}}\right)^2}\right)\|\mathbf{m}|_{\Omega^C}\|_2 \\
&\quad \overset{(47)-(51)}{\leq} \left(2 + \frac{\left(\frac{2T_{4S}}{\beta(4S)} + \frac{4T_{4S}}{\sqrt{\beta(4S)}}\right)}{T_{3S}\left(1 - \frac{1}{\sqrt{\beta(3S)}}\right)^2}\right) \\
&\qquad \times \left[\left(\frac{\left(\frac{T_{4S}}{\beta(4S)} + \frac{2T_{4S}}{\sqrt{\beta(4S)}} + \frac{T_{2S}}{\beta(2S)} + \frac{2T_{2S}}{\sqrt{\beta(2S)}}\right)}{T_{2S}\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2}\right)\|\mathbf{x} - \tilde{\mathbf{x}}^{i-1}\|_2 \right. \\
&\qquad \left. + \frac{\left(2 + \frac{2}{\sqrt{\beta(2S)}}\right)}{\sqrt{T_{2S}}\left(1 - \frac{1}{\sqrt{\beta(2S)}}\right)^2}\|\mathbf{w}\|_2\right] + \frac{2\|\mathbf{w}\|_2}{\sqrt{T_{3S}} - \frac{\sqrt{T_{3S}}}{\sqrt{\beta(3S)}}} \\
&\quad = c_1(N, S)\|\mathbf{x} - \tilde{\mathbf{x}}^{i-1}\|_2 + c_2(N, S)\|\mathbf{w}\|_2. \quad (52)
\end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank Dr. Francesco Renna for the helpful discussions throughout the development of this work.

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] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "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.

[3] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[4] 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.

[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, M. Sellathurai, and T. Ratnarajah, "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] C. Masouros and M. Matthaiou, "Space-constrained massive MIMO: Hitting the wall of favorable propagation," *IEEE Commun. Letters*, vol. 19, no. 5, pp. 771–774, May 2015.

[8] D. Ha, K. Lee, and J. Kang, "Energy efficiency analysis with circuit power consumption in massive MIMO systems," in *Proc. IEEE 24th Int. Symp. PIMRC*, Sep. 2013, pp. 938–942.

[9] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, 1st Quart. 2015.

[10] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.

[11] C. Masouros, "Improving the diversity of spatial modulation in MISO channels by phase alignment," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 729–732, May 2014.

[12] C. Masouros and L. Hanzo, "Constellation-randomization achieves transmit diversity for single-RF spatial modulation," *IEEE Trans. Veh. Technol.*, to be published.

[13] C. Masouros and L. Hanzo, "Dual layered downlink MIMO transmission for increased bandwidth efficiency," *IEEE Trans. Veh. Technol.*, to be published.
[14] A. Garcia-Rodriguez, C. Masouros, and L. Hanzo, "Pre-scaling optimization for space shift keying based on semidefinite relaxation," *IEEE Trans. Commun.*, to be published.
[15] A. Younis, S. Sinanovic, M. Di Renzo, R. Mesleh, and H. Haas, "Generalised sphere decoding for spatial modulation," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2805–2815, Jul. 2013.
[16] S. Sugiura, C. Xu, S. X. Ng, and L. Hanzo, "Reduced-complexity coherent versus non-coherent QAM-aided space-time shift keying," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3090–3101, Nov. 2011.
[17] C. Xu, S. Sugiura, S. X. Ng, and L. Hanzo, "Spatial modulation and space-time shift keying: Optimal performance at a reduced detection complexity," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 206–216, Jan. 2013.
[18] C.-M. Yu *et al.*, "Compressed sensing detector design for space shift keying in MIMO systems," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1556–1559, Oct. 2012.
[19] C.-H. Wu, W.-H. Chung, and H.-W. Liang, "OMP-based detector design for space shift keying in large MIMO system," in *Proc. IEEE GLOBECOM*, 2014, pp. 4072–4076.
[20] N. Serafimovski, S. Sinanović, M. Di Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, pp. 1–20, Sep. 2012.
[21] M. Di Renzo and H. Haas, "Bit error probability of space-shift keying MIMO over multiple-access independent fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3694–3711, Oct. 2011.
[22] L.-L. Yang, "Signal detection in antenna-hopping space-division multiple-access systems with space-shift keying modulation," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 351–366, Jan. 2012.
[23] T. Narasimhan, P. Raviteja, and A. Chockalingam, "Large-scale multiuser SM-MIMO versus massive MIMO," in *Proc. ITA*, Feb. 2014, pp. 1–9.
[24] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Academy Sci.*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.
[25] J. Zheng, "Low-complexity detector for spatial modulation multiple access channels with a large number of receive antennas," *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 2055–2058, Nov. 2014.
[26] S. Wang, Y. Li, M. Zhao, and J. Wang, "Energy efficient and low-complexity uplink transceiver for massive spatial modulation MIMO," *IEEE Trans. Veh. Technol.*, to be published.
[27] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2156–2168, Apr. 2015.
[28] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.
[29] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.
[30] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 2012, vol. 3.
[31] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. Conf. Rec. 44th ASILOMAR*, Nov. 2010, pp. 1498–1502.
[32] C. Peel, B. Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.
[33] R. Zhang, L.-L. Yang, and L. Hanzo, "Generalised pre-coding aided spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5434–5443, Nov. 2013.
[34] R. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
[35] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[36] E. Candès, "Compressive sampling," in *Proc. Int Congr. Math.*, Madrid, Spain, Aug. 22–30, 2006, pp. 1433–1452.
[37] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
[38] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
[39] X. Rao and V. K. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.
[40] M. Arakawa, "Computational workloads for commonly used signal processing kernels," Def. Tech. Inf. Center Document, Fort Belvoir, VA, USA, Tech. Rep., 2006.
[41] A. Björck, *Numerical Methods for Least Squares Problems*. Philadelphia, PA, USA: SIAM, 1996.
[42] R. Hunger, "Floating point operations in matrix-vector calculus," Technische Universitat Munchen, Munchen, Germany, Tech. Rep. [Online]. Available: https://mediatum.ub.tum.de/doc/625604/625604.pdf
[43] L. Vandenberghe, "Applied numerical computing," Univ. Calif., Los Angeles, CA, USA, Univ. Lecture, 2011.
[44] M. Matthaiou, M. McKay, P. Smith, and J. Nossek, "On the condition number distribution of complex Wishart matrices," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1705–1717, Jun. 2010.
[45] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Yale Univ., New Haven, CT, USA, 1988.
[46] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.
[47] G. Miao, "Energy-efficient uplink multi-user MIMO," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2302–2313, May 2013.
[48] K. Ntontin, M. Di Renzo, A. Perez-Neira, and C. Verikoukis, "Towards the performance and energy efficiency comparison of spatial modulation with conventional single-antenna transmission over generalized fading channels," in *Proc. IEEE Int. Workshop CAMAD*, Sep. 2012, pp. 120–124.
[49] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Maximizing energy efficiency in the vector precoded MU-MISO downlink by selective perturbation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4974–4984, Sep. 2014.
[50] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector perturbation based on symbol scaling for limited feedback MISO downlinks," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.
[51] A. Garcia-Rodriguez and C. Masouros, "Power-efficient Tomlinson-Harashima precoding for the downlink of multi-user MISO systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1884–1896, Jun. 2014.
[52] A. Garcia-Rodriguez and C. Masouros, "Power loss reduction for MMSE-THP with multidimensional symbol scaling," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1147–1150, Jul. 2014.
[53] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.
[54] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "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.
[55] A. Stavridis, S. Sinanovic, M. Di Renzo, and H. Haas, "Energy evaluation of spatial modulation at a multi-antenna base station," in *Proc. IEEE 78th VTC—Fall*, Sep. 2013, pp. 1–5.

Adrian Garcia-Rodriguez received the M.S. degree in telecommunications engineering from Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain, in 2012. He is currently working toward the Ph.D. degree in the Department of Electronic and Electrical Engineering, University College London, London, U.K. His research interests lie the field of signal processing and wireless communications, with emphasis on energy-efficient and multiantenna communications.

Christos Masouros (M'06–SM'14) received the Diploma degree in electrical and computer engineering from the University of Patras, Patras, Greece, in 2004 and the M.Sc. degree 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 is currently a Lecturer with the Department of Electronic and Electrical Engineering, University College London, London, U.K. He has been a Research Associate with the University of Manchester and a Research Fellow with Queen's University Belfast, Belfast, 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 lie in the field of 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.