

Compressed Sensing Improves the Performance of Subcarrier Index-Modulation Assisted OFDM

Hongming Zhang, *Student Member, IEEE*, Lie-Liang Yang, *Fellow, IEEE* and Lajos Hanzo, *Fellow, IEEE*

Abstract—In orthogonal frequency division multiplexing relying on subcarrier index modulation (OFDM-SIM), the information is conveyed by both the indices of the activated subcarriers and the conventional amplitude-phase modulated (APM) symbols. It has been shown that OFDM-SIM is capable of striking a tradeoff between the attainable spectral efficiency (SE) and energy efficiency (EE). In order to further increase the EE of the OFDM-SIM system, while potentially increasing its SE, we propose a compressed sensing (CS) assisted signalling strategy for the family of OFDM-SIM systems. Correspondingly, we first consider the joint maximum likelihood (JML) detection of the CS assisted index-modulated (CSIM) and of the classic APM symbols, despite its high complexity. Then, we propose a low complexity detection algorithm, which is termed as the iterative residual check (IRC) based detector. This is based on the Greedy Pursuit concept of CS, which makes locally optimal choices at each step. Finally, both analytical and simulation results are provided for characterizing the attainable system performance of our proposed OFDM-CSIM system. We demonstrate that in comparison to the conventional OFDM-SIM system, the proposed OFDM-CSIM arrangement is capable of achieving both a higher SE as well as an increased EE. We also show that the diversity gain provided by the OFDM-CSIM system is much higher than that of the OFDM-SIM system. Furthermore, our investigation of the detection performance shows that the proposed IRC detector is capable of providing an attractive detection performance at a low complexity.

Index Terms—OFDM, subcarrier index modulation, compressed sensing (CS), transmit diversity, spectral efficiency (SE), energy efficiency (EE)

NOTATIONS

$(\cdot)^T$ and $(\cdot)^*$	the transpose and the conjugate, respectively
$(\cdot)^H$ and $(\cdot)^{-1}$	the conjugate transpose and the inverse, respectively
$\mathbb{E}[\cdot]$	the expectation operator
\mathbb{Z}_+^N	a set of real integers containing $\{1, 2, \dots, N\}$
$\mathbb{C}^{M \times L}$	the set of $(M \times L)$ -element matrices in the complex field
\boldsymbol{x} and \boldsymbol{X}	vector in boldface lower-case and matrix in boldface upper-case, respectively
$x(l)$ and $x(m, l)$	l th block of symbols and m th symbol of l th block, respectively
\boldsymbol{I}_M	the $(M \times M)$ -element identity matrix
\boldsymbol{I}_i	the $(M \times M_i)$ -element mapping matrix whose columns are extracted from the \boldsymbol{I}_M according to an index set \mathcal{I}_i
$\boldsymbol{\mathcal{F}}_M$	the normalized $(M \times M)$ -element DFT matrix, which satisfies $\boldsymbol{\mathcal{F}}_M \boldsymbol{\mathcal{F}}_M^H = \boldsymbol{\mathcal{F}}_M^H \boldsymbol{\mathcal{F}}_M = \boldsymbol{I}_M$
$diag\{\boldsymbol{x}\}$	the diagonal matrix with elements in \boldsymbol{x} on its diagonal
$\ \boldsymbol{x}\ _p$	the ℓ_p -norm of a vector \boldsymbol{x} , where $p > 0$
$\ \boldsymbol{x}\ _0$	the pseudo- ℓ_0 -norm of a vector \boldsymbol{x} , which is defined as $\ \boldsymbol{x}\ _0 = Card(Supp(\boldsymbol{x}))$
$\{A\} \setminus \{B\}$	the difference of two sets $\{A\}$ and $\{B\}$

H. Zhang, L.-L. Yang and Lajos Hanzo are with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ, UK. (E-mail: zh1g11 lly lh@ecs.soton.ac.uk).

The financial support of the EPSRC projects EP/N004558/1 and EP/L018659/1, as well as of the European Research Council's Advanced Fellow Grant under the Beam-Me-Up project and of the Royal Society's Wolfson Research Merit Award is gratefully acknowledged.

I. INTRODUCTION

Recently, a paradigm shift took place from the development of spectrally efficient communication techniques to the conception of both spectral- and energy-efficient communication techniques, as detailed in [1–6]. Indeed, as pointed out in [1], striking a compelling compromise between the spectral efficiency (SE) (or the bandwidth efficiency) and energy efficiency (EE) (or the power efficiency) is essential for the design state-of-the-art networks. Hence, various joint SE and EE solutions have been proposed for the different protocol layers [4, 6]. Specifically, in the physical layer, more and more attention is paid to the SE and EE of communications systems, including both the digital signal processing and the analog front-end.

As the predominant transmission technique of broadband communications, at the time writing [7, 8], orthogonal frequency-division multiplexing (OFDM) is mainly characterized by its high degree of flexibility and high spectral efficiency [9]. Generally, a frequency-selective fading channel can be converted into a number of parallel flat-fading subchannels with the aid of OFDM, thereby, considerably reducing the receiver's complexity as a benefit of using single-tap frequency-domain (FD) equalization. Moreover, cyclic prefix (CP)-based OFDM can be employed for avoiding inter-OFDM-symbol interference. Recently, OFDM has been combined with index modulation (OFDM-IM) [10–12]. In [10, 11], the subcarrier IM (SIM) concept has been proposed based on a principle reminiscent of the implicit information conveyed by the activated antenna index in spatial modulation (SM) based multiple-input-multiple-output (MIMO) systems [2]. Then, a generalized IM scheme has been proposed and analyzed in [12]. Further details on the historical timeline of the SIM can be found in [13].

The most appealing aspect of SIM is its flexibility in terms of striking a tradeoff between the SE and EE, as analyzed in [14, 15]. Hence, SIM has the potential of representing a win-win alternative in scenarios, such as device-to-device (D2D) communications, in-vehicle communications [16], sensor networks, etc. However, since it is an emerging technique, there are also some open issues. Apart from the above-mentioned SE and EE aspects of SIM, we will also consider the associated diversity and complexity issues, which will also be addressed in this paper. Generally, the most powerful contributor to reliable communications over multipath propagation channels is diversity, including temporal diversity, frequency diversity and space diversity, as detailed in [17]. As demonstrated in [12, 13], OFDM-SIM systems are capable of outperforming the classic OFDM systems for a low throughput. In [18], a block interleaver has been employed for achieving a beneficial time diversity gain. Later in [19], the coordinate interleaved orthogonal design (CIOD) of [20] has been employed in OFDM-SIM systems for improving the attainable transmitter diversity gain. However, the attainable multipath diversity gain has not been fully exploited by these schemes and no generalized analytical results are available for the achievable diversity order of the OFDM-SIM system. The complexity of SIM has to be separately considered at the transmitter and receiver side. In [21], a generalized look-up table has

been proposed in order to improve the SE of SIM. However, in light of the limited SE improvement, the implementation complexity imposed may be deemed excessive. At the receiver side, since the information is conveyed in both the amplitude-phase modulated (APM) symbols and the SIM symbols, it is quite a challenge to design a detection scheme, which provides a good detection performance at a low complexity. In [12], the joint maximum likelihood (JML) detector has been conceived. As a further development, based on the fact that the FD symbols are either zero valued or non-zero valued, the authors of [12] have proposed a log-likelihood ratio (LLR) based detector, which has been shown to have the same detection performance as the JML detector, but at a lower complexity cost. Later, the JML and the LLR detectors have been modified in [19].

On the other hand, compressed sensing (CS), as an emerging theory, has attracted considerable research-attention. Initially, CS has been proposed for recovering vectors in high dimensions from vectors in low dimensions, as detailed in [22]. Then, it has been invoked for solving numerous problems in communications systems, such as channel estimation [23], narrowband interference mitigation [24], spectrum sensing [25], impulsive noise mitigation [26] and so on. However, to the best of the authors' knowledge, CS has not been applied to the family of OFDM-SIM systems. Specifically, when an OFDM-SIM system is designed for aforementioned EE communications scenarios, the number of activated subcarriers should be small. By taking advantage of this condition, each group of subcarriers can be potentially represented by a sparse vector. Hence, according to CS, it is also possible to detect (or recover) the corresponding sparse vector with the aid of algorithms such as the family of ℓ_1 -minimization algorithms [27–29] and Greedy pursuit algorithms [30–32].

To address the aforementioned issues, our contributions of this paper are summarized as follows.

- We propose a CS-assisted signalling strategy for OFDM-SIM systems. The general philosophy of our proposed CS-assisted IM (CSIM) scheme transpires from Fig. 1. The basic idea of CSIM is that the conventional IM is implemented in a high-dimensional virtual digital-domain, and then the high-dimensional IM symbols are compressed into the low-dimensional subcarriers in the FD with the aid of CS. In this way, both the SE and EE of our proposed CSIM becomes higher than that of conventional SIM.
- We provide both analytical and simulation results for characterizing the system performance of the OFDM-CSIM system. We first demonstrate that the attainable SE of the proposed OFDM-CSIM system is higher than that of the conventional OFDM-SIM system. Moreover, for a given SE, our OFDM-CSIM system is capable of achieving higher EE than the conventional OFDM-SIM system.
- We characterize the attainable diversity gain of both the conventional SIM and of our proposed CSIM schemes by both analytical and simulation results. Our investigations demonstrate that the OFDM-CSIM system is capable of achieving the maximum attainable multipath diversity

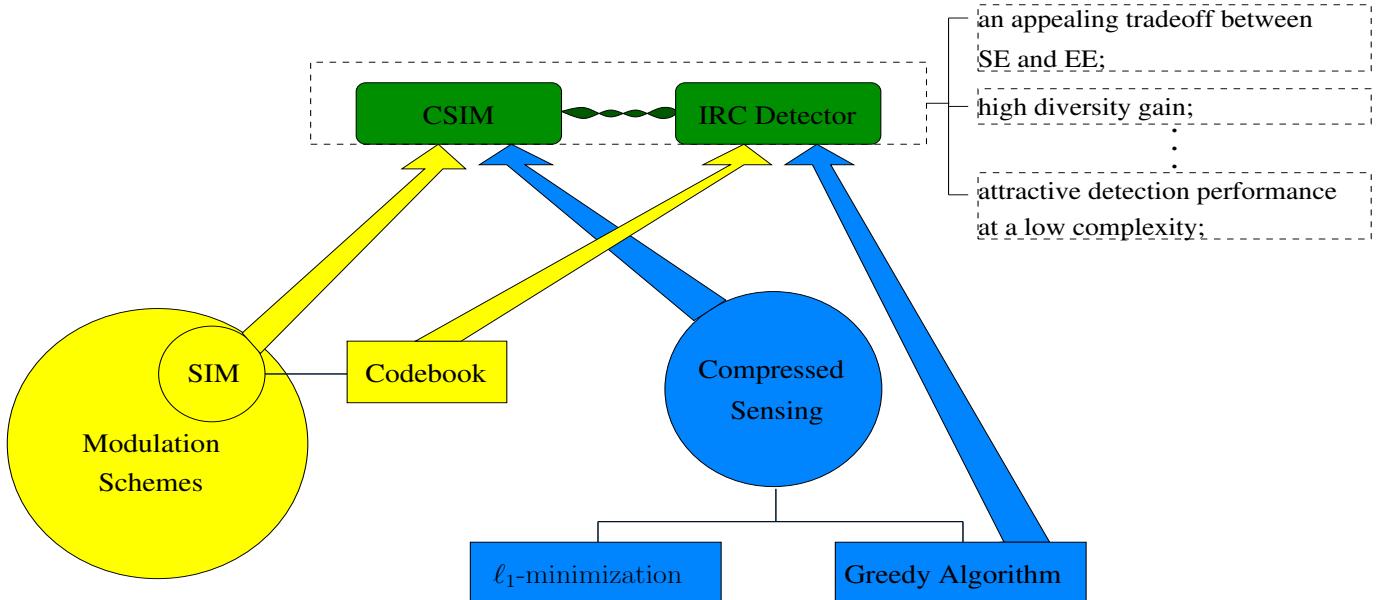


Fig. 1: Illustration of the relationships between compressed sensing (CS), subcarrier index modulation (SIM), CS-assisted index modulation (CSIM) as well as our iterative residual check (IRC) detector.

order, which is verified by our simulation results.

- We first invoke the JML detector for the OFDM-CSIM system for the sake of performance comparison. Since the complexity of the JML detector may be deemed excessive in practice, we propose a low-complexity detector, namely the iterative residual check (IRC) detector, for our system. As depicted in Fig. 1, the IRC detector is proposed based on the Greedy Pursuit concept of CS [33], which updates its estimates one step at a time by making locally optimal choices at each step. We demonstrate that an attractive detection performance can be attained by the IRC detector using as few as one or two iterations, yielding a low complexity.

The rest of the paper is organized as follows. In Section II, both our system model and the proposed CSIM signalling are detailed. Then, both the JML detector and the proposed IRC detector are described in Section III. In Section IV, the performance of the proposed system is analyzed. Our simulation results are discussed in Section V. Finally, we offer our conclusions in Section VI.

II. SYSTEM MODEL

A. Description of Transmitter

We assume a multicarrier system employing M subcarriers. The M subcarriers are divided into G groups, each of which contains $m = M/G$ subcarriers. Our proposed system is illustrated in Fig. 2. An L_b -length sequence of i.i.d. data bits is first split into G groups, each of which contains $L = L_b/G = L_1 + L_2$ bits. In each group, L_1 bits are mapped into an index symbol according to an index mapper $\mu_1: \{0, 1\}^{L_1} \rightarrow \mathcal{Z}$, where $\mathcal{Z} = \{\mathcal{Z}_1, \dots, \mathcal{Z}_C\}$ contains $C = 2^{L_1}$ index subsets, each of which is formed with the aid of K indices chosen from N available indices, as shown in [12]. Thereby, we have $L_1 = \log_2 C = \left\lfloor \log_2 \binom{N}{K} \right\rfloor$. Let the c th index subset of \mathcal{Z}

be denoted as $\mathcal{Z}_c = \{\mathcal{Z}_c(0), \dots, \mathcal{Z}_c(K-1)\} \subset \mathcal{Z}$, where $\mathcal{Z}_c(k) \in \mathbb{Z}_+^N$ for $k = 0, \dots, K-1$. Let us assume that the g th group of data bits is mapped into the c th candidate in \mathcal{Z} . Then, for the sake of simplicity, let the g th group of the index symbols be denoted as $\mathcal{I}_g = \mathcal{Z}_c \subset \mathcal{Z}$. On the other hand, the remaining L_2 bits are mapped onto K classic APM symbols according to a rotated Q -ary QAM/PSK constellation with alphabet $\mathcal{A}_\phi \triangleq \{a_1, \dots, a_Q : a_q = xe^{j\phi}, x \in \mathcal{A}\}$, where ϕ denotes the rotated angle and \mathcal{A} is the alphabet of the QAM/PSK scheme. Note that an appropriately rotated constellation scheme is invoked for ensuring that the signal space diversity is increased at no cost in terms of either power or bandwidth [34]. Specifically, let us denote the data symbols to be transmitted associated with the g th group as $\mathbf{x}_{d,g} = [x_{d,g}(0), \dots, x_{d,g}(K-1)]^T$, where we assume that $\mathbb{E}[|x_{d,g}(k)|^2] = 1, \forall x_{d,g}(k) \in \mathcal{A}_\phi$. Then, in order to further increase the attainable diversity gain, the CIOD of [20] is invoked for transforming $\mathbf{x}_{d,g}$ into $\mathbf{x}_{CI,g}$, which can be expressed as

$$\mathbf{x}_{CI,g} = f_{CI} \{ \mathbf{x}_{d,g} \}, \quad (1)$$

where $x_{CI,g}(k) = \Re\{x_{d,g}(k)\} + j\Im\{x_{d,g}((k+K/2) \bmod K)\}$ for $k = 0, \dots, K-1$. Next, as illustrated in Fig. 3, the K data symbols in $\mathbf{x}_{CI,g}$ are respectively assigned to the K active subcarrier indices in \mathcal{I}_g , yielding a block of virtual domain symbols $\mathbf{x}_g = [x_g(0), x_g(1), \dots, x_g(N-1)]^T$, which can be expressed as

$$\mathbf{x}_g = \mathbf{I}_g \mathbf{x}_{CI,g}, \quad (2)$$

where \mathbf{I}_g is a $(N \times K)$ -element mapping matrix obtained from an $(N \times N)$ -element identity matrix by choosing the columns having the column indices given by the indices in \mathcal{I}_g . Based on the above discussion, we can readily show that the symbol

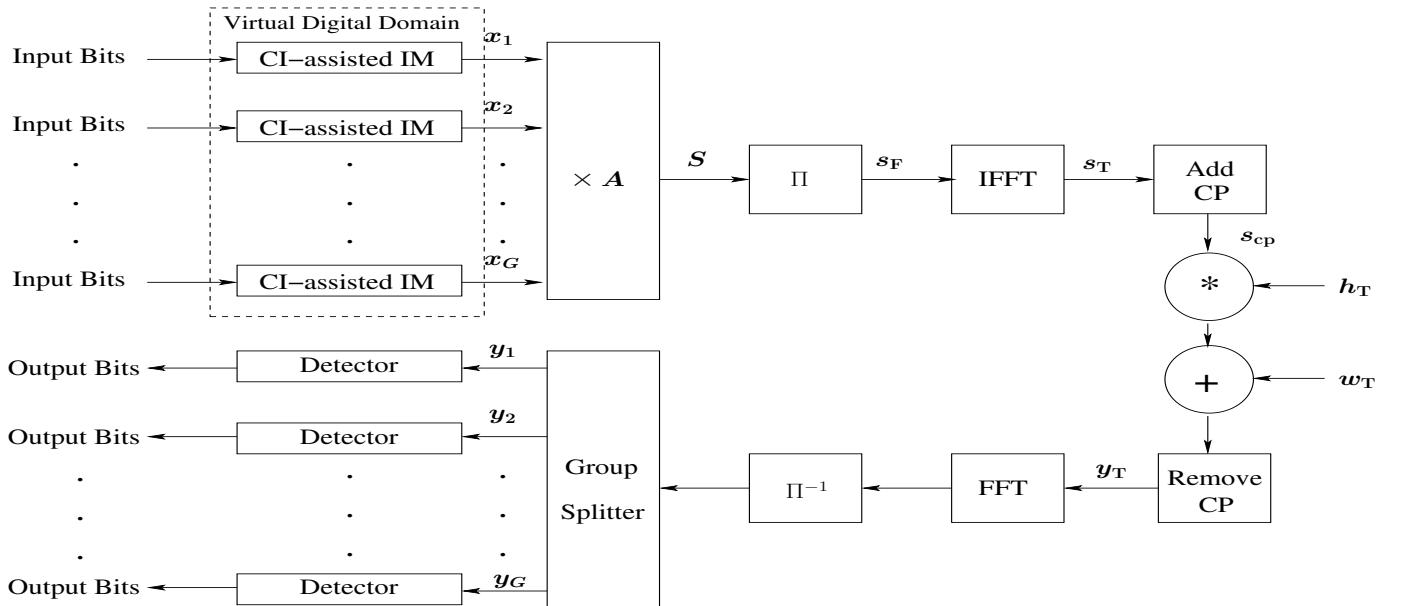


Fig. 2: Illustration of the OFDM-CSIM system employing CI and a depth- G interleaver Π .

vector of (2) conveys

$$L = L_1 + L_2 = \left\lfloor \log_2 \binom{N}{K} \right\rfloor + K \log_2 Q \quad (3)$$

data bits by each of the G groups. The total number of bits per symbol is GL . From (2), we can define a vector space \mathcal{V} for the digital virtual domain as $\mathcal{V} \triangleq \{\mathbf{v}_1, \dots, \mathbf{v}_{2^L}\}$, which is composed by all the legitimate versions of N -dimensional \mathbf{x}_g vectors with any $\mathcal{I}_g \subset \mathcal{Z}$ and $\mathbf{x}_{d,g} \in \mathcal{A}_\phi^K$. From another perspective, we can see that the g th group of L data bits is directly mapped into a symbol vector $\mathbf{x}_g \in \mathcal{V}$. Hence, the vector space \mathcal{V} contains a total of 2^L candidates, as it is defined. Here, the signalling relying on SIM using CI proposed in [19] is shown. Note that, when the CI is not used, the symbol vector $\mathbf{x}_{CI,g}$ in (2) can be replaced by $\mathbf{x}_{d,g}$, representing the conventional SIM of [12], which invokes SIM without any effort to obtain a diversity gain.

As shown in Fig. 3, a measurement matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ is employed for compressing the N -dimensional vector \mathbf{x}_g in the virtual digital domain into an m -dimensional vector \mathbf{s}_g in the FD, which can be expressed as

$$\mathbf{s}_g = \mathbf{A} \mathbf{x}_g, \quad (4)$$

where each column of \mathbf{A} is assumed to be of unity length to yield $\mathbb{E}[\|\mathbf{s}_g\|_2^2] = K$. As illustrated in Fig. 3, the dimension of the symbol vector \mathbf{x}_g is designed to be much higher than that of the subcarrier symbol vector \mathbf{s}_g to be transmitted in each group. By contrast, in the conventional SIM scheme, both these dimensions are the same, i.e. we have $m = N$. Clearly, more data bits can be transmitted by our proposed CSIM scheme than by the conventional SIM scheme, for fixed values of m, K, Q . For example, let $m = 8$, $N = 1024$, $K = 2$ and $Q = 2$, which results in 0.75 bits-per-channel-use for the conventional SIM scheme, while for our CSIM scheme, it can be readily shown that we have 2.5 bits-per-channel-use at no cost in terms of either power or bandwidth.

Furthermore, when the symbol vector \mathbf{x}_g is designed for satisfying $K \ll N$, \mathbf{x}_g is a sparse vector associated with a sparsity level of K . In this case, Eq. (4) represents a classical mathematic modelling of CS [22], whose recovery performance (or detection performance in our scenario) is determined by the specific characteristics of the measurement matrix \mathbf{A} , as it will be detailed in Section III.

Next, as seen in Fig. 2, in order to achieve multipath diversity, the G groups of compressed symbols denoted as $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_G] \in \mathbb{C}^{m \times G}$ are fed into a depth- G interleaver Π , yielding the interleaved FD symbols expressed as $\mathbf{s}_F = [s_F(0), s_F(1), \dots, s_F(M-1)]^T$, where we have $M = mG$. Here, as shown in [18], the depth- G interleaver Π permutes \mathbf{S} by reading column-by-column and then writing row-by-row, which can be expressed as $\mathbf{s}_F = \sum_{g=1}^G \mathbf{I}_{\Pi_g} \mathbf{s}_g$, where \mathbf{I}_{Π_g} is the $(M \times m)$ -element mapping matrix based on the depth- G interleaver Π for the g th group characterized by $s_g(i) = s_F(g-1+iG)$ for $g = 1, \dots, G$ and $i = 0, 1, \dots, m-1$. Furthermore, we can also show that $\mathbf{I}_{\Pi_g}^T \mathbf{I}_{\Pi_g} = \mathbf{I}_m$ and $\mathbf{I}_{\Pi_g}^T \mathbf{I}_{\Pi_{g'}} = \mathbf{0}$ for $g' \neq g$, which can be used for characterizing the de-interleaving process of Π^{-1} , as it will be detailed in Section II-B. Then, as shown in Fig. 2, the FD symbols in \mathbf{s}_F are input to the M -point IFFT, yielding the time-domain (TD) symbols $\mathbf{s}_T = \mathcal{F}_M^H \mathbf{s}_F$. Finally, after adding a CP of length L_{cp} , the transmitted baseband symbols can be expressed as $\mathbf{s}_{cp} = \mathbf{I}_{cp}^T \mathbf{s}_T$, where $\mathbf{I}_{cp} = [\mathbf{I}_{cp}, \mathbf{I}_M]$ and \mathbf{I}_{cp} is a $(M \times L_{cp})$ -element mapping matrix obtained from the last L_{cp} columns of an identity matrix \mathbf{I}_M .

B. Received Signals

We assume an L_h -tap frequency-selective Rayleigh fading channel, which has a channel impulse response (CIR) of $\mathbf{h}_T = [h_T(0), h_T(1), \dots, h_T(L_h-1)]^T$, where we have $h_T(l) \sim \mathcal{CN}(0, 1/L_h)$. We assume that $L_{cp} \geq L_h$ and that perfect

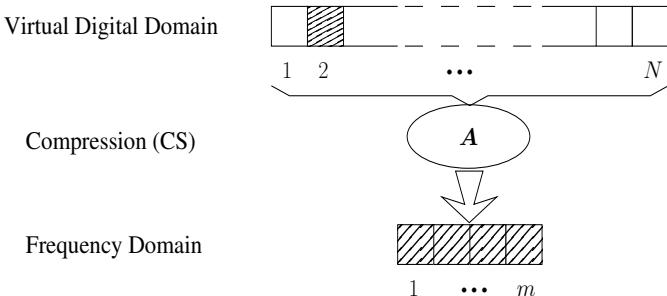


Fig. 3: Illustration of the CSIM scheme. The dimension N of virtual digital domain is higher than the dimension m of FD, i.e. we have $N > m$. The shaded box represents the activated index, while the blank box represents the in-activated index.

synchronization is achieved at the receiver. Furthermore, the bandwidth of each subchannel is assumed to be much lower than the channel's coherence bandwidth. Then, as shown in Fig. 2, after removing the CP, the received baseband equivalent observations $\mathbf{y}_T = [y_T(0), y_T(1), \dots, y_T(M-1)]^T$ can be formulated as

$$\mathbf{y}_T = \mathbf{H}_{\text{cir}} \mathbf{s}_T + \mathbf{w}_T, \quad (5)$$

where \mathbf{H}_{cir} is an $(M \times M)$ -element circulant matrix, which can be diagonalized by the DFT operations, giving $\mathbf{H}_{\text{cir}} = \mathcal{F}_M^H \mathbf{H} \mathcal{F}_M$, where $\mathbf{H} = \text{diag}\{\sqrt{M} \mathcal{F}_M \mathbf{I}_h \mathbf{h}_T\}$ with \mathbf{I}_h being a $(M \times L_h)$ -element mapping matrix formed by the first L_h columns of an identity matrix \mathbf{I}_M . In (5), \mathbf{w}_T is the additive white Gaussian noise (AWGN) with a mean of zero and a variance of N_0 . As shown in Fig. 2, after the FFT operation and de-interleaving, the g th group of received FD symbols can be expressed as

$$\begin{aligned} \mathbf{y}_g &= \mathbf{I}_{\Pi_g}^T (\mathcal{F}_M \mathbf{y}_T) \\ &= \mathbf{I}_{\Pi_g}^T (\mathbf{H} \mathbf{s}_F + \mathcal{F}_M \mathbf{w}_T) \\ &= \mathbf{I}_{\Pi_g}^T \mathbf{H} \sum_{g=1}^G \mathbf{I}_{\Pi_g} \mathbf{s}_g + \mathbf{I}_{\Pi_g}^T (\mathcal{F}_M \mathbf{w}_T) \end{aligned} \quad (6a)$$

$$= \mathbf{I}_{\Pi_g}^T \mathbf{H} \mathbf{I}_{\Pi_g} \mathbf{s}_g + \bar{\mathbf{w}}_g \quad (6b)$$

$$= \bar{\mathbf{H}}_g \mathbf{s}_g + \bar{\mathbf{w}}_g \quad (6c)$$

where by definition $\bar{\mathbf{w}}_g = \mathbf{I}_{\Pi_g}^T (\mathcal{F}_M \mathbf{w}_T)$ contains the g th group of de-interleaved FD noise samples. Note that, since \mathbf{H} is a diagonal matrix, we can readily show that $\mathbf{I}_{\Pi_g}^T \mathbf{H} \mathbf{I}_{\Pi_{g'}} = \mathbf{0}$ for $g \neq g'$, which is applied in (6a) to obtain (6b). In (6c), the diagonal matrix $\bar{\mathbf{H}}_g$ can be explicitly expressed as

$$\begin{aligned} \bar{\mathbf{H}}_g &= \mathbf{I}_{\Pi_g}^T \mathbf{H} \mathbf{I}_{\Pi_g} \\ &= \mathbf{I}_{\Pi_g}^T \text{diag}\left\{\sqrt{M} \mathcal{F}_M \mathbf{I}_h \mathbf{h}_T\right\} \mathbf{I}_{\Pi_g} \\ &= \text{diag}\left\{\sqrt{M} \mathbf{I}_{\Pi_g}^T \mathcal{F}_M \mathbf{I}_h \mathbf{h}_T\right\} \\ &= \text{diag}\left\{\mathbf{F}_{h, \Pi_g} \mathbf{h}_T\right\} = \text{diag}\{\bar{h}_g(0), \dots, \bar{h}_g(m-1)\}, \end{aligned} \quad (7)$$

where by definition $\mathbf{F}_{h, \Pi_g} = \sqrt{M} \mathbf{I}_{\Pi_g}^T \mathcal{F}_M \mathbf{I}_h \in \mathcal{C}^{m \times L_h}$. It can be shown that $\bar{h}_g(i) \sim \mathcal{CN}(0, 1)$ and $\bar{w}_g(i) \sim \mathcal{CN}(0, N_0)$, $\forall g, i$. Hence, we can observe from (6c) that \mathbf{y}_g having the

complex Gaussian distribution having the conditioned probability density function (PDF) of

$$p(\mathbf{y}_g | \mathbf{s}_g) = \frac{1}{(\pi N_0)^m} \exp\left\{-\frac{\|\mathbf{y}_g - \bar{\mathbf{H}}_g \mathbf{s}_g\|_2^2}{N_0}\right\}. \quad (8)$$

Then, as seen in Fig. 2, each group of the de-interleaved symbols is input into the detector discussed in Section III.

III. DETECTION OF THE OFDM-CSIM SIGNALS

In this section, we first consider the JML detector. Then, we propose an IRC detector for reducing the potentially excessive complexity of the JML detector. In our derivation, we stipulate the idealized simplifying assumption that perfect channel estimation is achieved at the receiver.

Firstly, upon substituting (4) into (6c), we arrive at

$$\mathbf{y}_g = \bar{\mathbf{H}}_g \mathbf{A} \mathbf{x}_g + \bar{\mathbf{w}}_g, \quad (9)$$

where the measurement matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ is assumed to be a sub-matrix of an $(N \times N)$ -element orthonormal dictionary. According to [35, 36], the measurement matrix \mathbf{A} should be designed for satisfying the mutual incoherence property (MIP)¹ or the restricted isometry property (RIP)². However, since computing the restricted isometry constant of the RIP has been shown to be NP-hard [37], the MIP is used in this paper. According to [35], the mutual coherence of \mathbf{A} should satisfy $\mu_{\mathbf{A}} < 1/(2K-1)$, which is the sufficient condition of guaranteeing that there exists at most one K -sparse vector \mathbf{x}_g such that $\mathbf{A} \mathbf{x}_g = \mathbf{s}_g$. Thus, in order to provide a guaranteed detection performance, the measurement matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ is designed for satisfying

$$\sqrt{\frac{N-m}{m(N-1)}} \leq \mu_{\mathbf{A}} < \frac{1}{2K-1}, \quad (10)$$

where the lower bound at the left hand side is the well-known Welch bound [38]. Therefore, with the aid of (8) and (10), the PDF conditioned on \mathbf{x}_g can be expressed as

$$p(\mathbf{y}_g | \mathbf{x}_g) = \frac{1}{(\pi N_0)^m} \exp\left\{-\frac{\|\mathbf{y}_g - \bar{\mathbf{H}}_g \mathbf{A} \mathbf{x}_g\|_2^2}{N_0}\right\}. \quad (11)$$

A. Joint Maximum Likelihood Detector

Typically, the optimum detector is the maximum *a posteriori* (MAP) detector, which solves the optimization problem of

$$\mathbf{x}_g^{\text{MAP}} = \arg \max_{\mathbf{v}_i \in \mathcal{V}} \{p(\mathbf{v}_i | \mathbf{y}_g)\}, \quad (12)$$

where $p(\mathbf{v}_i | \mathbf{y}_g)$ is the *a posteriori* probability of \mathbf{v}_i given the receiver's output of \mathbf{y}_g . Let us assume that both the SIM symbols in \mathcal{Z} and the classic APM symbols in \mathcal{A}_{ϕ} are equiprobable and independent. Then, the MAP detector of (12) becomes the ML detector, which can be described as

$$\mathbf{x}_g^{\text{ML}} = \arg \min_{\mathbf{v}_i \in \mathcal{V}} \left\{ \|\mathbf{y}_g - \bar{\mathbf{H}}_g \mathbf{A} \mathbf{v}_i\|_2^2 \right\}. \quad (13)$$

¹The MIP is defined as the mutual coherence of a matrix is very small [35].

²As defined in [36], the RIP is defined as the K th restricted isometry constant δ_K of matrix \mathbf{A} is the smallest number such that $(1 - \delta_K) \|\mathbf{x}_g\|_2^2 \leq \|\mathbf{A} \mathbf{x}_g\|_2^2 \leq (1 + \delta_K) \|\mathbf{x}_g\|_2^2$ is satisfied for any K -sparse vectors \mathbf{x}_g .

Furthermore, according to (1) and (2), the ML detector of (13) is equivalent to the JML detector, which can be written as

$$(\mathcal{I}_g^{\text{ML}}, \mathbf{x}_{\text{d},g}^{\text{ML}}) = \arg \min_{\mathcal{Z}_c \subset \mathcal{Z}, \mathbf{a} \in \mathcal{A}_{\phi}^K} \left\{ \|\mathbf{y}_g - \bar{\mathbf{H}}_g \mathbf{A} \mathbf{I}_c f_{\text{CI}} \{\mathbf{a}\}\|_2^2 \right\}, \quad (14)$$

where \mathbf{I}_c is a mapping matrix based on \mathcal{Z}_c , as seen in Section II-A. Explicitly, the JML detector of (14) has the complexity of $\mathcal{O}(2^{L_1} Q^K)$, which may become excessive in practice, when the constellation size Q is large and/or the dimensions of the measurement matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ are high. In order to reduce the detection complexity and to glean benefits from employing the CS principles, below we propose a lower-complexity IRC detector.

B. Iterative Residual Check Detector

Let us first rewrite (9) as

$$\mathbf{y}_g = \Phi_g \mathbf{x}_g + \bar{\mathbf{w}}_g \quad (15a)$$

$$\begin{aligned} &= \Phi_g \mathbf{I}_g f_{\text{CI}} \{\mathbf{x}_{\text{d},g}\} + \bar{\mathbf{w}}_g \\ &= \Phi_{g,\mathcal{I}_g} f_{\text{CI}} \{\mathbf{x}_{\text{d},g}\} + \bar{\mathbf{w}}_g, \end{aligned} \quad (15b)$$

where by definition we have $\Phi_g = \bar{\mathbf{H}}_g \mathbf{A}$ and $\Phi_{g,\mathcal{I}_g} = \Phi_g \mathbf{I}_g$. In (15a), the N -dimensional symbol vector $\mathbf{x}_g \in \mathcal{V}$ shown in Section II-A is exactly K -sparse, which is observed in the spaces of m dimensions. Hence, according to the CS principles, it can be recovered by solving the classic ℓ_0 -minimization problem described in [39, 40] as

$$\mathbf{x}_g^{\ell_0} = \arg \min_{\mathbf{v} \in \mathbb{C}^{N \times 1}} \|\mathbf{v}\|_0, \quad \text{s.t. } \mathbf{y}_g - \Phi_g \mathbf{v} \in \mathcal{B}, \quad (16)$$

where \mathcal{B} is a bounded set associated with the noise vector $\bar{\mathbf{w}}_g$, while \mathbf{v} is an N -length testing vector chosen from the complex vector space $\mathbb{C}^{N \times 1}$.

However, the ℓ_0 -minimization problem of (16) has been shown to be NP-hard [41]. For this reason, typically the ℓ_1 -minimization based solutions are employed in CS, since they are tractable to solve. Moreover, under the conditions shown in [22], the ℓ_0 -minimization problem of (16) is equivalent to solving the ℓ_1 -minimization problem of

$$\mathbf{x}_g^{\ell_1} = \arg \min_{\mathbf{v} \in \mathbb{C}^{N \times 1}} \|\mathbf{v}\|_1, \quad \text{s.t. } \mathbf{y}_g - \Phi_g \mathbf{v} \in \mathcal{B}. \quad (17)$$

Note that the ℓ_1 -minimization problem is a well-known convex optimisation problem, which can be solved in polynomial time, for example, by the interior point method [42]. However, for our detection problem, the computational cost of solving the ℓ_1 -minimization problem may still be excessive. Moreover, in the OFDM-CSIM system of Section II, the sparsity level K of the symbol vector \mathbf{x}_g is determined by the number of active indices, which is a fixed and known value. In this case, the lower-complexity family of greedy pursuit algorithms [30, 43] is preferred. However, the existing greedy algorithms can only provide relatively good detection performance, when the SNR is relatively high. Another concern is that for the existing greedy algorithms, the sparse estimation is operated without any restrictions on the signal space. Thus, the existing greedy algorithms cannot be directly invoked for detection in our scenario.

Based on the philosophy of greedy pursuits [33], which provides an approximated estimation via making locally optimal choices at each step, we propose a Greedy-like detector, namely the IRC detector, for detecting both the IM and the classic APM symbols in the sparse vector \mathbf{x}_g . The proposed IRC detector has the advantage of not only achieving a low detection complexity, but also guaranteeing a good detection performance even in the low to medium SNR regions. Our IRC detector is divided into two stages, an initialization stage and an iterative detection stage, as detailed below.

Firstly, at the initialization stage, the IRC detector invokes the minimum mean square error (MMSE) processing to obtain an N -dimensional soft estimate of $\hat{\mathbf{x}}_g \in \mathbb{C}^{N \times 1}$ for \mathbf{x}_g from the m -dimensional observations \mathbf{y}_g seen in (15a), yielding

$$\hat{\mathbf{x}}_g = \left(\Phi_g^H \Phi_g + \frac{1}{\gamma_s} \mathbf{I}_N \right)^{-1} \Phi_g^H \mathbf{y}_g, \quad (18)$$

where $\gamma_s = \mathbb{E}[\|\mathbf{s}_T\|_2^2]/\mathbb{E}[\|\mathbf{w}_T\|_2^2] = K/(mN_0)$ is the average SNR per symbol. In principle, the active elements in \mathbf{x}_g should generate relatively high magnitudes in $\hat{\mathbf{x}}_g$ with high probability. Therefore, we order the elements in $\hat{\mathbf{x}}_g$ in descending order according to their magnitudes as

$$|\hat{x}_g(i_1)|^2 \geq |\hat{x}_g(i_2)|^2 \geq \dots \geq |\hat{x}_g(i_N)|^2, \quad (19)$$

which more or less reflects the reliabilities of the IM detection. Since the set $\{i_1, i_2, \dots, i_N\}$ contains K active indices and $(N - K)$ inactive indices, the leftmost $\hat{x}_g(i_1)$ has the highest probability to be one of the active indices, while the rightmost $\hat{x}_g(i_N)$ has the highest probability to be one of the inactive indices. With the aid of this information, the IRC detector then proceeds to the second stage, where both the IM and the classic APM symbols can be detected using the principle of greedy pursuits, as detailed next.

During the second stage, the IRC detector carries out iterative detection of both the IM symbols in \mathcal{I}_g and the classic APM symbols in $\mathbf{x}_{\text{d},g}$ by first identifying a candidate set, and then testing the candidates based on the likelihood principle of (14). In detail, during the first iteration, the IRC detector first derives a candidate set from the index i_1 of $\hat{x}_g(i_1)$, since it has the highest probability to be an active index. Let this candidate set be expressed as

$$\mathcal{Z}^1 = \{\mathcal{Z}_1^1, \mathcal{Z}_2^1, \dots, \mathcal{Z}_{C_1}^1\} \subset \mathcal{Z} \quad (20)$$

where i_1 is a member of each of the candidates, i.e., we have $\bigcap_{c_1=1}^{C_1} \mathcal{Z}_{c_1}^1 = i_1$. Here, we can readily show that the number of candidates in \mathcal{Z}^1 of (20) obeys $C_1 \leq \binom{N-1}{K-1} < C$, where $C = 2^{L_1}$ denotes the total number of subsets of \mathcal{Z} , as shown in Section II-A. Then, for the candidate $\mathcal{Z}_{c_1}^1$, we can obtain the Moore-Penrose pseudoinverse matrix

$$\Phi_{g,\mathcal{Z}_{c_1}^1}^{\dagger} = \left(\Phi_{g,\mathcal{Z}_{c_1}^1}^H \Phi_{g,\mathcal{Z}_{c_1}^1} \right)^{-1} \Phi_{g,\mathcal{Z}_{c_1}^1}^H, \quad \mathcal{Z}_{c_1}^1 \subset \mathcal{Z}^1 \subset \mathcal{Z}, \quad (21)$$

where $\Phi_{g,\mathcal{Z}_{c_1}^1} = \Phi_g \mathbf{I}_{c_1}^1 \in \mathbb{C}^{m \times K}$ is of full column-rank when $K < m$, while $\mathbf{I}_{c_1}^1$ is a $(N \times K)$ -element mapping matrix formed based on $\mathcal{Z}_{c_1}^1$. Then, with the aid of (21) and (15b),

we can form the estimation of the APM symbols as

$$\begin{aligned}\check{\mathbf{x}}_{d,g}(c_1) &= f_{CI}^{-1} \left\{ \Phi_{g,\mathcal{Z}_{c_1}^1}^\dagger \mathbf{y}_g \right\} \\ &= f_{CI}^{-1} \left\{ \Phi_{g,\mathcal{Z}_{c_1}^1}^\dagger \Phi_{g,\mathcal{I}_g} f_{CI} \{ \mathbf{x}_{d,g} \} + \Phi_{g,\mathcal{Z}_{c_1}^1}^\dagger \bar{\mathbf{w}}_g \right\} \\ &= \mathbf{x}_{d,g} + \mathbf{r}_{\mathcal{I}_g, \mathcal{Z}_{c_1}^1} + \check{\mathbf{w}}_g,\end{aligned}\quad (22)$$

where $\mathbf{r}_{\mathcal{I}_g, \mathcal{Z}_{c_1}^1}$ is the potential residual interference caused by the indices in \mathcal{I}_g , but not included in $\mathcal{Z}_{c_1}^1$, i.e., $\mathcal{Z}_{c_1}^1 \neq \mathcal{I}_g$. Otherwise, if $\mathcal{Z}_{c_1}^1 = \mathcal{I}_g$, i.e., both \mathcal{I}_g and $\mathcal{Z}_{c_1}^1$ contain the same set of indices representing a correct detection of the index symbols, then we have $\mathbf{r}_{\mathcal{I}_g, \mathcal{Z}_{c_1}^1} = \mathbf{0}$. In (22), the noise vector $\check{\mathbf{w}}_g = f_{CI}^{-1} \{ \Phi_{g,\mathcal{Z}_{c_1}^1}^\dagger \bar{\mathbf{w}}_g \} \in \mathbb{C}^{K \times 1}$ is the linear combination of Gaussian noise samples, which is also a Gaussian noise vector. Based on (22), a simple symbol-by-symbol ML detection can be carried out, yielding

$$a^1(k, c_1) = \arg \min_{a_q \in \mathcal{A}_\phi} \| \check{\mathbf{x}}_{d,g}(k, c_1) - a_q \|^2, \quad k = 0, \dots, K-1. \quad (23)$$

So far, we have obtained the estimations of the classic APM symbols $\mathbf{a}^1(c_1) = [\mathbf{a}^1(0, c_1), \dots, \mathbf{a}^1(K-1, c_1)]^T$ from the candidate set $\mathcal{Z}_{c_1}^1 \in \mathcal{Z}^1$. Similarly, we can obtain the estimates of the classic APM symbols from the other candidates in \mathcal{Z}^1 of (20). Based on the estimates, let us define a set $\mathcal{A}^1 = \{ \mathbf{a}^1(1), \dots, \mathbf{a}^1(C_1) \} \subset \mathcal{A}_\phi^K$ for all the possible APM symbols detected using the indices in \mathcal{Z}^1 . Then, the best one among them after the first iteration can be found by solving the optimization of

$$(\mathcal{I}_g^{[1]}, \mathbf{x}_{d,g}^{[1]}) = \arg \min_{\mathcal{Z}_{c_1}^1 \in \mathcal{Z}^1, \mathbf{a}^1(c_1) \in \mathcal{A}^1} \| \mathbf{y}_g - \Phi_{g,\mathcal{Z}_{c_1}^1} f_{CI} \{ \mathbf{a}^1(c_1) \} \|^2. \quad (24)$$

Correspondingly, the residual error $\varepsilon^{[1]}$ is given by

$$\varepsilon_g^{[1]} = \| \mathbf{y}_g - \Phi_{g,\mathcal{I}_g^{[1]}} f_{CI} \{ \mathbf{x}_{d,g}^{[1]} \} \|^2. \quad (25)$$

Let ε_{TS} be a threshold, which is used for terminating the search process, when the detection is believed to have achieved the required reliability. Then, if $\varepsilon_g^{[1]} \leq \varepsilon_{TS}$, the search process is completed, and the results given in (24) are taken as the final estimates of the received IM and the classic APM symbols. Otherwise, the search process forwards to the next iteration.

Similarly to the first iteration, during the second iteration, the IRC detector first identifies a candidate set based on the index i_2 of the second largest in magnitude of the variables given in (19). It can be shown that the number of new candidates containing i_2 as a member satisfies $C_2 \leq \binom{N-1}{K-1} - \binom{N-2}{K-2} = \binom{N-2}{K-1} < C$, after excluding those having been considered in association with the candidates derived from index i_1 . Having obtained the candidate set for the second iteration, the other processes are repeated in the same way as in the first iteration. In general, when the IRC detector enters into the t th iteration, the number of new candidates including i_t as a member satisfies $C_t \leq \binom{N-t}{K-1} < C$. Based on the above-mentioned method, when all the N indices are checked, we can readily show that $\sum_{t=1}^N C_t = C = 2^{L_1}$. This implies that

all the candidates in $\mathcal{Z} = \{ \mathcal{Z}_1, \dots, \mathcal{Z}_C \}$ have been checked by the IRC detector.

Finally, if the condition of $\varepsilon^{[t]} \leq \varepsilon_{TS}$ cannot be fulfilled after the maximum affordable number of iterations, the detection result giving the minimum number of residual errors is output. In summary, our proposed IRC detector can be stated as in Algorithm 1.

The complexity of our proposed algorithm is determined by the number of iterations as well as the number of candidates in each iteration. Explicitly, the best case is that the detection is completed within a single iteration by testing only one candidate. In this case, we can readily show that the complexity is on the order of $\mathcal{O}_{IRC}^{\text{best}}(QK)$. On the other hand, the worst case is that all the candidates in \mathcal{Z} are checked. In this case, the corresponding complexity is given by $\mathcal{O}_{IRC}^{\text{worst}}(2^{L_1} QK)$. Note that for $K > 1$ active indices, the probability that i_1 seen in (19) is not an active index is usually very small. Therefore, in most cases, our IRC detector requires a single iteration for testing the $C_1 \leq \binom{N-1}{K-1} < C$ candidates. In this case, the detection complexity is given by $\mathcal{O}_{IRC}^{\text{usually}}(C_1 QK)$. In fact, our IRC detector can be constrained to simply carry out one or two iterations, as it will be shown in Section V.

IV. PERFORMANCE ANALYSIS

In this section, the system performance of the OFDM-CSIM system is analyzed. In our analysis, the JML detector is assumed. We commence by deriving the SE and EE of the proposed system. Then, the analytical diversity order of the OFDM-CSIM system is detailed.

A. Spectral Efficiency and Energy Efficiency

Let us assume that the proposed OFDM-CSIM system is operated within a frequency band of f_B Hertz. The total power consumed by transmitting the data at a reliable rate of R in bit-per-second within a symbol duration of T_s is denoted as P_T Watts. We consider the specific case when the total power consumed is equal to the total power of the transmitted signal, i.e. we have $P_T = P_{Tx}$. Then, the SE of $\eta_{SE} = R/f_B$ can be measured in bit-per-second-per-Hertz. Since each group is independent, the SE of the system shown in Section II can be expressed in terms of the channel SNR as [44]

$$\begin{aligned}f_{SE}(\gamma_s) &= \frac{R}{f_B} \\ &= \frac{\sum_{g=1}^G I(\mathbf{x}_g; \mathbf{y}_g)}{T_s f_B} = \frac{\sum_{g=1}^G I(\mathbf{x}_g; \mathbf{y}_g)}{M},\end{aligned}\quad (26)$$

where $I(\mathbf{x}_g; \mathbf{y}_g)$ denotes the mutual information [45] associated with the transmission of the g th group. Here, the rate reduction caused by the CP is ignored. Then, upon substituting

(A.6) into (26), we have

$$\begin{aligned}
f_{\text{SE}}(\gamma_s) &= \frac{1}{M} \sum_{g=1}^G \left\{ \mathbb{E}_{\bar{\mathbf{H}}_g, \bar{\mathbf{w}}_g} [-\log_2 \epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s)] \right. \\
&\quad \left. - m \log_2 \left(\frac{\pi e K}{m \gamma_s} \right) \right\} \\
&= \frac{G}{M} \left\{ \mathbb{E}_{\bar{\mathbf{H}}_g, \bar{\mathbf{w}}_g} [-\log_2 \epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s)] \right. \\
&\quad \left. - m \log_2 \left(\frac{\pi e K}{m \gamma_s} \right) \right\} \\
&= \frac{\mathbb{E}_{\bar{\mathbf{H}}_g, \bar{\mathbf{w}}_g} [-\log_2 \epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s)]}{m} - \log_2 \left(\frac{\pi e K}{m \gamma_s} \right), \tag{27}
\end{aligned}$$

Algorithm 1 IRC Detector

Require: $\mathbf{y}_g, \Phi_g = \bar{\mathbf{H}}_g \mathbf{A}, \gamma_s, \mathcal{Z}, \varepsilon_{\text{TS}}$

1: **Initialization:** Set the maximum number of iterations to T , a check space of $\mathcal{Z}_{\text{check}} = \mathcal{Z}$, $\varepsilon_0 = \infty$, $\mathcal{I}_g^{\text{IRC}} = \emptyset$ and $\mathbf{x}_{\text{d},g}^{\text{IRC}} = \emptyset$.

2: MMSE detection based on (18), expressed as

$$\hat{\mathbf{x}}_g = \left(\Phi_g^H \Phi_g + \frac{1}{\gamma_s} \mathbf{I}_N \right)^{-1} \Phi_g^H \mathbf{y}_g;$$

3: Order the elements in $\hat{\mathbf{x}}_g$ in descending order according to their magnitudes as

$$|\hat{x}_g(i_1)|^2 \geq |\hat{x}_g(i_2)|^2 \geq \dots \geq |\hat{x}_g(i_N)|^2.$$

4: **for** $t = 1$ to T **do**

5: Obtain $\mathcal{Z}^t = \{\mathcal{Z}_1^t, \mathcal{Z}_2^t, \dots, \mathcal{Z}_{C_t}^t\}$ from $\mathcal{Z}_{\text{check}}$, where $\bigcap_{c_t=1}^{C_t} \mathcal{Z}_{c_t}^t = i_t$;

6: **if** $\mathcal{Z}^t = \emptyset$ **then**

7: **else**

8: For $c_t = 1, \dots, C_t$, calculate

$$\check{\mathbf{x}}_{\text{d},g}(c_t) = f_{\text{CI}}^{-1} \left\{ \Phi_{g, \mathcal{Z}_{c_t}^t}^{\dagger} \mathbf{y}_g \right\};$$

9: Generate the estimates for the APM symbols associated with the candidates of $c_t = 1, \dots, C_t$ according to (23) as

$$a^t(k, c_t) = \arg \min_{a_q \in \mathcal{A}_\phi} \left\| \check{\mathbf{x}}_{\text{d},g}(k, c_t) - a_q \right\|^2, \quad k = 0, 1, \dots, K-1$$

10: Find the best estimates according to (24) as

$$(\mathcal{I}_g^{[t]}, \mathbf{x}_{\text{d},g}^{[t]}) = \arg \min_{\mathcal{Z}_{c_t}^t \in \mathcal{Z}^t, \mathbf{a}^t(c_t) \in \mathcal{A}^t} \left\| \mathbf{y}_g - \Phi_{g, \mathcal{Z}_{c_t}^t} f_{\text{CI}} \left\{ \mathbf{a}^t(c_t) \right\} \right\|^2$$

11: Calculate the residual error according to (25) as

$$\varepsilon_g^{[t]} = \left\| \mathbf{y}_g - \Phi_{g, \mathcal{I}_g^{[t]}} f_{\text{CI}} \left\{ \mathbf{x}_{\text{d},g}^{[t]} \right\} \right\|^2;$$

12: **if** $\varepsilon_g^{[t]} < \varepsilon_{\text{TS}}$ **then**

13: Update $\mathcal{I}_g^{\text{IRC}} = \mathcal{I}_g^{[t]}$ and $\mathbf{x}_{\text{d},g}^{\text{IRC}} = \mathbf{x}_{\text{d},g}^{[t]}$.

14: **exit**

15: **else if** $\varepsilon_g^{[t]} < \varepsilon_0$ **then**

16: Update $\mathcal{Z}_{\text{check}} = \mathcal{Z}_{\text{check}} \setminus \mathcal{Z}^t$, $\varepsilon_0 = \varepsilon_g^{[t]}$, $\mathcal{I}_g^{\text{IRC}} = \mathcal{I}_g^{[t]}$ and $\mathbf{x}_{\text{d},g}^{\text{IRC}} = \mathbf{x}_{\text{d},g}^{[t]}$.

17: **else**

18: **end if**

19: **end if**

20: **end for**

21: **return** $\mathcal{I}_g^{\text{IRC}}$ and $\mathbf{x}_{\text{d},g}^{\text{IRC}}$.

where $\epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s)$ is the *implementation function*, as defined in (A.7). It should be noted that it is not easy to further simplify (27) to a closed-form expression. Nevertheless, according to (A.7), its first term can be evaluated by Monte Carlo simulations. By contrast, since the second term of (27) is determined by the SNR per symbol γ_s , the SE of (27) can be obtained accordingly, as it will be shown in Section V. Furthermore, we can readily show with the aid of (A.10) that the SE of (27) is upper bounded by

$$\begin{aligned}
f_{\text{SE}}(\gamma_s) &= \sum_{g=1}^G \frac{I(\mathbf{x}_g; \mathbf{y}_g)}{M} \\
&\leq \sum_{g=1}^G \frac{L}{M} = \frac{L}{m}, \tag{28}
\end{aligned}$$

where the upper bound L/m is attainable, when the SNR per symbol γ_s is sufficiently high.

On the other hand, the EE can be defined in terms of Joule-per-bit per noise level [44], or bit-per-Joule [3, 46], or bit-per-Joule per noise level [1]. In this paper, only the signal's transmission power is considered, hence the EE definition of [1] is adopted. In this case, given the SE value of $f_{\text{SE}}(\gamma_s) = \eta_{\text{SE}}$, the corresponding EE can be expressed as [1]

$$\begin{aligned}
\eta_{\text{EE}} &= \frac{\eta_{\text{SE}}}{P_{\text{T}}/N_0} = \frac{\eta_{\text{SE}}}{P_{\text{Tx}}/N_0} \\
&= \frac{\eta_{\text{SE}}}{\gamma_s} = \frac{\eta_{\text{SE}}}{f_{\text{SE}}^{-1}(\eta_{\text{SE}})}, \tag{29}
\end{aligned}$$

where the last equality is arrived at by expressing γ_s as an inverse function of the SE η_{SE} , which gives the SNR per symbol required for achieving the SE of η_{SE} . Again, since it is an open challenge to derive a closed-form expression for the EE shown in (29), the tradeoff between the SE and EE of our proposed system is studied by Monte Carlo simulations, as it will be shown in Section V.

Furthermore, let us define the received SNR per bit (or SNR per reliable bit) as $\gamma_b^{\text{Rx}} \triangleq E_b^{\text{Rx}}/N_0$, where $E_b^{\text{Rx}} = P_{\text{Tx}}/\eta_{\text{SE}}$ is the received energy per bit in Joule-per-bit. Hence, we can explicitly show that γ_b^{Rx} can be formulated in terms of the EE

defined in (29) as

$$\begin{aligned}\gamma_b^{\text{Rx}} &= \frac{\gamma_s}{\eta_{\text{SE}}} = \frac{f_{\text{SE}}^{-1}(\eta_{\text{SE}})}{\eta_{\text{SE}}} \\ &= \frac{1}{\eta_{\text{EE}}},\end{aligned}\quad (30)$$

which can be used for quantifying the minimum received signal energy per information bit required for reliable communication, as it will be studied by simulation results shown in Section V.

B. Diversity and Coding Gains

Let us first define the pairwise error event as $\{\mathbf{x}_g^c \rightarrow \mathbf{x}_g^e\}$, where $\mathbf{x}_g^c = \mathbf{v}_i$ is the transmitted symbol vector according to (2), while $\mathbf{x}_g^e = \mathbf{v}_j$ denotes the corresponding incorrect detection results of the ML detection shown in (13), i.e. we have $\mathbf{v}_i \neq \mathbf{v}_j$ for $\mathbf{v}_i, \mathbf{v}_j \in \mathcal{V}$. Based on (4), let $\mathbf{s}_g^c = \mathbf{A}\mathbf{x}_g^c$ and $\mathbf{s}_g^e = \mathbf{A}\mathbf{x}_g^e$. Then, when the measurement matrix \mathbf{A} of the mutual coherence of (10) is used, similar to [47], the conditional pairwise error probability (PEP) can be expressed as

$$\begin{aligned}P(\mathbf{x}_g^c \rightarrow \mathbf{x}_g^e | \bar{\mathbf{H}}_g) &= P(\mathbf{s}_g^c \rightarrow \mathbf{s}_g^e | \bar{\mathbf{H}}_g) \\ &= Q\left(\sqrt{\frac{m\gamma_s \|\bar{\mathbf{H}}_g(\mathbf{s}_g^c - \mathbf{s}_g^e)\|^2}{2K}}\right) \\ &= \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \exp\left(-\frac{m\gamma_s d^2(\mathbf{z}_g^c, \mathbf{z}_g^e)}{4K \sin^2 \theta}\right) d\theta,\end{aligned}\quad (31)$$

where $d^2(\mathbf{z}_g^c, \mathbf{z}_g^e) = \|\mathbf{z}_g^c - \mathbf{z}_g^e\|^2$ denotes the squared Euclidean distance between $\mathbf{z}_g^c = \bar{\mathbf{H}}_g \mathbf{s}_g^c$ and $\mathbf{z}_g^e = \bar{\mathbf{H}}_g \mathbf{s}_g^e$. With the aid of (7), we have

$$\begin{aligned}d^2(\mathbf{z}_g^c, \mathbf{z}_g^e) &= \|\bar{\mathbf{H}}_g(\mathbf{s}_g^c - \mathbf{s}_g^e)\|^2 \\ &= \|\text{diag}\{\mathbf{F}_{\text{h},\Pi_g} \mathbf{h}_{\text{T}}\}(\mathbf{v}_i - \mathbf{v}_j)\|^2 \\ &= \|\text{diag}\{\mathbf{F}_{\text{h},\Pi_g} \mathbf{h}_{\text{T}}\}\mathbf{e}\|^2 \\ &= \|\mathbf{E}\mathbf{F}_{\text{h},\Pi_g} \mathbf{h}_{\text{T}}\|^2 \\ &= \mathbf{h}_{\text{T}}^H \mathbf{D}_g \mathbf{h}_{\text{T}},\end{aligned}\quad (32)$$

where we have $\mathbf{E} = \text{diag}\{\mathbf{e}\} \in \mathcal{C}^{m \times m}$ with $\mathbf{e} \in \mathcal{E}$ as defined in (A.5) and $\mathbf{D}_g = \mathbf{F}_{\text{h},\Pi_g}^H \mathbf{E}^H \mathbf{E} \mathbf{F}_{\text{h},\Pi_g} \in \mathcal{C}^{L_{\text{h}} \times L_{\text{h}}}$. Furthermore, let the nonzero eigenvalues of \mathbf{D}_g be denoted as λ_i for $i = 1, \dots, V_{\text{D}}$, where

$$\begin{aligned}V_{\text{D}} &= \text{rank}(\mathbf{D}_g) = \text{rank}(\mathbf{F}_{\text{h},\Pi_g}^H \mathbf{E}^H \mathbf{E} \mathbf{F}_{\text{h},\Pi_g}) \\ &= \min\{\text{rank}(\mathbf{E}), \text{rank}(\mathbf{F}_{\text{h},\Pi_g})\}.\end{aligned}\quad (33)$$

We can readily show that we have $\text{rank}(\mathbf{E}) = \|\mathbf{e}\|_0 \leq m$ and $\text{rank}(\mathbf{F}_{\text{h},\Pi_g}) = \min\{m, L_{\text{h}}\}$, which gives

$$V_{\text{D}} = \begin{cases} \|\mathbf{e}\|_0 & \text{if } m < L_{\text{h}} \\ L_{\text{h}} & \text{if } m \geq L_{\text{h}}. \end{cases}\quad (34)$$

We should emphasize here that since the measurement matrix $\mathbf{A} \in \mathcal{C}^{m \times N}$ is used in our proposed system, the number of non-zero elements in $\mathbf{e} \in \mathcal{E}$ is $\|\mathbf{e}\|_0 = m$ for most cases. As shown in [19], the maximum diversity order achieved by the

TABLE I: Parameters for All Simulations

K	Q	ϕ	M	L_{cp}	L_{h}
2	4	$\pi/12$	256	16	10
		m	N	$\mu_{\mathbf{A}}$	$1/(2K-1)$
$\mathbf{A} \in \mathcal{C}^{m \times N}$	8	15	0.2500	0.3333	
	16	31	0.1768		
	32	61	0.1229		

conventional OFDM-SIM system is $2K$. By contrast, in our proposed system, since (10) has to be satisfied for assisting the detection, it can be readily shown that for a large N , we have $2K < \sqrt{m}$, resulting in $m > (2K)^2$. Therefore, the diversity order of our proposed system may be much higher than that of the conventional OFDM-SIM system, provided that $L_{\text{h}} > 2K$ is satisfied.

Since \mathbf{h}_{T} is assumed to be an i.i.d. Gaussian distributed vector, after integrating the conditional PEP of (31) with respect to the distribution of the squared Euclidean distance of $d^2(\mathbf{z}_g^c, \mathbf{z}_g^e)$ by following the approach shown in [8], the average PEP can be expressed as

$$P(\mathbf{x}_g^c \rightarrow \mathbf{x}_g^e) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{i=1}^{V_{\text{D}}} \left(1 + \frac{\lambda_i m \gamma_s}{4K \sin^2 \theta}\right)^{-1} d\theta. \quad (35)$$

Furthermore, for sufficiently high SNRs, (35) is upper bounded by [17, 48]

$$P(\mathbf{x}_g^c \rightarrow \mathbf{x}_g^e) \leq \left(V_{\text{C}} \frac{m \gamma_s}{4K}\right)^{-V_{\text{D}}}, \quad (36)$$

where V_{D} gives the diversity order of the detection, while $V_{\text{C}} = \left(\prod_{i=1}^{V_{\text{D}}} \lambda_i\right)^{1/V_{\text{D}}}$ is the coding gain of the OFDM-CSIM transmission scheme. We can see from (36) that the diversity order V_{D} determines how rapidly the average PEP decreases with the increase of the SNR, while the coding gain V_{C} determines the SNR-shift of this PEP curve relative to a benchmark error rate curve of $(\gamma_s/4)^{-V_{\text{D}}}$.

From the above analysis, we can infer the following observations. Firstly, when $m < L_{\text{h}}$, the diversity order that can be achieved by the proposed OFDM-CSIM system is dependent on the number of non-zero elements in $\mathbf{e} \in \mathcal{E}$, which is m in most cases. Secondly, provided that $m \geq L_{\text{h}}$, the diversity order achieved by the proposed OFDM-CSIM system is L_{h} , which is in this case as high as the number of multipath components. Finally, since L given in (3) is independent of m in our proposed system, the higher the value of m , the lower the achievable SE of the system, as shown in (28). Hence, the diversity order of the proposed system can be increased at the cost of decreasing the achievable SE of the system.

V. SIMULATION RESULTS

In this section, simulation results are provided for characterizing the achievable performance of the proposed OFDM-CSIM system. The system setup and the parameters used in our simulations are summarized in Table I. For all simulations, a ten-path (i.e. $L_{\text{h}} = 10$) slow-varying Rayleigh fading channel

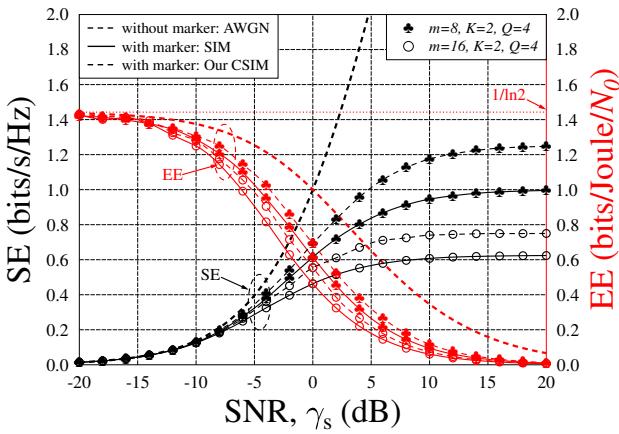


Fig. 4: SE and EE as a function of the SNR in dB for both the OFDM-SIM and the OFDM-CSIM systems communicating over an $L_h = 10$ -path Rayleigh fading channel. The system parameters are given in Table I.

is considered, which is based on Section II-B. We assume an OFDM system employing $M = 256$ subcarriers and a CP of length of $L_{cp} = 16$. Moreover, QPSK modulation associated with a rotated angle of $\phi = \pi/12$ is used for modulating the activated subcarriers. For both the OFDM-SIM and our OFDM-CSIM systems, the number of active indices is chosen to be $K = 2$. Specifically, in order to satisfy the condition of (10), the partial Fourier matrix constructed by the search algorithm of [26] is employed as the measurement matrix \mathbf{A} for the OFDM-CSIM system, as shown in Table I³.

In Fig. 4, both the SE and EE are investigated as a function of the SNR in dB for both the OFDM-SIM and the OFDM-CSIM systems, when communicating over Rayleigh fading channels. In this figure, the unrotated QPSK constellation (i.e., $\phi = 0$) is assumed, but the CI and the depth- G interleaver are not employed. Furthermore, both the OFDM-SIM and the OFDM-CSIM systems employ the JML detector of Section III-A for the sake of comparison. Firstly, we observe from Fig. 4 that the achievable SE of the proposed CSIM scheme is higher than that of the conventional SIM scheme. This observation can be explained with the aid of the analytical results of (3) and (28). Explicitly, as shown in (28), when the value of m is fixed, the achievable SE of both the SIM and the CSIM schemes depend on the number of data bits L transmitted by each group. Furthermore, when the values of K and Q are fixed, more data bits can be transmitted by the CSIM associated with $N > m$ than by the SIM using $N = m$, as exemplified in Section II-A. Thus, the attainable SE of the CSIM is higher than that of the conventional SIM. Moreover, we can also show that the corresponding SE is reduced both for the CSIM and for the SIM, as the value of m is increased. Secondly, as seen in Fig. 4, for a given SNR, our proposed CSIM scheme is capable of achieving a higher EE than the conventional SIM scheme. Meanwhile, the SE of our

³It should be noted that for $m = 4, K = 2$, the measurement matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ satisfying (10) does not exist. Hence, the case of $m = 4$ is not considered in our simulations

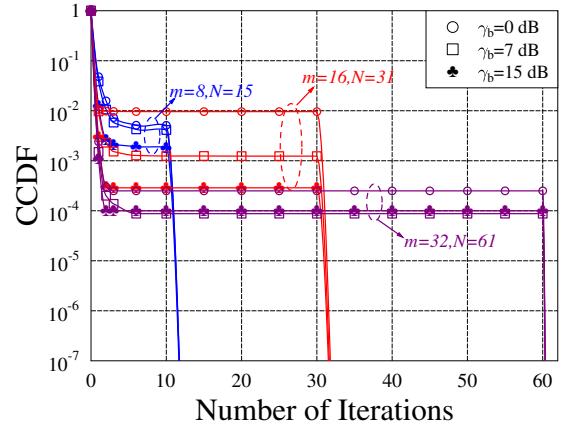


Fig. 5: CCDF of the number of iterations of the IRC detector for the CSIM employing CI and a depth- G interleaver operates at the SNR per bit value of $\gamma_b = 0$ dB, 7 dB or 15 dB, when communicating over an $L_h = 10$ -path Rayleigh fading channel.

CSIM scheme is also higher than that of the conventional SIM scheme. In other words, our proposed CSIM scheme is capable of striking a more appealing tradeoff between the SE and EE than the conventional SIM scheme. This observation can be explained with the aid of the analytical results of (29) and the numerical SE results shown in this figure. As observed in Fig. 4, for a given SE, the SNR required by our CSIM scheme is lower than that of the SIM scheme. Since the EE defined by (29) is inversely proportional to the SNR for a given SE, the EE of our CSIM is in turn higher than that of the conventional SIM. Thirdly, as shown in Fig. 4, the maximum EE that can be attained by both schemes is about $1/\ln 2 \approx 1.4427$. This result inferred from (30) represents the minimum energy per bit required for reliable communication, which is given by $\gamma_{b,\min}^{\text{Rx}} = \ln 2$. This confirms the analytical result given in [44]. As demonstrated in [44], when the SNR is low, the minimum energy per bit required for reliable communication is $\ln 2$, which is independent of the fading. Hence, for OFDM systems having a fixed transmission power, bandwidth and IFFT/FFT size, our proposed CSIM scheme is capable of striking a more appealing tradeoff between the SE and EE than the conventional SIM scheme.

Fig. 5 plots the complementary cumulative distribution function (CCDF) of the number of iterations for the proposed IRC detector of our CSIM scheme, when communicating over an $L_h = 10$ -path Rayleigh fading channel. Generally, the CCDF is the probability that a random variable X will take a value higher than x , i.e. we have $\text{CCDF}_X(x) \triangleq P(X > x)$. In this figure, the OFDM-CSIM system employing CI and the depth- G interleaver operates at the SNR per bit value of $\gamma_b = 0$ dB, 7 dB or 15 dB. As seen in Fig. 5, the CCDF curves first exhibit a sharp decay and then become flat as the number of iterations increases. This observation implies that with a near-unity probability, the IRC detector completes the detection in just a single iteration. Furthermore, when just a

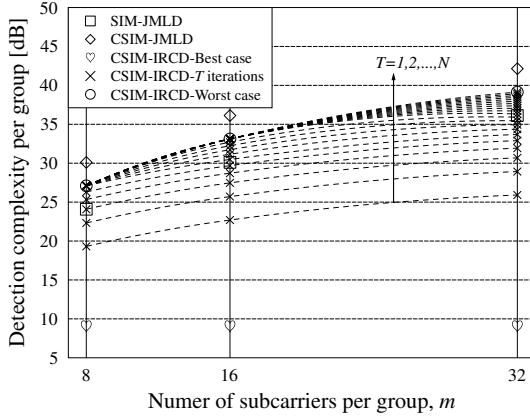


Fig. 6: Detection complexity required for detecting each group of symbols, when the JML detector (JMLD) and the IRC detector (IRCD) are employed. For the IRCD of Algorithm 1, the maximum number of iterations is chosen to be $T = 1, \dots, N$.

single iteration is considered, the difference of the probabilities for a given CSIM operating at different SNR per bit values is unnoticeable. Thus, it can be inferred that the performance of the IRC detector operating at just one iteration is stable at all SNR per bit values. Finally, we can also see that as the EE of the CSIM system increases, the probability of completing detection in just one iteration increases. It should be noted that the fewer iterations the IRC detector invokes, the lower its complexity becomes, as analyzed in Section III-B.

In Fig. 6, we investigate the detection complexity imposed by detecting each group of symbols, when both the JML and IRC detectors are employed. In this figure, the IRC detector of Algorithm 1 relying on $T = 1, \dots, N$ iterations is considered for our CSIM scheme. We observe that when JML detectors are used, the detection complexity of the CSIM is higher than that of the conventional SIM. However, as seen in Fig. 6, the detection complexity can be significantly reduced by using our proposed IRC detector, especially when the maximum number of iterations is set to $T = 1$ or 2 . Furthermore, we can observe that the detection complexity of our CSIM using the IRC detector is lower than that of the same scheme using the JML detector, even for the worst case of the IRC detector. There are two reasons for this trend. Firstly, each index is uniquely checked, as shown in Algorithm 1. In this way, the repeated reliability computations of the same index are avoided. The other reason can be readily inferred from comparing (23) to (14), where the APM symbols are detected on a symbol-by-symbol basis by the IRC detector, while the joint group detection is carried out by the JML detector. Finally, we can show with the aid of Fig. 5 that our proposed IRC detector is an efficient one for the CSIM considered.

The BER performance of the classic OFDM, OFDM-SIM and OFDM-CSIM systems is compared in Fig. 7, Fig. 8 and Fig. 9. In these figures, both the OFDM-SIM and the OFDM-CSIM systems employ CI, the depth- G interleaver as well as JML detectors. Explicitly, for the OFDM-CSIM systems, Fig. 7, Fig. 8 and Fig. 9, respectively, are characterized by

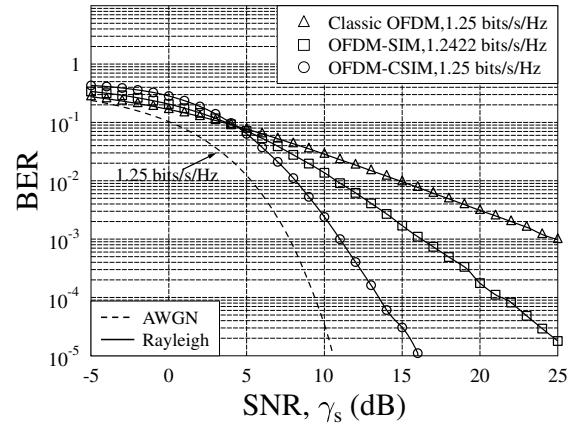


Fig. 7: BER performance of the classic OFDM, OFDM-SIM and OFDM-CSIM systems communicating over an $L_h = 10$ -path Rayleigh fading channel. Both the OFDM-SIM and the OFDM-CSIM systems employ CI, a depth- G interleaver as well as JML detectors.

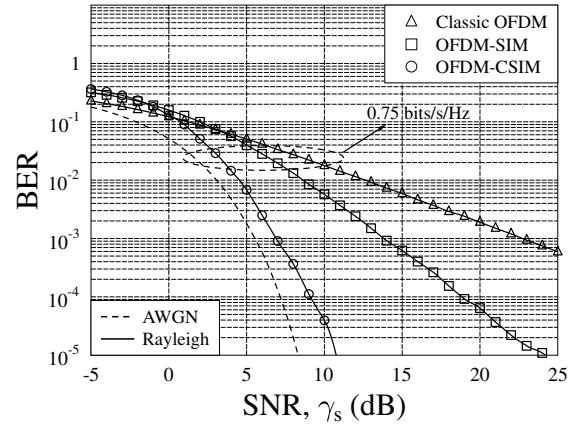


Fig. 8: BER performance of the classic OFDM, OFDM-SIM and OFDM-CSIM systems communicating over an $L_h = 10$ -path Rayleigh fading channel. Both the OFDM-SIM and the OFDM-CSIM systems employ CI, a depth- G interleaver as well as JML detectors.

$m = 8$, $m = 16$ and $m = 32$, as well as the corresponding system parameters given in Table I. For the sake of comparison, only 160 subcarriers are used for the transmission of QPSK symbols in the classic OFDM system, yielding a transmission rate of 1.25 bits/s/Hz. Similarly, the OFDM-SIM system employs 212 subcarriers for carrying SIM- $(m = 4, K = 2, Q = 4)$ symbols, giving a transmission rate of 1.2422 bits/s/Hz. Here, we should point out that reducing the number of transmission subcarriers results in a certain power gain reduction, rather than in a BER curve slope-change. We can observe from these figures that the OFDM-CSIM significantly outperforms both the classic OFDM and the OFDM-SIM.

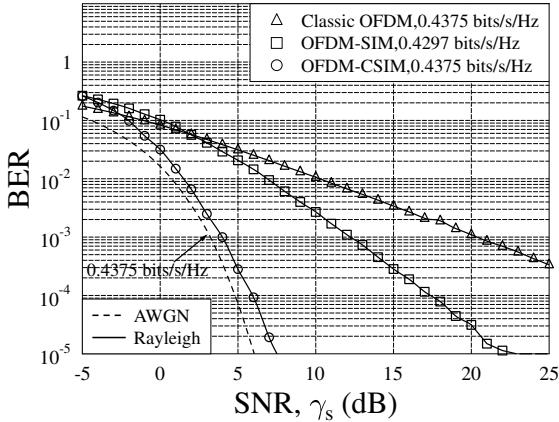


Fig. 9: BER performance of the classic OFDM, OFDM-SIM and OFDM-CSIM systems communicating over an $L_h = 10$ -path Rayleigh fading channel. Both the OFDM-SIM and the OFDM-CSIM systems employ CI, a depth- G interleaver as well as JML detectors.

Furthermore, in the high-SNR region, the performance gain of the systems associated with $m = 16$ (0.75 bits/s/Hz) in Fig. 8 are similar to those associated with $m = 32$ (0.4375 bits/s/Hz) in Fig. 9, both of which are higher than that of the systems associated with $m = 8$ in Fig. 10. For example, at a BER of 10^{-3} , about 16 dB gain is observed both in Fig. 8 and Fig. 9 for our OFDM-CSIM systems. By contrast, about 14 dB gain is observed in Fig. 7. These observations can be explained with the aid of our analytical results shown in Section IV-B. As analyzed in Section IV-B, when the measurement matrix \mathbf{A} associated with a mutual coherence satisfying (10) is used, the proposed OFDM-CSIM system becomes capable of achieving V_D -order diversity, where V_D is given in (34). Since the diversity order of the OFDM-CSIM system is much higher than that of the conventional OFDM-SIM, a better BER performance can be attained by the OFDM-CSIM system. Furthermore, as shown in (34), the diversity order found for the case of $m < L_h$ (Fig. 7) is lower than that for the case of $m \geq L_h$ (Fig. 8 and Fig. 9). Hence, a higher performance gain can be achieved by the OFDM-CSIM system associated with $m = 16, 32$ than that with $m = 8$. Meanwhile, for the case of $m \geq L_h$, the maximum multipath diversity gain is achieved by the OFDM-CSIM system. Thus, the performance gains achieved by the OFDM-CSIM system having $m = 16$ and $m = 32$ are the same. Finally, as inferred from (34) and the above observations, we can show that for a target throughput higher than 2 bits/s/Hz, our CS-assisted OFDM-SIM system is capable of attaining a similar performance gain to that of the classic OFDM systems (as seen in (34), the maximum attainable diversity order is independent of the throughput).

In Fig. 10, Fig. 11 and Fig. 12, we investigate the BER performance of both the JML detector and the IRC detector for the OFDM-CSIM systems communicating over an $L_h = 10$ -path Rayleigh fading channel. In these figures, the OFDM-CSIM

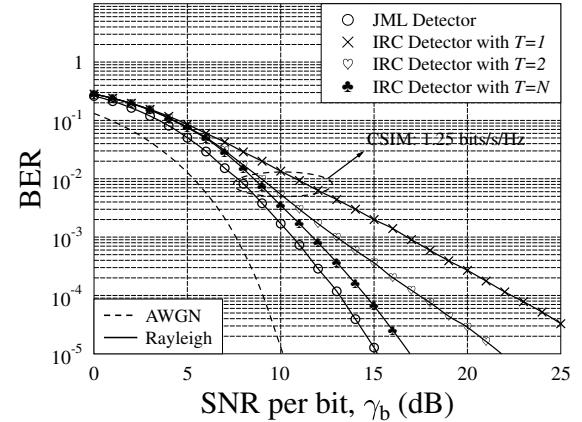


Fig. 10: BER versus SNR per bit for the OFDM-CSIM using the JML detector and the IRC detector. The OFDM-CSIM with ($N = 15, m = 8$) employs CI and a depth- G interleaver. For the IRC detector shown in Algorithm 1, the maximum number of iterations is chosen to be $T = 1, 2$ or $N = 15$.

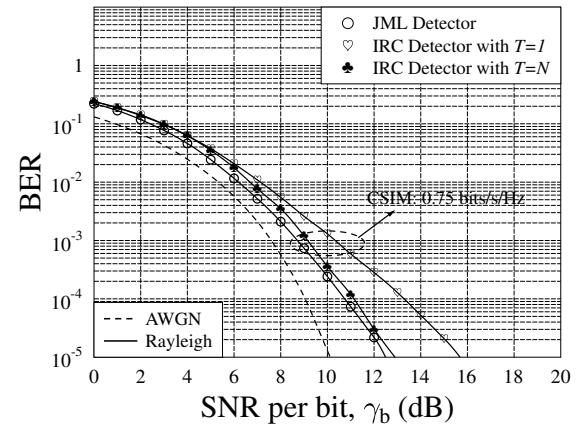


Fig. 11: BER versus SNR per bit for the OFDM-CSIM using the JML detector and the IRC detector. The OFDM-CSIM with ($N = 31, m = 16$) employs CI and a depth- G interleaver. For the IRC detector shown in Algorithm 1, the maximum number of iterations is chosen to be $T = 1$, or $N = 31$.

system employs CI and the depth- G interleaver. Explicitly, Fig. 10, Fig. 11 and Fig. 12 are respectively characterized by $m = 8, m = 16$ and $m = 32$, while the corresponding system parameters are given in Table I. We observe from the results of Fig. 10, Fig. 11 and Fig. 12 that the BER performance of the OFDM-CSIM system using our IRC detector invoking a single iteration is stable for all SNR per bit values. This observation verifies our inference obtained from the results of Fig. 5. Furthermore, as seen in Fig. 10, Fig. 11 and Fig. 12, a good BER performance is attainable for the IRC detector using just a single iteration (Fig. 11 and Fig. 12) or two iterations (Fig. 10). Bearing in mind the complexity results of Fig. 6,

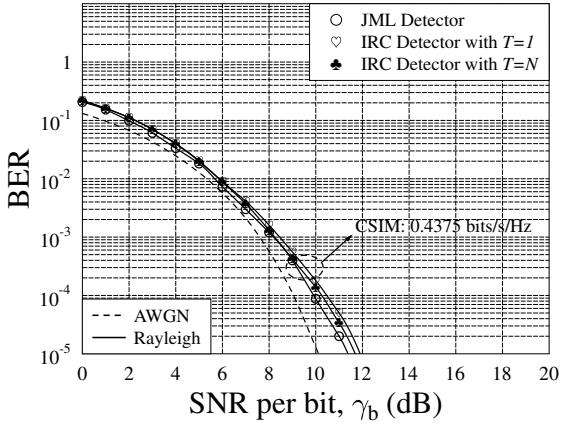


Fig. 12: BER versus SNR per bit for the OFDM-CSIM using the JML detector and the IRC detector. The OFDM-CSIM with ($N = 61$, $m = 32$) employs CI and a depth- G interleaver. For the IRC detector shown in Algorithm 1, the maximum number of iterations is chosen to be $T = 1$, or $N = 61$.

it can be readily shown that the complexity of the proposed IRC detector can be further reduced by limiting the maximum number of iterations at the cost of a few dBs performance loss. Therefore, when the IRC detector is employed by the OFDM-CSIM system, the maximum number of iterations may be appropriately adjusted for striking a flexible tradeoff between the performance gain attained and the complexity imposed.

VI. CONCLUSIONS

In this paper, a CSIM scheme has been proposed. Our analytical results have shown that in comparison to the conventional SIM, our proposed scheme is capable of achieving a higher SE, which has also been verified by our simulation results. Moreover, we have shown that the CSIM is capable of striking a more appealing tradeoff between the SE and EE than the conventional SIM. Then, the JML detector has been considered for our CSIM scheme. Based on the JML detector, the diversity order of the CSIM has been first analyzed mathematically, and then, it was verified by simulations. Both the analytical and the simulation results show that in comparison to the SIM system, our proposed system is capable of providing a significantly higher diversity gain. Finally, the reduced-complexity IRC detector was proposed for our CSIM scheme. Simulations results have been provided for investigating its detection performance. Our investigations demonstrated that in comparison to the JML detector, the complexity of our IRC detector is much lower, while its performance is close to that of the JML detector.

APPENDIX A DERIVATION OF SE

According to Section III, the channel state information (CSI) is assumed to be perfectly known at the receiver. Hence, based

on (9), the mutual information $I(\mathbf{x}_g; \mathbf{y}_g)$ can be expressed as [45]

$$\begin{aligned} I(\mathbf{x}_g; \mathbf{y}_g) &= \mathbb{E}_{\bar{\mathbf{H}}_g} [I(\mathbf{x}_g; \mathbf{y}_g | \bar{\mathbf{H}}_g)] \\ &= \mathbb{E}_{\bar{\mathbf{H}}_g} [\mathcal{H}(\mathbf{y}_g | \bar{\mathbf{H}}_g) - \mathcal{H}(\mathbf{y}_g | \mathbf{x}_g, \bar{\mathbf{H}}_g)] \\ &= \mathbb{E}_{\bar{\mathbf{H}}_g} [\mathcal{H}(\mathbf{y}_g | \bar{\mathbf{H}}_g)] - \mathcal{H}(\bar{\mathbf{w}}_g), \end{aligned} \quad (\text{A.1})$$

where the differential entropy of the m -dimensional complex Gaussian vector with a mean vector $\mathbf{0}$ and a covariance matrix of $N_0 \mathbf{I}_m$ is given by [45]

$$\begin{aligned} \mathcal{H}(\bar{\mathbf{w}}_g) &= \log_2(\pi e N_0)^m = m \log_2(\pi e N_0) \\ &= m \log_2 \left(\frac{\pi e K}{m \gamma_s} \right). \end{aligned} \quad (\text{A.2})$$

On the other hand, the conditional entropy $\mathcal{H}(\mathbf{y}_g | \bar{\mathbf{H}}_g)$ can be formulated as [45]

$$\begin{aligned} \mathcal{H}(\mathbf{y}_g | \bar{\mathbf{H}}_g) &= - \int_{\mathbf{y}_g} p(\mathbf{y}_g | \bar{\mathbf{H}}_g) \log_2 p(\mathbf{y}_g | \bar{\mathbf{H}}_g) d\mathbf{y}_g \\ &= \mathbb{E}_{\mathbf{y}_g} [-\log_2 p(\mathbf{y}_g | \bar{\mathbf{H}}_g)]. \end{aligned} \quad (\text{A.3})$$

According to Section II-A, the vector space \mathcal{V} is defined so that each group of L data bits can be regarded as being directly mapped into a symbol vector in \mathcal{V} . Since the input data bits are i.i.d, it is reasonable to assume that the candidates in \mathcal{V} are equiprobable, i.e. we have $p(\mathbf{v}_i) = 1/2^L$ for $i = 1, \dots, 2^L$. Hence, with the aid of (11) and by invoking Bayes' rule, the conditional probability $p(\mathbf{y}_g | \bar{\mathbf{H}}_g)$ can be expressed as

$$\begin{aligned} p(\mathbf{y}_g | \bar{\mathbf{H}}_g) &= \sum_{i=1}^{2^L} p(\mathbf{y}_g | \mathbf{v}_i, \bar{\mathbf{H}}_g) p(\mathbf{v}_i) = \frac{1}{2^L} \sum_{i=1}^{2^L} p(\mathbf{y}_g | \mathbf{v}_i, \bar{\mathbf{H}}_g) \\ &= \frac{1}{2^L} \sum_{i=1}^{2^L} \frac{1}{(\pi N_0)^m} \exp \left\{ - \frac{\|\mathbf{y}_g - \bar{\mathbf{H}}_g \mathbf{A} \mathbf{v}_i\|_2^2}{N_0} \right\} \\ &= \frac{1}{2^L} \sum_{i=1}^{2^L} \frac{1}{(\pi N_0)^m} \exp \left\{ - \frac{\|\bar{\mathbf{H}}_g \mathbf{A}(\mathbf{x}_g - \mathbf{v}_i) + \bar{\mathbf{w}}_g\|_2^2}{N_0} \right\}. \end{aligned} \quad (\text{A.4})$$

Additionally, let us define a set of

$$\mathcal{E} \triangleq \{\mathbf{e} = \mathbf{A}(\mathbf{v}_j - \mathbf{v}_i) : \mathbf{v}_i \neq \mathbf{v}_j, \forall \mathbf{v}_i, \mathbf{v}_j \in \mathcal{V}\}. \quad (\text{A.5})$$

Then, with the aid of (A.5), Eq. (A.4) can be formulated as

$$\begin{aligned} p(\mathbf{y}_g | \bar{\mathbf{H}}_g) &= \frac{1}{2^L} \left(\frac{1}{(\pi N_0)^m} \exp \left\{ - \frac{\|\bar{\mathbf{w}}_g\|_2^2}{N_0} \right\} \right. \\ &\quad \left. + \mathbb{E}_{\mathbf{e}} \left[\frac{1}{(\pi N_0)^m} \exp \left\{ - \frac{\|\bar{\mathbf{H}}_g \mathbf{e} + \bar{\mathbf{w}}_g\|_2^2}{N_0} \right\} \right] \right) \\ &= \frac{1}{2^L} \left(\frac{m \gamma_s}{\pi K} \right)^m \left(\exp \left\{ - \frac{\|\bar{\mathbf{w}}_g\|_2^2}{K} m \gamma_s \right\} \right. \\ &\quad \left. + \mathbb{E}_{\mathbf{e}} \left[\exp \left\{ - \frac{\|\bar{\mathbf{H}}_g \mathbf{e} + \bar{\mathbf{w}}_g\|_2^2}{K} m \gamma_s \right\} \right] \right), \end{aligned} \quad (\text{A.6})$$

which is dependent on $\bar{\mathbf{w}}_g$ given $\bar{\mathbf{H}}_g$ and γ_s . Hence, for the sake of convenience, we define the implementation function of

$$\epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s) \triangleq p(\mathbf{y}_g | \bar{\mathbf{H}}_g) \quad (\text{A.7})$$

to explicitly emphasize the dependence on the channel parameters. Then, Eq.(A.3) can be alternatively expressed as

$$\mathcal{H}(\mathbf{y}_g | \bar{\mathbf{H}}_g) = \mathbb{E}_{\bar{\mathbf{w}}_g} [-\log_2 \epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s)]. \quad (\text{A.8})$$

Finally, upon substituting (A.2) and (A.8) into (A.1), we arrive at

$$I(\mathbf{x}_g; \mathbf{y}_g) = \mathbb{E}_{\bar{\mathbf{H}}_g, \bar{\mathbf{w}}_g} [-\log_2 \epsilon(\bar{\mathbf{w}}_g, \bar{\mathbf{H}}_g, \gamma_s)] - m \log_2 \left(\frac{\pi e K}{m \gamma_s} \right). \quad (\text{A.9})$$

It should be noted that the upper bound of $I(\mathbf{x}_g; \mathbf{y}_g)$ in our scenario can be derived as

$$\begin{aligned} I(\mathbf{x}_g; \mathbf{y}_g) &= \mathbb{E}_{\bar{\mathbf{H}}_g} [I(\mathbf{x}_g; \mathbf{y}_g | \bar{\mathbf{H}}_g)] \\ &= \mathbb{E}_{\bar{\mathbf{H}}_g} [\mathcal{H}(\mathbf{x}_g | \bar{\mathbf{H}}_g) - \mathcal{H}(\mathbf{x}_g | \mathbf{y}_g, \bar{\mathbf{H}}_g)] \\ &= \mathcal{H}(\mathbf{x}_g) - \mathbb{E}_{\bar{\mathbf{H}}_g} [\mathcal{H}(\mathbf{x}_g | \mathbf{y}_g, \bar{\mathbf{H}}_g)] \\ &\leq \mathcal{H}(\mathbf{x}_g) = - \sum_{i=1}^{2^L} p(\mathbf{v}_i) \log_2 p(\mathbf{v}_i) \\ &= - \sum_{i=1}^{2^L} \frac{1}{2^L} \log_2 \frac{1}{2^L} = L, \end{aligned} \quad (\text{A.10})$$

where the upper bound is achieved, when γ_s is sufficiently high.

REFERENCES

- [1] J. Akhtman and L. Hanzo, "Power versus bandwidth-efficiency in wireless communications: The economic perspective," in *2009 IEEE 70th Vehicular Technology Conference Fall (VTC 2009-Fall)*, Sept 2009, pp. 1–5.
- [2] M. Di Renzo, H. Haas, A. Ghayeb, S. Sugiura and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proceedings of the IEEE*, vol. 102, no. 1, pp. 56–103, Jan 2014.
- [3] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, June 2011.
- [4] G. Y. Li *et al.*, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28–35, December 2011.
- [5] C.-L. I *et al.*, "Toward green and soft: a 5G perspective," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 66–73, February 2014.
- [6] Y. Wu *et al.*, "Green transmission technologies for balancing the energy efficiency and spectrum efficiency trade-off," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 112–120, Nov 2014.
- [7] L. Hanzo, M. Münster, B. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-user Communications, WLANs and Broadcasting*. John Wiley & Sons, 2003.
- [8] Lie-Liang Yang, *Multicarrier Communications*. John Wiley & Sons, Ltd, 2009.
- [9] L. Hanzo, M. El-Hajjar and O. Alamri, "Near-capacity wireless transceivers and cooperative communications in the MIMO era: Evolution of standards, waveform design, and future perspectives," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1343–1385, Aug 2011.
- [10] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2009, pp. 177–181.
- [11] D. Tsonev, S. Sinanovic and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, Dec 2011, pp. 728–732.
- [12] E. Basar, U. Aygolu, E. Panayirci and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5536–5549, Nov 2013.
- [13] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM - will it work?" *IEEE Access*, vol. 4, pp. 2580–2593, 2016.
- [14] M. Wen, X. Cheng and L. Yang, "Optimizing the energy efficiency of OFDM with index modulation," in *2014 IEEE International Conference on Communication Systems (ICCS)*, Nov 2014, pp. 31–35.
- [15] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with index modulation," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 1919–1932, April 2016.
- [16] H. Zhang, L. L. Yang, and L. Hanzo, "LDPC-coded index-modulation aided OFDM for in-vehicle power line communications," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–5.
- [17] V. Tarokh, N. Seshadri and A.R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, Mar 1998.
- [18] Y. Xiao *et al.*, "OFDM with interleaved subcarrier-index modulation," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1447–1450, Aug 2014.
- [19] E. Basar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Communications Letters*, vol. PP, no. 99, pp. 1–1, 2015.
- [20] A. Khan, M.Z. and B. S. Rajan, "Single-symbol maximum likelihood decodable linear STBCs," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2062–2091, May 2006.
- [21] R. Fan, Y. Yu and Y. Guan, "Orthogonal frequency division multiplexing with generalized index modulation," in *2014 IEEE Global Communications Conference (GLOBECOM)*, Dec 2014, pp. 3880–3885.
- [22] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [23] W. U. Bajwa, J. Haupt, A.M. Sayeed and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.
- [24] A. Gomaa and N. Al-Dhahir, "A sparsity-aware approach for NBI estimation in MIMO-OFDM," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1854–1862, June 2011.
- [25] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007.*, vol. 4, April 2007, pp. IV-1357–IV-1360.
- [26] H. Zhang, L. L. Yang, and L. Hanzo, "Compressed impairment sensing-assisted and interleaved-double-FFT-aided modulation improves broadband power line communications subjected to asynchronous impulsive noise," *IEEE Access*, vol. 4, pp. 81–96, 2016.
- [27] S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comp.*, vol. 20, no. 1, pp. 33–61, 1999.
- [28] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [29] E. Candes and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n ," *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, 2007.
- [30] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, Oct 2004.
- [31] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.
- [32] T. Blumensath and M. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.
- [33] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.
- [34] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1453–1467, Jul 1998.
- [35] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, Nov 2001.
- [36] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec 2005.
- [37] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1248–1259, Feb 2014.
- [38] L. Welch, "Lower bounds on the maximum cross correlation of signals (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 397–399, May 1974.

- [39] D. L. Donoho, M. Elad and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, Jan 2006.
- [40] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, July 2011.
- [41] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, pp. 227–234, 1995.
- [42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [43] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.
- [44] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, Jun 2002.
- [45] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience, 1991.
- [46] J. Joung, C. K. Ho, and S. Sun, "Spectral efficiency and energy efficiency of OFDM systems: Impact of power amplifiers and countermeasures," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 208–220, February 2014.
- [47] Lie-Liang Yang, "Signal detection in antenna-hopping space-division multiple-access systems with space-shift keying modulation," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 351–366, Jan 2012.
- [48] Z. Liu, Y. Xin and G. B. Giannakis, "Linear constellation precoding for OFDM with maximum multipath diversity and coding gains," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 416–427, March 2003.