Cooperative randomized MIMO-OFDM downlink for multicell networks: design and analysis

Francesco Verde, Donatella Darsena, *Member, IEEE,* and Anna Scaglione, *Senior Member, IEEE*

***Abstract*—This paper proposes a low-complexity physical (PHY) layer design to introduce cooperation in the downlink of an infrastructure-based multicell multiple-input multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) network, aimed at supporting future high-throughput broadband wireless Internet access with large-scale coverage. In such a system, several multi-antenna base stations (BSs) are organized in a cellular architecture to serve multiantenna mobile stations (MSs) and are connected to a central service unit via a high-speed wired backbone. To improve the network performance, a novel PHY layer design is proposed that allows cooperation among an arbitrary and unknown number of BSs, by suitably randomizing the MIMO-OFDM block codes used by the BSs. Such a randomized MIMO-OFDM code renders the encoding/decoding rule independent of the number of actual BSs cooperating and works without any channel feedback, which greatly simplifies the protocol as well as the MS design. To provide performance insights and develop PHY layer designs, the paper provides analytical upper bounds on the symbol error probability for linear receivers, which allow to accurately evaluate the diversity order and the coding gain achievable through the proposed scheme. Finally, we present numerical results that validate the theory and highlight the performance gain and the coverage expansion attainable with our cooperative transceiver.**

***Index Terms*—Cooperative downlink, infrastructured multicell networks, minimum mean-square error (MMSE) linear decoding, multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) technology, physical (PHY) layer design, randomized space-time block coding (STBC), zero-forcing (ZF) linear decoding.**

I. INTRODUCTION

NEXT-GENERATION wireless communication systems are expected to deliver a wide variety of very high data rate multimedia services (such as voice over Internet protocol, video conferencing, streaming data, multiplayer interactive gaming, web browsing, instant messaging, and media content downloading) and, at the same time, provide wide-rage coverage to mobile users. The IEEE 802.16e-2005 standard [1], also referred to as mobile worldwide interoperability for microwave access (WiMAX), is one step towards the specification of a cost-effective communication system capable of offering fixed and mobile broadband wireless access over both local and metropolitan area networks. A key physical (PHY) layer technology of mobile WiMAX is multiple-input multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM). MIMO provides multiplexing and diversity gains (see, e.g., [2]–[4]) and, at the same time, the OFDM air interface mitigates effectively the intersymbol interference (ISI) caused by multipath fading and asynchronism (see, e.g., [5]). While in principle mobile WiMAX can support alternative architectures based on multihop ad hoc networking in multipoint-to-multipoint operation mode, the conventional cellular deployment remains an integral part of IEEE 802.16e medium access control (MAC) sublayer in point-to-multipoint (PMP) operation mode. WiMAX PMP mode provides wide-area coverage by organizing nodes into a cellular infrastructure, where mobile stations (MSs) within a given cell are wirelessly served by a single base station (BS); all (or a subset of) the BSs are connected by a high-speed wired backbone that allows information to be reliably exchanged among them. As a relevant example, WiMAX PMP mode can provide last-mile unicast or multicast access to a broadband Internet service provider (ISP), which will clearly be the major demand in WiMAX networks. In such applications, the uplink traffic is much less than the downlink traffic. Thus, in this paper, the attention is focused on enhancing the throughput of *downlink* communications for cell-dependent infrastructured networks.

Traditionally, downlink transmissions in cellular systems are carried out through single-cell processing, where a MS only communicates with its own BS, while a BS serves multiple MSs. Recently, motivated by the requirement for more spectrally efficient wireless systems and, at the same time, pressed by the desire to keep MSs simple and cost-effective by shifting processing complexity to the BSs, there has been increasing interest in enhancing the performance of multicell networks by exploiting *cooperative processing at the BSs* [6]–[14]. In BS cooperative approaches, a MS or, as in multicast applications (e.g., group communications) a group of MSs, can be served by a set of different BSs which mutually exchange information via a central service unit (e.g., ISP). In practice, the set of BSs assigned to a given MS is dynamically updated on the basis of the channel quality, i.e.,

Manuscript received December 1, 2008; revised July 25, 2009. The work of F. Verde has been partially supported by the Italian National Project: Wireless multiplatfOrm mimo active access netwoRks for QoS-demanding muLtimedia Delivery (WORLD), under grant number 2007R989S. The work of A. Scaglione has been supported by the NSF grant NeTS-ProWin no. 0626751. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Wing-Kin Ma.

F. Verde is with the Department of Biomedical, Electronic and Telecommunication Engineering, University Federico II, Naples I-80125, Italy (e-mail: f.verde@unina.it). D. Darsena is with the Department for Technologies, Parthenope University, Naples I-80143, Italy (e-mail: darsena@uniparthenope.it). A. Scaglione is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616-5294, USA (e-mail: ascaglione@ucdavis.edu).

Publisher Item Identifier S 0000-0000(00)00000-3.

the signal-to-noise ratio (SNR), between the served MS and each potentially available BS. The benefits of node cooperation in wireless networks have theoretical foundation [15]–[19]. In the specific context of a multicell downlink, BS cooperation provides significant capacity and coverage improvements, as well as efficient solutions for handover management.

A. Related work

BS cooperation can be realized in more than one way. The cooperative multi-cell downlink sum capacity is identical to that of a MIMO broadcast channel and can be achieved by resorting to dirty paper (DP) precoding [6]–[8]. However, the implementation of DP precoders, which are intrinsically nonlinear, might entail in practice prohibitive complexity. For this reason, suboptimal linear precoding or beamforming designs have been proposed and analyzed in, e.g., [9]–[14]. These approaches make use of channel state information (CSI) for the link towards the served MS, which is obtained either by reciprocity in the uplink and downlink channels or by using a feedback uplink channel. High mobility and the need of providing multicast services, which are included in the WiMAX standard, may reduce the appeal of these solutions, while precoding schemes that do not require CSI at each BS but extract transmit diversity and provide spatial multiplexing are more desirable.

Unlike a system with a set of co-located array elements, a cooperative downlink has a set of cooperating BSs that dynamically varies while the served MS moves, i.e., when the SNR over the link between the MS and each BSs changes. Conventional space-time block codes (STBC) proposed for cooperative schemes (see, e.g., [16], [18], [20]), do not account for the time-varying nature of the set of cooperating BSs and the receiver needs to be informed of the specific STBC used at any particular time. Furthermore, the complexity of the channel acquisition and receiver increases as the number of cooperative nodes grows. Clearly, cooperative multicell downlink transmissions can benefit from coding rules that are agnostic of the actual PHY layer resources involved in transmission, and that require a fixed receiver complexity. A good analogy in this context is that of fading, which is caused by a random recombination of signal paths. The question is whether the cooperative BSs can emulate the same effect. To simplify the recruitment of cooperative nodes, some authors proposed to use delay diversity, generating a virtual multipath channel by naturally staggering the transmissions of the cooperative nodes [21], [22], while others tried to optimize the resulting gains [23]. Randomized STBC (RSTBC) was proposed in [24] much in the same spirit: this technique allows for a decentralized implementation (i.e., neither internode communication nor central control unit is required) of distributed multiple-input single-output systems. In [25] it was shown that delay diversity can be studied in exactly the same framework as in [24]. Furthermore, [24], [25] indicated that a soft-randomization rule, that randomly combines codes, rather than the random selection of codes, attains better diversity and coding gains overall. In fact, such a strategy allows achieving full diversity, equal to the minimum between the number of nodes cooperating and the code diversity order, irrespective of the number of relays; moreover, it approaches the performance of a centralized scheme [16], [18] both in diversity and coding gain, under the maximum likelihood (ML) detection rule. However, the analyses and the design guidelines provided in [24] and [25] cannot apply directly to the infrastructure-based multicell downlink framework at hand for the following three main reasons: (i) they focus on the benefits of RSTBC in cell-independent decentralized systems; (ii) they refer to ML receiving structure, whose computational complexity might be highly unmanageable for broadband WiMAX applications; (iii) they are targeted at single-antenna nodes.

B. Our contribution

This paper proposes and analyzes a low complexity BS cooperation strategy for broadband wireless Internet access, with linear processing both at the transmit and receiver sides. The target system has an infrastructure multicell architecture, as in a WiMAX PMP network. Since cooperative transmissions are especially effective when the BSs and the MSs are equipped with multiple antennae, according to the IEEE 802.16 standard, the PHY layer uses MIMO-OFDM with linear precoding at the transmitter and, due to its simplicity, we envision either zero-forcing (ZF) or minimum mean-square error (MMSE) linear equalization at the receiver. Our idea to simplify the design of the MSs is to have each BS transmits an *independent random linear transformation of a single MIMO-OFDM symbol*, which will be referred to as *randomized MIMO-OFDM coding (R-MIMO-OFDM)*. The key contributions of the paper are as follows: (i) the novel R-MIMO-OFDM cooperative coding, a design that does not need CSI at each BS and, more importantly, does not require any knowledge at the MS regarding the set of the cooperating BSs; (ii) a performance analysis of the proposed R-MIMO-OFDM coding rule when linear ZF and MMSE equalizers are used at the MS, which provides an upper bound for the achievable performance with BS cooperation and, as a by-product, allows one to recognize and discuss some interesting tradeoffs between the main parameters of the network, as well as to assess the coverage improvement with respect to a noncooperative system; (iii) numerical simulation results provided to further demonstrate the advantages of the proposed PHY design, by showing in particular how network performance is influenced by different choices of the statistics of the randomized coding.

C. Outline of the paper

In Section II, we give a brief description of the multicell MIMO-OFDM system architecture with BS cooperation. Section III introduces the novel R-MIMO-OFDM cooperative PHY design for multicell cooperative downlinks, and describes the system model. In Section IV, we analyze the symbol error probability (SEP) of the proposed R-MIMO-OFDM cooperative downlink to evaluate the corresponding attainable diversity and coding gains, whose insightfulness is subsequently come up for discussion. Numerical results are presented in Section V in terms of average bit-error-rate (ABER), followed by concluding remarks in Section VI.

D. Notations

Upper- and lower-case bold letters denote matrices and vectors; the superscripts $*$, T, H, -1 and $\dagger$ denote the conjugate, the transpose, the Hermitian (conjugate transpose), the inverse and the Moore-Penrose generalized inverse of a matrix; $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Z}$ are the fields of complex, real and integer numbers; $\mathbb{C}^n$ $[\mathbb{R}^n]$ denotes the vector-space of all n-column vectors with complex [real] coordinates; similarly, $\mathbb{C}^{n\times m}$ $[\mathbb{R}^{n\times m}]$ denotes the vector-space of all the $n \times m$ matrices with complex [real] elements; the operators $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ stand for real and imaginary parts of any complex-valued quantity, whereas $j \triangleq \sqrt{-1}$ denotes the imaginary unit; the continuous-time function $\delta(t)$ denotes the Dirac impulse; $\{\mathbf{A}\}_{i,k}$ and $\{\mathbf{a}\}_i$ indicate the $(i+1,k+1)$th element of matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ and the $(i+1)$th entry of vector $\mathbf{a} \in \mathbb{C}^m$, with $i \in \{0,1,\ldots,m-1\}$ and $k \in \{0,1,\ldots,n-1\}$; $\mathbf{0}_n$, $\mathbf{O}_{n\times m}$ and $\mathbf{I}_n$ denote the n-column zero vector, the $n \times m$ zero matrix and the $n \times n$ identity matrix; for any $\mathbf{A} \in \mathbb{C}^{n\times m}$, $\mathrm{rank}(\mathbf{A})$, $\mathrm{trace}(\mathbf{A})$ and $\det(\mathbf{A})$ denote the rank, the trace and the determinant (when $n = m$) of $\mathbf{A}$; for any $\mathbf{a} \in \mathbb{C}^n$, $\|\mathbf{a}\|$ denotes the Euclidean norm; the symbol $*$ stands for (linear) convolution and $\otimes$ denotes Kronecker product of matrices; matrix $\mathbf{A} = \mathrm{diag}(a_0, a_1, \ldots, a_{p-1}) \in \mathbb{C}^{p\times p}$ is diagonal; $\mathrm{E}[\cdot]$ denotes ensemble averaging and, finally, a circularly symmetric complex Gaussian random vector $\mathbf{x} \in \mathbb{C}^n$ with mean $\boldsymbol{\mu} \in \mathbb{C}^n$ and covariance matrix $\mathbf{C} \in \mathbb{C}^{n\times n}$ is denoted as $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$, whereas a Gamma random variable X with shape parameter α and scale parameter β is denoted as $X \sim \mathrm{gamma}(\alpha, \beta)$.

II. The multicell cooperative MIMO-OFDM downlink

The system considered in this work is shown in Fig. 1. There are TX_c cooperative BSs (Fig. 1 has $TX_\mathrm{c} = 3$) that collaborate to serve one MS, which will be also referred to as the *destination* D. The set $\mathcal{S}$ of adjacent BSs serving D is dynamically updated depending on the SNR over the link between the MS and each surrounding distributed BS, which requires the feedback of the SNR associated with the served user. However, this only needs a minimum amount of signaling in the uplink direction to provide timely SNR estimates. We assume that $\mathcal{S}$ has already been defined and, moreover, both $\mathcal{S}$ and its cardinality TX_c are deterministic but *unknown* at the MS.[1] All the participants are equipped with multiple antennae: specifically, each BS has TX_a transmit antennae and the served MS is equipped with RX receive antennae. Moreover, each BS employs a MIMO-OFDM modulation with (fixed) bandwidth $W \approx 1/T$, number of subcarriers M and cyclic prefix (CP) of length M_cp. The symbol rate is adjusted by selecting the dimension Q of a quadrature amplitude modulation (QAM) square constellation, which is fixed for all the subcarriers. In such a framework, the *actual* number TX of transmit antennae includes the number TX_a of local antennae at each BS times the number TX_c of cooperative BSs, that is, $TX = TX_\mathrm{a} \cdot TX_\mathrm{c}$. Consequently, the *multiplexing rate*, which is defined as the number N of symbols transmitted per subcarrier, is upper bounded as $N \leq \min(TX, RX)$. The smaller N the greater is the attainable energy per symbol (so-called diversity-multiplexing tradeoff [26]).

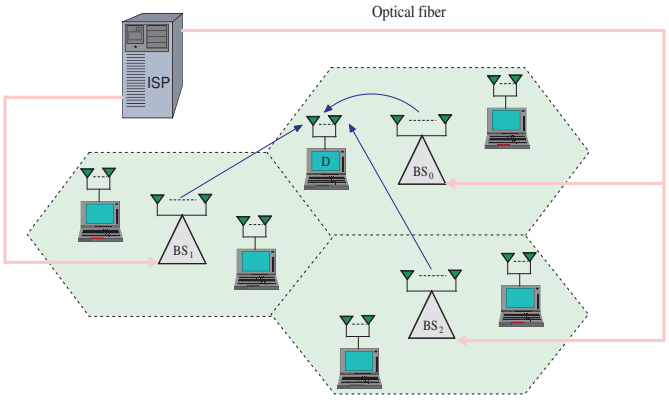

Fig. 1. A multicell cooperative MIMO downlink: the adjacent cooperative base stations $\mathcal{S} = \{BS_0, BS_1, BS_2\}$, which are equipped with TX_a transmit antennae, cooperatively serve the mobile station D, which has RX receive antennae.

As can be seen from Fig. 1, there is also a central service unit, e.g., a broadband ISP in last-mile Internet access, to which are connected all the BSs via very high bandwidth wired links, e.g., optical fiber. All the BSs belonging to $\mathcal{S}$ receive reliably information from the central service unit by means of a high-speed wired backbone: we assume that they share the same sequence $s[n]$ $(n \in \mathbb{Z})$ of Q-QAM symbols and they are perfectly synchronized by the central unit. The channel between the BSs and the served MS is assumed to be frequency-selective and quasi-static, which means that the channel impulse responses remain fixed during the transmission of an OFDM symbol but are allowed to change from one symbol to another. More precisely, let $T_s \triangleq PT$ denote the overall duration of the OFDM block, where $P \triangleq M + M_\mathrm{cp}$, we assume that, over the kth symbol interval $[kT_s, (k+1)T_s)$, the baseband equivalent time-varying channel impulse response $c_{r,i,\alpha}(t,\tau)$ between the local transmit antenna $\alpha \in \{0,1,\ldots,TX_\mathrm{a}-1\}$ of the BS $i \in \{0,1,\ldots,TX_\mathrm{c}-1\}$ and the local receive antenna $r \in \{0,1,\ldots,RX-1\}$ of D remains (approximatively) constant, i.e., $c_{r,i,\alpha}(t,\tau) \approx c_{r,i,\alpha}(\tau)$ for $t \in [kT_s, (k+1)T_s)$ and $\tau \in \mathbb{R}$, but might change with $k \in \mathbb{Z}$. Furthermore, we assume that CSI is known only at the MS via training but is unknown at each BS.

III. The proposed PHY layer with R-MIMO-OFDM coding

In this section, we present a proposal addressing some of the challenges associated with the critical design issues of BS cooperation for next-generation cellular wireless PHY layer, by providing a detailed description of both the transmitter and receiver design (see Fig. 2 as a reference).

[1] A different approach consists of modeling $\mathcal{S}$ as a random set and TX_c as a random variable, with given probability density functions. However, such densities depend on the grouping strategy for the collaborating BSs, which involves higher layers of the protocol stack, and, moreover, are highly scenario dependent, since they depend on several physical parameters, such as location, speed and mobility model of the served user, geometrical form and radius of the cells, and so forth.

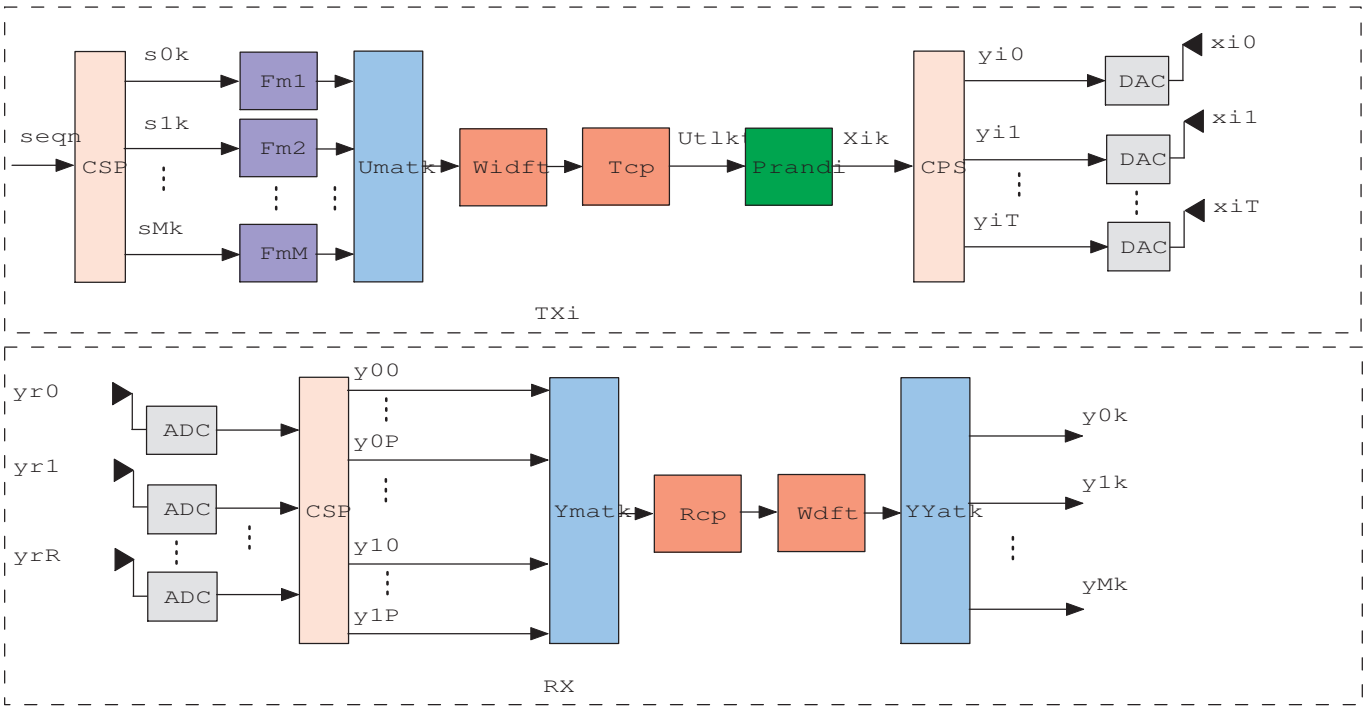

Fig. 2. Block diagram of the proposed baseband R-MIMO-OFDM transceiver.

A. Transmitter design

For each BS, the symbol sequence $s[n]$ is subject to serial-to-parallel (S/P) conversion and, subsequently, is parsed in the matrix $\mathbf{S}[k] = (\mathbf{s}_0[k], \mathbf{s}_1[k], \ldots, \mathbf{s}_{M-1}[k]) \in \mathbb{C}^{N\times M}$, whose mth column $\mathbf{s}_m[k]$ is the vector of N symbols transmitted on the mth subcarrier during the kth OFDM symbol, with $m \in \{0, 1, \ldots, M-1\}$ and $k \in \mathbb{Z}$. Since the actual number TX of transmit antennae is unknown, each BS uses as encoding parameter the number L of *sought* antennae, which is fixed irrespective of TX. While linear designs are by no mean the only alternative, they are the simple coding rule we are going to consider to multiplex symbols in space and adapt to rate requirements. More specifically, each cooperating BS precodes spatially the symbol sequence in the same way by means of the linear transformation $\mathbf{F}_m\,\mathbf{s}_m[k]$, where $\mathbf{F}_m \in \mathbb{C}^{L\times N}$ is a semi-unitary precoder, i.e., $L \geq N$ and $\mathbf{F}_m^H\mathbf{F}_m = \frac{1}{N}\,\mathbf{I}_N$.[2] Equivalently, the matrix $\mathbf{S}[k]$ is mapped into the matrix $\mathbf{U}[k] = (\mathbf{F}_0\,\mathbf{s}_0[k], \mathbf{F}_1\,\mathbf{s}_1[k], \ldots, \mathbf{F}_{M-1}\,\mathbf{s}_{M-1}[k]) \in \mathbb{C}^{L\times M}$. Subsequently, each BS performs the OFDM precoding in the temporal dimension. Namely, each row of the matrix $\mathbf{U}[k]$ is temporally precoded with an OFDM filterbank, and the obtained data matrix is $\widetilde{\mathbf{U}}[k] = \mathbf{U}[k]\,\boldsymbol{\mathcal{W}}_{\text{IDFT}}\,\mathbf{T}_{\text{cp}} \in \mathbb{C}^{L\times P}$, where $\boldsymbol{\mathcal{W}}_{\text{IDFT}} \in \mathbb{C}^{M\times M}$ is the *unitary* symmetric IDFT matrix and $\mathbf{T}_{\text{cp}} \triangleq [\mathbf{I}_{\text{cp}}^T, \mathbf{I}_M] \in \mathbb{R}^{M\times P}$, with $\mathbf{I}_{\text{cp}} \in \mathbb{R}^{M_{\text{cp}}\times M}$ collecting the last M_{cp} rows of $\mathbf{I}_M$. Each row of $\widetilde{\mathbf{U}}[k]$ undergoes parallel-to-serial (P/S) conversion and the resulting sequence $\widetilde{u}_\ell[n]$ is defined as follows: $\widetilde{u}_\ell[kP+p] = \{\widetilde{\mathbf{U}}[k]\}_{\ell,p} \triangleq \widetilde{u}_{\ell,p}[k]$, for $\ell \in \{0, 1, \ldots, L-1\}$ and $p \in \{0, 1, \ldots, P-1\}$.

[2]The precoding matrices $\{\mathbf{F}_m\}_{m=0}^{M-1}$ are assumed to be nonrandom and known at the MS. Indeed, since the relevant BSs have no CSI, the precoder is fixed and used to adjust the multiplexing rate.

Since each BS has only TX_{a} antennae at its disposal, the question that naturally arises is how the L streams $\{\widetilde{u}_\ell[n]\}_{\ell=0}^{L-1}$ are transmitted. The idea of randomized cooperation is very simple: the ith BS, with $i \in \{0, 1, \ldots, TX_{\text{c}}-1\}$, transmits via each local antenna α, for $\alpha \in \{0, 1, \ldots, TX_{\text{a}}-1\}$, a *random* linear combination of these streams with its own unique set of random coefficients $p_{i,\alpha,l}$, thereby obtaining the TX_{a} different code sequences

$$x_{i,\alpha}[n] = \sum_{\ell=0}^{L-1} p_{i,\alpha,\ell}\,\widetilde{u}_\ell[n]\,. \tag{1}$$

By introducing $\mathbf{P}_i \in \mathbb{C}^{TX_{\text{a}}\times L}$ and $\mathbf{X}_i[k] \in \mathbb{C}^{TX_{\text{a}}\times P}$ such that $\{\mathbf{P}_i\}_{\alpha,\ell} = p_{i,\alpha,l}$ and $x_{i,\alpha,p}[k] \triangleq \{\mathbf{X}_i[k]\}_{\alpha,p} = x_{i,\alpha}[kP+p]$, with $\alpha \in \{0, 1, \ldots, TX_{\text{a}}-1\}$, $\ell \in \{0, 1, \ldots, L-1\}$ and, finally, $p \in \{0, 1, \ldots, P-1\}$, the randomization coding performed by BS i can be expressed in matrix form as $\mathbf{X}_i[k] = \mathbf{P}_i\,\widetilde{\mathbf{U}}[k]$. We have the five-steps mapping summarized in (2) at the top of the next page (see also Fig. 2). The sequence $x_{i,\alpha}[n]$ feeds a digital-to-analog converter (DAC) at rate $1/T = P/T_s$, and, after up-conversion, the ith BS transmits from its local antenna α the signal ($t \in \mathbb{R}$):

$$z_{i,\alpha}(t) = \text{Re}\{x_{i,\alpha}(t)\,e^{j(\omega_0+\varepsilon_i)t}\}\,, \tag{3}$$

with

$$x_{i,\alpha}(t) = \sum_{n=-\infty}^{+\infty}\sum_{p=0}^{P-1} x_{i,\alpha,p}[n]\,g(t-pT-nT_s), \tag{4}$$

$$s[n] \overset{\text{S/P conversion}}{\longmapsto} \mathbf{S}[k] = (\mathbf{s}_0[k], \ldots, \mathbf{s}_{M-1}[k]) \overset{\text{spatial precoding}}{\longmapsto} \mathbf{U}[k] = (\mathbf{F}_0\, \mathbf{s}_0[k], \ldots, \mathbf{F}_{M-1}\, \mathbf{s}_{M-1}[k])$$
$$\overset{\text{temporal (OFDM) precoding}}{\longmapsto} \widetilde{\mathbf{U}}[k] = \mathbf{U}[k]\, \boldsymbol{\mathcal{W}}_{\text{IDFT}}\, \mathbf{T}_{\text{cp}} \overset{\text{randomized coding}}{\longmapsto} \mathbf{X}_i[k] = \mathbf{P}_i\, \widetilde{\mathbf{U}}[k] \overset{\text{parallel-to-serial (P/S) conversion}}{\longmapsto} x_{i,\alpha}[n]\,. \quad (2)$$

where ω_0 is the carrier frequency, ε_i is the carrier offset,[3] and $g(\tau)$ is a raised cosine pulse ($\tau \in \mathbb{R}$).[4]

B. Receiver design

Let τ_i account for the transmission and propagation delays of BS $i \in \{0, 1, \ldots, TX_{\text{c}} - 1\}$, over the kth symbol interval $[k\,T_s, (k+1)\,T_s)$, the complex envelope $\widetilde{y}_r(t)$ of the received signal at the rth local receive antenna of D, for $r \in \{0, 1, \ldots, RX - 1\}$, is the mixture of all the contributions received from the BSs, each convolved with the respective baseband equivalent *composite* channel impulse response $c_{r,i,\alpha}(t, \tau) \approx c_{r,i,\alpha}(\tau)$ (encompassing the cascade of the interpolation filter, the physical channel between the local transmit antenna α of the ith BS and the local receive antenna r of D and the anti-aliasing filter), modulated by the residual carrier offset, plus the filtered noise, i.e.,

$$\widetilde{y}_r(t) = \sum_{i=0}^{TX_{\text{c}}-1} \sum_{\alpha=0}^{TX_{\text{a}}-1} e^{j\,\varepsilon_i\,t}\, [c_{r,i,\alpha}(t) * x_{i,\alpha}(t - \tau_i)] + \widetilde{v}_r(t)\,, \quad (5)$$

for $t \in [k\,T_s, (k+1)\,T_s)$. From (1)–(5), after sampling at the Nyquist rate $1/T$, we have the discrete-time model at the output of the analog-to-digital converter (ADC):

$$\widetilde{y}_{r,q}[k] \triangleq \widetilde{y}_r(k\,T_s + q\,T) = \sum_{\ell=0}^{L-1} \sum_{n=-\infty}^{+\infty} \sum_{p=0}^{P-1} \widetilde{u}_{\ell,p}[n] \cdot \widetilde{c}_{r,\ell}[k, q, (k-n)P + (q-p)] + \widetilde{v}_{r,q}[k]\,, \quad (6)$$

with $k \in \mathbb{Z}$ and $q \in \{0, 1, \ldots, P-1\}$, where

$$\widetilde{c}_{r,\ell}[k, q, h] \triangleq \sum_{i=0}^{TX_{\text{c}}-1} \sum_{\alpha=0}^{TX_{\text{a}}-1} e^{j\,\varepsilon_i\,(kP+q)\,T}\, p_{i,\alpha,\ell}\, c_{r,i,\alpha}(h\,T - \tau_i) \quad (7)$$

represents the channel impulse response between the *fictitious* transmit antenna ℓ of the distributed array and the local receive antenna r of D, and $\widetilde{v}_{r,q}[k] \triangleq \widetilde{v}_r(k\,T_s + q\,T)$. Eqs. (6) and (7) deserve two remarks. First, even if the channels are assumed to be originally time-invariant over each symbol period, the residual carrier offset ε_i causes time variations in the kth OFDM block, that is, the impulse response in (7) also depends on both the variables k and q; the Doppler spread of the effective channel is in the order of the frequency offset among the local oscillators of the cooperative BSs. Second, even if the channels were to be flat over the kth symbol period, i.e., $c_{r,i,\alpha}(\tau) = c_{r,i,\alpha}\,\delta(\tau)$, the time offsets τ_i cause dispersive effects and, thus, interblock interference (IBI).

In the following we assume that:

- **(a1)** the combined effect of the channel and relative delays of the BSs is such that the support in h of $\widetilde{c}_{r,\ell}[k, q, h]$ does not exceed the CP length M_{cp}, i.e., $\widetilde{c}_{r,\ell}[k, q, h] \equiv 0$ for any $h \notin \{0, 1, \ldots, M_{\text{cp}}\}$;
- **(a2)** the duration $T_s = P\,T$ of the OFDM block is such that $|\varepsilon_i|\,T_s \ll 1$.

These assumptions imply that over one OFDM symbol the channel can be considered approximately a linear time invariant finite impulse response (FIR) filter of order less than or equal to M_{cp}, i.e.,

$$\widetilde{c}_{r,\ell}[k, q, h] \approx \widetilde{c}_{r,\ell}[k, h] \triangleq \sum_{i=0}^{TX_{\text{c}}-1} \sum_{\alpha=0}^{TX_{\text{a}}-1} p_{i,\alpha,\ell}\, c_{r,i,\alpha}[k, h]\,, \quad (8)$$

where $c_{r,i,\alpha}[k, h] \triangleq e^{j\,\varepsilon_i k\,T_s}\, c_{r,i,\alpha}[h]$, with $c_{r,i,\alpha}[h] \triangleq c_{r,i,\alpha}(h\,T - \tau_i)$. The processing of the received signal occurs in blocks (see Fig. 2): first, within the kth OFDM symbol, the samples (6) are collected in the matrix $\widetilde{\mathbf{Y}}[k]$ such that $\{\widetilde{\mathbf{Y}}[k]\}_{r,q} = \widetilde{y}_{r,q}[k]$, for $r \in \{0, 1, \ldots, RX - 1\}$ and $q \in \{0, 1, \ldots, P-1\}$; second, to suppress IBI, the first M_{cp} columns are removed from $\widetilde{\mathbf{Y}}[k]$ by means of the CP removal matrix $\mathbf{R}_{\text{cp}} \triangleq (\mathbf{O}_{M \times M_{\text{cp}}}, \mathbf{I}_M)^T \in \mathbb{R}^{P \times M}$; finally, the resulting matrix $\widetilde{\mathbf{Y}}[k]\,\mathbf{R}_{\text{cp}}$ is subject to discrete Fourier transform (DFT), thus obtaining the frequency-domain matrix $\mathbf{Y}[k] \triangleq \widetilde{\mathbf{Y}}[k]\,\mathbf{R}_{\text{cp}}\,\boldsymbol{\mathcal{W}}_{\text{DFT}} \in \mathbb{C}^{RX \times M}$, with $\boldsymbol{\mathcal{W}}_{\text{DFT}} \triangleq \boldsymbol{\mathcal{W}}_{\text{IDFT}}^{-1}$. Under assumptions (a1) and (a2), accounting for (6)–(8) and partitioning $\mathbf{Y}[k]$ as $\mathbf{Y}[k] = (\mathbf{y}_0[k], \mathbf{y}_1[k], \ldots, \mathbf{y}_{M-1}[k])$, we obtain the *per carrier* simple vector model

$$\mathbf{y}_m[k] = \boldsymbol{\mathcal{C}}_m[k]\,\mathbf{P}\,\mathbf{F}_m\,\mathbf{s}_m[k] + \mathbf{v}_m[k]\,, \quad (9)$$

for $m \in \{0, 1, \ldots, M-1\}$, where $\boldsymbol{\mathcal{C}}_m[k] \triangleq (\boldsymbol{\mathcal{C}}_{m,0}[k], \boldsymbol{\mathcal{C}}_{m,1}[k], \ldots, \boldsymbol{\mathcal{C}}_{m,TX_{\text{c}}-1}[k]) \in \mathbb{C}^{RX \times TX}$ collects the DFT samples corresponding to the mth subcarrier of all the MIMO channels $\{\boldsymbol{\mathcal{C}}_{m,i}[k]\}_{i=0}^{TX_{\text{c}}-1}$ from the cooperative BSs to the MS, with $\{\boldsymbol{\mathcal{C}}_{m,i}[k]\}_{r,\alpha} = \sum_{h=0}^{M_{\text{cp}}} c_{r,i,\alpha}[k, h]\, e^{-j\,\frac{2\pi}{M}\,h\,m}$, for $r \in \{0, 1, \ldots, RX - 1\}$ and $\alpha \in \{0, 1, \ldots, TX_{\text{a}} - 1\}$, the matrix $\mathbf{P} \triangleq (\mathbf{P}_0^T, \mathbf{P}_1^T, \ldots, \mathbf{P}_{TX_{\text{c}}-1}^T)^T \in \mathbb{C}^{TX \times L}$ is the collection of the random matrices used by all the BSs, and, finally, the vector $\mathbf{v}_m[k] \in \mathbb{C}^{RX}$ denotes the mth column of the noise matrix $\widetilde{\mathbf{V}}[k]\,\mathbf{R}_{\text{cp}}\,\boldsymbol{\mathcal{W}}_{\text{DFT}}$, with $\{\widetilde{\mathbf{V}}[k]\}_{r,q} = \widetilde{v}_{r,q}[k]$, for $r \in \{0, 1, \ldots, RX - 1\}$ and $q \in \{0, 1, \ldots, P-1\}$. Since $\mathbf{P}$ is chosen at random locally, our coding method is completely distributed and can be performed without agreement of the BSs that actually cooperate. In fact, the MS has to merely estimate the equivalent channel matrix

[3] The phase offset can be simply included in the channel impulse response.

[4] This filter is effectively the result of the convolution of the signal at the receiver with the filter that is matched to the pulse shaping characteristic, which in reality will be a root-squared raised cosine. Since we assume that the residual Doppler varies at a very slow pace compared to the effective duration of the pulse, we can approximately say that the imperfect down-conversion has no effect over the receive filtering operation and, hence, equivalently assume that the convolution with the matched root raised cosine filter brings a combined pulse shape equal to the raised cosine response. Thus, we include the operation in the definition of the transmit signal.

$\boldsymbol{\mathcal{E}}_m[k] \triangleq \boldsymbol{\mathcal{C}}_m[k]\,\mathbf{P} \in \mathbb{C}^{RX\times L}$ to decode $\mathbf{s}_m[k]$, and this can be accomplished completely ignoring what the actual channel parameters or randomization matrices are. Specifically, estimation of $\boldsymbol{\mathcal{E}}_m[k]$ can be performed by resorting to standard training-based identification methods [27]. In this case, each data transmission is preceded by a training period, wherein all the cooperating BSs transmit a symbol sequence known to the MS, by using certain randomization matrices that will be maintained until a new training phase is initiated.

In the following, we assume that not only $\mathbf{P}$ but also $\boldsymbol{\mathcal{C}}_m[k]$ is random and, for the moment, we do not make any specific assumption on their statistical models, assuming only that:

(a3) $\boldsymbol{\mathcal{H}}_m[k] \triangleq \boldsymbol{\mathcal{E}}_m[k]\,\mathbf{F}_m = \boldsymbol{\mathcal{C}}_m[k]\,\mathbf{P}\,\mathbf{F}_m \in \mathbb{C}^{RX\times N}$ is full-column rank, i.e., $\mathrm{rank}(\boldsymbol{\mathcal{H}}_m[k]) = N$, with probability one.

Since $N \leq \min(TX, RX)$ by construction, the necessary condition $N \leq RX$ needed for satisfying assumption (a3) is met automatically. Moreover, fulfillment of assumption (a3) necessarily requires that $\mathbf{P}\,\mathbf{F}_m \in \mathbb{C}^{TX\times N}$ is full-column rank, i.e., $\mathrm{rank}(\mathbf{P}\,\mathbf{F}_m) = N$, with probability one.[5]

Hereinafter, the following additional assumptions are made:

(a4) for any $m \in \{0, 1, \ldots, M-1\}$, $\mathbf{s}_m[k]$ is a zero-mean random vector, whose entries are independent and identically distributed (i.i.d.) Q-QAM equiprobable symbols having $\mathrm{E}[|s[n]|^2] = 1$, with $\mathbf{s}_{m_1}[k_1]$ and $\mathbf{s}_{m_2}[k_2]$ statistically independent of each other for $k_1 \neq k_2$ and $m_1 \neq m_2$;

(a5) for any $m \in \{0, 1, \ldots, M-1\}$, $\mathbf{v}_m[k] \sim \mathcal{CN}(\mathbf{0}_{RX}, \sigma_v^2\,\mathbf{I}_{RX})$ is statistically independent of $\mathbf{s}_m[k]$, with $\mathbf{v}_{m_1}[k_1]$ and $\mathbf{v}_{m_2}[k_2]$ statistically independent of each other for $k_1 \neq k_2$ and $m_1 \neq m_2$.

Under these assumptions, given $\boldsymbol{\mathcal{E}}_m[k]$, the minimum-error-probability detection of $\mathbf{s}_m[k]$ can be based only on $\mathbf{y}_m[k]$ (*one-shot* detection) and, thus, we will omit in the sequel the OFDM symbol index k. Moreover, symbol detection can be carried out separately for different subcarriers, i.e., a *per frequency tone* equalizing scheme can be applied. ML detection could be performed via, e.g., sphere decoding method [28]. However, it might exhibit a prohibitively large computational complexity for moderate-to-large values of the multiplexing rate N. Therefore, we focus attention on linear zero-order FIR decoding structures, whose input-output relationship can be expressed as $\boldsymbol{\xi}_m = \mathbf{G}_m\,\mathbf{y}_m$, with $\mathbf{G}_m \in \mathbb{C}^{N\times RX}$, which offer a good tradeoff between performance and complexity. The decision on the symbols transmitted on the mth subcarrier is obtained by quantizing each entry of $\boldsymbol{\xi}_m$ to the nearest (in terms of Euclidean distance) Q-QAM symbol. We consider the popular ZF and MMSE equalization criteria [29]. Under assumption (a3), the minimum-norm ZF equalizer is given by $\mathbf{G}_m^{\mathrm{zf}} = \boldsymbol{\mathcal{H}}_m^{\dagger}$ and assures perfect symbol recovery in the absence of noise, i.e., $\mathbf{G}_m^{\mathrm{zf}}\,[\mathbf{y}_m - \mathbf{v}_m] = \mathbf{s}_m$. In noisy environments, the ZF detector perfectly recovers the desired symbol block at the price of noise enhancement. To better counteract the noise, one can resort to the MMSE equalizer given by $\mathbf{G}_m^{\mathrm{mmse}} = \boldsymbol{\mathcal{H}}_m^H\,(\boldsymbol{\mathcal{H}}_m\,\boldsymbol{\mathcal{H}}_m^H + \sigma_v^2\,\mathbf{I}_{RX})^{-1}$, which,[6] under assumptions (a4) and (a5), minimizes the conditional mean square error $\mathrm{MSE}_m(\mathbf{G}_m) \triangleq \mathrm{E}[\|\boldsymbol{\xi}_m - \mathbf{s}_m\|^2 \,|\, \boldsymbol{\mathcal{E}}_m]$, given $\boldsymbol{\mathcal{E}}_m$.

[5]Inequality $N \leq \min(TX, RX)$ also implies that $N \leq TX$.

IV. Performance analysis and discussion

In this section, we evaluate the diversity order and coding gain [26] achieved by the R-MIMO-OFDM coding scheme with either ZF or MMSE detection. Let $\gamma \triangleq 1/\sigma_v^2$ denote the average SNR per QAM symbol expended by each BS, the diversity order $\mathcal{G}^{\mathrm{d}}$ is defined as the slope of the SEP curve $P_{m,n}(e)$ versus γ (in dB/dB scale), as $\gamma \to +\infty$, that is,

$$\mathcal{G}^{\mathrm{d}} \triangleq \lim_{\gamma\to+\infty} -\frac{\log P_{m,n}(e)}{\log \gamma}\,, \tag{10}$$

where $P_{m,n}(e)$, which is a function of γ, denotes the SEP at the output of the minimum distance detector of the nth entry of $\mathbf{s}_m$. On the other hand, the coding gain $\mathcal{G}_{m,n}^{\mathrm{c}}$ (in dB) measures the shift at high SNR of the SEP curve $P_{m,n}(e)$ relative to the benchmark SEP curve $\gamma^{-\mathcal{G}^{\mathrm{d}}}$, i.e.,

$$P_{m,n}(e) \approx (\mathcal{G}_{m,n}^{\mathrm{c}}\,\gamma)^{-\mathcal{G}^{\mathrm{d}}}\,, \quad \text{for high values of } \gamma\,. \tag{11}$$

If the FIR ZF equalizer $\mathbf{G}_m^{\mathrm{zf}}$ is used at the receiver and assumption (a3) holds, the equalizer output becomes $\boldsymbol{\xi}_m^{\mathrm{zf}} = \mathbf{s}_m + \mathbf{G}_m^{\mathrm{zf}}\,\mathbf{v}_m$, i.e., perfect intracarrier interference (ICI) suppression is achieved. Therefore, by virtue of assumption (a5) and conditioned on $\boldsymbol{\mathcal{E}}_m$, it follows that $\mathbf{d}_m^{\mathrm{zf}} \triangleq \mathbf{G}_m^{\mathrm{zf}}\,\mathbf{v}_m \sim \mathcal{CN}(\mathbf{0}_N, \sigma_v^2\,(\boldsymbol{\mathcal{H}}_m^H\,\boldsymbol{\mathcal{H}}_m)^{-1})$. The SEP in the detection of the nth entry of $\mathbf{s}_m$ conditioned on $\boldsymbol{\mathcal{E}}_m = \boldsymbol{\mathcal{C}}_m\,\mathbf{P}$ is given [30] by

$$P_{m,n}^{\mathrm{zf}}(e\,|\,\boldsymbol{\mathcal{E}}_m) \approx 2\left(1 - \frac{1}{\sqrt{Q}}\right)\mathrm{erfc}\left(\sqrt{\frac{3}{2(Q-1)}\,\mathrm{SINR}_{m,n}^{\mathrm{zf}}}\right), \tag{12}$$

with $n \in \{0, 1, \ldots, N-1\}$, where $\mathrm{erfc}(x) \triangleq (2/\sqrt{\pi})\int_x^{+\infty} e^{-u^2}\,du$ is the complementary error function and, under assumptions (a4) and (a5), the signal-to-interference-plus-noise ratio (SINR) at the output of the ZF equalizer is

$$\mathrm{SINR}_{m,n}^{\mathrm{zf}} = \frac{1}{\sigma_v^2\,\{(\boldsymbol{\mathcal{H}}_m^H\,\boldsymbol{\mathcal{H}}_m)^{-1}\}_{n,n}}\,. \tag{13}$$

In the case of the MMSE receiver, the equalizer output is $\boldsymbol{\xi}_m^{\mathrm{mmse}} = \mathbf{s}_m + \mathbf{d}_m^{\mathrm{mmse}}$, where, due to the presence of the ICI term,[7] the residual error vector $\mathbf{d}_m^{\mathrm{mmse}} \triangleq \mathbf{G}_m^{\mathrm{mmse}}\,\mathbf{y}_m - \mathbf{s}_m$ is non-Gaussian given $\boldsymbol{\mathcal{E}}_m$. However, under assumptions (a4) and (a5) and for large enough dimensions, conditional on $\boldsymbol{\mathcal{E}}_m$, the output disturbance can be well-modeled [31] as $\mathbf{d}_m^{\mathrm{mmse}} \sim \mathcal{CN}(\mathbf{0}_N, \sigma_v^2\,(\sigma_v^2\,\mathbf{I}_N + \boldsymbol{\mathcal{H}}_m^H\,\boldsymbol{\mathcal{H}}_m)^{-1})$. Hence, the SEP in the detection of the nth entry of $\mathbf{s}_m$ conditioned on $\boldsymbol{\mathcal{E}}_m$ is well approximated as follows (see [30])

$$P_{m,n}^{\mathrm{mmse}}(e\,|\,\boldsymbol{\mathcal{E}}_m) \approx 2\left(1 - \frac{1}{\sqrt{Q}}\right)\mathrm{erfc}\left(\sqrt{\frac{3}{2(Q-1)}\,\mathrm{SINR}_{m,n}^{\mathrm{mmse}}}\right), \tag{14}$$

[6]Matrix $\boldsymbol{\mathcal{H}}_m\,\boldsymbol{\mathcal{H}}_m^H + \sigma_v^2\,\mathbf{I}_{RX}$ turns out to be the autocorrelation matrix $\mathrm{E}[\mathbf{y}_m\,\mathbf{y}_m^H\,|\,\boldsymbol{\mathcal{E}}_m]$ of the vector $\mathbf{y}_m$, given $\boldsymbol{\mathcal{E}}_m$, which can be consistently estimated from the received data without any *a priori* knowledge of which BS is collaborating.

[7]Under assumption (a3), the ICI contribution becomes zero for $\gamma \to +\infty$.

with $n \in \{0,1,\ldots,N-1\}$, where, under assumptions (a4) and (a5), the SINR at the output of the MMSE equalizer is

$$\mathrm{SINR}_{m,n}^{\mathrm{mmse}} = \frac{1}{\sigma_v^2 \left\{(\sigma_v^2\,\mathbf{I}_N + \boldsymbol{\mathcal{H}}_m^H\,\boldsymbol{\mathcal{H}}_m)^{-1}\right\}_{n,n}} - 1\,. \quad (15)$$

In order to analyze the performance of the proposed R-MIMO-OFDM coding scheme, we have to evaluate the *average* SEPs $P_{m,n}^{\mathrm{zf}}(e)$ and $P_{m,n}^{\mathrm{mmse}}(e)$, which are defined as the expected value of (12) and (14), respectively, over the sample space of the pairs $\{\boldsymbol{\mathcal{C}}_m, \mathbf{P}\}$. In a unified perspective, let $P_{m,n}(e)$ and $P_{m,n}(e\,|\,\boldsymbol{\mathcal{E}}_m)$ denote the average and conditional SEPs, respectively, of either the ZF equalizer or the MMSE one, by virtue of the conditional expectation rule [32], one has

$$\begin{aligned} P_{m,n}(e) &\triangleq \mathrm{E}_{\boldsymbol{\mathcal{C}}_m,\mathbf{P}}\Big[P_{m,n}(e\,|\,\boldsymbol{\mathcal{E}}_m)\Big] \\ &= \mathrm{E}_{\mathbf{P}}\Big\{\underbrace{\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P}}\Big[P_{m,n}(e\,|\,\boldsymbol{\mathcal{E}}_m)\Big]}_{P_{m,n}(e\,|\,\mathbf{P})}\Big\} = \mathrm{E}_{\mathbf{P}}\Big[P_{m,n}(e\,|\,\mathbf{P})\Big]\,, \end{aligned} \quad (16)$$

where $\mathrm{E}_{\mathbf{P}}[\cdot]$ denotes the expectation over $\mathbf{P}$ and $\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P}}[\cdot]$ is the expectation over $\boldsymbol{\mathcal{C}}_m$ given $\mathbf{P}$.

In the following, we assume that:

(a6) $\forall r \in \{0,1,\ldots,RX-1\}$, $i \in \{0,1,\ldots,TX_{\mathrm{c}}-1\}$ and $\alpha \in \{0,1,\ldots,TX_{\mathrm{a}}-1\}$, the channel impulse response $c_{r,i,\alpha}(\tau)$ is modeled as a circularly symmetric complex Gaussian process with zero mean and autocorrelation function $\mathrm{E}[c_{r,i,\alpha}(\tau_1)\,c_{r,i,\alpha}^*(\tau_2)] = \frac{\rho_{m,i}}{M_{\mathrm{cp}}+1}\,\delta(\tau_1-\tau_2)$, where $c_{r_1,i_1,\alpha_1}(\tau)$ and $c_{r_2,i_2,\alpha_2}(\tau)$ are statistically independent of each other for $r_1 \neq r_2$, $i_1 \neq i_2$ and $\alpha_1 \neq \alpha_2$, whereas $\rho_{m,i} > 0$ accounts for the path loss associated with the ith BS.

The (deterministic) path loss in assumption (a6) is modeled as $\rho_{m,i} = A_m\,d_i^{-\xi}$ [33], where A_m is a (positive) propagation constant, d_i is the distance from the MS to the ith BS and ξ is the path loss exponent which can range from 3 to 6 depending on the operative environment. Observe that assumption (a6) states a Rayleigh model for fading, which is widely accepted for typical cellular environments with a large delay spread of the power delay profile and a relatively large number of small reflectors [26], [30]. Moreover, note that the channel model stated in assumption (a6) does not take into account both transmit and receive correlation. Such a model is reasonable in richly scattering environments [26]. On the other hand, if there are very few or no paths in some angular directions, while the transmit correlation can still be assumed negligible since BS antennae and inter-cell spacing are typically quite large [12], terminal antennae might be correlated. The effects of receive correlation on the performance of the proposed R-MIMO-OFDM transceiver are investigated in Section V through computer simulations.

It follows from assumption (a6) that, for any $\tau_i \in \mathbb{R}$, the processes $c_{r,i,\alpha}[h] = c_{r,i,\alpha}(h\,T-\tau_i)$, with $h \in \{0,1,\ldots,M_{\mathrm{cp}}\}$, is a sequence of mutually independent zero-mean i.i.d. circularly symmetric complex Gaussian random variables having variance $\frac{\rho_{m,i}}{M_{\mathrm{cp}}+1}$. Since $c_{r,i,\alpha}[h]$ is circularly symmetric complex Gaussian, then $c_{r,i,\alpha}[h]$ and $c_{r,i,\alpha}[k,h] = e^{j\,\varepsilon_i k\,T_s}\,c_{r,i,\alpha}[h]$ have the same probability distribution [34], for any $\varepsilon_i \in \mathbb{R}$ and $k \in \mathbb{Z}$. Thus, matrix $\boldsymbol{\mathcal{C}}_{m,i}$ can decomposed as $\boldsymbol{\mathcal{C}}_{m,i} = \sqrt{\rho_{m,i}}\,\boldsymbol{\mathcal{N}}_{m,i}$, where the entries of $\boldsymbol{\mathcal{N}}_{m,i} \in \mathbb{C}^{RX\times TX_{\mathrm{a}}}$ are i.i.d. circularly symmetric complex Gaussian random variables with zero mean and unit variance, where $\boldsymbol{\mathcal{N}}_{m,i_1}$ and $\boldsymbol{\mathcal{N}}_{m,i_2}$ are statistically independent of each other for $i_1 \neq i_2$.[8] Consequently, one has that $\boldsymbol{\mathcal{C}}_m = \boldsymbol{\mathcal{N}}_m\,(\boldsymbol{\Omega}_m \otimes \mathbf{I}_{TX_{\mathrm{a}}})$, with $\boldsymbol{\mathcal{N}}_m \triangleq (\boldsymbol{\mathcal{N}}_{m,0}, \boldsymbol{\mathcal{N}}_{m,1}, \ldots, \boldsymbol{\mathcal{N}}_{m,TX_{\mathrm{c}}-1}) \in \mathbb{C}^{RX\times TX}$ and $\boldsymbol{\Omega}_m \triangleq \mathrm{diag}(\sqrt{\rho_{m,0}}, \sqrt{\rho_{m,1}}, \ldots, \sqrt{\rho_{m,TX_{\mathrm{c}}-1}}) \in \mathbb{R}^{TX_{\mathrm{c}}\times TX_{\mathrm{c}}}$.

A. SEP analysis for a given matrix $\mathbf{P}$

On the basis of (16), the first step of our analysis is to keep the randomization matrix $\mathbf{P}$ fixed and evaluate the SEP $P_{m,n}(e\,|\,\mathbf{P})$ by averaging (12) and (14) over the matrix $\boldsymbol{\mathcal{C}}_m$. Unfortunately, the probabilities (12) and (14) are nontrivial functions of the statistics of $\boldsymbol{\mathcal{C}}_m$ and, thus, to obtain insightful results, in the following theorem we resort in Appendix A to an alternative definition of the complementary error function in order to derive upper bounds on the average SEPs (with respect to $\boldsymbol{\mathcal{C}}_m$) for a given realization of the randomization matrix $\mathbf{P}$ in both the ZF and MMSE cases:

Theorem 4.1 (Deterministic $\mathbf{P}$): Let $\boldsymbol{\mathcal{H}}_m^{(-n)} \in \mathbb{C}^{RX\times(N-1)}$ denote the matrix $\boldsymbol{\mathcal{H}}_m$ deprived of its nth column. Under assumptions (a3) and (a6), the probability $P_{m,n}^{\mathrm{zf}}(e\,|\,\mathbf{P})$, which is the expectation of (12) over $\boldsymbol{\mathcal{C}}_m$ for a given realization of $\mathbf{P}$, can be upper bounded as follows:

$$\begin{aligned} P_{m,n}^{\mathrm{zf}}(e\,|\,\mathbf{P}) \leq\; & 2\,\Theta(RX-N+1) \\ & \cdot\left(1-\frac{1}{\sqrt{Q}}\right)\left[\frac{3\,\gamma\,\Sigma_{m,n}}{2(Q-1)}\right]^{-(RX-N+1)}, \end{aligned} \quad (17)$$

where $\Theta(\alpha) \triangleq (2/\pi)\int_0^{\pi/2}(\sin^2\theta)^{\alpha}\,\mathrm{d}\theta$, with $\alpha \in \mathbb{R}$, whereas $\Sigma_{m,n} \triangleq 1/\{\mathbf{Z}_m^{-1}\}_{n,n} > 0$, with the deterministic matrix $\mathbf{Z}_m \triangleq \mathbf{F}_m^H\,\mathbf{P}^H\,(\boldsymbol{\Omega}_m^2 \otimes \mathbf{I}_{TX_{\mathrm{a}}})\,\mathbf{P}\,\mathbf{F}_m \in \mathbb{C}^{N\times N}$ being nonsingular as a consequence of assumption (a3).

Under assumptions (a3) and (a6), for moderate-to-high values of the SNR, i.e., when σ_v^2 is much smaller than the largest eigenvalue of $[\boldsymbol{\mathcal{H}}_m^{(-n)}]^H\,\boldsymbol{\mathcal{H}}_m^{(-n)}$, the probability $P_{m,n}^{\mathrm{mmse}}(e\,|\,\mathbf{P})$, which is the expectation of (14) over $\boldsymbol{\mathcal{C}}_m$ for a given realization of $\mathbf{P}$, can be upper bounded as follows

$$\begin{aligned} P_{m,n}^{\mathrm{mmse}}(e\,|\,\mathbf{P}) \lesssim\; & 2\,\psi_{m,n}(\mathbf{P})\,\Theta(RX-N+1) \\ & \cdot\left(1-\frac{1}{\sqrt{Q}}\right)\left[\frac{3\,\gamma\,\Sigma_{m,n}}{2(Q-1)}\right]^{-(RX-N+1)}, \end{aligned} \quad (18)$$

with

$$\begin{aligned} \psi_{m,n}(\mathbf{P}) \triangleq\; & \binom{RX}{N-1}\frac{N-1}{\Theta(RX-N+1)} \\ & \cdot\left\{\sum_{i_1=0}^{N-2}\sum_{i_2=1}^{RX-i_1}\binom{N-2}{i_1}\frac{(-1)^{i_2+N-RX-2}\,(i_2-1)!}{(RX-i_1)!}\right. \\ & \cdot\Theta(i_1+i_2-N+1)\left[\frac{2(Q-1)\,\rho_{m,\max}}{3\,N}\,\frac{\lambda_{\max}(\mathbf{P}^H\mathbf{P})}{\Sigma_{m,n}}\right]^{i_1+i_2-RX} \\ & \left.+\,(-1)^{N-RX-2}\sum_{i_1=0}^{N-2}\binom{N-2}{i_1}\frac{\Theta(i_1-N+2)}{(RX-i_1)!}\right. \end{aligned}$$

[8] In such a case, the squared Frobenius norm of $\boldsymbol{\mathcal{C}}_{m,i}$ is given by $\mathrm{E}[\mathrm{trace}(\boldsymbol{\mathcal{C}}_{m,i}^H\,\boldsymbol{\mathcal{C}}_{m,i})] = \mathrm{E}[\mathrm{trace}(\boldsymbol{\mathcal{C}}_{m,i}\,\boldsymbol{\mathcal{C}}_{m,i}^H)] = \rho_{m,i}\,TX_{\mathrm{a}}\,RX$.

$$\cdot \left[\frac{3\,N}{2(Q-1)\,\rho_{m,\max}}\,\frac{\Sigma_{m,n}}{\lambda_{\max}(\mathbf{P}^H\mathbf{P})}\right]^{RX-i_1-1}\Bigg\}\,, \quad (19)$$

where $\rho_{m,\max} \triangleq \max_{i\in\{0,1,\ldots,TX_c-1\}} \rho_{m,i}$ is the path loss associated with the BS that is farthest from the MS, whereas $\lambda_{\max}(\mathbf{P}^H\mathbf{P}) > 0$ is the largest eigenvalue of $\mathbf{P}^H\,\mathbf{P} \in \mathbb{C}^{L\times L}$.

Proof: See Appendix A. ∎

Our simulation results (see Section V) confirm that, for moderate-to-high values of the SNR, the bounds (17) and (18) become tight, and can be used to accurately predict the network performance when the BSs keep the randomization matrices $\{\mathbf{P}_i\}_{i=0}^{TX_c-1}$ fixed over the transmission of a given data packet. Two remarks about Theorem 4.1 are in order:

Remark 1: Since $P_{m,n}^{\text{zf}}(e\,|\,\mathbf{P}) = O(\gamma^{-(RX-N+1)})$ and $P_{m,n}^{\text{mmse}}(e\,|\,\mathbf{P}) = O(\gamma^{-(RX-N+1)})$,[9] independently of the choice of the matrix $\mathbf{P}$, the diversity order of the considered scheme, with linear (L) equalization at the receiver, is equal to $\mathcal{G}^{\text{d,L}} = RX - N + 1$. It is worth observing that, when $TX > RX$, which is a situation of practical relevance, then $N \le \min(TX, RX) = RX$ and, consequently, $\mathcal{G}^{\text{d,L}} = RX - N + 1 \ge 1$. As a benchmark, note that, for a space-frequency coded broadband MIMO-OFDM systems with ML detection at the receiver, under our assumptions, the maximum achievable diversity order is $\mathcal{G}^{\text{d,ML}} = M_{\text{cp}}\,TX\,RX$ [35].

Remark 2: The upper bounds (17) and (18) differ just by the factor $\psi_{m,n}(\mathbf{P})$. In particular, Theorem 4.1 shows that, at high SNR and for any fixed matrix $\mathbf{P}$ ensuring that $\mathbf{Z}_m$ is nonsingular, the coding gains of the ZF and MMSE equalizers are respectively given by

$$\mathcal{G}_{m,n}^{\text{c,zf}}(\mathbf{P}) = \frac{3\,\Sigma_{m,n}}{2(Q-1)} \cdot \left[2\,\Theta(RX-N+1)\,\left(1-\frac{1}{\sqrt{Q}}\right)\right]^{-\frac{1}{RX-N+1}}, \quad (20)$$

$$\mathcal{G}_{m,n}^{\text{c,mmse}}(\mathbf{P}) = \mathcal{G}_{m,n}^{\text{c,zf}}(\mathbf{P})\;[\psi_{m,n}(\mathbf{P})]^{-\frac{1}{RX-N+1}}\,, \quad (21)$$

where $\Sigma_{m,n}$ depends on the choice of the matrix $\mathbf{P}$, the pre-coding matrix $\mathbf{F}_m$ and the path loss coefficients $\{\rho_{m,i}\}_{i=0}^{TX_c-1}$ which are related to the distances $\{d_i\}_{i=0}^{TX_c-1}$ from the MS to the cooperating BSs.

It is interesting to derive the condition under which the matrix $\mathbf{Z}_m$ maximizes the coding gains of the proposed transceivers, i.e., the SEP is minimized in the high SNR regime. To this aim, for the sake of brevity, we restrict ourselves to the ZF equalizer, which assures perfect symbol recovery in the absence of noise. Applying the Cauchy-Schwartz inequality [38], it is easy to show that $\{\mathbf{Z}_m^{-1}\}_{n,n} \ge \{\mathbf{Z}_m\}_{n,n}$, where the equality holds for any $n \in \{0,1,\ldots,N-1\}$ if and only if $\mathbf{Z}_m = \mu\,\mathbf{I}_N$, for some constant $\mu > 0$. Consequently, the achievable coding gain (20) can be upper bounded as

$$\mathcal{G}_{m,n}^{\text{c,zf}}(\mathbf{P}) \le \frac{3}{2(Q-1)\{\mathbf{Z}_m\}_{n,n}} \cdot \left[2\,\Theta(RX-N+1)\,\left(1-\frac{1}{\sqrt{Q}}\right)\right]^{-\frac{1}{RX-N+1}}, \quad (22)$$

[9]The expression $O(x)$ denotes the Landau symbol: for a function $f(x) = O(x)$, the ratio $f(x)/x$ is bounded, as $x \to +\infty$.

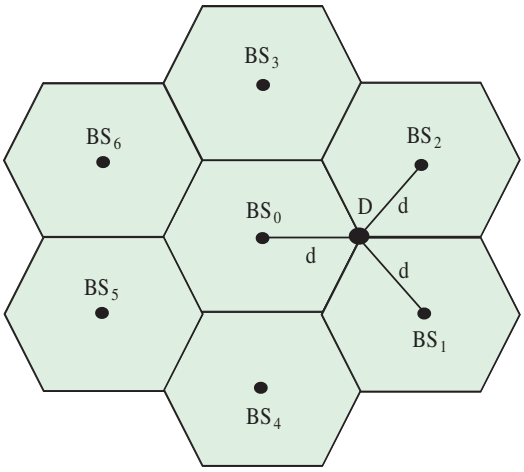

Fig. 3. Example of a uniform hexagonal topology. The destination D (i.e., the MS) is at equal distance d from the three cooperative base stations BS_0, BS_1 and BS_2.

where the equality holds for any $n \in \{0,1,\ldots,N-1\}$ if and only if $\mathbf{F}_m^H\,\mathbf{P}^H\,(\mathbf{\Omega}_m^2 \otimes \mathbf{I}_{TX_a})\,\mathbf{P}\,\mathbf{F}_m = \mu\,\mathbf{I}_N$, for some constant $\mu > 0$. For given matrices $\mathbf{F}_m$ and $\mathbf{P}$, it is also readily seen that $\{\mathbf{Z}_m\}_{n,n} \ge \rho_{m,\max}\,\{\overline{\mathbf{Z}}_m\}_{n,n}$, with $\overline{\mathbf{Z}}_m \triangleq \mathbf{F}_m^H\,\mathbf{P}^H\,\mathbf{P}\,\mathbf{F}_m \in \mathbb{C}^{N\times N}$ and $\rho_{m,\max}$ already defined in Theorem 4.1, where equality holds if and only if $\rho_{m,i} = \rho_m$, $\forall i \in \{0,1,\ldots,TX_c-1\}$. In other words, for given matrices $\mathbf{F}_m$ and $\mathbf{P}$, the maximum coding gain is achieved in the ZF case when the MS is at equal distance from the cooperating BSs. For instance, this is the case of Fig. 3, where there are three active BSs (namely, BS_0, BS_1 and BS_2) and the MS is at the cell edge positioned at the common vertex of the three hexagonal cells, which is the worst-case scenario for a non-cooperative system. In such a scenario, the path loss coefficient ρ_m is given by $\rho_m = A_m\,d^{-\xi}$. A case of interest in which $\mathbf{Z}_m$ ends up to scaled identity matrix is the centralized MIMO-OFDM coding strategy, which requires an *a priori* agreement among the BSs involved in the cooperative transmission, as well as a preliminary identification at the destination of the set of cooperating BSs. Indeed, for a centralized MIMO-OFDM coding scheme, we have that $L = TX$ and $\mathbf{P} = \sqrt{TX_c}\;\mathbf{I}_{TX}$,[10] and, hence, $\lambda_{\max}(\mathbf{P}^H\mathbf{P}) = TX_c$. Moreover, in the special case of equal distance of cooperating BSs and the destination, i.e., $\mathbf{\Omega}_m = \sqrt{\rho_m}\;\mathbf{I}_{TX_c}$, it follows that $\mathbf{Z}_m = \frac{TX_c\,\rho_m}{N}\,\mathbf{I}_N$, which in turn implies that $\Sigma_{m,n} = (TX_c\,\rho_m)/N$ for a centralized MIMO-OFDM coding rule. In such a case, we can infer from (17)–(19) the following corollary:

Corollary 4.2 (Centralized MIMO-OFDM coding): In the case wherein $\mathbf{\Omega}_m = \sqrt{\rho_m}\;\mathbf{I}_{TX_c}$, with $\rho_m = A_m\,d^{-\xi}$, under the same hypotheses of Theorem 4.1, the performance of the centralized scheme for the ZF and MMSE receivers are respectively given by

$$P_m^{\text{zf, central}}(e) \le 2\,\Theta(RX-N+1)$$

[10]The factor $\sqrt{TX_c}$ ensures that the total power per carrier transmitted by the ith BS $\mathrm{E}[\|\mathbf{P}_i\,\mathbf{F}_m\,\mathbf{s}_m\|^2]$ is equal to one.

$$\cdot \left(1-\frac{1}{\sqrt{Q}}\right)\left[\frac{3\,\gamma\,TX_{\rm c}\,\rho_m}{2\,N(Q-1)}\right]^{-(RX-N+1)}, \tag{23}$$

$$P_m^{\text{mmse, central}}(e) \le 2\,\psi_{\text{central}}\,\Theta(RX-N+1) \cdot \left(1-\frac{1}{\sqrt{Q}}\right)\left[\frac{3\,\gamma\,TX_{\rm c}\,\rho_m}{2\,N(Q-1)}\right]^{-(RX-N+1)}, \tag{24}$$

with $\Theta(\cdot)$ already defined in Theorem 4.1 and

$$\psi_{\text{central}} \triangleq \binom{RX}{N-1}\frac{N-1}{\Theta(RX-N+1)} \cdot \left\{\sum_{i_1=0}^{N-2}\sum_{i_2=1}^{RX-i_1}\binom{N-2}{i_1}\frac{(-1)^{i_2+N-RX-2}\,(i_2-1)!}{(RX-i_1)!} \cdot \left[\frac{2(Q-1)}{3}\right]^{i_1+i_2-RX}\Theta(i_1+i_2-N+1) + (-1)^{N-RX-2}\sum_{i_1=0}^{N-2}\binom{N-2}{i_1}\frac{\Theta(i_1-N+2)}{(RX-i_1)!} \cdot \left[\frac{3}{2(Q-1)}\right]^{RX-i_1-1}\right\}. \tag{25}$$

Corollary 4.2 shows that, at high SNR, the maximum coding gains of the centralized scheme in the ZF and MMSE cases can be respectively expressed as

$$\mathcal{G}_m^{\text{c, zf, central}} = \frac{3\,TX_{\rm c}\,\rho_m}{2\,N(Q-1)} \cdot \left[2\,\Theta(RX-N+1)\,\left(1-\frac{1}{\sqrt{Q}}\right)\right]^{-\frac{1}{RX-N+1}}, \tag{26}$$

$$\mathcal{G}_m^{\text{c, mmse, central}} = \mathcal{G}_m^{\text{c, zf, central}}\,(\psi_{\text{central}})^{-\frac{1}{RX-N+1}}. \tag{27}$$

It can be seen that both $\mathcal{G}_m^{\text{c, zf, central}}$ and $\mathcal{G}_m^{\text{c, mmse, central}}$ increase linearly as the number of cooperative BSs increases and, additionally, the performance gain of the MMSE equalizer over the ZF one is dictated by ψ_{central}, which depends only on RX, Q and N. We end this subsection with two remarks:

Remark 3: In a noncooperative cellular system, the MS is exclusively served by its own BS. Such an approach can be seen as the trivial case of a centralized cooperative scheme with only one active BS, i.e., $TX_{\rm c}=1$. Therefore, upper bounds on the performances of a MIMO-OFDM noncooperative system with linear ZF and MMSE equalization at the receiver can be obtained from (23) and (24) by simply setting $TX_{\rm c}=1$, provided that $N \le \min(TX_{\rm a}, RX)$.

Remark 4: Let us consider the simple cooperative scheme where each BS transmits the same spatially-precoded block $\mathbf{F}_m\,\mathbf{s}_m$, without any coordination amongst themselves, which will be referred to as repetition coding rule. This can be accomplished by setting in (9) $L=TX_{\rm a}$ and $\mathbf{P}_i=\mathbf{I}_{TX_{\rm a}}$, $\forall i \in \{0,1,\ldots,TX_{\rm c}-1\}$. In this case, it is easily seen that $\mathbf{P}^H\mathbf{P}=TX_{\rm c}\,\mathbf{I}_{TX_{\rm a}}$, which leads to $\lambda_{\max}(\mathbf{P}^H\mathbf{P})=TX_{\rm c}$, and, hence, in the special case of $\mathbf{\Omega}_m=\sqrt{\rho_m}\,\mathbf{I}_{TX_{\rm c}}$ and provided that $N \le \min(TX_{\rm a}, RX)$, one has $\mathbf{Z}_m=\frac{TX_{\rm c}\,\rho_m}{N}\,\mathbf{I}_N$, which in its turn implies that, exactly as for a centralized MIMO-OFDM coding rule, $\Sigma_{m,n}=(TX_{\rm c}\,\rho_m)/N$. Thus, when $TX_{\rm a} \ge RX$, the simplest decentralized strategy, consisting of transmitting the same block $\mathbf{F}_m\,\mathbf{s}_m$ from each BS, exhibits the same maximum coding gains (26) and (27) of the centralized scheme, under the same number $N \le RX$ of QAM symbols transmitted per subcarrier. However, when $RX > TX_{\rm a}$, such a repetition coding rule does not exploit the degrees of freedom available in the channel effectively, by enabling the transmission of only $N \le TX_{\rm a}$ QAM symbols per subcarrier, whereas the centralized scheme can reliably transmit up to RX QAM symbols per subcarrier, provided that $TX=TX_{\rm a}\cdot TX_{\rm c} \ge RX$.

B. SEP analysis with complex Gaussian matrix $\mathbf{P}$

At this point, according to (16), we are in the position of providing an upper bound on the average performance of our R-MIMO-OFDM coding rule, by averaging (17) and (18) with respect to the randomization matrix $\mathbf{P}$. The obtained results are representative of the receiver average performance when the BSs change the randomization matrices $\{\mathbf{P}_i\}_{i=0}^{TX_{\rm c}-1}$ several times over the transmission, so that each data packet experiences a large number of realizations of $\mathbf{P}$. To complete the analysis, we assume that:

(a7) $\forall i \in \{0,1,\ldots,TX_{\rm c}-1\}$, the entries of $\mathbf{P}_i \in \mathbb{C}^{TX_{\rm a}\times L}$ are i.i.d. circularly symmetric complex Gaussian random variables with zero mean and variance $1/TX_{\rm a}$, where $\mathbf{P}_{i_1}$ and $\mathbf{P}_{i_2}$ are statistically independent of each other for $i_1 \ne i_2$.

Such an assumption implies that $\text{rank}(\mathbf{P}_i)=\min\{TX_{\rm a}, L\}$ with probability one and, accounting also for assumption (a4), the total power per carrier transmitted by the ith BS is fixed to one, i.e., $\text{E}[\|\mathbf{P}_i\,\mathbf{F}_m\,\mathbf{s}_m\|^2]=1$. Note that the complex Gaussian randomization is used only for the sake of mathematical tractability and other choices are possible, as discussed in Section V. Moreover, to keep the analysis reasonably simple, we focus attention on the special case of nearly equal distance between the cooperating BSs and the MS, i.e., $\mathbf{\Omega}_m=\sqrt{\rho_m}\,\mathbf{I}_{TX_{\rm c}}$, with $\rho_m=A_m\,d^{-\xi}$. Notwithstanding its simplicity, this model captures the essential features of a multicell cooperative downlink and it allows us to gain insight into the network performance. Capitalizing on Theorem 4.1, the following theorem provides an upper bound on the average (with respect to both $\boldsymbol{\mathcal{C}}_m$ and $\mathbf{P}$) SEPs at the output of the linear ZF and MMSE equalizers:

Theorem 4.3 (Complex Gaussian randomization): Assume that $TX > RX$ and $\mathbf{\Omega}_m=\sqrt{\rho_m}\,\mathbf{I}_{TX_{\rm c}}$, with $\rho_m=A_m\,d^{-\xi}$. At high SNR, i.e., as $\sigma_v^2 \to 0$, under assumptions (a3), (a6) and (a7), the probability $P_{m,n}^{\text{zf}}(e)$, which is the expectation of $P_{m,n}^{\text{zf}}(e\,|\,\mathbf{P})$ over the randomization matrix $\mathbf{P}$, depends neither on the symbol index n nor on the subcarrier index m and can be upper bounded as

$$P_{m,n}^{\text{zf}}(e)=P^{\text{zf}}(e) \lesssim 2\,\Theta(RX-N+1)\,\left(1-\frac{1}{\sqrt{Q}}\right) \cdot \frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)}\left[\frac{3\,\gamma\,\rho_m}{2\,N\,TX_{\rm a}\,(Q-1)}\right]^{-(RX-N+1)}, \tag{28}$$

with $\Theta(\cdot)$ already defined in Theorem 4.1 and $\Gamma(\alpha) \triangleq \int_0^{+\infty} t^{\alpha-1}\,\exp(-t)\,\text{d}t$ defining the gamma function ($\alpha>0$).

Under assumptions (a3), (a6) and (a7), and for sufficiently large values of TX with $L/TX \neq 0$, the probability $P_{m,n}^{\text{mmse}}(e)$, which is the expectation of $P_{m,n}^{\text{mmse}}(e \,|\, \mathbf{P})$ over the randomization matrix $\mathbf{P}$, depends neither on the symbol index n nor on the subcarrier index m and can be upper bounded as follows

$$P_{m,n}^{\text{mmse}}(e) = P^{\text{mmse}}(e) \lesssim 2\,\psi\,\Theta(RX-N+1)\left(1-\frac{1}{\sqrt{Q}}\right) \cdot \frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)}\left[\frac{3\,\gamma\,\rho_m}{2\,N\,TX_{\text{a}}(Q-1)}\right]^{-(RX-N+1)}, \quad (29)$$

with

$$\begin{aligned}\psi \triangleq{}& \binom{RX}{N-1}\frac{(N-1)}{\Gamma(TX-RX)\,\Theta(RX-N+1)} \cdot \left\{\sum_{i_1=0}^{N-2}\sum_{i_2=1}^{RX-i_1}\binom{N-2}{i_1}\frac{(-1)^{i_2+N-RX-2}\,(i_2-1)!}{(RX-i_1)!}\right. \\ &\cdot\left[\frac{2(Q-1)}{3\,N}\right]^{i_1+i_2-RX}\Theta(i_1+i_2-N+1) \\ &\cdot\left(1+\sqrt{\frac{L}{TX}}\right)^{2(i_1+i_2-RX)}\Gamma(TX-i_1-i_2) \\ &\cdot(N\,TX)^{i_1+i_2-RX}+(-1)^{N-RX-2} \\ &\cdot\sum_{i_1=0}^{N-2}\binom{N-2}{i_1}\frac{\Theta(i_1-N+2)}{(RX-i_1)!}\left[\frac{3\,N}{2(Q-1)}\right]^{RX-i_1-1} \\ &\left.\cdot\left(1+\sqrt{\frac{L}{TX}}\right)^{2(i_1-RX+1)}\Gamma(TX-i_1-1)\,(N\,TX)^{i_1-RX+1}\right\}. \end{aligned} \quad (30)$$

Proof: See Appendix B. ∎

Observe that, since TX is equal to the product of the number TX_{a} of local antennae at each BS and the number TX_{c} of cooperative BSs, assumption $TX > RX$ is very likely satisfied in practice. Moreover, inequalities (28) and (29) can be further explained by observing that $\Gamma(n) = (n-1)!$, for any positive integer n. Remarkably, eqs. (28) and (29) render explicit the effect of the randomization on the coding gain; the coding gains of the R-MIMO-OFDM scheme in the ZF and MMSE cases are respectively given by

$$\mathcal{G}_m^{\text{c, zf}} = \frac{3\,\rho_m}{2\,N\,TX_{\text{a}}\,(Q-1)}\left[2\,\Theta(RX-N+1)\cdot\left(1-\frac{1}{\sqrt{Q}}\right)\frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)}\right]^{-\frac{1}{RX-N+1}}, \quad (31)$$

$$\mathcal{G}_m^{\text{c, mmse}} = \mathcal{G}_{\text{c, zf}}\,\psi^{-\frac{1}{RX-N+1}}. \quad (32)$$

Remark 5: The upper bounds (28) and (29), as well as the coding gains (31) and (32), are proportionally related by means of the factor ψ, which only depends on the system parameters TX, RX, L, N and Q and, thus, it is easily computable. Specifically, the performance gain of the MMSE equalizer over the ZF in terms of coding gain is determined by ψ.

The results of Theorems 4.1 and 4.3 lead to a direct and fruitful comparison between the randomized coding rule and its centralized counterpart, which is the subject of Subsection IV-C. More importantly, they allow us to estimate the cell coverage expansion provided by the R-MIMO-OFDM scheme compared to the traditional noncooperative scenario.

C. Discussion

First, it is interesting to compare (31) and (32) with their centralized counterparts (26) and (27), respectively. To this aim, let us consider in particular the ZF case. By comparing (31) and (26), we can infer that the coding gain for the randomized scheme with linear ZF receiver turns out to be equal to the coding gain of its centralized counterpart scaled by the factor

$$\theta^{\text{zf}} \triangleq \frac{\mathcal{G}_m^{\text{c, zf}}}{\mathcal{G}_m^{\text{c, zf, central}}} = \frac{1}{TX}\left[\frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)}\right]^{-\frac{1}{RX-N+1}}, \quad (33)$$

where, when $TX > RX$, one has $N \leq RX$. It is useful to specialize (33) for the particular case where $N = RX$, for which the diversity order is equal to $\mathcal{G}^{\text{d,L}} = 1$. In such a case, the scaling factor (33) assumes the simpler form

$$\theta^{\text{zf}} = 1 - \frac{RX}{TX}, \quad (34)$$

which shows that $\theta^{\text{zf}} < 1$, i.e., as expected, the centralized scheme exhibits better performance compared with the randomized coding rule. However, we can infer from (34) that, when the linear ZF equalizer is used at the receiver, the coding gain of the proposed R-MIMO-OFDM coding rule approaches that of the centralized scheme as $RX/TX \to 0$. In particular, if the number of local transmit antennae TX_{a} and the number of local receive antennae RX are fixed, the randomized scheme performs comparably to the centralized one as the number TX_{c} of cooperative BSs increases. As shown by simulation results in Section V, similar conclusions also hold for the MMSE equalizer.

Let us now examine the coverage expansion provided by the cooperative R-MIMO-OFDM approach compared to the traditional noncooperative scenario with single-cell signaling, where each MS can exclusively communicate with its own BS. Cell coverage is defined as the maximum distance at which the MS can be located from the serving BS while still maintaining a certain quality of service (QoS). To measure the QoS of the MS, we resort to the average SEP $P_{m,n}(e)$ defined in (16) and we say that the MS can reliably communicate with its own BS if $P_{m,n}(e) \leq SEP_{m,\text{target}}$, where $SEP_{m,\text{target}}$ is a predetermined average SEP threshold regarding the mth subcarrier. Therefore, the cell coverage is the distance of the MS from its own BS where the average SEP $P_{m,n}(e)$ equals the maximum allowable value $SEP_{m,\text{target}}$. For the proposed cooperative scheme with complex Gaussian randomization, the cell coverage can be approximated analytically by resorting to the upper bounds (28) and (29), where we remember that $TX > RX$ and $\rho_m = A_m\,d^{-\xi}$. Specifically, considering ZF equalization [eq. (28)], the cell coverage $D^{\text{cell, zf}}$ of the R-

MIMO-OFDM transceiver is the value of d satisfying

$$2\,\Theta(RX-N+1)\left(1-\frac{1}{\sqrt{Q}}\right)\frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)} \cdot\left[\frac{3\,\gamma\,A_m\,d^{-\xi}}{2\,N\,TX_{\rm a}\,(Q-1)}\right]^{-(RX-N+1)} = SEP_{m,\text{target}}\,, \quad (35)$$

whose solution is given by

$$D^{\text{cell, zf}} = \left\{\frac{2\,N\,TX_{\rm a}\,(Q-1)}{3\,\gamma\,A_m}\,[2\,\Theta(RX-N+1) \cdot\left(1-\frac{1}{\sqrt{Q}}\right)\frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)}\,\frac{1}{SEP_{m,\text{target}}}\right]^{\frac{1}{RX-N+1}}\right\}^{-\frac{1}{\xi}}. \quad (36)$$

In order to approximate analytically the cell coverage for the noncooperative system, we recall that such an approach can be seen as a special case of a centralized cooperative scheme with only one active BS (see Remark 3). Hence, the cell coverage in the noncooperative scenario can be obtained from the upper bounds (23) and (24) by setting $TX_{\rm c}=1$. In the specific case of ZF equalization [see (23)], the cell coverage $D^{\text{cell, zf, nocoop}}$ is the value of d satisfying

$$2\,\Theta(RX-N+1)\left(1-\frac{1}{\sqrt{Q}}\right) \cdot\left[\frac{3\,\gamma\,A_m\,d^{-\xi}}{2\,N(Q-1)}\right]^{-(RX-N+1)} = SEP_{m,\text{target}}\,, \quad (37)$$

whose solution can be expressed as

$$D^{\text{cell, zf, nocoop}} = \left\{\frac{2\,N\,(Q-1)}{3\,\gamma\,A_m}\,[2\,\Theta(RX-N+1) \cdot\left(1-\frac{1}{\sqrt{Q}}\right)\frac{1}{SEP_{m,\text{target}}}\right]^{\frac{1}{RX-N+1}}\right\}^{-\frac{1}{\xi}}. \quad (38)$$

To quantify the improvement in cell coverage with linear ZF equalization, we define the coverage expansion (CE) coefficient as follows

$$\text{CE}^{\text{zf}} \triangleq \frac{D^{\text{cell, zf}} - D^{\text{cell, zf, nocoop}}}{D^{\text{cell, zf, nocoop}}}\cdot 100\%\,, \quad (39)$$

which, accounting for (36) and (38), takes the form

$$\text{CE}^{\text{zf}} = \left\{\left[TX_{\rm a}\left(\frac{\Gamma(TX-RX)}{\Gamma(TX-N+1)}\right)^{\frac{1}{RX-N+1}}\right]^{-\frac{1}{\xi}} - 1\right\} \cdot 100\%\,. \quad (40)$$

Assuming that $TX_{\rm a}\geq RX$ and $TX_{\rm c}>1$, it is again interesting to specialize (40) when $N=RX$. In such a case, the coverage expansion is simply given by

$$\text{CE}^{\text{zf}} = \left[\left(TX_{\rm c}-\frac{RX}{TX_{\rm a}}\right)^{1/\xi} - 1\right]\cdot 100\%\,, \quad (41)$$

which depends only on the number $TX_{\rm c}$ of cooperative BSs, on the ratio $RX/TX_{\rm a}$ between the number of transmit and receive local antennae, and on the path loss exponent ξ. For instance, let us consider the case of Fig. 3, where there are $TX_{\rm c}=3$ cooperative BSs, and set $\xi=3$. If the number of receive antennae is equal to the number of local transit antennae, i.e., $RX/TX_{\rm a}=1$, the proposed R-MIMO-OFDM ZF-based transceiver with complex Gaussian randomization is able to assure a coverage expansion of about 26%. If each BS employs a number of antennae that is twice the number of receive antennae, i.e., $RX/TX_{\rm a}=1/2$, the coverage expansion becomes about 36%. In the limiting case when $RX/TX_{\rm a}\ll TX_{\rm c}$, the proposed cooperative PHY layer assures a cell coverage of little more than 44%, with respect to a traditional noncooperative cellular system. The coverage analysis we just completed can be similarly developed for the MMSE receiver as well.

V. Numerical results

The Monte Carlo simulations presented in this section are aimed at showing the performance of the R-MIMO-OFDM coding technique proposed in Section III and, at the same time, are useful for validating and extending the performance analysis carried out in Section IV.

In all the examples, the number of subcarriers in the OFDM air interface is $M=64$ with QPSK signaling, i.e., $Q=4$, and the CP length is $M_{\rm cp}=16$. All the channels are generated according to assumptions (a1), (a2) and, unless otherwise specified, (a6), with $A_m=1.35\cdot 10^7$ and $\xi=3$ [33]. For the triplet $(TX_{\rm a},RX,N)$ of transmit antennae $TX_{\rm a}$, receive antennae RX and the multiplexing rate N, we considered the $(1,2,2)$ and $(4,2,2)$ system configurations. Observe that, in all the cases, the diversity order of a MIMO-OFDM transceiver with linear equalization is $\mathcal{G}^{\rm d,L}=RX-N+1=1$. We refer to the scenario depicted in Fig. 3, wherein the three base stations BS_0, BS_1 and BS_2 cooperate to serve D, which is at equal distance $d=250$ meters from the cooperating BSs. Besides numerically evaluating the performance of the R-MIMO-OFDM scheme in the case of complex Gaussian randomization, where $\mathbf{P}$ is built as in assumption (a7), we consider four additional randomized coding schemes:

1) Uniform phase randomization: $\forall i\in\{0,1,\ldots,TX_{\rm c}-1\}$, the entries of the ith randomization matrix $\mathbf{P}_i\in\mathbb{C}^{TX_{\rm a}\times L}$ are i.i.d. random variables, with the $(\alpha+1,\ell+1)$th element generated as $\{\mathbf{P}_i\}_{\alpha,\ell}=\frac{e^{j\,\vartheta_{\alpha,\ell}}}{\sqrt{TX_a}}$, for $\alpha\in\{0,1,\ldots,TX_{\rm a}-1\}$ and $\ell\in\{0,1,\ldots,L-1\}$, where $\vartheta_{\alpha,\ell}$ is uniformly distributed in the interval $(0,2\pi)$, and $\mathbf{P}_{i_1}$ and $\mathbf{P}_{i_2}$ are statistically independent of each other for $i_1\neq i_2$. The normalization by $1/\sqrt{TX_a}$ ensures that the total power per carrier transmitted by the ith BS is $\text{E}[\|\mathbf{P}_i\,\mathbf{F}_m\,\mathbf{s}_m\|^2]=1$.

2) Haar randomization: $\forall i\in\{0,1,\ldots,TX_{\rm c}-1\}$, the ith matrix $\mathbf{P}_i\in\mathbb{C}^{TX_{\rm a}\times L}$ is a Haar-distributed unitary random matrix [36], with $\mathbf{P}_{i_1}$ and $\mathbf{P}_{i_2}$ statistically independent of each other for $i_1\neq i_2$. Specifically, let $\mathbf{\Theta}_i\in\mathbb{C}^{TX_{\rm a}\times L}$ (with $L>TX_{\rm a}$ in our simulation setting) be a random matrix whose entries are i.i.d. circularly symmetric complex Gaussian random variables with zero mean and unit variance, which is full-row rank, i.e., $\text{rank}(\mathbf{\Theta}_i)=TX_{\rm a}$, with probability one, we get $\mathbf{P}_i=\frac{1}{\sqrt{\epsilon_i}}\,(\mathbf{\Theta}_i\,\mathbf{\Theta}_i^H)^{-1/2}\,\mathbf{\Theta}_i$, with $\epsilon_i\triangleq\text{trace}[\mathbf{F}_m^H\,\mathbf{\Theta}_i^H\,(\mathbf{\Theta}_i\,\mathbf{\Theta}_i^H)^{-1}\,\mathbf{\Theta}_i\,\mathbf{F}_m]$. The normalization by $1/\sqrt{\epsilon_i}$ assures that $\text{E}[\|\mathbf{P}_i\,\mathbf{F}_m\,\mathbf{s}_m\|^2]=1$.

3) Delay randomization: it is a modified version of the scheme proposed in [22] for single-antenna single-carrier ad hoc networks, where the cooperating BSs introduce intentional transmission delays, by randomly selecting the delays from a pool. For $i \in \{0,1,\ldots,TX_{\rm c}-1\}$, such a scheme can be recasted in the signal model (9) by setting $L = TX_{\rm a}$, $\mathbf{P}_i = \mathbf{I}_{TX_{\rm a}}$ and by decomposing the delay τ_i in (5) as $\tau_i = \tau_{i,1}+\tau_{i,2}$, where $\tau_{i,1}$ accounts for the *unintentional* transmission and propagation delays, whereas $\tau_{i,2}$ is the *intentional* delay introduced by the ith BS.[11] We assume that $\tau_{i,2} = \Delta_i\, T$, with $\Delta_0, \Delta_1, \ldots, \Delta_{TX_{\rm c}-1}$ being numbers chosen randomly and independently from $\{0,1,\ldots,\Delta_{\max}\}$, where $\Delta_{\max}$ must be chosen such that assumption (a1) is fulfilled. Note that, based on the aforementioned decomposition of τ_i, after CP removal and DFT, the $RX \times TX_{\rm a}$ frequency-domain channel matrix of the ith BS assumes the form $e^{-j\,\frac{2\pi}{M}\Delta_i\, m}\,\boldsymbol{\mathcal{C}}_{m,i}$, with $\boldsymbol{\mathcal{C}}_{m,i}$ generated in accordance with assumption (a6). We set $\Delta_{\max} = 4$ and,[12] assuming $N \leq \min\{TX_{\rm a}, RX\}$, the channel-independent precoding matrix $\mathbf{F}_m \in \mathbb{C}^{TX_{\rm a}\times N}$ is obtained in this case by picking up the first N columns of a $(TX_{\rm a})$-point IDFT matrix which is properly normalized such that $\mathrm{trace}(\mathbf{F}_m^H\,\mathbf{F}_m) = 1$.

4) Antenna selection: instead of transmitting via its $TX_{\rm a}$ local antennae a random linear combination of the data streams $\{\widetilde{u}_\ell[n]\}_{\ell=0}^{L-1}$ with random coefficients $\{p_{i,\alpha,\ell}\}_{\alpha=0}^{TX_{\rm a}-1}$, each BS chooses randomly and independently to transmit over $TX_{\rm a}$ out of the L virtual antennae. Assume that $L/TX_{\rm a}$ is an integer number χ and, moreover, let $\overline{\ell}_0, \overline{\ell}_1, \ldots, \overline{\ell}_{TX_{\rm c}-1}$ be numbers chosen randomly and independently from $\{0,1,\ldots,\chi-1\}$, random antenna selection is accomplished by setting

$$\mathbf{P}_i = \sqrt{\chi}\,[\overbrace{\mathbf{O}_{TX_{\rm a}\times TX_{\rm a}},\ldots,\mathbf{O}_{TX_{\rm a}\times TX_{\rm a}}}^{\overline{\ell}_i},\mathbf{I}_{TX_{\rm a}},\mathbf{O}_{TX_{\rm a}\times TX_{\rm a}},\ldots,\mathbf{O}_{TX_{\rm a}\times TX_{\rm a}}] \in \mathbb{R}^{TX_{\rm a}\times L}, \quad (42)$$

for $i \in \{0,1,\ldots,TX_{\rm c}-1\}$, the factor $\sqrt{\chi}$ ensures that $\mathrm{E}[\|\mathbf{P}_i\,\mathbf{F}_m\,\mathbf{s}_m\|^2] = 1$. Note that, differently from the previous randomization rules, the matrix $\mathbf{P}$ is drawn from a discrete distribution.

For all the coding schemes using L as encoding parameter, the channel-independent precoding matrix $\mathbf{F}_m \in \mathbb{C}^{L\times N}$ is obtained by picking up the first N columns of a L-point IDFT matrix, which is properly normalized such that $\mathrm{trace}(\mathbf{F}_m^H\,\mathbf{F}_m) = 1$. Irrespective of the antenna arrangements, each BS employs $L = 28$ sought antennae for the proposed randomized coding rule.[13] Moreover, we additionally report the performance of the centralized MIMO-OFDM scheme, for which $L = TX$, $\mathbf{P} = \sqrt{TX_{\rm c}}\,\mathbf{I}_{TX}$ and the channel-independent precoding matrix $\mathbf{F}_m \in \mathbb{C}^{TX\times N}$ is obtained by picking up the first N columns of a TX-point IDFT matrix, which is properly normalized such that $\mathrm{trace}(\mathbf{F}_m^H\,\mathbf{F}_m) = 1$; we also report the performance of the noncooperative system, for which $L = TX_{\rm a}$, $TX_{\rm c} = 1$, $\mathbf{P} = \mathbf{P}_0 = \mathbf{I}_{TX_{\rm a}}$ and the channel-independent precoding matrix $\mathbf{F}_m \in \mathbb{C}^{TX_{\rm a}\times N}$ is obtained by picking up the first N columns of a $(TX_{\rm a})$-point IDFT matrix, which is properly normalized such that $\mathrm{trace}(\mathbf{F}_m^H\,\mathbf{F}_m) = 1$. With reference to the R-MIMO-OFDM coding rule with Gaussian randomization and the centralized MIMO-OFDM scheme, we compare the simulation results with their corresponding analytical upper bounds given in Theorem 4.3 and Corollary 4.2, respectively.

As performance measure, we resort to the ABER at the output of the minimum distance detector on the generic subcarrier m which, if a Gray code is used to map the information bits into QPSK symbols and the SNR is sufficiently high, can be defined as $\mathrm{ABER}_m \triangleq \frac{1}{2N}\sum_{n=0}^{N-1} P_{m,n}(e)$, where $P_{m,n}(e)$ is the SEP in the detection of the nth entry of symbol vector $\mathbf{s}_m$ transmitted on subcarrier m. In each example, we carry out 10^6 independent Monte Carlo runs, with each run employing a different set of channel and randomization matrices.

A. ABER versus SNR

Figs. 4, 5, 6 and 7 report the ABER as a function of γ for both the ZF and MMSE cases, with reference to the $(1,2,2)$ and $(4,2,2)$ configurations. For the centralized scheme and the Gaussian randomization, there is a good agreement between simulation and analytical results for moderate-to-high values of the SNR, thus demonstrating the tightness in the high SNR regime of the upper bounds given in Corollary 4.2 and Theorem 4.3. Remarkably, the upper bounds (23) and (28) derived for the ZF receivers almost perfectly match the corresponding simulated ABER curves.

Looking in detail at Figs. 4 and 5, we point out that, in the $(1,2,2)$ scenario where the number of local transmit antennae is smaller than the number of receive antennae, the performance of the noncooperative system and delay randomization are not reported since such approaches cannot reliably transmit two QAM symbols per subcarrier: indeed, when $TX_{\rm a} = 1$ and $RX = 2$, both methods can reliably transmit at least one QAM symbol per subcarrier, since $N \leq \min\{TX_{\rm a}, RX\} = TX_{\rm a} = 1$. On the contrary, R-MIMO-OFDM transceivers do not suffer of this limitation. In fact, it can be observed from Figs. 4 and 5 that, except for the antenna selection method, all the randomization rules exhibit satisfactory performance, by paying an acceptable price in term of coding gain compared to the centralized scheme. In particular, it is worth noting that the Haar randomization outperforms both the uniform phase and Gaussian randomization schemes, which in turn perform comparably. We underline that the centralized scheme requires *a priori* coordination among the cooperating BSs and, moreover, identification at the MS of the cooperating BSs. On the contrary, such coordination and identification are unnecessary for all the proposed randomized coding rules.

Figs. 6 and 7 refer to the $(4,2,2)$ system configuration. In such a figure, we also report the performance of the

[11] If $\tau_{i,2} = 0$ for each BS, the delay randomization scheme boils down to the repetition coding rule (see Remark 5).

[12] The value of $\Delta_{\max}$ is limited to be no larger than the difference between the CP length and the allowed channel order. For highly time-dispersive channels, such a difference cannot expand excessively, since it could lead to a significant reduction of the transmission data rate.

[13] As shown by our analysis, the performance of the proposed R-MIMO-OFDM scheme with complex Gaussian randomization does not critically depend on the choice of L, provided that $L \geq N$. Numerical results not reported here show that such a property also holds for the other randomization rules relying on continuous distributions (uniform phase, Haar and delay randomizations). On the contrary, as intuitively expected, for a fixed $TX_{\rm a}$, the performance of the antenna selection method degrades as L increases.

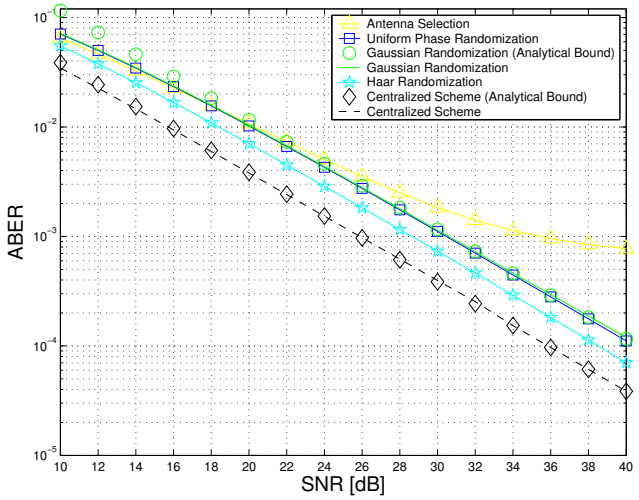

Fig. 4. ZF equalization: ABER versus γ for the scenario $(1, 2, 2)$.

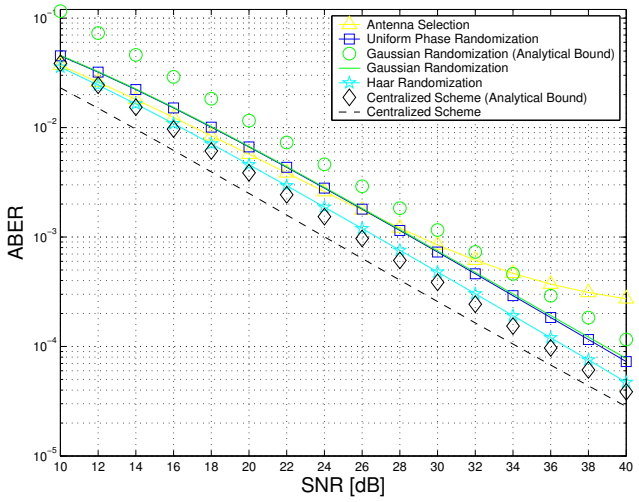

Fig. 5. MMSE equalization: ABER versus γ for the scenario $(1, 2, 2)$.

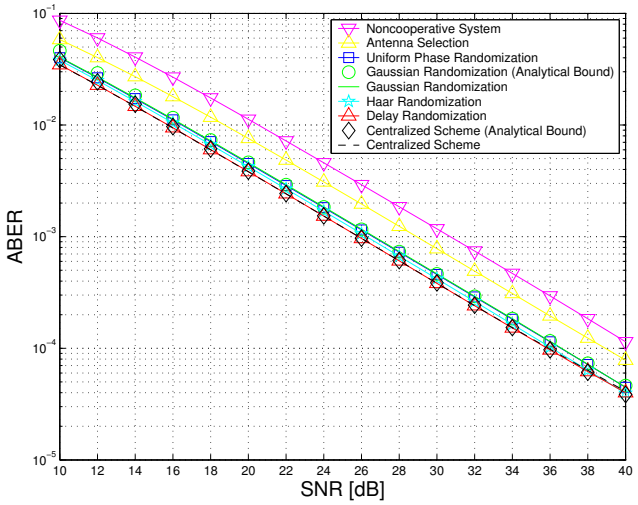

Fig. 6. ZF equalization: ABER versus γ for the scenario $(4, 2, 2)$.

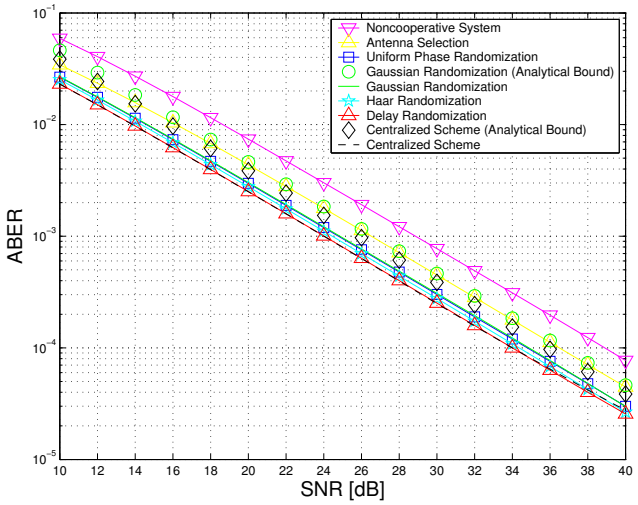

Fig. 7. MMSE equalization: ABER versus γ for the scenario $(4, 2, 2)$.

noncooperative system and delay randomization since, where the number of local transmit antennae is greater than the number of receive antennae, these schemes can be employed to transmit two QAM symbols per subcarrier, indeed, $N \leq \min\{TX_{\rm a}, RX\} = RX = 2$. From our performance analysis [see the discussion in Subsection IV-C and, in particular, eq. (34) for the ZF case], when RX/TX approaches zero, the coding gain performance of Gaussian randomization becomes close to that of the centralized scheme. As we can see, this statement is supported by the results in Figs. 6 and 7, which show that, when the number of transmit antennae increases from one to four, i.e., the ratio RX/TX decreases, the curves of Gaussian, uniform phase and Haar randomization rules closely approach that of the centralized scheme, thereby assuring a significant performance advantage with respect to the noncooperative system. Additionally, we observe that, in this scenario, delay randomization works very well, by slightly outperforming at high SNR all the other randomization rules and exhibiting essentially the same performance of the centralized scheme. This result is not surprising since delay randomization can be seen as a repetition coding rule (see Remark 4), which additionally takes advantage of the further diversity obtained by artificially delaying at each BS the transmitted signals. As pointed out in Remark 4, without introducing delay diversity, the simple repetition strategy can achieve the same worst-case performance of the centralized scheme under the same number N of QAM symbols transmitted per subcarrier, when $TX_{\rm a} \geq RX$. However, it is noteworthy that, similarly to the repetition coding rule, when $RX > TX_{\rm a}$, this scheme can reliably transmit a smaller number of QAM symbols per subcarrier, compared to the centralized scheme and the randomization approaches. Furthermore, the price to pay for introducing random delays at the transmitter and, at the same time, obtaining perfect IBI cancellation after the CP removal at the receiver is the inevitable increase in the CP length, which leads to an inherent reduction of the transmission data rate. In truth, the delay randomization can be also employed in the centralized scheme, as well as it can be used in conjunction with other forms of randomization, e.g., Gaussian, uniform phase or Haar randomization, such that to further improve the

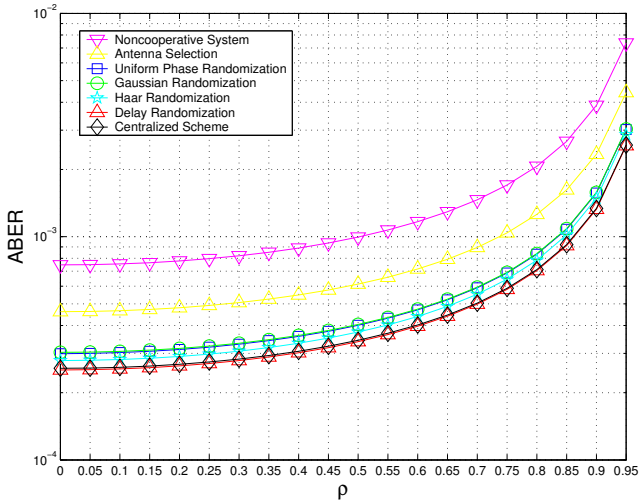

Fig. 8. ZF equalization: ABER versus ϱ for the scenario $(4, 2, 2)$.

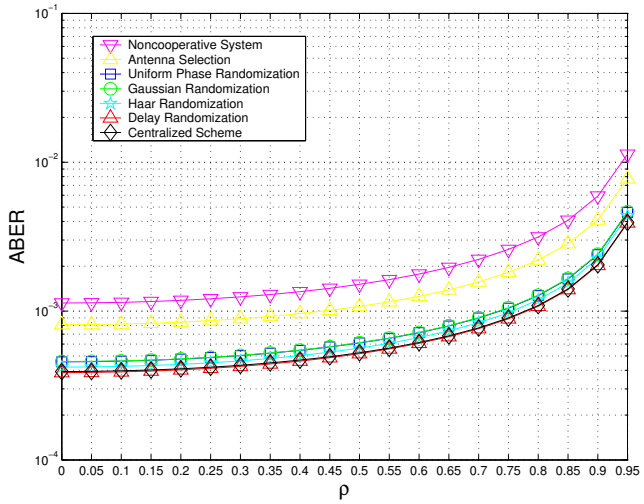

Fig. 9. MMSE equalization: ABER versus ϱ for the scenario $(4, 2, 2)$.

performance of the corresponding transceivers.

B. ABER versus degree of fading correlation at the receiver

In this example, we refer to the same simulation setting described in Subsection V-A, by considering in particular the $(4, 2, 2)$ system configuration. However, differently from the previous example where the channels are generated according to assumption (a6), the matrix $\boldsymbol{\mathcal{C}}_{m,i}$ is herein modeled as

$$\boldsymbol{\mathcal{C}}_{m,i} = \sqrt{\rho_{m,i}}\, \boldsymbol{\mathcal{R}}^{1/2}\, \boldsymbol{\mathcal{N}}_{m,i}\,, \quad \text{where} \quad \boldsymbol{\mathcal{R}} = \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix} \tag{43}$$

is the nonsingular correlation matrix at the MS, with the degree of receive correlation measured by $0 \leq \varrho < 1$.[14]

It can be argued from Figs. 8 and 9 that receive correlation does not affect the relative performances of the considered randomized MIMO-OFDM coding rule, by substantially corroborating the curve arrangement of Figs. 8 and 9 obtained in the case of independent path gains between different antenna pairs. Moreover, the performances of all the methods under comparison gracefully degrade as ϱ increases.

VI. Conclusions

We proposed a novel technology for the PHY layer design of a cooperative downlink multicell network, which is based on the idea that the cooperating BSs transmit an independent random linear transformation of the same MIMO-OFDM symbol. Specifically, we developed a detailed mathematical model of the PHY layer employing randomized MIMO-OFDM coding at the transmitter and low-complexity linear ZF and MMSE equalization at the receiver. Our design accounts for various impairments of the wireless channel and is relatively simple since not only it does not require CSI at each BS, but it also lets the MS reliably demodulate the desired packet by completely ignoring both the number and the identities of the cooperating BSs. Moreover, we investigated the diversity and coding gain achievable for the proposed transceiver, by deriving analytical upper bounds on the SEP when the randomization matrices are either fixed during the transmission of a given packet or randomly varying in accordance with a complex Gaussian distribution. Such results allowed us to gain insight on the performance gain provided by cooperation in terms of coverage expansion, even assuming a very simple transceiver architecture. Finally, besides considering the Gaussian model, we verified through Monte Carlo simulations the effectiveness of our PHY layer proposal for other different statistical models of the randomization matrices.

[14] Also in this case, the squared Frobenius norm of $\boldsymbol{\mathcal{C}}_{m,i}$ is given by $\mathrm{E}[\mathrm{trace}(\boldsymbol{\mathcal{C}}_{m,i}^H\, \boldsymbol{\mathcal{C}}_{m,i})] = \mathrm{E}[\mathrm{trace}(\boldsymbol{\mathcal{C}}_{m,i}\, \boldsymbol{\mathcal{C}}_{m,i}^H)] = \rho_{m,i}\, TX_{\mathrm{a}}\, RX$.

Appendix A
Proof of Theorem 4.1

Let $\mathrm{SINR}_{m,n}$ denote the output SINR of either the ZF equalizer or the MMSE one and, for notational simplicity, set $b \triangleq 2\left(1 - \frac{1}{\sqrt{Q}}\right)$ and $u \triangleq \frac{3}{2(Q-1)}$. After a change of variables from rectangular to polar coordinates in the integral defining the complementary error function, one readily has

$$\begin{aligned} P_{m,n}(e\,|\,\mathbf{P}) &= \mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P}}\left[P_{m,n}(e\,|\,\boldsymbol{\mathcal{E}}_m)\right] = \\ &= \frac{2\,b}{\pi}\int_0^{\pi/2} \mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P}}\left[\exp\left(-\frac{u}{\sin^2\theta}\,\mathrm{SINR}_{m,n}\right)\right]\mathrm{d}\theta \\ &= \frac{2\,b}{\pi}\int_0^{\pi/2} \Phi_{\mathrm{SINR}_{m,n}}\left(-\frac{u}{\sin^2\theta}\right)\mathrm{d}\theta\,, \end{aligned} \tag{44}$$

where $\Phi_{\mathrm{SINR}_{m,n}}(s) \triangleq \mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P}}\left[\exp\left(s\,\mathrm{SINR}_{m,n}\right)\right]$ is the moment generating function (MGF) of the random variable $\mathrm{SINR}_{m,n}$, conditional on $\mathbf{P}$.

A. Preliminaries

Before considering the ZF and MMSE cases separately, it is useful to introduce some notations. Let $\boldsymbol{\mathcal{H}}_m^{(-n)} \in \mathbb{C}^{RX\times(N-1)}$ denote the matrix $\boldsymbol{\mathcal{H}}_m$ deprived of its nth column $\mathbf{h}_{m,n}$. Under assumption (a3), one has $\mathrm{rank}[\boldsymbol{\mathcal{H}}_m^{(-n)}] = N - 1$ with probability one. Hence, the economic-size singular value decomposition $\boldsymbol{\mathcal{H}}_m^{(-n)} = \mathbf{W}_{m,n}\,\mathbf{D}_{m,n}\,\mathbf{V}_{m,n}^H$ holds, where the

semi-unitary matrix $\mathbf{W}_{m,n} \in \mathbb{C}^{RX\times(N-1)}$ and the unitary matrix $\mathbf{V}_{m,n} \in \mathbb{C}^{(N-1)\times(N-1)}$ collect the left and right singular vectors, respectively, associated with the nonzero singular values $d_{m,n,0}, d_{m,n,1}, \ldots, d_{m,n,N-2}$ of $\boldsymbol{\mathcal{H}}_m^{(-n)}$, and $\mathbf{D}_{m,n} \triangleq \mathrm{diag}(d_{m,n,0}, d_{m,n,1}, \ldots, d_{m,n,N-2}) \in \mathbb{R}^{(N-1)\times(N-1)}$. Define the Hermitian matrix $\mathbf{Z}_m \triangleq \mathbf{F}_m^H \mathbf{P}^H (\boldsymbol{\Omega}_m^2 \otimes \mathbf{I}_{TX_a}) \mathbf{P} \mathbf{F}_m \in \mathbb{C}^{N\times N}$, which, under assumption (a3), is positive definite with probability one. We denote with $\mathbf{Z}_m^{(-n,-n)} \in \mathbb{C}^{(N-1)\times(N-1)}$ the matrix $\mathbf{Z}_m$ deprived of its nth column and nth row, whereas $\mathbf{z}_m^{(-n)} \in \mathbb{C}^{N-1}$ represents the nth column $\mathbf{z}_m \in \mathbb{C}^N$ of $\mathbf{Z}_m$ with its nth entry removed. It is seen that $\mathbf{Z}_m^{(-n,-n)} = [\mathbf{F}_m^{(-n)}]^H \mathbf{P}^H (\boldsymbol{\Omega}_m^2 \otimes \mathbf{I}_{TX_a}) \mathbf{P} \mathbf{F}_m^{(-n)}$, where $\mathbf{F}_m^{(-n)} \in \mathbb{C}^{L\times(N-1)}$ denotes the matrix $\mathbf{F}_m$ deprived of its nth column. We also define the eigenvalue decomposition of $\mathbf{Z}_m^{(-n,-n)}$ as $\mathbf{Z}_m^{(-n,-n)} = \mathbf{L}_{m,n} \boldsymbol{\Lambda}_{m,n} \mathbf{L}_{m,n}^H$, where $\mathbf{L}_{m,n} \in \mathbb{C}^{(N-1)\times(N-1)}$ is the unitary matrix of the eigenvectors and $\boldsymbol{\Lambda}_{m,n} \in \mathbb{R}^{(N-1)\times(N-1)}$ is the diagonal matrix of the eigenvalues. Under assumption (a6), the rows of $\boldsymbol{\mathcal{H}}_m$ are i.i.d. with each row $\sim \mathcal{CN}(\mathbf{0}_N, \mathbf{Z}_m)$; consequently, the rows of $\boldsymbol{\mathcal{H}}_m^{(-n)}$ are i.i.d. with each row $\sim \mathcal{CN}(\mathbf{0}_{N-1}, \mathbf{Z}_m^{(-n,-n)})$.

B. ZF equalization

Let us start from the ZF case, for which $\mathrm{SINR}_{m,n} = \mathrm{SINR}_{m,n}^{\mathrm{zf}}$ is given by (13). Following [37], it can be shown that $\mathrm{SINR}_{m,n}^{\mathrm{zf}} = \gamma \|\mathbf{a}_{m,n}\|^2$, with $\mathbf{a}_{m,n} \triangleq \boldsymbol{\mathcal{O}}_{m,n}^H \mathbf{h}_{m,n} \in \mathbb{C}^{RX-N+1}$, where $\boldsymbol{\mathcal{O}}_{m,n} \in \mathbb{C}^{RX\times(RX-N+1)}$ is the orthogonal complement of $\mathbf{W}_{m,n}$, i.e., $\mathbf{W}_{m,n}^H \boldsymbol{\mathcal{O}}_{m,n} = \mathbf{O}_{(N-1)\times(RX-N+1)}$ and $(\mathbf{W}_{m,n}, \boldsymbol{\mathcal{O}}_{m,n})^H (\mathbf{W}_{m,n}, \boldsymbol{\mathcal{O}}_{m,n}) = \mathbf{I}_{RX}$. Under assumption (a6), conditional on $\mathbf{P}$, it results [37] that $\mathbf{a}_{m,n} \sim \mathcal{CN}(\mathbf{0}_{RX-N+1}, \Sigma_{m,n} \mathbf{I}_{RX-N+1})$, where $\Sigma_{m,n} \triangleq 1/\{\mathbf{Z}_m^{-1}\}_{n,n} > 0$. Consequently, one has $\|\mathbf{a}_{m,n}\|^2 \sim \mathrm{gamma}(RX-N+1, \Sigma_{m,n})$, conditional on $\mathbf{P}$. In its turn, given $\mathbf{P}$, it follows that $\mathrm{SINR}_{m,n}^{\mathrm{zf}} \sim \mathrm{gamma}(RX-N+1, \gamma\Sigma_{m,n})$, and, thus, by remembering the expression of the MGF for a Gamma random variable [32], one obtains

$$\Phi_{\mathrm{SINR}_{m,n}^{\mathrm{zf}}}\left(-\frac{u}{\sin^2\theta}\right) = \frac{1}{\left(1+\frac{u\,\gamma\,\Sigma_{m,n}}{\sin^2\theta}\right)^{RX-N+1}} \leq (\sin^2\theta)^{RX-N+1} (u\,\gamma\,\Sigma_{m,n})^{-(RX-N+1)}. \quad (45)$$

Substituting (45) in (44) and remembering the expressions of b and u, eq. (17) easily follows.

C. MMSE equalization

Let us now consider the more cumbersome MMSE case, for which $\mathrm{SINR}_{m,n} = \mathrm{SINR}_{m,n}^{\mathrm{mmse}}$ is given by (15). Following again [37], one has $\mathrm{SINR}_{m,n}^{\mathrm{mmse}} = \mathrm{SINR}_{m,n}^{\mathrm{zf}} + \gamma\,\boldsymbol{\beta}_{m,n}^H (\mathbf{I}_{N-1} + \gamma\,\mathbf{D}_{m,n}^2)^{-1} \boldsymbol{\beta}_{m,n}$, where $\boldsymbol{\beta}_{m,n} \triangleq (\beta_{m,n,0}, \beta_{m,n,1}, \ldots, \beta_{m,n,N-2})^T = \mathbf{W}_{m,n}^H \mathbf{h}_{m,n} \in \mathbb{C}^{N-1}$. If assumption (a6) holds, $\mathrm{SINR}_{m,n}^{\mathrm{zf}} \sim \mathrm{gamma}(RX-N+1, \gamma\Sigma_{m,n})$, given $\mathbf{P}$, and is independent of the random variable $T_{m,n} \triangleq \gamma\,\boldsymbol{\beta}_{m,n}^H (\mathbf{I}_{N-1} + \gamma\,\mathbf{D}_{m,n}^2)^{-1} \boldsymbol{\beta}_{m,n}$. Since $\mathrm{SINR}_{m,n}^{\mathrm{zf}}$ and $T_{m,n}$ are independent, one obtains

$$\Phi_{\mathrm{SINR}_{m,n}^{\mathrm{mmse}}}\left(-\frac{u}{\sin^2\theta}\right) = \Phi_{\mathrm{SINR}_{m,n}^{\mathrm{zf}}}\left(-\frac{u}{\sin^2\theta}\right) \cdot \Phi_{T_{m,n}}\left(-\frac{u}{\sin^2\theta}\right) \leq (\sin^2\theta)^{RX-N+1} \cdot (u\,\gamma\,\Sigma_{m,n})^{-(RX-N+1)}\,\Phi_{T_{m,n}}\left(-\frac{u}{\sin^2\theta}\right), \quad (46)$$

where we have used (45) and $\Phi_{T_{m,n}}(s) \triangleq \mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P}}[\exp(s\,T_{m,n})]$ is MGF of the random variable $T_{m,n}$, conditional on $\mathbf{P}$, whose evaluation is more challenging than that of $\Phi_{\mathrm{SINR}_{m,n}^{\mathrm{zf}}}(s)$. Therefore, to obtain useful results, we provide an upper bound on $\Phi_{T_{m,n}}(-u/\sin^2\theta)$. To this aim, by writing $T_{m,n}$ as a function of the entries of $\boldsymbol{\beta}_{m,n}$ and the diagonal entries of $\mathbf{D}_{m,n}$, application of the conditional expectation rule [32] yields (47) at the top of the next page. where $\mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}[\cdot]$ denotes the expectation over $\boldsymbol{\mathcal{H}}_m^{(-n)}$ with $\mathbf{P}$ fixed, and $\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P},\boldsymbol{\mathcal{H}}_m^{(-n)}}[\cdot]$ is the expectation over $\boldsymbol{\mathcal{C}}_m$ given both $\mathbf{P}$ and $\boldsymbol{\mathcal{H}}_m^{(-n)}$. Conditional on $\mathbf{P}$ and $\boldsymbol{\mathcal{H}}_m^{(-n)}$, the diagonal entries of $\mathbf{D}_{m,n}$ are deterministic quantities, whereas [37] $\beta_{m,n,i} \sim \mathcal{CN}(d_{m,n,i}\{\mathbf{V}_{m,n}^H (\mathbf{Z}_m^{(-n,-n)})^{-1} \mathbf{z}_m^{(-n)}\}_i, \Sigma_{m,n})$. Consequently, the random variable $|\beta_{m,n,i}|^2$ has a noncentral chi-square distribution [32] with two degrees of freedom and noncentrality parameter $d_{m,n,i}^2 \left|\{\mathbf{V}_{m,n}^H (\mathbf{Z}_m^{(-n,-n)})^{-1} \mathbf{z}_m^{(-n)}\}_i\right|^2$. Hence,

$$\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P},\boldsymbol{\mathcal{H}}_m^{(-n)}}\left[\exp\left(-\frac{u\,\gamma}{\sin^2\theta}\cdot\frac{|\beta_{m,n,i}|^2}{1+\gamma\,d_{m,n,i}^2}\right)\right] = \frac{1}{1+\frac{u\,\gamma\,\Sigma_{m,n}}{\sin^2\theta\,(1+\gamma\,d_{m,n,i}^2)}} \cdot \exp\left(-\frac{d_{m,n,i}^2\left|\{\mathbf{V}_{m,n}^H (\mathbf{Z}_m^{(-n,-n)})^{-1}\mathbf{z}_m^{(-n)}\}_i\right|^2}{\Sigma_{m,n}+\frac{\sin^2\theta\,(1+\gamma\,d_{m,n,i}^2)}{u\,\gamma}}\right) \leq \frac{1}{1+\frac{u\,\gamma\,\Sigma_{m,n}}{\sin^2\theta\,(1+\gamma\,d_{m,n,i}^2)}}, \quad (48)$$

where we have observed that $\exp(-a\,x) \leq 1$, for any $x \geq 0$ and $a > 0$. For moderate-to-high values of the SNR, i.e., when $\sigma_v^2 \ll \max_{i\in\{0,1,\ldots,N-2\}} d_{m,n,i}^2$, one has $\gamma\,d_{m,n,i}^2 \gg 1$ and (48) becomes[15]

$$\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P},\boldsymbol{\mathcal{H}}_m^{(-n)}}\left[\exp\left(-\frac{u\,\gamma}{\sin^2\theta}\cdot\frac{|\beta_{m,n,i}|^2}{1+\gamma\,d_{m,n,i}^2}\right)\right] \lesssim \frac{d_{m,n,i}^2}{d_{m,n,i}^2+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}}, \quad (49)$$

which can be substituted in (47), thus yielding (50) as shown at the top of the next page (remember that $\boldsymbol{\mathcal{H}}_m^{(-n)} = \mathbf{W}_{m,n}\,\mathbf{D}_{m,n}\,\mathbf{V}_{m,n}^H$). where $\boldsymbol{\mathcal{K}}_{m,n} \triangleq \boldsymbol{\mathcal{H}}_m^{(-n)}\,\mathbf{L}_{m,n}\,\boldsymbol{\Lambda}_{m,n}^{-1/2} \in \mathbb{C}^{RX\times(N-1)}$ and we have used the determinant properties.[16]

[15]Note that $d_{m,n,0}^2, d_{m,n,1}^2, \ldots, d_{m,n,N-2}^2$ are the nonzero eigenvalues of the Hermitian matrix $(\boldsymbol{\mathcal{H}}_m^{(-n)})^H \boldsymbol{\mathcal{H}}_m^{(-n)}$.

[16]Given $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times n}$ and $s \in \mathbb{C}$, then $\det(\mathbf{A}\,\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(s\,\mathbf{A}) = s^n \det(\mathbf{A})$ [38].

$$\Phi_{T_{m,n}}\left(-\frac{u}{\sin^2\theta}\right) = \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P},\boldsymbol{\mathcal{H}}_m^{(-n)}}\left[\exp\left(-\frac{u\,\gamma}{\sin^2\theta}\sum_{i=0}^{N-2}\frac{|\beta_{m,n,i}|^2}{1+\gamma\,d_{m,n,i}^2}\right)\right]\right\} \cdot \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\prod_{i=0}^{N-2}\mathrm{E}_{\boldsymbol{\mathcal{C}}_m|\mathbf{P},\boldsymbol{\mathcal{H}}_m^{(-n)}}\left[\exp\left(-\frac{u\,\gamma}{\sin^2\theta}\cdot\frac{|\beta_{m,n,i}|^2}{1+\gamma\,d_{m,n,i}^2}\right)\right]\right\}, \quad (47)$$

$$\begin{aligned}\Phi_{T_{m,n}}\left(-\frac{u}{\sin^2\theta}\right) &\lesssim \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\prod_{i=0}^{N-2}\frac{d_{m,n,i}^2}{d_{m,n,i}^2+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}}\right\} = \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\frac{\det(\mathbf{D}_{m,n}^2)}{\det\left(\mathbf{D}_{m,n}^2+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\mathbf{I}_{N-1}\right)}\right\}\\ &= \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\frac{\det[\mathbf{V}_{m,n}^H\,(\boldsymbol{\mathcal{H}}_m^{(-n)})^H\,\boldsymbol{\mathcal{H}}_m^{(-n)}\,\mathbf{V}_{m,n}]}{\det\left\{\mathbf{V}_{m,n}^H\left[(\boldsymbol{\mathcal{H}}_m^{(-n)})^H\,\boldsymbol{\mathcal{H}}_m^{(-n)}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\mathbf{I}_{N-1}\right]\mathbf{V}_{m,n}\right\}}\right\}\\ &= \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\frac{\det[(\boldsymbol{\mathcal{H}}_m^{(-n)})^H\,\boldsymbol{\mathcal{H}}_m^{(-n)}]}{\det\left[(\boldsymbol{\mathcal{H}}_m^{(-n)})^H\,\boldsymbol{\mathcal{H}}_m^{(-n)}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\mathbf{I}_{N-1}\right]}\right\}\\ &= \mathrm{E}_{\boldsymbol{\mathcal{H}}_m^{(-n)}|\mathbf{P}}\left\{\frac{\det[\mathbf{L}_{m,n}\,\boldsymbol{\Lambda}_{m,n}^{1/2}\,\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}\,\boldsymbol{\Lambda}_{m,n}^{1/2}\,\mathbf{L}_{m,n}^H]}{\det\left[\mathbf{L}_{m,n}\,\boldsymbol{\Lambda}_{m,n}^{1/2}\,\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}\,\boldsymbol{\Lambda}_{m,n}^{1/2}\,\mathbf{L}_{m,n}^H+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\mathbf{I}_{N-1}\right]}\right\}\\ &= \mathrm{E}_{\boldsymbol{\mathcal{K}}_{m,n}|\mathbf{P}}\left\{\frac{\det(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})}{\det\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right]}\right\}, \end{aligned} \quad (50)$$

Observe that the entries of $\boldsymbol{\mathcal{K}}_{m,n}$ are i.i.d. circularly symmetric complex Gaussian random variables with zero mean and unit variance. To further upper bound $\Phi_{T_{m,n}}(-u/\sin^2\theta)$, we use the lemma:

Lemma A.1: The determinant of the matrix $\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}$ can be lower bounded as follows

$$\begin{aligned}&\det\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right]\\ &\geq \det\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{N\,u\,\Sigma_{m,n}}{\rho_{m,\max}\,\sin^2\theta}\,\lambda_{\max}^{-1}(\mathbf{P}^H\,\mathbf{P})\,\mathbf{I}_{N-1}\right], \end{aligned} \quad (51)$$

where $\rho_{m,\max}\triangleq\max_{i\in\{0,1,\ldots,TX_c-1\}}\rho_{m,i}$ is the path loss associated with the BS that is farthest from the MS, and $\lambda_{\max}(\mathbf{P}^H\mathbf{P})>0$ is the largest eigenvalue of the Hermitian matrix $\mathbf{P}^H\,\mathbf{P}\in\mathbb{C}^{L\times L}$.

Proof: Let $\lambda_0(\mathbf{A})\leq\lambda_1(\mathbf{A})\leq\cdots\leq\lambda_{n-1}(\mathbf{A})$ denote the eigenvalues of any matrix $\mathbf{A}\in\mathbb{C}^{n\times n}$, arranged in increasing order. By virtue of the Weyl's inequality [38], one has

$$\begin{aligned}\lambda_i\left(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right) &\geq \lambda_i(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})\\ &\quad+\lambda_0\left(\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right), \end{aligned} \quad (52)$$

for $i\in\{0,1,\ldots,N-2\}$, where it is easily seen that

$$\lambda_0\left(\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right)=\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\lambda_{N-2}^{-1}[\mathbf{Z}_m^{(-n,-n)}]\,. \quad (53)$$

Furthermore, remembering that $\mathbf{Z}_m^{(-n,-n)}=[\mathbf{F}_m^{(-n)}]^H\,\mathbf{P}^H\,(\boldsymbol{\Omega}_m^2\otimes\mathbf{I}_{TX_a})\,\mathbf{P}\,\mathbf{F}_m^{(-n)}$ and observing that, for arbitrary matrices $\mathbf{A}\in\mathbb{C}^{n\times m}$ and $\mathbf{B}\in\mathbb{C}^{m\times n}$, the nonzero characteristic roots of $\mathbf{A}\,\mathbf{B}$ are the nonzero characteristic roots of $\mathbf{B}\,\mathbf{A}$ [38], we get

$$\begin{aligned}\lambda_{N-2}[\mathbf{Z}_m^{(-n,-n)}] &= \lambda_{N-2}\{(\boldsymbol{\Omega}_m\otimes\mathbf{I}_{TX_a})\,\mathbf{P}\,\mathbf{F}_m^{(-n)}\\ &\quad\cdot[\mathbf{F}_m^{(-n)}]^H\,\mathbf{P}^H\,(\boldsymbol{\Omega}_m\otimes\mathbf{I}_{TX_a})\}\\ &\leq \rho_{m,\max}\,\lambda_{N-2}[\mathbf{P}\,\mathbf{F}_m^{(-n)}\,\{\mathbf{F}_m^{(-n)}]^H\,\mathbf{P}^H\}\\ &= \rho_{m,\max}\,\lambda_{N-2}\{[\mathbf{F}_m^{(-n)}]^H\,\mathbf{P}^H\mathbf{P}\,\mathbf{F}_m^{(-n)}\}\\ &\leq \frac{\rho_{m,\max}}{N}\,\lambda_{N-2}(\mathbf{P}^H\mathbf{P})\,, \end{aligned} \quad (54)$$

where the first inequality comes from the Ostrowski's theorem [38], whereas the second inequality follows [38] by noticing that $[\mathbf{F}_m^{(-n)}]^H\,\mathbf{F}_m^{(-n)}=\frac{1}{N}\,\mathbf{I}_{N-1}$. Observe that, as a consequence of assumption (a3), the eigenvalues of $[\mathbf{F}_m^{(-n)}]^H\,\mathbf{P}^H\mathbf{P}\,\mathbf{F}_m^{(-n)}$ and, thus, those of $\mathbf{P}^H\mathbf{P}$, are strictly positive with probability one. Substituting (54) in (53) and the result back in (52), one obtains

$$\begin{aligned}&\lambda_i\left(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right)\\ &\quad\geq\lambda_i(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})+\frac{N\,u\,\Sigma_{m,n}}{\rho_{m,\max}\,\sin^2\theta}\,\lambda_{N-2}^{-1}(\mathbf{P}^H\mathbf{P})\\ &\quad=\lambda_i\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{N\,u\,\Sigma_{m,n}}{\rho_{m,\max}\sin^2\theta}\,\lambda_{N-2}^{-1}(\mathbf{P}^H\mathbf{P})\,\mathbf{I}_{N-1}\right], \end{aligned} \quad (55)$$

for $i \in \{0,1,\ldots,N-2\}$. Consequently, inequality (51) follows by observing that

$$\begin{aligned}
&\det\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right]\\
&\qquad=\prod_{i=0}^{N-2}\lambda_i\left(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{u\,\Sigma_{m,n}}{\sin^2\theta}\,\boldsymbol{\Lambda}_{m,n}^{-1}\right)\\
&\geq\prod_{i=0}^{N-2}\lambda_i\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{N\,u\,\Sigma_{m,n}}{\rho_{m,\max}\,\sin^2\theta}\,\lambda_{N-2}^{-1}(\mathbf{P}^H\mathbf{P})\,\mathbf{I}_{N-1}\right]\\
&=\det\left[\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n}+\frac{N\,u\,\Sigma_{m,n}}{\rho_{m,\max}\,\sin^2\theta}\,\lambda_{N-2}^{-1}(\mathbf{P}^H\mathbf{P})\,\mathbf{I}_{N-1}\right].
\end{aligned}\tag{56}$$

■

At this point, substituting (51) in (50) and using again the determinant properties, we get

$$\begin{aligned}
&\Phi_{T_{m,n}}\left(-\frac{u}{\sin^2\theta}\right)\\
&\quad\lesssim\zeta^{N-1}\,\mathrm{E}_{\boldsymbol{\mathcal{K}}_{m,n}|\mathbf{P}}\left\{\frac{\det(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})}{\det(\mathbf{I}_{N-1}+\zeta\,\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})}\right\},
\end{aligned}\tag{57}$$

where we have set $\zeta\triangleq\frac{\rho_{m,\max}\,\lambda_{\max}(\mathbf{P}^H\mathbf{P})\,\sin^2\theta}{N\,u\,\Sigma_{m,n}}$. For moderate-to-high values of the SNR, the expectation in (57) can be well-approximated as (see [39])

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{\mathcal{K}}_{m,n}|\mathbf{P}}\left\{\frac{\det(\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})}{\det(\mathbf{I}_{N-1}+\zeta\,\boldsymbol{\mathcal{K}}_{m,n}^H\,\boldsymbol{\mathcal{K}}_{m,n})}\right\}\\
&\qquad\approx\binom{RX}{N-1}(N-1)\overbrace{\int_0^{+\infty}\frac{t^{N-2}\,\exp(-t)}{(1+\zeta\,t)^{RX+1}}\,\mathrm{d}t}^{\mathcal{I}}.
\end{aligned}\tag{58}$$

After lengthy calculations, the integral $\mathcal{I}$ in (58) can be evaluated in closed form as follows

$$\begin{aligned}
\mathcal{I}=&\sum_{i_1=0}^{N-2}\sum_{i_2=1}^{RX-i_1}\binom{N-2}{i_1}\frac{(-1)^{i_2+N-RX-2}\,(i_2-1)!}{(RX-i_1)!}\\
&\cdot\zeta^{i_1+i_2-RX-N+1}+(-1)^{N-RX-2}\,\frac{\exp(1/\zeta)\;E_1(1/\zeta)}{\zeta^{N-1}}\\
&\cdot\sum_{i_1=0}^{N-2}\binom{N-2}{i_1}\frac{1}{\zeta^{RX-i_1}\,(RX-i_1)!}.
\end{aligned}\tag{59}$$

where $E_1(x)=\int_x^{+\infty}\frac{\exp(-t)}{t}\,\mathrm{d}t$ defines the exponential integral function [40]. The upper bound (18) directly follows from (59), (58), (57) and (46), by substituting the expressions of b, u and ζ, and, additionally, using the fact [40] that $E_1(x)<\exp(-x)\,\ln(1+\frac{1}{x})\leq\frac{1}{x}\,\exp(-x)$, for $x>0$.

APPENDIX B
PROOF OF THEOREM 4.3

As done in Appendix A, we set $b\triangleq 2\left(1-\frac{1}{\sqrt{Q}}\right)$ and $u\triangleq\frac{3}{2(Q-1)}$ for notational convenience.

A. ZF equalization

Let us start from the ZF case, by considering the upper bound in (17). By averaging both hands of (17) with respect to $\mathbf{P}$, we get

$$\begin{aligned}
P_{m,n}^{\mathrm{zf}}(e)&=\mathrm{E}_{\mathbf{P}}\left[P_{m,n}^{\mathrm{zf}}(e\,|\,\mathbf{P})\right]\\
&\leq\frac{b\,\Theta(RX-N+1)}{(u\,\gamma)^{RX-N+1}}\,\mathrm{E}_{\mathbf{P}}\left[\left(\frac{1}{\Sigma_{m,n}}\right)^{RX-N+1}\right],
\end{aligned}\tag{60}$$

where $\Sigma_{m,n}=1/\{\mathbf{Z}_m^{-1}\}_{n,n}$ and, in the case of $\boldsymbol{\Omega}_m=\sqrt{\rho_m}\,\mathbf{I}_{TX_{\mathrm{c}}}$, one has $\mathbf{Z}_m=\rho_m\,\mathbf{F}_m^H\,\mathbf{P}^H\mathbf{P}\,\mathbf{F}_m$. Under assumption (a7), the rows of the matrix $\mathbf{P}\,\mathbf{F}_m$ are i.i.d. with each row $\sim\mathcal{CN}(\mathbf{0}_N,\frac{1}{NTX_{\mathrm{a}}}\,\mathbf{I}_N)$. Therefore, using the results of [37], it can be proven that $\Sigma_{m,n}\sim\mathrm{gamma}(TX-N+1,\rho_m/(NTX_{\mathrm{a}}))$. Consequently, the probability density function (pdf) of the random variable $Y=1/\Sigma_{m,n}$ is given by

$$\begin{aligned}
f_Y(y)=&\frac{\left(\frac{NTX_{\mathrm{a}}}{\rho_m}\right)^{TX-N+1}}{\Gamma(TX-N+1)}\left(\frac{1}{y}\right)^{TX-N+2}\\
&\cdot\exp\left(-\frac{NTX_{\mathrm{a}}}{\rho_m}\,\frac{1}{y}\right),\quad\text{for } 0<y<+\infty,
\end{aligned}\tag{61}$$

where $\Gamma(\alpha)\triangleq\int_0^{+\infty}t^{\alpha-1}\,\exp(-t)\,\mathrm{d}t$ defines the gamma function ($\alpha>0$). Function (61) is the inverse gamma pdf [32]. Using (61), after some straightforward manipulations, one has

$$\begin{aligned}
\mathrm{E}_{\mathbf{P}}\left[\left(\frac{1}{\Sigma_{m,n}}\right)^{RX-N+1}\right]&=\int_0^{+\infty}y^{RX-N+1}\,f_Y(y)\,\mathrm{d}y\\
&=\frac{\left(\frac{NTX_{\mathrm{a}}}{\rho_m}\right)^{RX-N+1}}{\Gamma(TX-N+1)}\underbrace{\int_0^{+\infty}t^{TX-RX-1}\exp(-t)\,\mathrm{d}t}_{\mathcal{J}}
\end{aligned}\tag{62}$$

Since $TX>RX$, the integral $\mathcal{J}$ in (62) is finite and is equal to $\mathcal{J}=\Gamma(TX-RX)$. The upper bound (28) follows by substituting (62) in (60) and accounting for the expressions of b and u.

B. MMSE equalization

Let us focus attention on the MMSE case, by considering the upper bound in (18) and (19). Since $\rho_{m,\max}=\rho_m$ when $\boldsymbol{\Omega}_m=\sqrt{\rho_m}\,\mathbf{I}_{TX_{\mathrm{c}}}$, by averaging both hands of (18) with respect to $\mathbf{P}$, after some algebraic rearrangements, we have

$$\begin{aligned}
P_{m,n}^{\mathrm{mmse}}(e)=\mathrm{E}_{\mathbf{P}}\left[P_{m,n}^{\mathrm{mmse}}(e\,|\,\mathbf{P})\right]&\lesssim\frac{b\,\Theta(RX-N+1)}{(u\,\gamma)^{RX-N+1}}\\
&\cdot\mathrm{E}_{\mathbf{P}}\left[\frac{\psi_{m,n}(\mathbf{P})}{(\Sigma_{m,n})^{RX-N+1}}\right],
\end{aligned}\tag{63}$$

where

$$\mathrm{E}_{\mathbf{P}}\left[\frac{\psi_{m,n}(\mathbf{P})}{(\Sigma_{m,n})^{RX-N+1}}\right] = \binom{RX}{N-1}\frac{N-1}{\Theta(RX-N+1)} \cdot \left\{\sum_{i_1=0}^{N-2}\sum_{i_2=1}^{RX-i_1}\binom{N-2}{i_1}\frac{(-1)^{i_2+N-RX-2}\,(i_2-1)!}{\left(\frac{N\,u}{\rho_m}\right)^{i_1+i_2-RX}(RX-i_1)!} \cdot\Theta(i_1+i_2-N+1) \cdot\mathrm{E}_{\mathbf{P}}\left[\lambda_{\max}^{i_1+i_2-RX}(\mathbf{P}^H\mathbf{P})\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1+i_2-N+1}\right] + (-1)^{N-RX-2}\sum_{i_1=0}^{N-2}\binom{N-2}{i_1} \cdot\left(\frac{N\,u}{\rho_m}\right)^{RX-i_1-1}\frac{\Theta(i_1-N+2)}{(RX-i_1)!} \cdot\mathrm{E}_{\mathbf{P}}\left[\lambda_{\max}^{i_1-RX+1}(\mathbf{P}^H\mathbf{P})\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1-N+2}\right]\right\}. \quad (64)$$

Evaluations of the two expectations in (64) are complicated by the fact that $\Sigma_{m,n} = 1/\{\mathbf{Z}_m^{-1}\}_{n,n}$, with $\mathbf{Z}_m = \rho_m\,\mathbf{F}_m^H\,\mathbf{P}^H\mathbf{P}\,\mathbf{F}_m$, also depends on $\lambda_{\max}(\mathbf{P}^H\mathbf{P})$. Therefore, to obtain simple and useful approximations, we resort to the following lemma:

Lemma B.1: Under assumption (a7), for $TX \to +\infty$, the largest eigenvalue of the (central) complex Wishart matrix $\mathbf{P}^H\mathbf{P}$ converges almost surely to $TX_{\mathrm{c}}\,(1+\sqrt{\eta})^2$, with $(L/TX) \to \eta \neq 0$.

Proof: This lemma is the complex counterpart of the result given in [41]. ∎

Lemma B.1 states that $\lambda_{\max}(\mathbf{P}^H\mathbf{P})$ converges almost surely to a limiting nonrandom constant. Hence, for sufficiently large values of TX with $L/TX > 0$, we can approximatively write

$$\mathrm{E}_{\mathbf{P}}\left[\lambda_{\max}^{i_1+i_2-RX}(\mathbf{P}^H\mathbf{P})\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1+i_2-N+1}\right] \approx \left[TX_{\mathrm{c}}\left(1+\sqrt{\frac{L}{TX}}\right)^2\right]^{i_1+i_2-RX} \cdot\mathrm{E}_{\mathbf{P}}\left[\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1+i_2-N+1}\right], \quad (65)$$

where expectation in (65) can be calculated by using the inverse gamma pdf (61), thus yielding

$$\mathrm{E}_{\mathbf{P}}\left[\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1+i_2-N+1}\right] = \int_0^{+\infty} y^{i_1+i_2-RX}\,f_Y(y)\,\mathrm{d}y = \frac{\left(\frac{N\,TX_{\mathrm{a}}}{\rho_m}\right)^{i_1+i_2-N+1}}{\Gamma(TX-N+1)}\,\Gamma(TX-i_1-i_2)\,, \quad (66)$$

with $TX - i_1 - i_2 > 0$, for $i_1 \in \{0,1,\ldots,N-2\}$ and $i_2 \in \{1,2,\ldots,RX-i_1\}$, since $TX > RX$. Similarly, the second expectation in (64) can be approximatively upper bounded as

$$\mathrm{E}_{\mathbf{P}}\left[\lambda_{\max}^{i_1-RX+1}(\mathbf{P}^H\mathbf{P})\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1-N+2}\right] \lesssim \left[TX_{\mathrm{c}}\left(1+\sqrt{\frac{L}{TX}}\right)^2\right]^{i_1-RX+1} \cdot\mathrm{E}_{\mathbf{P}}\left[\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1-N+2}\right], \quad (67)$$

where, using again (61), the expectation in (67) assumes the expression

$$\mathrm{E}_{\mathbf{P}}\left[\left(\frac{1}{\Sigma_{m,n}}\right)^{i_1-N+2}\right] = \int_0^{+\infty} y^{i_1-N+2}\,f_Y(y)\,\mathrm{d}y = \frac{\left(\frac{N\,TX_{\mathrm{a}}}{\rho_m}\right)^{i_1-N+2}}{\Gamma(TX-N+1)}\,\Gamma(TX-i_1-1)\,, \quad (68)$$

with $TX - i_1 - 1 > 0$, for $i_1 \in \{0,1,\ldots,N-2\}$, since $TX \geq N$. The upper bound (29) comes from substituting (66) and (68) in (65) and (67), respectively, and the results back in (64) together with the expressions of b and u.

REFERENCES

[1] IEEE 802.16e–2005, "Local and metropolitan networks – Part 16: Air interface for fixed and mobile broadband wireless access systems, Amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands and corrigendum 1," 2006.

[2] J. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, pp. 41–59, 1996.

[3] I.E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[4] L. Zheng and D.N.C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, May 2003.

[5] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications – Where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, pp. 29–48, May 2000.

[6] S. Shamai and B. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. Veh. Technol. Conf.*, Rhodes, Greece, May 2001, pp. 1745–1749.

[7] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *Eurasip J. Wireless Commun. and Networking*, pp. 222–235, Dec. 2004.

[8] S. Jafar, G. Foschini, and A. Goldsmith, "Phantomnet: exploring optimal multicellular multiple antenna systems," *Eurasip J. App. Sig. Proc.*, pp. 591–604, 2004.

[9] N.H. Dawod, I.D. Marsland, and R.H.M. Hafez, "Improved transmit steering for MIMO-OFDM downlinks with distributed base station antenna arrays," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 419–426, Mar. 2006.

[10] M.K. Karakayali, G.J. Foschini, and R.A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun. Mag.*, vol. 3, pp. 56–61, Aug. 2006.

[11] O. Somekh, O. Simeone, Y. Bar-Ness, and M. Haimovich, "Distributed multi-cell zero-forcing beamforming in cellular downlink channels," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, CA, Nov. 2006.

[12] H. Skjevling, D. Gesbert, A. Hjørungnes, "Precoded distributed space-time block codes in cooperative diversity-based downlink," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 4209–4214, Dec. 2007.

[13] H. Zhang, N.B. Mehta, A.F. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 155–165, Jan. 2008.

[14] A. Tölli, M. Codreanu, and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1428–1440, Apr. 2008.
[15] T.M. Cover and A.A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Sept. 1979.
[16] J.N. Laneman and G.W. Wornell, "Distributed space-time block coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.
[17] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity – Part I and II: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.
[18] J.N. Laneman, D. Tse, and G.W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Sept. 2004.
[19] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, Sept. 2005.
[20] H. Wang, X.-G. Xia, and Q. Yin, "Distributed space-frequency codes for cooperative communication systems with multiple carrier frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1045–1055, Feb. 2009.
[21] A. Scaglione and Y.-W. Hong, "Opportunistic large arrays: cooperative transmission in wireless multihop ad hoc networks to reach far distances", *IEEE Trans. Signal Process.*, vol. 51, pp. 2082-2092, Aug. 2003.
[22] S. Wei, D.L. Goeckel, and M.C. Valenti, "Asynchronous cooperative diversity," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1547–1557, June 2006.
[23] H. El Gamal and D. Aktas, "Distributed space-time filtering for cooperative wireless networks," in *Proc. Global Telecommun. Conf.*, San Francisco, CA, Dec. 2003.
[24] B. Sirkeci-Mergen and A. Scaglione, "Randomized space-time coding for distributed cooperative communication", *IEEE Trans. Signal Process.*, vol. 55, pp. 5003-5017, Oct. 2007.
[25] M. Sharp, A. Scaglione and B. Sirkeci-Mergen, "Randomized cooperation in asynchronous dispersive links", *IEEE Trans. Commun.*, vol. 57, pp. 64–68, Jan. 2009.
[26] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
[27] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1158–1168, May 2006.
[28] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels", *IEEE Trans. Inf. Theory*, vol. 45, pp. 1639-1642, July 1999.
[29] A. Scaglione, G. B. Giannakis and S. Barbarossa, "Redundant filterbank precoders and equalizers – Part I & II," *IEEE Trans. Signal Process.*, vol. 47, pp. 1988–2006, July 1999.
[30] J.G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2001.
[31] H.V. Poor and S. Verdù, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inf. Theory*, vol. 43, pp. 868–871, May 1997.
[32] A. Papoulis. *Probability, Random variables, and Stochastics Processes (3rd ed.).* McGraw-Hill, Singapore, 1991.
[33] T.S. Rappaport. *Wireless communications: principles and practice.* Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.
[34] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, pp. 3473–3482, Dec. 1994.
[35] W. Su, Z. Safar, and K.J.R. Liu, "Towards maximum achievable diveristy in space, time, and frequency: performance analysis and code design," *IEEE Trans. Commun.*, vol. 4, pp. 1847–1857, July 2005.
[36] P.R. Halmos. *Measure Theory.* Van Nostrand, Princeton, NJ, 1950.
[37] P. Li, D. Paul, R. Narasimhan and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Trans. Inf. Theory*, vol. 52, pp. 271–286, Jan. 2006.
[38] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.
[39] M. Kiessling and J. Speidel, "Unifying performance analysis of linear MIMO receivers in correlated Rayeigh fading environments," in *Proc. Int. Symp. Spread Spectrum Techniques*, Sydney, Australia, Aug.-Sep. 2004, pp. 634–638.
[40] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1972.
[41] S. Geman, "A limit theorem for the norm of random matrices," *The Annals of Probability*, vol. 8, pp. 252–261, 1980.

Francesco Verde was born in Santa Maria Capua Vetere, Italy, on June 12, 1974. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1998 from the Second University of Napoli, and the Ph.D. degree in information engineering in 2002, from the University of Napoli Federico II.

Since 2002, he has been an Assistant Professor of Signal Theory with the Department of Biomedical, Electronic and Telecommunication Engineering, University of Napoli Federico II. He also held teaching positions at the Second University of Napoli. His research activities lie in the broad area of statistical signal processing, digital communications, and communication systems. In particular, his current interests include cyclostationarity-based techniques for blind identification, equalization and interference suppression for narrowband modulation systems, code-division multiple-access systems, multicarrier modulation systems, and space-time processing for cooperative communications systems.

Donatella Darsena (M'06) was born in Napoli, Italy, on December 11, 1975. She received the Dr. Eng. degree *summa cum laude* in telecommunications engineering in 2001, and the Ph.D. degree in electronic and telecommunications engineering in 2005, both from the University of Napoli Federico II, Italy.

Since 2005, she has been an Assistant Professor with the Department for Technologies, University of Napoli Parthenope, Italy. Her research activities lie in the area of statistical signal processing, digital communications, and communication systems. In particular, her current interests are focused on equalization, channel identification, narrowband-interference suppression for multicarrier systems, and space-time processing for cooperative communications systems.

Anna Scaglione (SM'09) received the Laurea (M.Sc. degree) in 1995 and the Ph.D. degree in 1999 from the University of Rome, "La Sapienza". She is currently Associate Professor in Electrical and Computer Engineering at University of California at Davis, since July 2008. She was previously at Cornell University, Ithaca, NY, from 2001 where became Associate Professor in 2006; prior to joining Cornell she was Assistant Professor in the year 2000-2001, at the University of New Mexico.

She served as Associate Editor for the IEEE Transactions on Wireless Communications from 2002 to 2005, and serves since 2008 the Editorial Board of the IEEE Transactions on Signal Processing from 2008. She has been in the Signal Processing for Communication Committee since 2004. She was general chair of the workshop SPAWC 2005 and keynote speaker in SPAWC 2008. Dr. Scaglione is the first author of the paper that received the 2000 IEEE Signal Processing Transactions Best Paper Award; she has also received the NSF Career Award in 2002 and she is co-recipient of the Ellersick Best Paper Award (MILCOM 2005). Her expertise is in the broad area of signal processing for communication systems and networks. Her current research focuses on cooperative wireless networks and sensors' systems for monitoring and control applications.