This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TWC.2016.2554112

Coordinated Shared Spectrum Precoding with Distributed CSIT

Miltiades C. Filippou, *Member, IEEE,* Paul de Kerret, *Member, IEEE,* David Gesbert, *Fellow, IEEE,* Tharmalingam Ratnarajah, *Senior Member, IEEE,* Adriano Pastore, *Member, IEEE,* and George A. Ropokis, *Member, IEEE,*

***Abstract*—In this paper, the operation of a Licensed Shared Access (LSA) system is investigated, considering downlink communication. The system comprises a Multiple-Input-Single-Output (MISO) incumbent transmitter (TX) - receiver (RX) pair, which offers a spectrum sharing opportunity to a MISO licensee TX-RX pair. Our main contribution is the design of a *coordinated* transmission scheme, inspired by the *underlay* Cognitive Radio (CR) approach, with the aim of maximizing the average rate of the licensee, subject to an average rate constraint for the incumbent. In contrast to most prior works on underlay CR, the coordination of the two TXs takes place under a realistic Channel State Information (CSI) scenario, where each TX has solely access to the instantaneous direct channel of its served terminal. Such a CSI knowledge setting brings about a formulation based on the *theory of Team Decisions*, whereby the TXs aim at optimizing a common objective given the same constraint set, on the basis of individual channel information. Consequently, a novel set of applicable precoding schemes consisting in letting the two TXs cooperate on the basis of the statistical information is proposed. We verify by simulations that this novel, practically relevant, coordinated precoding scheme outperforms the standard underlay CR approach.**

***Index terms*—**Spectrum sharing, coordination, precoding, local CSI, QoS, cognitive radio, team decision

I. Introduction

The utilization of the radio spectrum is internationally regulated by governments, with the aim of providing wireless communication services that can be efficiently protected from harmful interference. Nevertheless, the tremendous spread of wireless services has given rise to a great need for bandwidth, which cannot be satisfied by an exclusivity of spectral allocation. On the other hand, such an exclusivity has created the phenomenon of *spectrum under-utilization* i.e., the low exploitation of large parts of the spectrum. The latter topic has been widely discussed throughout the wireless communications fora (see, for instance the report by the Federal Communications Commission (FCC) in 2002 [1]). As an answer, the principle of *Cognitive Radio* (CR) has been suggested as a promising technology in view of increasing wireless spectral efficiency by exploiting the existing *spectrum holes* in time, frequency or space [2], [3].

Focusing on the *underlay* CR approach, a primary network allows the simultaneous use of its spectral resources by a new-coming (unlicensed) secondary network, given the condition that the latter will utilize the available resources in a way that the interference created by a secondary transmitter (TX) towards a primary receiver (RX) is below a threshold predefined by the primary network [4], [5]. Under such a setup, efficient schemes, mainly exploiting multiple antennas at the terminals, have been proposed with the aim of maximizing the information rate of the secondary system, subject to given constraints over the harmful interference suffered by primary terminals [6]–[12]. However, the ability of the secondary TX to acquire global, multi-user Channel State Information (CSI) in practice is very limited, leading to the fact that most of these works in the literature are not applicable in most of the cases.

As a result, an extensive literature has focused on designing transmission schemes being robust to imperfect CSI or merely requiring local channel knowledge [See [13] and references therein]. In addition, iterative schemes, based on game theory, have been also investigated as a way to avoid the need for global multi-user CSI exchange, with respect to spectrum sharing scenarios [14]–[16].

Yet, all these works focus on designing a transmission scheme for the secondary TX, while the primary TX remains unaware of the presence of the secondary system. However, standardization bodies have lately focused on the design of Authorized or Licensed Shared Access systems (termed as ASA and LSA) [17], [18]. The key difference between the latter systems and underlay CR systems is that the incumbents (equivalent to primary nodes in a CR system) can share the spectrum with the licensees (the licensed equivalent of secondary nodes in a CR system), *provided* that Quality-of-Service (QoS) metrics, that have been negotiated prior to licensing, are satisfied for all involved entities. Motivated by this new framework, it is evident that a major drawback of

This work is supported by the Seventh Framework Programme for Research of the European Commission under grant number ADEL-619647. D. Gesbert and P. de Kerret are supported by the ERC under the European Union's Horizon 2020 research and innovation program (Agreement no. 670896).

M. C. Filippou was with the Institute for Digital Communications, University of Edinburgh, Edinburgh EH9 3FG, U.K. He is now with Intel Germany GmbH, Am Campeon 10-12, 85579, Neubiberg, Germany (e-mail: miltiades.filippou.de@ieee.org).

P. de Kerret was with Télécom Bretagne, IMT, UMR CNRS 3192 Lab-STICC, France. He is now with EURECOM, Campus SophiaTech, 450 Route des Chappes, 06410 Biot, France (e-mail: paul.dekerret@eurecom.fr).

D. Gesbert is with EURECOM, Campus SophiaTech, 450 Route des Chappes, 06410 Biot, France (e-mail: gesbert@eurecom.fr).

T. Ratnarajah is with the Institute for Digital Communications, University of Edinburgh, Edinburgh EH9 3FG, U.K. (e-mail: t.ratnarajah@ed.ac.uk).

A. Pastore is with Laboratory for Information in Networked Systems, École polytechnique fédérale de Lausanne, Route Cantonale, 1015 Lausanne, Switzerland (e-mail: adriano.pastore@epfl.ch).

G.A. Ropokis was with Computer Technology Institute and Press "Diophantus", 26500, Rio-Patras, Greece. He is now with CONNECT, Trinity College Dublin, Dublin 2, Ireland (e-mail: ropokisg@tcd.ie).

standard interference temperature-based underlay CR systems consists in the lack of coordination between the primary and the secondary systems. As a result, the primary system tends to overspend its available resources, leading to poor throughput performance at the secondary side.

Given this situation, we propose, in this work, a new coordination scheme for the two TXs based on commonly available, slow-varying statistical information. Such a coordination scheme does not require the exchange of any quickly varying CSI and can, therefore, be implemented to practical scenarios with only low requirements for the communication links between the TXs. Each TX exploits its locally available CSI, relevant to its served user, in a way that this transmission falls within the paradigm of *Team Decision* theory [19]–[23]: Both TXs (incumbent and licensee) aim at jointly maximizing a common utility (the licensee average rate), subject to a common constraint related to the incumbent average rate.

Preliminary results have been presented in [24]. In this work, the analysis was restricted to spatially uncorrelated direct channel links and solely to two strategies. In contrast, we focus here on the performance of an extended set of joint precoding schemes, with the assumption of correlated Rayleigh fading for all the involved MISO channels. The adoption of the correlated Rayleigh fading channel model follows from the modeling of signal propagation in heavily built-up environments, which are the ones where our analysis aims to find application [25], [26]. More particularly, our main contributions are the following:

- We formulate a novel framework of coordination between an incumbent TX and a licensee TX consisting in coordinating on the basis of the available long term information. This kind of coordination can be practically feasible for many scenarios thanks to the fact that the statistical (covariance) information is slowly varying.
- Within this framework, we design a statistically coordinated precoding scheme for a MISO spectrum sharing system, which can be applicable to a shared spectrum access system (ASA or LSA).
- We show through extensive simulations that the proposed scheme outperforms the standard interference temperature-based underlay CR approach.

Throughout the paper, the following notations are adopted: all boldface letters indicate vectors (lower case) or matrices (upper case). $\mathbf{A}^{\mathrm{H}}$, $\mathrm{tr}(\mathbf{A})$ and $[\mathbf{A}]_{m,n}$ denote the Hermitian transpose of matrix $\mathbf{A}$, its trace, and its (m,n)-th entry, respectively, whereas $\lambda_j(\mathbf{A})$ stands for its j-th eigenvalue. Also, $\mathrm{diag}\,(\alpha_1,\ldots,\alpha_n)$ symbolizes a diagonal matrix, the elements of which are $\alpha_1,\ldots,\alpha_n$. Additionally, $\mathbb{E}[\cdot]$ symbolizes the expectation operator and $\|\cdot\|$ denotes the Euclidean norm, while $\mathbf{0}_n$ denotes the all-zero vector of dimension n. The identity matrix of dimension $n\times n$ is denoted by $\mathbf{I}_n$, while $\bar{i}$ denotes the complementary index of i, when the cardinality of the considered set is equal to two, i.e., $\bar{i} = i \bmod 2+1$. For a random vector $\boldsymbol{x}$, $\boldsymbol{x}\sim\mathcal{CN}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes that $\boldsymbol{x}$ follows a Circularly Symmetric Complex Gaussian (CSCG) distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Finally, $E_1(\cdot)$ represents the exponential integral function, which is defined in [27, eq. (5.1.1)], while $\gamma\approx 0.5772$ stands for the Euler-Mascheroni constant, as it is defined in [27, eq. (4.1.32)].

II. System and Channel Model

The spectrum sharing system, which is illustrated in Fig. 1, is composed of a MISO incumbent system, comprising of a TX, TX 1, equipped with M_1 antennas, along with its assigned single-antenna terminal, RX 1. Focusing on downlink communication, the incumbent system is willing to share its resources with a MISO licensee system. The latter system consists of a multiple antenna TX, TX 2, equipped with M_2 antennas, as well as of a licensee terminal, RX 2, assigned to TX 2.

Considering the involved channels, spatially correlated Rayleigh fading is assumed for both direct and interfering channel links. As a consequence, for the channel between TX j and RX i, we have: $\boldsymbol{h}_{i,j}\sim\mathcal{CN}(\mathbf{0}_{M_j},\mathbf{R}_{i,j})$.

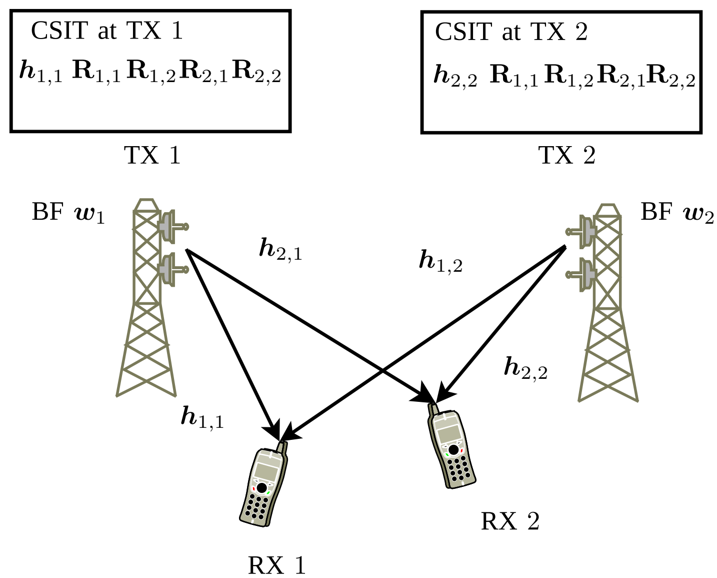

Fig. 1. The examined LSA system (post-licensing phase).

The signal received at RX i, $i\in\{1,2\}$, can be expressed as

$$y_i = \boldsymbol{h}_{i,i}^{\mathrm{H}}\boldsymbol{w}_i s_i + \boldsymbol{h}_{i,\bar{i}}^{\mathrm{H}}\boldsymbol{w}_{\bar{i}} s_{\bar{i}} + n_i, \quad (1)$$

where, $\boldsymbol{w}_i$ denotes the transmit beamforming vector at TX i and it is assumed that $\boldsymbol{w}_i=\sqrt{P_i}\boldsymbol{u}_i$, with $P_i\le P_i^{\max}$ and $\|\boldsymbol{u}_i\|=1$, where $P_i^{\max}$ is a maximum instantaneous power level at TX i. Also, Gaussian noise is considered at RX i, i.e., $n_i\sim\mathcal{CN}(0,N_0)$ and we assume that the information symbols for transmission are taken from a standard complex Gaussian codebook, i.e., $s_i\sim\mathcal{CN}(0,1)$, $i\in\{1,2\}$. By analyzing (1), the instantaneous information rate of RX i, $i\in\{1,2\}$ is given by [28]

$$R_i = \log_2\left(1+\frac{P_i|\boldsymbol{h}_{i,i}^{\mathrm{H}}\boldsymbol{u}_i|^2}{N_0+P_{\bar{i}}|\boldsymbol{h}_{i,\bar{i}}^{\mathrm{H}}\boldsymbol{u}_{\bar{i}}|^2}\right). \quad (2)$$

In the section that follows, the problem of joint downlink precoding with combined, local CSIT, is formulated.

III. Problem Formulation

A. Initial Optimization Problem

Focusing on the described system model, a realistic CSI at the TX (CSIT) assumption that can be made, is that TX $i, \quad i \in \{1,2\}$, has both instantaneous and statistical (covariance) knowledge of its direct links (i.e., TX 1 has instantaneous knowledge of direct link $\boldsymbol{h}_{1,1}$ and TX 2 has instantaneous knowledge of direct link $\boldsymbol{h}_{2,2}$), whereas, the interference cross-links are merely statistically known via knowledge of their covariance matrices. The second order statistics of the involved channels constitute slow-varying information that can be realistically collected by each TX through low capacity/high delay links.

Capitalizing on the available CSIT at TX $i, \quad i \in \{1,2\}$, the optimization problem of maximizing the average rate of the licensee system, subject to an average rate constraint for RX 1 can be formulated as a *functional* optimization problem, with functional dependencies related to the available CSI. Hence, the resulting optimization problem can be described as follows

$$\begin{aligned}&\left(\boldsymbol{w}_1^*, \boldsymbol{w}_2^*\right) = \arg\max \mathbb{E}\left[R_2\left(\boldsymbol{w}_1(\boldsymbol{h}_{1,1}), \boldsymbol{w}_2(\boldsymbol{h}_{2,2})\right)\right]\\&\text{subject to } \mathbb{E}\left[R_1\left(\boldsymbol{w}_1(\boldsymbol{h}_{1,1}), \boldsymbol{w}_2(\boldsymbol{h}_{2,2})\right)\right] \geq \tau_1 > 0,\\&0 \leq \|\boldsymbol{w}_1(\boldsymbol{h}_{1,1})\|^2 \leq P_1^{\max}, \quad 0 \leq \|\boldsymbol{w}_2(\boldsymbol{h}_{2,2})\|^2 \leq P_2^{\max},\end{aligned} \tag{P1}$$

where τ_1 stands for the QoS demand of RX 1, in terms of average rate.

Remark 1. The key difference between optimization problem (P1) and other approaches from the literature that focus on the case of *centralized CSIT*, comes from the fact that we aim at optimizing *precoding functions* at the TXs:

$$\begin{aligned}\boldsymbol{w}_i: \quad &\mathbb{C}^{M_i} \to \mathbb{C}^{M_i} \qquad ,\forall i \in \{1,2\}\\ &\boldsymbol{h}_{i,i} \mapsto \boldsymbol{w}_i(\boldsymbol{h}_{i,i}).\end{aligned} \tag{3}$$

The need to optimize over precoding functions instead of simply considering a vector optimization problem is a direct consequence of the *distributed CSIT* assumption. Indeed, as TX 1 is not aware of the channel realization $\boldsymbol{h}_{2,2}$, and reciprocally TX 2 is not aware of $\boldsymbol{h}_{1,1}$, it is necessary to consider the expectation over the precoding functions taken at the other TX, which requires the knowledge of the precoder for every channel realization, i.e., the *precoding function*. This makes it impossible to simply consider the optimization problem for a given channel realization as it is usually done in the literature of precoder optimization with centralized CSIT. □

For the sake of clarity, we will omit to mention explicitly the dependencies of the precoders in the following.

A reasonable assumption is that the QoS threshold, τ_1, can be achieved in the absence of any licensee. This comes down to considering that

$$\mathbb{E}\left[\log_2\left(1 + \frac{P_1^{\max}\|\boldsymbol{h}_{1,1}\|^2}{N_0}\right)\right] \geq \tau_1. \tag{4}$$

B. Approximated Optimization Problem

The expectation over the interfering channels makes the optimization difficult to handle. However, exploiting the convexity of function $\log_2\left(1+\frac{1}{x}\right)$, it becomes possible to apply Jensen's inequality [28] over the interfering channels. This significantly simplifies the optimization problem, while preserving its important features. The average rate expression for RX i, thus, becomes

$$\begin{aligned}\mathbb{E}\left[R_i\right] &= \mathbb{E}_{\boldsymbol{h}_{i,i},\boldsymbol{h}_{i,\bar{i}}}\left[\log_2\left(1 + \frac{P_i|\boldsymbol{h}_{i,i}^{\mathrm{H}}\boldsymbol{u}_i|^2}{N_0 + P_{\bar{i}}|\boldsymbol{h}_{i,\bar{i}}^{\mathrm{H}}\boldsymbol{u}_{\bar{i}}|^2}\right)\right]\\&\geq \mathbb{E}_{\boldsymbol{h}_{i,i}}\left[\log_2\left(1 + \frac{P_i|\boldsymbol{h}_{i,i}^{\mathrm{H}}\boldsymbol{u}_i|^2}{N_0 + \mathbb{E}_{\boldsymbol{h}_{i,\bar{i}}}\left[P_{\bar{i}}|\boldsymbol{h}_{i,\bar{i}}^{\mathrm{H}}\boldsymbol{u}_{\bar{i}}|^2\right]}\right)\right]\\&= \mathbb{E}_{\boldsymbol{h}_{i,i}}\left[\log_2\left(1 + \frac{P_i|\boldsymbol{h}_{i,i}^{\mathrm{H}}\boldsymbol{u}_i|^2}{N_0 + P_{\bar{i}}\boldsymbol{u}_{\bar{i}}^{\mathrm{H}}\mathbf{R}_{i,\bar{i}}\boldsymbol{u}_{\bar{i}}}\right)\right]\\&\triangleq \mathbb{E}\left[\tilde{R}_i(\boldsymbol{w}_i, \boldsymbol{w}_{\bar{i}})\right].\end{aligned} \tag{5}$$

Remark 2. It should be noted that applying Jensen's inequality in such a way is only possible thanks to the fact that the precoders $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are independent of the instantaneous cross-channels (as only the direct links are instantaneously known). □

With the aim of deriving a practical solution, slow power control depending on the long term statistical channel information, is assumed. Hence, instead of (instantaneous) power levels P_1 and P_2, we can use slow power allocation levels $\bar{P}_1$ and $\bar{P}_2$, where $0 \leq \bar{P}_i \leq P_i^{\max}, \quad i \in \{1,2\}$.

Altogether, in the remainder of the paper, we will work on the following optimization problem:

$$\begin{aligned}(\bar{P}_1^*, \boldsymbol{u}_1^*, \bar{P}_2^*, \boldsymbol{u}_2^*) &= \arg\max \ \mathbb{E}\left[\tilde{R}_2(\bar{P}_1, \boldsymbol{u}_1, \bar{P}_2, \boldsymbol{u}_2)\right]\\&\text{subject to } \mathbb{E}\left[\tilde{R}_1(\bar{P}_1, \boldsymbol{u}_1, \bar{P}_2, \boldsymbol{u}_2)\right] \geq \tau_1,\\&0 \leq \bar{P}_1 \leq P_1^{\max}, \quad 0 \leq \bar{P}_2 \leq P_2^{\max},\\&\|\boldsymbol{u}_1\|^2 = 1, \quad \|\boldsymbol{u}_2\|^2 = 1.\end{aligned} \tag{P2}$$

IV. Preliminary Results

The following two propositions provide some characteristics of the optimal solution of problem (P2) which will prove useful for designing the novel precoding scheme.

Proposition 1. *The ergodic rate constraint of RX* 1 *is satisfied with equality by any optimal solution* $(\bar{P}_1^\star, \boldsymbol{u}_1^\star, \bar{P}_2^\star, \boldsymbol{u}_2^\star)$ *of* (P2), *i.e.,*

$$\mathbb{E}\left[\tilde{R}_1(\bar{P}_1^\star, \boldsymbol{u}_1^\star, \bar{P}_2^\star, \boldsymbol{u}_2^\star)\right] = \tau_1. \tag{6}$$

Proof. The objective $\mathbb{E}\left[\tilde{R}_2(\bar{P}_1, \boldsymbol{u}_1, \bar{P}_2, \boldsymbol{u}_2)\right]$ is monotonically decreasing with respect to $\bar{P}_1$, while, on the other hand, the constraint $\mathbb{E}\left[\tilde{R}_1(\bar{P}_1, \boldsymbol{u}_1, \bar{P}_2, \boldsymbol{u}_2)\right]$ is monotonically increasing and continuous in $\bar{P}_1$. As a result, one can increase the objective by reducing power level $\bar{P}_1$ up to the point, where the average rate constraint of RX 1 will be satisfied with equality. This is always feasible because $\tau_1 > 0$ implies that $\bar{P}_1^\star > 0$. □

The second proposition yields some insight with respect to the optimal power allocation scheme.

Proposition 2. *An optimal solution of problem* (P2) *satisfies that either TX* 1 *or TX* 2 *transmits with full power, i.e., when* $\bar{P}_1^\star = P_1^{\max}$ *or* $\bar{P}_2^\star = P_2^{\max}$.

Proof. Considering an optimal solution, one can write $\bar{P}_1^\star = \alpha_1^\star \bar{P}$, for some $\alpha_1^\star \geq 0$ and $\bar{P}_2^\star = \alpha_2^\star \bar{P}$, for some $\alpha_2^\star \geq 0$, where $\bar{P} > 0$. Then, taking every term of the objective and dividing the numerator and the denominator of its Signal to Interference plus Noise Ratio (SINR) by $\bar{P}$, one obtains

$$\mathbb{E}\left[\tilde{R}_2(\boldsymbol{w}_1^\star, \boldsymbol{w}_2^\star)\right] = \mathbb{E}\left[\log_2\left(1 + \frac{\alpha_2^\star |\boldsymbol{h}_{2,2}^{\mathrm{H}} \boldsymbol{u}_2^\star|^2}{\frac{N_0}{\bar{P}} + \alpha_1^\star (\boldsymbol{u}_1^\star)^{\mathrm{H}} \mathbf{R}_{2,1} \boldsymbol{u}_1^\star}\right)\right], \quad (7)$$

which is a monotonically increasing function of $\bar{P}$. Similarly, the achievable average rate at RX 1 becomes

$$\mathbb{E}\left[\tilde{R}_1(\boldsymbol{w}_1^\star, \boldsymbol{w}_2^\star)\right] = \mathbb{E}\left[\log_2\left(1 + \frac{\alpha_1^\star |\boldsymbol{h}_{1,1}^{\mathrm{H}} \boldsymbol{u}_1^\star|^2}{\frac{N_0}{\bar{P}} + \alpha_2^\star (\boldsymbol{u}_2^\star)^{\mathrm{H}} \mathbf{R}_{1,2} \boldsymbol{u}_2^\star}\right)\right], \quad (8)$$

which is a monotonically increasing function of $\bar{P}$, as well.

If none of the two TXs transmits with full power, it means that it is possible to transmit with $\bar{P}' > \bar{P}$. Thus, the transmission using $(\alpha_1^\star \bar{P}', \boldsymbol{u}_1^\star, \alpha_2^\star \bar{P}', \boldsymbol{u}_2^\star)$ is feasible and leads to a larger objective, which contradicts the optimality of $(\alpha_1^\star \bar{P}, \boldsymbol{u}_1^\star, \alpha_2^\star \bar{P}, \boldsymbol{u}_2^\star)$. □

V. Statistically Coordinated Precoding

We now present our main contribution which is a new transmission scheme constituting a possible solution for optimization problem (P1). Indeed, it is important to note that, although possibly suboptimal, our approach is able to *guarantee* the incumbent rate constraint and is, therefore, a solution to the initial optimization problem.

A. General Approach

Since the derivation of closed-form expressions for the optimal precoders is hardly tractable due to the functional nature of optimization problem (P2) (which requires optimizing over an infinite dimensional space), we discretize the functional space and restrict the space of possible precoding solutions to a *set of transmission strategies*. Such a restriction, allows for every transmission strategy (i.e., joint precoding scheme), to be evaluated both in terms of feasibility and in terms of performance. Also, it provides a simple and practical method for coordinating the TXs.

The set of transmission strategies is obtained by the following steps. Note that to avoid breaking the flow of the description and for the sake of clarity, the detailed computations of the expectations can be found in the Appendix. They are provided as 3 lemmas that are used throughout the description of the precoding scheme.

1) Beamforming Design: The first step consists in designing the beamforming schemes that can be potentially applicable by each of the TXs. Although any beamforming scheme could be chosen in theory, a good heuristic choice is key to the tractability and the efficiency of the approach. In this work, we restrict our analysis to the Matched Filter (MF) and the statistical Zero-Forcing (sZF) strategies, as they represent the extreme approaches between which it will be necessary to strike a trade-off.

MF precoding corresponds to the *egoistic* beamforming scheme, where TX i transmits using

$$\boldsymbol{u}_{i,\mathrm{MF}} \triangleq \frac{\boldsymbol{h}_{i,i}}{\|\boldsymbol{h}_{i,i}\|}. \quad (9)$$

This beamformer (BF) maximizes the strength of the direct link without any consideration of the interference.

In contrast, sZF corresponds to an *altruistic* beamforming scheme, where TX i transmits using

$$\boldsymbol{u}_{i,\mathrm{sZF}} = \arg\max_{\boldsymbol{u} \in \mathbb{C}^{M_i \times 1}} \boldsymbol{u}^{\mathrm{H}} \mathbf{R}_{\bar{i},i}^{-\frac{1}{2}} \mathbf{R}_{i,i} \mathbf{R}_{\bar{i},i}^{-\frac{1}{2}} \boldsymbol{u}. \quad (10)$$

The sZF beamforming scheme consists in exploiting the statistical information of the cross-links to reduce the created interference, while also taking into consideration the statistical information of the direct links. This strategy has the advantage of using only statistical information available at both TXs and, hence, enforces perfect *coordination* between them, which will prove critical in terms of realizing an efficient joint transmission scheme.

2) Power Control Policy: Power control is a key ingredient to ensure that the average rate constraint for the incumbent RX is not violated. Furthermore, it is shown in Section IV that, in optimality, the incumbent QoS constraint is fulfilled with equality and that one of the two TXs emits with full power, while the other reduces its power to respect the incumbent constraint. Therefore, we denote by $\mathcal{P}_1$ the joint power policy where TX 1 emits with full power and by $\mathcal{P}_2$ the joint power policy where TX 2 transmits with full power.

3) Choice of the Transmission Policy: Considering the potential applicability of the two power control policies for each of the joint beamforming solutions, such a formulation leads to a joint transmission strategy set, which consists of 8 possible transmission schemes. However, the incumbent constraint is only fulfilled with probability one for some of the strategies and has to be verified otherwise. It is, hence, necessary to compute for each of these 8 transmission schemes the power emitted by one of the TXs and then evaluate the ergodic rate of both RXs. Once this is done, the best solution, in terms of average throughput for the licensee RX, is directly obtained.

Remark 3. It is critical to understand that coherent decisions upon transmission will be made, as the TXs are *statistically coordinated*: only statistical information is necessary to evaluate the ergodic rates and select the best strategy. □

B. Computation of the Ergodic Rates for each Strategy

The ergodic rates for each of the 8 strategies need to be evaluated. However, the expressions are practically the same in the sense that the 8 possible strategies come from the combination of only a few parameters. We will, hence, only present in full detail two strategies: MF-MF-$\mathcal{P}_1$ and sZF-sZF-$\mathcal{P}_2$, where the first two acronyms stand for the beamforming schemes applied by TX 1 and TX 2, respectively, while the third one denotes the followed power policy. The expressions for the other strategies can be trivially deduced.

Remark 4. The feasibility of a given strategy has to be verified. However, the feasibility of the optimization problem is preserved as the feasibility is guaranteed for strategy MF-MF-$\mathcal{P}_1$. Indeed, it contains the case where TX 1 transmits using MF and full power, while TX 2 does not transmit at all. □

1) Strategy MF-MF-$\mathcal{P}_1$: The TXs transmit using the beamforming vectors $\boldsymbol{u}_{1,\text{MF}}$ and $\boldsymbol{u}_{2,\text{MF}}$. Furthermore, TX 1 transmits using $\bar{P}_1 = P_1^{\max}$. It, thus, remains to determine how TX 2 controls its power to ensure that the incumbent ergodic rate constraint is fulfilled, i.e., that

$$\mathbb{E}\left[\tilde{R}_1\right] \geq \tau_1. \tag{11}$$

This can then be rewritten as

$$\begin{aligned}
&\mathbb{E}\left[\tilde{R}_1\right] \\
&= \mathbb{E}\left[\log_2\left(1 + \frac{P_1^{\max}\left|\boldsymbol{h}_{1,1}^{\text{H}}\boldsymbol{u}_{1,\text{MF}}\right|^2}{N_0 + \bar{P}_2\boldsymbol{u}_{2,\text{MF}}^{\text{H}}\mathbf{R}_{1,2}\boldsymbol{u}_{2,\text{MF}}}\right)\right] \\
&= \mathbb{E}_{\boldsymbol{h}_{1,1},\boldsymbol{h}_{2,2}}\left[\log_2\left(1 + \frac{P_1^{\max}\|\boldsymbol{h}_{1,1}\|^2}{N_0 + \bar{P}_2\frac{\boldsymbol{h}_{2,2}^{\text{H}}\mathbf{R}_{1,2}\boldsymbol{h}_{2,2}}{\|\boldsymbol{h}_{2,2}\|^2}}\right)\right] \\
&\overset{(a)}{\geq} \mathbb{E}_{\boldsymbol{h}_{1,1}}\left[\log_2\left(1 + \frac{P_1^{\max}\|\boldsymbol{h}_{1,1}\|^2}{N_0 + \bar{P}_2\mathbb{E}_{\boldsymbol{h}_{2,2}}\left[\frac{\boldsymbol{h}_{2,2}^{\text{H}}\mathbf{R}_{1,2}\boldsymbol{h}_{2,2}}{\|\boldsymbol{h}_{2,2}\|^2}\right]}\right)\right] \geq \tau_1,
\end{aligned} \tag{12}$$

where (a) holds by applying Jensen's inequality to convex function $\log_2\left(1+\frac{1}{x}\right)$ and the expectation in the denominator can then be computed using Lemma 3 in the Appendix with $\mathbf{A} = \mathbf{R}_{2,2}$ and $\mathbf{B} = \mathbf{R}_{2,2}^{\frac{1}{2}}\mathbf{R}_{1,2}\mathbf{R}_{2,2}^{\frac{1}{2}}$.

Finally, a closed form expression for the ergodic rate is obtained with Lemma 1. Hence, the value of $\bar{P}_2$ can be deduced by bisection, in order for the lower bound derived in (12) to be equal to τ_1.

It remains to evaluate the corresponding achievable average rate of RX 2. Following a similar approach as the one for the ergodic rate of the incumbent, we can obtain the following lower bound:

$$\begin{aligned}
&\mathbb{E}\left[\tilde{R}_2\right] \\
&= \mathbb{E}\left[\log_2\left(1 + \frac{\bar{P}_2\left|\boldsymbol{h}_{2,2}^{\text{H}}\boldsymbol{u}_{2,\text{MF}}\right|^2}{N_0 + P_1^{\max}\boldsymbol{u}_{1,\text{MF}}^{\text{H}}\mathbf{R}_{2,1}\boldsymbol{u}_{1,\text{MF}}}\right)\right] \\
&= \mathbb{E}_{\boldsymbol{h}_{1,1},\boldsymbol{h}_{2,2}}\left[\log_2\left(1 + \frac{\bar{P}_2\|\boldsymbol{h}_{2,2}\|^2}{N_0 + P_1^{\max}\frac{\boldsymbol{h}_{1,1}^{\text{H}}\mathbf{R}_{2,1}\boldsymbol{h}_{1,1}}{\|\boldsymbol{h}_{1,1}\|^2}}\right)\right] \\
&\geq \mathbb{E}_{\boldsymbol{h}_{2,2}}\left[\log_2\left(1 + \frac{\bar{P}_2\|\boldsymbol{h}_{2,2}\|^2}{N_0 + P_1^{\max}\mathbb{E}_{\boldsymbol{h}_{1,1}}\left[\frac{\boldsymbol{h}_{1,1}^{\text{H}}\mathbf{R}_{2,1}\boldsymbol{h}_{1,1}}{\|\boldsymbol{h}_{1,1}\|^2}\right]}\right)\right].
\end{aligned} \tag{13}$$

Once more, the expectation in the denominator is obtained using Lemma 3, while a closed form expression for the ergodic rate is obtained with Lemma 1.

2) Strategy sZF-sZF-$\mathcal{P}_2$: In this strategy, the TXs transmit using the BFs $\boldsymbol{u}_{1,\text{sZF}}$ and $\boldsymbol{u}_{2,\text{sZF}}$, while TX 2 transmits using $\bar{P}_2 = P_2^{\max}$. It remains then to determine $\bar{P}_1$. In that setting, the rate of RX 1 can be lower bounded as

$$\begin{aligned}
\mathbb{E}\left[\tilde{R}_1\right] &= \mathbb{E}_{\boldsymbol{h}_{1,1}}\left[\log_2\left(1 + \frac{\bar{P}_1|\boldsymbol{h}_{1,1}^{\text{H}}\boldsymbol{u}_{1,\text{sZF}}|^2}{N_0 + P_2^{\max}\boldsymbol{u}_{2,\text{sZF}}^{\text{H}}\mathbf{R}_{1,2}\boldsymbol{u}_{2,\text{sZF}}}\right)\right] \\
&\geq \tau_1.
\end{aligned} \tag{14}$$

This rate can be directly computed in closed form using Lemma 2. Finally, the power $\bar{P}_1$, such that the ergodic rate constraint for RX 1 is met by the derived lower bound, can be obtained by bisection.

It now remains to evaluate the corresponding ergodic rate of the licensee RX. This is given by the following expression

$$\mathbb{E}\left[\tilde{R}_2\right] = \mathbb{E}_{\boldsymbol{h}_{2,2}}\left[\log_2\left(1 + \frac{P_2^{\max}|\boldsymbol{h}_{2,2}^{\text{H}}\boldsymbol{u}_{2,\text{sZF}}|^2}{N_0 + \bar{P}_1\boldsymbol{u}_{1,\text{sZF}}^{\text{H}}\mathbf{R}_{2,1}\boldsymbol{u}_{1,\text{sZF}}}\right)\right]. \tag{15}$$

The latter expression can be computed in closed form by applying Lemma 2.

VI. Reference Precoding Schemes

In this section, we present two schemes which will be used to evaluate the efficiency of our statistically coordinated precoding approach.

The first one, denoted as "interference temperature-based" precoding, is an adaptation of the approaches in the literature to allow for a fair comparison. Intuitively, it corresponds to the conventional underlay CR paradigm, where solely the secondary TX adapts its strategy in order for the interference received by the primary RX to be below a given threshold [5].

The second one constitutes a coordination benchmark and it is a priori not reachable. It, hence, represents an upperbound which allows to evaluate the sub-optimality of the proposed approach.

A. Interference Temperature-Based Precoding

The interference temperature approach, extensively used in the CR literature, consists in forcing the secondary TX to create less interference to the primary user, than a given interference threshold, which is here denoted by $\mathcal{I}$.

Considering that the secondary TX aims at minimizing the interference created and transmits using $\boldsymbol{u}_{2,\text{sZF}}$, the power emitted by the secondary TX is then given by

$$\bar{P}_2 = \min\left\{\frac{\mathcal{I}}{\boldsymbol{u}_{2,\text{sZF}}^{\text{H}}\mathbf{R}_{1,2}\boldsymbol{u}_{2,\text{sZF}}}, P_2^{\max}\right\}. \tag{16}$$

In order to conduct a fair comparison with the designed statistically coordinated precoding scheme, we need to determine the interference temperature, $\mathcal{I}$, such that the ergodic rate constraint of RX 1 is satisfied with equality, i.e.,

$$\mathbb{E}\left[\log_2\left(1 + \frac{P_1^{\max}\|\boldsymbol{h}_{1,1}\|^2}{N_0 + \mathcal{I}}\right)\right] = \tau_1. \tag{17}$$

The expectation appearing in (17) can be computed by applying Lemma 1. The interference temperature threshold, $\mathcal{I}$, can be then easily found by bisection.

B. Coordination Benchmark

When designing the BFs, we can observe a clear trade-off between maximizing the desired signal (using MF) and minimizing the interference created. Hence, if we assume that one can achieve both goals at the same time, the following optimization problem is obtained for the power control, and leads to an, a priori, infeasible performance upperbound.

$$\begin{aligned} \max_{\bar{P}_1, \bar{P}_2} \ & \mathbb{E}\left[\log_2\left(1 + \frac{\bar{P}_2 \|\boldsymbol{h}_{2,2}\|^2}{N_0 + \bar{P}_1 \lambda_{\min}(\mathbf{R}_{2,1})}\right)\right] \\ \text{subject to} \quad & \mathbb{E}\left[\log_2\left(1 + \frac{\bar{P}_1 \|\boldsymbol{h}_{1,1}\|^2}{N_0 + \bar{P}_2 \lambda_{\min}(\mathbf{R}_{1,2})}\right)\right] \geq \tau_1, \\ & 0 \leq \bar{P}_1 \leq P_1^{\max}, \quad 0 \leq \bar{P}_2 \leq P_2^{\max}. \end{aligned} \tag{P3}$$

The ergodic rate expressions appearing in (P3) can be computed in closed form by applying Lemma 1. The optimal slow power control values are obtained by exploiting Proposition 2. Indeed, one of the two TXs transmits with full power, while the other one controls its power by bisection. Comparing the performance and the feasibility of both solutions leads to the solution of optimization problem (P3).

VII. Numerical Evaluation

With the aim of evaluating the performance of the proposed statistically coordinated precoding scheme, extensive Monte Carlo simulations have been performed and, more specifically, 20000 channel realizations have been simulated. We choose $M_1 = M_2 = M = 4$ antennas at each TX. Furthermore, we consider unit noise variance ($N_0 = 1$) and a QoS threshold $\tau_1 = 1$ bps/Hz.

We consider a classical exponential channel correlation model [29], in which the covariance matrices $\mathbf{R}_{i,j}$ are given by

$$\mathbf{R}_{i,j} = \beta_{i,j} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{M-1} \\ \rho & 1 & \rho & \dots & \rho^{M-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{M-1} & \rho^{M-2} & \rho^{M-3} & \dots & 1 \end{bmatrix}, \tag{18}$$

where, $\beta_{i,j}, \ \ i,j \in \{1,2\}$, represents the pathloss and is chosen here equal to 1 when $i = j$ and to 0.3 otherwise. In the investigated scenario the antenna correlation factor, ρ, is equal to 0.5.

In Fig. 2 and Fig. 3, the average rate of RX 1 and the average rate of RX 2 are depicted as a function of the system's transmit Signal-to-Noise Ratio (SNR). The three curves represent the throughput performance achieved by the proposed statistically coordinated precoding scheme, the interference temperature-based precoding scheme, as well as the described coordination benchmark. Focusing on RX 2, the coordination benchmark outperforms both the proposed precoding scheme, as well as the interference temperature-based scheme, as expected. By observing Fig. 2, it should be noted that, in contrast with the coordination benchmark, the proposed precoding scheme fails to satisfy the incumbent average rate constraint with equality. This occurs because we resort to tackling optimization problem (P2), which involves a lower bound of the average rate of RX 1. Nevertheless, the proposed algorithm successfully manages to control the average rate of RX 1 and this capability can be translated to a significant throughput gain for the licensee, in comparison to the one achieved by the interference temperature-based precoding scheme.

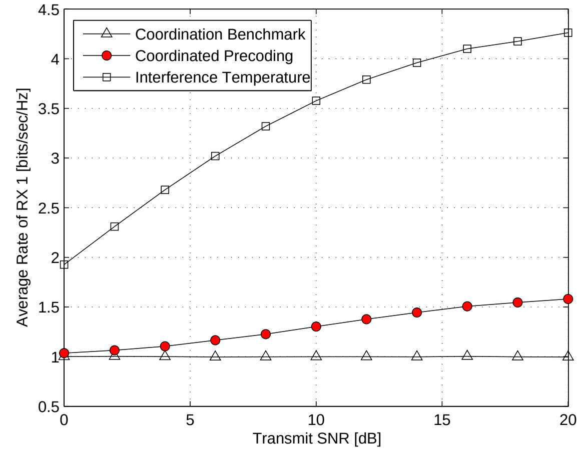

Fig. 2. Ergodic rate of RX 1 vs. transmit SNR, when incumbent QoS threshold $\tau_1 = 1$bps/Hz.

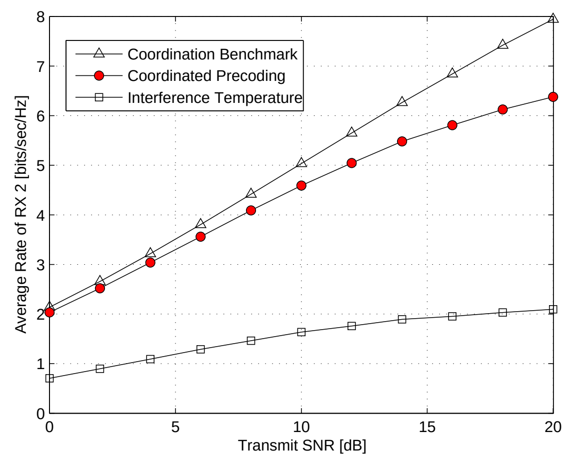

Fig. 3. Ergodic rate of RX 2 vs. transmit SNR, when incumbent QoS threshold $\tau_1 = 1$bps/Hz.

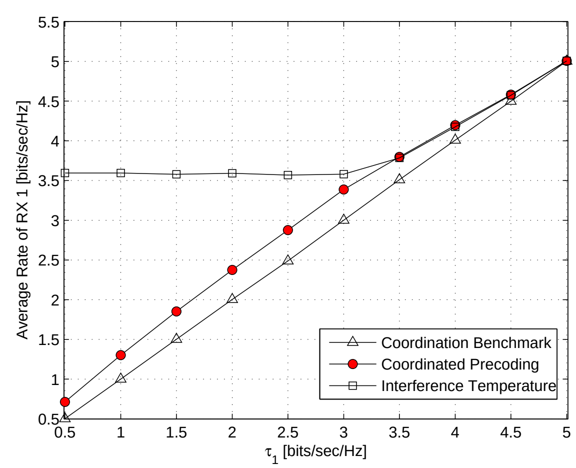

Fig. 4. Ergodic rate of RX 1 vs. threshold τ_1, SNR=10dB.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TWC.2016.2554112

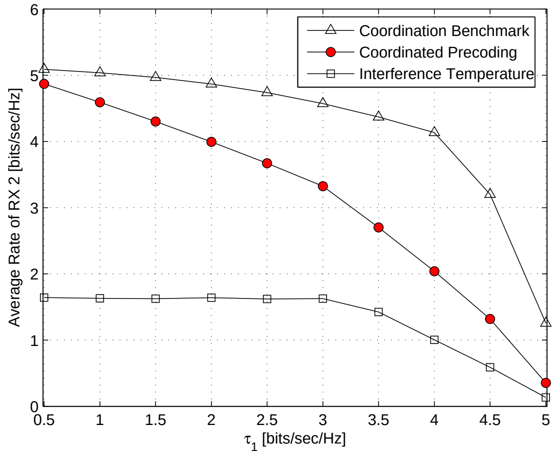

Fig. 5. Ergodic rate of RX 2 vs. threshold τ_1, SNR=10dB.

The achievable average rates of RX 1 and RX 2, by applying the proposed precoding algorithm, along with the ones achieved by the two reference precoding schemes, are depicted in Fig. 4 and Fig. 5, respectively, as a function of QoS threshold, τ_1, when the transmit SNR of the system is equal to 10 dB. The average rate constraint for RX 1 is fulfilled by all three schemes for the whole examined range of τ_1. Also, the proposed precoding scheme outperforms the interference temperature-based one, especially when the average rate constraint of the incumbent is loose, which occurs due to the fact that under this regime there is more to gain for the licensee by means of an efficient coordination.

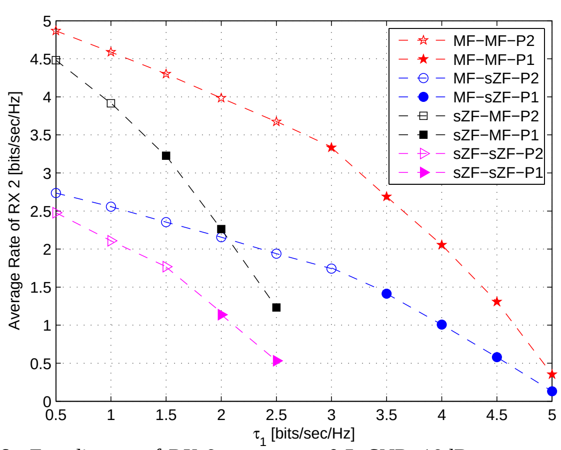

Fig. 6. Ergodic rate of RX 2 vs. τ_1, ρ = 0.5, SNR=10dB.

In Fig. 6, the average licensee rate for each of the *feasible* joint precoding schemes is depicted as a function of QoS constraint, τ_1, when the transmit SNR is equal to 10 dB and ρ = 0.5. The term "feasible" is used here to characterize the joint precoding schemes, for which both optimality conditions, i.e., Proposition 1 and Proposition 2 are satisfied. In other words, there can be found power levels such that the constraint on the average incumbent rate is satisfied with *equality*. It can be observed that when $\tau_1 \in [0.5 \quad 2.5]$ bits/sec/Hz, strategy MF-MF-$\mathcal{P}_2$ is the rate-optimal one, whereas, for stricter QoS constraints on the incumbent, strategy MF-MF-$\mathcal{P}_1$ has to be selected, exactly because the system focuses primarily on preserving the average rate of the incumbent RX. It is also worth mentioning that as τ_1 increases, only the subset of the most "conservative" joint precoding schemes is feasible and can, thus, be put into comparison by means of the achievable average licensee rate.

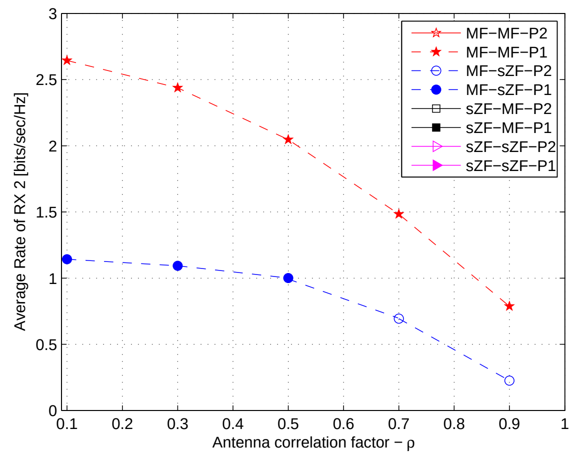

Fig. 7. Ergodic rate of RX 2 vs. antenna correlation factor, ρ, SNR=15dB, τ_1 = 4bits/sec/Hz.

Finally, in Fig. 7, the performance of the *feasible* (in the same sense as above) joint precoding schemes is evaluated by means of the achievable average licensee rate, as a function of the antenna correlation factor, ρ, when the transmit SNR is equal to 15 dB and the QoS threshold on incumbent communication is $\tau_1 = 4$bits/sec/Hz. One can first observe that only a subset (i.e., 3) of the 8 joint precoding strategies is feasible under the selected system scenario, as the bisection methods applied to the other schemes could not return a power level within the intervals $[0, P_1^{\max}]$ regarding power $\bar{P}_1$ or $[0, P_2^{\max}]$ considering power $\bar{P}_2$. Hence, the applicable (in terms of feasibility) schemes for the examined system scenario are: MF-MF-$\mathcal{P}_1$ for the whole examined range of ρ, as well as MF-sZF-$\mathcal{P}_1$ and MF-sZF-$\mathcal{P}_2$ for subsets of that value interval. Also importantly, it is observed that, for all depicted schemes, the throughput performance of the licensee degrades, as parameter ρ increases. This can be justified by the fact that, as the transmit antennas become more correlated, the property of spatial diversity cannot be exploited efficiently, with the aim of maximizing the power received by the licensee RX.

VIII. Conclusions

In this work, we propose a novel, joint precoding scheme, with reference to a shared spectrum access system, where the two TXs coordinate on the basis of statistical knowledge of the global multi-user channel. Our approach consists in formulating a Team Decision problem, the solution of which is approached by reducing the transmission strategy space to a finite number of strategies. This method is key to enforcing coordination between the TXs and obtaining a practical solution to the intricate Team Decision problem. Such an approach allows to improve over the conventional underlay CR approach, at the price of low CSI and communication requirements, as the coordination can be realized offline. Approaching the global optimum is both a difficult and challenging problem that will be further tackled in the future. The proposed scheme has also a strong potential for other more complex scenarios with multiple incumbent and/or licensee networks.

Appendix

Lemma 1. *[30, eq. (37)] Let $\bar{\gamma} \in \mathbb{R}^+$ and $\boldsymbol{h} \sim \mathcal{CN}(\mathbf{0}_n, \mathbf{R}_{\boldsymbol{h}})$, where covariance matrix $\mathbf{R}_{\boldsymbol{h}}$ has n distinct eigenvalues $\{\lambda_j\}_{j=1}^n$. It then holds*

$$\mathbb{E}_{\boldsymbol{h}}\left[\log_2(1+\bar{\gamma}\|\boldsymbol{h}\|^2)\right] = \frac{1}{\ln(2)\bar{\gamma}\prod_{j=1}^n \lambda_j}\sum_{j=1}^n \frac{\bar{\gamma}\lambda_j e^{\frac{1}{\bar{\gamma}\lambda_j}} E_1\left(\frac{1}{\bar{\gamma}\lambda_j}\right)}{\prod_{m=1,m\neq j}^n\left(\frac{1}{\lambda_m}-\frac{1}{\lambda_j}\right)}. \tag{19}$$

Lemma 2. *[31, eq. (75)-(76)] Let $\bar{\gamma} \in \mathbb{R}^+$ and $\boldsymbol{w} \in \mathbb{C}^{n\times 1}$ be deterministic, and $\boldsymbol{h} \sim \mathcal{CN}(\mathbf{0}_n, \mathbf{R}_{\boldsymbol{h}})$. It then holds*

$$\mathbb{E}_{\boldsymbol{h}}\left[\log_2\left(1+\bar{\gamma}|\boldsymbol{h}^{\mathrm{H}}\boldsymbol{w}|^2\right)\right] = \frac{e^{\frac{1}{\bar{\gamma}\lambda_1(\mathbf{R}_{\mathit{eff}})}}}{\ln(2)} E_1\left(\frac{1}{\bar{\gamma}\lambda_1(\mathbf{R}_{\mathit{eff}})}\right), \tag{20}$$

where $\lambda_1(\mathbf{R}_{\mathit{eff}})$ is the unique non-zero (positive) eigenvalue of matrix $\mathbf{R}_{\mathit{eff}} = \mathbf{R}_{\boldsymbol{h}}^{\frac{1}{2}}\boldsymbol{w}\boldsymbol{w}^{\mathrm{H}}\mathbf{R}_{\boldsymbol{h}}^{\frac{1}{2}}$.

Lemma 3. *Let us consider two positive semi-definite matrices $\mathbf{A}$ and $\mathbf{B}$ in $\mathbb{C}^{n\times n}$ with eigenvalues denoted as $\lambda_1(\mathbf{A}),\ldots,\lambda_n(\mathbf{A})$ and $\lambda_1(\mathbf{B}),\ldots,\lambda_n(\mathbf{B})$, respectively, where it is assumed that $\mathbf{A}$ is of full rank and has no multiple eigenvalues. We also assume that matrix $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{U}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{A}}\mathbf{U}_{\mathbf{A}}^{\mathrm{H}}$, where $\mathbf{U}_{\mathbf{A}}$ is a unitary matrix and $\mathbf{\Lambda}_{\mathbf{A}} = \mathrm{diag}\left(\lambda_1(\mathbf{A}),\ldots,\lambda_n(\mathbf{A})\right)$, where it holds that $0 < \lambda_1(\mathbf{A}) < \ldots < \lambda_n(\mathbf{A})$. Let $\boldsymbol{x} \in \mathbb{C}^{n\times n}$ be a standard complex Gaussian random vector, such that $\boldsymbol{x} \sim \mathcal{CN}(\mathbf{0}_n, \mathbf{I}_n)$. It then holds*

$$\begin{aligned}\mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{B}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{A}\boldsymbol{x}}\right] &= \sum_{i=1}^n [\tilde{\mathbf{B}}]_{i,i}\Bigg\{\frac{\lambda_i(\mathbf{A})^{n-2}\left((n-1)\left(\ln(\lambda_i(\mathbf{A}))-\gamma\right)+1\right)}{\prod_{j\neq i}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)} \\ &- \frac{\lambda_i(\mathbf{A})^{n-1}\left(\ln(\lambda_i(\mathbf{A}))-\gamma\right)\sum_{r=1,r\neq i}^n\prod_{j\neq i,r}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)}{\left(\prod_{j\neq i}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)\right)^2} \\ &+ \sum_{k=1,k\neq i}^n \frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)\prod_{j\neq k,i}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)}{\left(\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)\right)^2}\Bigg\},\end{aligned} \tag{21}$$

where $\tilde{\mathbf{B}} = \mathbf{U}_{\mathbf{A}}^{\mathrm{H}}\mathbf{B}\mathbf{U}_{\mathbf{A}}$.

Proof. We prove this result in two steps. Firstly, we show that considering two matrices $\mathbf{A}$ and $\mathbf{B}$ with different eigenbases, we can come back to the case of matrices having the same eigenbasis. We then prove the lemma for this case.

Let us assume that $\mathbf{A}$ and $\mathbf{B}$ have *different* eigenbases. We consider their eigendecompositions $\mathbf{A} = \mathbf{U}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{A}}\mathbf{U}_{\mathbf{A}}^{\mathrm{H}}$ and $\mathbf{B} = \mathbf{U}_{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{B}}\mathbf{U}_{\mathbf{B}}^{\mathrm{H}}$, where the diagonal entries of $\mathbf{\Lambda}_{\mathbf{A}}$ are sorted in an increasing order and the diagonal entries of $\mathbf{\Lambda}_{\mathbf{B}}$ are non-decreasingly sorted. Introducing matrix $\tilde{\mathbf{B}} = \mathbf{U}_{\mathbf{A}}^{\mathrm{H}}\mathbf{B}\mathbf{U}_{\mathbf{A}}$, the expectation in question becomes

$$\begin{aligned}\mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{B}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{A}\boldsymbol{x}}\right] &= \mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{U}_{\mathbf{A}}^{\mathrm{H}}\mathbf{B}\mathbf{U}_{\mathbf{A}}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{A}}\boldsymbol{x}}\right] = \mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\tilde{\mathbf{B}}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{A}}\boldsymbol{x}}\right] \\ &= \mathbb{E}\left[\frac{\sum_{i=1}^n\sum_{j=1}^n[\tilde{B}]_{i,j}x_i^*x_j}{\sum_{k=1}^n\lambda_k(\mathbf{A})|x_k|^2}\right].\end{aligned} \tag{22}$$

Exploiting the fact that $\boldsymbol{x} \sim \mathcal{CN}(\mathbf{0}_n, \mathbf{I}_n)$, if we write each $x_i, \quad i = 1,\ldots,n$ in polar representation, i.e., $x_i = |x_i|e^{j\phi_i}$, then we have that all phases $\phi_i, \quad i=1,\ldots,n$ and amplitudes $|x_i|, \quad i=1,\ldots,n$ are mutually independent and the phases are uniformly distributed. As a result, the expectation takes the following form

$$\begin{aligned}\mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{B}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{A}\boldsymbol{x}}\right] &= \mathbb{E}\left[\frac{\sum_{i=1}^n\sum_{j=1}^n[\tilde{B}]_{i,j}|x_i||x_j|e^{j(\phi_j-\phi_i)}}{\sum_{k=1}^n\lambda_k(\mathbf{A})|x_k|^2}\right] \\ &= \mathbb{E}\left[\frac{\sum_{i=1}^n[\tilde{B}]_{i,i}|x_i|^2}{\sum_{k=1}^n\lambda_k(\mathbf{A})|x_k|^2}\right],\end{aligned} \tag{23}$$

or, equivalently

$$\mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{B}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{A}\boldsymbol{x}}\right] = \mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{U}_{\mathbf{A}}\,\mathrm{diag}\left([\tilde{B}]_{1,1},\ldots,[\tilde{B}]_{n,n}\right)\mathbf{U}_{\mathbf{A}}^{\mathrm{H}}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{A}\boldsymbol{x}}\right]. \tag{24}$$

Hence, the case of *equal* eigenbases is recovered.

Consequently, we proceed by considering, without loss of generality, that matrices $\mathbf{A}$ and $\mathbf{B}$ have the same eigenbases. However, it should be noted that elements $[\tilde{B}]_{i,i}, \quad i = 1,\ldots,n$ are not sorted in any particular order.

Focusing, now, on the derivation of a closed form expression of the expectation, we have that

$$\begin{aligned}\mathbb{E}\left[\frac{\boldsymbol{x}^{\mathrm{H}}\mathbf{B}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\mathbf{A}\boldsymbol{x}}\right] &= \mathbb{E}\left[\frac{\sum_{i=1}^n[\tilde{B}]_{i,i}|x_i|^2}{\sum_{j=1}^n\lambda_j(\mathbf{A})|x_j|^2}\right] \\ &= \sum_{i=1}^n[\tilde{B}]_{i,i}\mathbb{E}\left[\frac{|x_i|^2}{\sum_{j=1}^n\lambda_j(\mathbf{A})|x_j|^2}\right] \\ &= \sum_{i=1}^n[\tilde{B}]_{i,i}\mathbb{E}\left[\frac{\partial}{\partial\lambda_i(\mathbf{A})}\ln\left(\sum_{j=1}^n\lambda_j(\mathbf{A})|x_j|^2\right)\right] \\ &= \sum_{i=1}^n[\tilde{B}]_{i,i}\frac{\partial}{\partial\lambda_i(\mathbf{A})}\mathbb{E}\left[\ln\left(\sum_{j=1}^n\lambda_j(\mathbf{A})|x_j|^2\right)\right].\end{aligned} \tag{25}$$

Let us define random variable (RV) $X \triangleq \sum_{j=1}^n\lambda_j(\mathbf{A})|x_j|^2$. Using [32, eq. (8)], it is shown by induction that the Probability Density Function (PDF) of X is the following

$$p_X(x) = \sum_{k=1}^n\prod_{j\neq k}\frac{\lambda_k(\mathbf{A})^{n-2}}{\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})}e^{-\frac{x}{\lambda_k(\mathbf{A})}}. \tag{26}$$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TWC.2016.2554112

As a result, the expectation of RV $\ln(X)$ is given by the expression that follows

$$\mathbb{E}\left[\ln(X)\right]=\sum_{k=1}^{n}\int_{0}^{\infty}\ln(x)e^{-\frac{x}{\lambda_k(\mathbf{A})}}dx\prod_{j\neq k}\frac{\lambda_k(\mathbf{A})^{n-2}}{\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})}. \tag{27}$$

Exploiting [33, eq. (4.331.1)], expression (27) becomes

$$\mathbb{E}\left[\ln(X)\right]=\sum_{k=1}^{n}\frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)}{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)}. \tag{28}$$

Taking, now, the partial derivative of (28), with respect to $\lambda_i(\mathbf{A})$, we obtain

$$\begin{aligned}\frac{\partial}{\partial\lambda_i(\mathbf{A})}\mathbb{E}\left[\ln(X)\right]&=\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\frac{\lambda_i(\mathbf{A})^{n-1}\left(\ln(\lambda_i(\mathbf{A}))-\gamma\right)}{\prod_{j\neq i}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\right\}\\&+\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\sum_{k=1,k\neq i}^{n}\frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)}{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\right\}.\end{aligned} \tag{29}$$

For the first term of (29), the following expression is obtained

$$\begin{aligned}&\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\frac{\lambda_i(\mathbf{A})^{n-1}\left(\ln(\lambda_i(\mathbf{A}))-\gamma\right)}{\prod_{j\neq i}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\right\}\\&=\frac{\lambda_i(\mathbf{A})^{n-2}\left((n-1)\left(\ln(\lambda_i(\mathbf{A}))-\gamma\right)+1\right)}{\prod_{j\neq i}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\\&-\frac{\lambda_i(\mathbf{A})^{n-1}\left(\ln(\lambda_i(\mathbf{A}))-\gamma\right)\sum_{r=1,r\neq i}^{n}\prod_{j\neq i,r}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)}{\left(\prod_{j\neq i}\left(\lambda_i(\mathbf{A})-\lambda_j(\mathbf{A})\right)\right)^2}.\end{aligned} \tag{30}$$

It now remains to find the second term of (29) in closed form. We, thus, obtain the following

$$\begin{aligned}&\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\sum_{k=1,k\neq i}^{n}\frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)}{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\right\}\\&=\sum_{k=1,k\neq i}^{n}\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)}{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\right\}.\end{aligned} \tag{31}$$

Given that $i\neq k$, the partial derivative appearing in the right hand side of (31), is given by the following expression

$$\begin{aligned}&\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)}{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)}\right\}=\\&-\frac{\lambda_k(\mathbf{A})^{n-1}\left(\ln(\lambda_k(\mathbf{A}))-\gamma\right)\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)\right\}}{\left(\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)\right)^2},\end{aligned} \tag{32}$$

where

$$\frac{\partial}{\partial\lambda_i(\mathbf{A})}\left\{\prod_{j\neq k}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right)\right\}=-\prod_{j\neq k,i}\left(\lambda_k(\mathbf{A})-\lambda_j(\mathbf{A})\right). \tag{33}$$

This concludes the proof. □

References

[1] Federal Communications Commission, "Spectrum policy task force report, FCC 02-155," 2002.

[2] J. Mitola and J. Maguire, G.Q., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[3] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[4] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[5] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. Cambridge University Press, 2012.

[6] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[7] S.-J. Kim and G. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.

[8] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 2112–2120, Apr. 2009.

[9] Y. Huang and D. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[10] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5306–5315, Dec. 2008.

[11] Y. Y. He and S. Dey, "Sum rate maximization for cognitive MISO broadcast channels: Beamforming design and large systems analysis," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2383–2401, May 2014.

[12] J. Zhang, C.-K. Wen, C. Yuen, S. Jin, and X. Gao, "Large system analysis of cognitive radio network via partially-projected regularized zero-forcing precoding," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 4934–4947, Sep. 2015.

[13] K.-Y. Wang, N. Jacklin, Z. Ding, and C.-Y. Chi, "Robust MISO transmit optimization under outage-based QoS constraints in two-tier heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1883–1897, Apr. 2013.

[14] B. Wang, Y. Wu, and K. R. Liu, "Game theory for cognitive radio networks: An overview," *Computer Networks*, vol. 54, no. 14, pp. 2537 – 2561, 2010.

[15] G. Scutari and D. Palomar, "MIMO cognitive radio: A game theoretical approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 761–780, Feb. 2010.

[16] W. Zhong, Y. Xu, and H. Tianfield, "Game-theoretic opportunistic spectrum sharing strategy selection for cognitive MIMO multiple access channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2745–2759, Jun. 2011.

[17] E. CEPT, "Licensed shared access (LSA)," *ECC Report*, vol. 205, 2014.

[18] J. Holdren and E. Lander, "Realizing the full potential of government-held spectrum to spur economic growth," Technical Report, Tech. Rep., 2012.

[19] R. Radner, "Team decision problems," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. pp. 857–881, 1962.

[20] Y.-C. Ho, "Team decision theory and information structures," *Proceedings of the IEEE*, vol. 68, no. 6, pp. 644–654, 1980.

[21] R. Zakhour and D. Gesbert, "Team decision for the cooperative MIMO channel with imperfect CSIT sharing," in *Information Theory and Applications Workshop (ITA), 2010*, Jan. 2010, pp. 1–6.

[22] P. de Kerret and D. Gesbert, "CSI sharing strategies for transmitter cooperation in wireless networks," *IEEE Wireless Communications*, vol. 20, no. 1, pp. 43–49, Feb. 2013.

[23] M. C. Filippou, G. A. Ropokis, and D. Gesbert, "A team decisional beamforming approach for underlay cognitive radio networks," in *IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2013*, Sep. 2013, pp. 575–579.

[24] P. de Kerret, M. C. Filippou, and D. Gesbert, "Statistically coordinated precoding for the MISO cognitive radio channel," in *48th Asilomar Conference on Signals, Systems and Computers, 2014*, Nov. 2014, pp. 1083–1087.

[25] J. G. Proakis, "Digital communications." *McGraw-Hill, New York*, 1995.

[26] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. I. Characterization," *IEEE Communications Magazine*, vol. 35, no. 7, pp. 90–100, 1997.

[27] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*. Dover publications, 1965.

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

[29] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

[30] R. Mallik, M. Win, J. Shao, M.-S. Alouini, and A. Goldsmith, "Channel capacity of adaptive transmission with maximal ratio combining in correlated Rayleigh fading," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1124–1133, Jul. 2004.

[31] J. Wang, S. Jin, X. Gao, K.-K. Wong, and E. Au, "Statistical eigenmode-based SDMA for two-user downlink," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5371–5383, Oct. 2012.

[32] E. Abbe, S.-L. Huang, and E. Telatar, "Proof of the outage probability conjecture for MISO channels," in *2012 IEEE Information Theory Workshop (ITW)*, Sep. 2012, pp. 65–69.

[33] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Elsevier/Academic Press, Amsterdam, 2007.

Miltiades C. Filippou (S'12-M'15) was born in Athens, Greece in 1984. He received his Dipl. Eng. degree in Electrical and Computer Engineering from the National Technical University of Athens (NTUA), Greece, in 2007. From November 2007 until July 2009 he was occupied at the Wireless Communications laboratory of the same school as a postgraduate researcher in the area of satellite communications. In July 2014, he received his Ph.D. degree in Electronics and Telecommunications from Telecom ParisTech. Between September 2014 and October 2015, he was a Research Fellow at the Institute for Digital Communications (IDCOM), University of Edinburgh, U.K. Since November 2015 he works as a Standards and Research Engineer at Intel Deutschland GmbH, Neubiberg, Germany. Dr. Filippou is a member of the IEEE and of the Technical Chamber of Greece. His current research interests include: 5G air interface design, millimeter-wave and cooperative communications, as well as spectrum sharing systems.

Paul de Kerret graduated from Télécom Bretagne in France, and obtained a diploma degree in Electrical Engineering and Information Technology from the Munich University of Technology (TUM) in 2010. In December 2013, he obtained his doctorate from Telecom ParisTech. From December 2014 to August 2015, he worked as assistant Professor in Télécom Bretagne. He is currently a research engineer at EURECOM working in a project funded by the European Research Council to investigate distributed coordination problems within the future 5G networks. He has worked in several major European collaborative projects on mobile communications, co-presented several tutorials in IEEE international conferences, and published more than 25 papers in the most selective international IEEE conferences and journals.

David Gesbert (IEEE Fellow) is Professor and Head of the Communication Systems Department, EURECOM, France, where he also heads the Communications Theory Group. He obtained the Ph.D. degree from Ecole Nationale Superieure des Telecommunications, France, in 1997. From 1997 to 1999 he has been with the Information Systems Laboratory, Stanford University. In 1999, he was a founding engineer of Iospan Wireless Inc., San Jose, Ca., a startup company pioneering MIMO-OFDM (now Intel). Between 2001 and 2003 he has been with the Department of Informatics, University of Oslo as an adjunct professor. D. Gesbert has published about 250 papers and several patents all in the area of signal processing, communications, and wireless networks. He was named in the 2014 Thomson-Reuters List of Highly Cited Researchers in Computer Science. Since 2015, he holds an ERC Advanced grant on the topic of "Smart Device Communications". D. Gesbert was a co-editor of several special issues on wireless networks and communications theory, for JSAC (2003, 2007, 2009), EURASIP Journal on Applied Signal Processing (2004, 2007), Wireless Communications Magazine (2006). He served on the IEEE Signal Processing for Communications Technical Committee, 2003-2008. He was an associate editor for IEEE Transactions on Wireless Communications and the EURASIP Journal on Wireless Communications and Networking. He authored or co-authored papers winning the 2015 IEEE Best Tutorial Paper Award (Communications Society), 2012 SPS Signal Processing Magazine Best Paper Award, 2004 IEEE Best Tutorial Paper Award (Communications Society), 2005 Young Author Best Paper Award for Signal Proc. Society journals, and paper awards at conferences 2011 IEEE SPAWC, 2004 ACM MSWiM workshop. He co-authored the book "Space time wireless communications: From parameter estimation to MIMO systems", Cambridge Press, 2006. He is a Technical Program Chair for upcoming IEEE ICC 2017, to be held in Paris.

Tharmalingam Ratnarajah (A'96-M'05-SM'05) is currently with the Institute for Digital Communications, University of Edinburgh, Edinburgh, U.K., as a Professor in Digital Communications and Signal Processing. His research interests include signal processing and information theoretic aspects of 5G wireless networks, full-duplex radio, mmWave communications, random matrices theory, interference alignment, statistical and array signal processing and quantum information theory. He has published over 280 publications in these areas and holds four U.S. patents. He is currently the coordinator of the FP7 projects ADEL (€3.7M) in the area of licensed shared access for 5G wireless networks. Previously, he was the coordinator of the FP7 project HARP (€3.2M) in the area of highly distributed MIMO and FP7 Future and Emerging Technologies projects HIATUS (€2.7M) in the area of interference alignment and CROWN (€2.3M) in the area of cognitive radio networks. Dr. Ratnarajah is a Fellow of Higher Education Academy (FHEA), U.K., and an associate editor of the IEEE Transactions on Signal Processing.

Adriano Pastore was born in Munich (Germany). He received a Diplôme de l'École Centrale Paris in 2006 as well as a Dipl.-Ing. degree in electrical engineering from the Technical University of Munich (TUM) in 2009. From 2009 to 2014, he has been working as a Research Assistant and obtained his Ph.D. from the Department of Signal Theory and Communications at the Universitat Politècnica de Catalunya (BarcelonaTech). He is currently a postdoctoral researcher at Ecole Polytechnique Fédérale de Lausanne (EPFL) in the group of Michael Gastpar. His main topics of interest are in wireless communications, network information theory and channel coding.

George A. Ropokis was born in Athens, Greece, on February 22, 1982. He received the Diploma degree in Computer Engineering and Informatics from the University of Patras, Greece in 2004, the M.Sc. degree in Mobile and Satellite Communications from the University of Surrey, U.K. and the Ph.D. degree in Wireless Communications form the University of Patras in 2010. Between 2012 and 2015 he was affiliated with the Signal Processing and Communications Research Unit of the Computer Technology Institute and Press "Diophantus", Greece. Between 2012 and 2014 he was a Visiting Scientist with the Mobile Communications Department, EURECOM, France. He is now a Research Fellow with CONNECT centre, Trinity College Dublin, Ireland. His research interests are focused on the performance analysis and optimization of Wireless Communications systems with emphasis on Spectrum Sharing systems.