D2D Cooperation to Avoid Instantaneous Feedback in Nonreciprocal Massive MIMO Systems

Samah A. M. Ghanem and Laura Cottatellucci
Eurecom, Sophia Antipolis, France
Email: {samah.ghanem; laura.cottatellucci}@eurecom.fr

1. Introduction

A system where the number antennas at the access point (AP) is much larger than the number of users is referred to as very large antenna array or massive MIMO system [1], [2]. Such a technology is widely believed to be a key enabler of the future 5G networks and fuels intense research activities.

When the channel reciprocity can be exploited, massive MIMO systems have impressive advantages in combatting interference without the costly signalling for cooperation and/or coordination of current cooperative multi-point (CoMP) architectures. In [2] it is shown that massive MIMO systems are not limited by data interference from adjacent cells, allow for a vanishing power per bit as the number of antennas grows, and simple linear detectors such as matched filters can be implemented at the receiver without substantial loss in performance.

However, as well known, an effective design of the downlink communications is strongly sensitive to CSI knowledge. Then, it is of paramount importance the CSI acquisition, which may cause a huge feedback overhead, particularly in fast fading channels with short coherence time. In [2], Marzetta limited the applicability of massive MIMO networks to time division duplex (TDD) mode since heavy load feedback can be avoided by resorting channel reciprocity and adopting open loop feedback.

When uplink-downlink channel reciprocity does not hold, as for example in FDD systems, a closed-loop CSI feedback is required. In a closed loop feedback scheme, each user needs to retransmit an amount of information proportional to the number of antennas at the AP. This task becomes unfeasible during the channel coherence time when the number of antennas grows large. Promising schemes for FDD mode have been proposed in [3], [4], and [5]. Most of these contributions leverage on the reduced rankness of the channel covariance matrices of user terminals in a massive MIMO system. In [6] and [7], the authors pointed out the low rankness of the channel covariance matrices for wide classes of antenna array configurations and exploited it for channel estimation. In [3] and [8], the authors exploit the reduced rankness of the correlation structure of UTs channels in the proposed Joint Spatial Division and Multiplexing (JSDM) scheme. The image of a user channel covariance matrix, whose dimension is substantially lower than the number of antennas at the AP, is referred to as *effective channel subspace* (ECS). The JSDM approach clusters users that share the same ECS and schedules simultaneous transmissions to clusters with orthogonal ECS. The signals for a certain cluster are projected onto the cluster ECS before transmission. Within a cluster, the signal is precoded based on the feedback for the ECS. Further feedback reductions are also possible by restricting the pre-beamformers to a subset of the ECS called *reduced-dimension effective channel subspace.* Recently, in [5], the authors capitalize on D2D communications and allow the users in a cluster to share their CSI, obtaining significant reduction in the CSI feedback. In [9], an iterative compensation algorithm, that reduces the complexity of a two-tier precoding for massive MIMOs has been proposed. Additionally, in [4], the authors introduce a sparse model representation of multi-cell massive MIMO systems and exploit compressive sensing (CS) techniques to reduce the training as well as the feedback overhead for the CSI acquisition at the AP.

Under realistic conditions, with arbitrary configurations of large antenna arrays and randomly located users, the ECSs of different users are in general different and the dimensions of a cluster subspace, union of the ECSs of the cluster users, might be still too large. The selection of a reduced dimension ECS is not obvious and an arbitrary subspace selection can severely impair the system performance [8].

In this paper, we address this issue and we provide a solution that leverages on D2D communications to create virtual MIMOs. As in the JSDM protocol, we cluster n users and project the signals to be transmitted onto the ECS via a beamformer. In contrast to the JSDM that requires instantaneous CSI feedback, we design the precoder as for a point-to-point MIMO system with only statistical CSI of the ECS at the transmitter. At each channel use, n information data streams for a target user in the cluster are precoded and subsequently beamformed for transmission. Each non-target user in the cluster amplifies and forwards the received signal such that the target user receives $n-1$ additional independent versions of the transmit signal to create a virtual MIMO system.

This article is organized as follows. Section 2 describes the system model. In Section 3 we define the proposed solution and derive the corresponding performance in terms of achievable total rate of a cluster. In Section 4 we assess the performance of the proposed system by numerical simulations and compare it with the benchmark JSDM protocol in [8].

The following notation has been used throughout the paper: Boldface uppercase and lowercase letters denote matrices and vectors, respectively. Scalars are in italic. $\mathbf{I}_n$ is the identity matrix of size $n \times n$. The Hermitian operator of a matrix $\mathbf{X}$ is denoted by $\mathbf{X}^H$. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a complex Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. $\mathbb{E}\{\cdot\}$ is the expectation operator; $\mathrm{tr}(\cdot)$ denotes the traces of the matrix argument. Finally, $\mathrm{diag}(\mathbf{v})$ is the square diagonal matrix having the element of vector $\mathbf{v}$ as diagonal elements.

2. System Model

We consider a single-cell massive MIMO system with the access point equipped with M antennas and serving single antenna users. The channel is fast fading and non-reciprocal, for example because the system operates in FDD mode. Let $\mathbf{R}_k$ denote the covariance matrix of user k channel and let us refer to its image as ECS of user k. As in [3], users with almost overlapping ECSs are clustered together. For a cluster $\mathcal{C}$, we introduce the cluster covariance matrix defined as $\mathbf{R}_c = \sum_{k\in\mathcal{C}} \mathbf{R}_k$ and we refer to its image as the cluster ECS. Simultaneous transmissions are scheduled to clusters with orthogonal (or quasi-orthogonal) cluster ECSs. Under strict orthogonality of the cluster ECSs, by projecting the signals meant for a single cluster onto its ECS by beamforming, the cluster signals do not cause interference to the other simultaneously scheduled clusters. Thus, in the following we can focus on a single cluster with n users.

Down-link Transmission

The downlink transmission to a single cluster $\mathcal{C}$ is modeled by:

$$\mathbf{y} = \mathbf{H}_c^H \mathbf{B}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is the n-dimensional complex column vector of received signals at all the cluster users; $\mathbf{s}$ denotes the vector of i.i.d. Gaussian signals with zero-mean and unit-variance; and $\mathbf{n}$ represents the spatially and temporally white additive Gaussian noise (AWGN) with zero-mean and element-wise variance σ_n^2. Finally, $\mathbf{B}$ is the down-link beamformer such that $\mathrm{tr}\{\mathbf{B}\mathbf{B}^H\} = P_{max}$, if P_{max} is the total transmit power constraint. The down-link channel between the access point and the k-th user in the cluster $\mathcal{C}$ is denoted by the M-dimensional complex vector $\mathbf{h}_k$. Therefore,

$$\mathbf{H}_c = \left[\mathbf{h}_1, \mathbf{h}_2, \dots \mathbf{h}_n\right] \in \mathbb{C}^{M\times n} \tag{2}$$

and $\mathbf{R}_k = \mathbb{E}\{\mathbf{h}_k\mathbf{h}_k^H\}$. By leveraging on the low rankness of $\mathbf{R}_c$ and assuming its rank equals b, the cluster covariance matrix can be expressed as

$$\mathbf{R}_c = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}^H \tag{3}$$

where $\widetilde{\mathbf{\Lambda}}$ is the $b\times b$ matrix of the nonzero eigenvalues of $\mathbf{R}_c$ in non increasing order and $\widetilde{\mathbf{U}}$ is the $M \times b$ matrix whose columns are the normalized eigenvectors of $\mathbf{R}_c$. Similarly, $\mathbf{R}_k = \widetilde{\mathbf{U}}_k\widetilde{\mathbf{\Lambda}}_k\widetilde{\mathbf{U}}_k^H$, with analogous meaning for b_k, $\widetilde{\mathbf{\Lambda}}_k$ and $\widetilde{\mathbf{U}}_k$. Since, by construction, the subspace spanned by the column vectors of $\widetilde{\mathbf{U}}_k$ lies into the subspace spanned by the column vectors of $\widetilde{\mathbf{U}}$, then $\widetilde{\mathbf{U}}_k = \widetilde{\mathbf{U}}\mathbf{T}_k$, where $\mathbf{T}_k$ is a $b \times b$ matrix whose i-th column elements are the coefficients of the i-th column of $\widetilde{\mathbf{U}}_k$ in the basis $\widetilde{\mathbf{U}}$. Additionally, the Gaussian vector $\mathbf{h}_k$ can be expressed as $\mathbf{h}_k = \widetilde{\mathbf{U}}_k\widetilde{\mathbf{\Lambda}}_k^{1/2}\widetilde{\mathbf{h}}_k$ being $\widetilde{\mathbf{h}}_k$ a b_k-dimensional vector of zero-mean, unit variance, independent Gaussian elements. Then,

$$\mathbf{H}_c = \widetilde{\mathbf{U}}\left[\mathbf{T}_1\widetilde{\mathbf{\Lambda}}_1^{1/2}\widetilde{\mathbf{h}}_1, \cdots \mathbf{T}_n\widetilde{\mathbf{\Lambda}}_n^{1/2}\widetilde{\mathbf{h}}_n\right] = \widetilde{\mathbf{U}}\mathbf{A}_c, \tag{4}$$

where $\mathbf{A}_c$ is a matrix of Gaussian elements, in general, column-wise correlated. In the following we adopt the notation $\mathbf{a}_{ck} = \mathbf{T}_k\widetilde{\mathbf{\Lambda}}_k^{1/2}\widetilde{\mathbf{h}}_k$.

Finally, by projecting the signals to be transmitted onto the cluster ECS, which implies to structure the matrix $\mathbf{B}$ as $\mathbf{B} = \widetilde{\mathbf{U}}\widetilde{\mathbf{B}}$, being the first operator $\widetilde{\mathbf{U}}$ the projection beamformer, we obtain an equivalent system model in the ECS with reduced dimensions

$$\mathbf{y} = \mathbf{A}_c^H\widetilde{\mathbf{B}}\mathbf{s} + \mathbf{n}. \tag{5}$$

The received signal at user k is given by

$$y_k = \mathbf{a}_{ck}^H\widetilde{\mathbf{B}}\mathbf{s} + \mathbf{n} \tag{6}$$

and the corresponding averaged received power is

$$P_k = \mathrm{tr}\{\mathbf{B}^H\mathbf{R}_k\mathbf{B}\} + \sigma_n^2. \tag{7}$$

Intra-cluster D2D Communications

By D2D communications, the users in a cluster retransmit the received signals in orthogonal time intervals. User ℓ amplifies and forwards its received signal y_ℓ such that its transmitted signal is

$$x_\ell = \sqrt{\frac{P_r}{P_\ell}}y_\ell$$

where P_r is the average transmit power constraint as user ℓ acts as relay. To keep the notation easy, we assume that P_r is equal to all the users.

As likely from physical considerations and the analysis in [7], users within a cluster are closely located and far apart from users in other clusters. Then, we can assume that the simultaneous D2D transmissions in other clusters do not interfere with the intra-cluster transmissions. Thus, the signal received by user k from user ℓ is given by

$$\begin{aligned} r_k &= g_{k\ell}x_\ell + w_k \\ &= \sqrt{\frac{P_r}{P_\ell}}g_{k,\ell}\mathbf{a}_{ck}^H\widetilde{\mathbf{B}}\mathbf{s} + \sqrt{\frac{P_r}{P_\ell}}g_{k\ell}n_\ell + w_k \end{aligned} \tag{8}$$

where $g_{k,\ell}$ is the channel coefficient of the fast fading link from user ℓ to user k, realization of a Gaussian random process with variance $\gamma_{k,\ell}$. Finally, w_k is the additive Gaussian noise with unit variance at user k when it acts as receiver in the D2D communications.

At the end of the relaying phase, user k has n independent received versions of the original transmitted signal and can act as a virtual MIMO. The corresponding system model is given by

$$\mathbf{r}_k = \mathbf{\Gamma}_k\mathbf{A}_c^H\widetilde{\mathbf{B}}\mathbf{s} + \mathbf{z} \tag{9}$$

where $\mathbf{\Gamma}_k$ is an $n\times n$ diagonal matrix given by

$$\begin{aligned} \mathbf{\Gamma}_k = \mathrm{diag}\Bigg[&\sqrt{\frac{P_r}{P_1}}g_{k,1}, \cdots, \sqrt{\frac{P_r}{P_{k-1}}}g_{k,k-1}, 1, \\ &\sqrt{\frac{P_r}{P_{k+1}}}g_{k,k+1} \cdots \sqrt{\frac{P_r}{P_n}}g_{k,n}\Bigg] \end{aligned}$$

and $\mathbf{z}$ is the equivalent Gaussian noise with diagonal covariance matrix, at UT k, given by:

$$\boldsymbol{\Sigma}_k = \text{diag}\left[\frac{P_r}{P_1}\gamma_{k,1}\sigma_n^2+1,\cdots,\frac{P_r}{P_{k-1}}\gamma_{k,k-1}\sigma_n^2+1,\right.$$
$$\left.\sigma_n^2,\frac{P_r}{P_{k+1}}\gamma_{k,k+1}\sigma_n^2+1\cdots\frac{P_r}{P_n}\gamma_{k,n}\sigma_n^2+1\right].$$

3. Design and Analysis of the Proposed System

In this section we detail the proposed massive MIMO system which combats intra-cluster interference via virtual MIMO systems in the cluster. The AP transmits at each channel use information streams to the same final user in the cluster refered to as target user throughout this work. With a simple round robin algorithm, a transmission to each of the users in the cluster is performed each n time intervals. In contrast to the JSDM communication scheme, instantaneous CSI is not fed back to the access point to optimally design a precoder. Additionally, in this work, we assume that also the statistics that characterize the D2D communications are not fed back to the access point. Each user simply feeds back the rate at which its own data have to be transmitted. As detailed later, this rate can be computed locally by each user terminal. Then, the access point is substantially oblivious of the presence of varying virtual MIMO systems at the receiver side and the precoder design at the transmitter can be optimized for the fictitious point-to-point MIMO system in (1) or, equivalently, for the MIMO system in the cluster ECS in (5) by keeping in mind that the global precoder $\mathbf{B}$ consists of the cascade of the precoder $\widetilde{\mathbf{B}}$ in the cluster ECS followed by the beamformer $\widetilde{\mathbf{U}}$.

Although, in general, an optimal design of the system to maximize the rate at the fictitious point-to-point MIMO system would require the transmission of the information over b symbol streams and a careful design of the precoder $\widetilde{\mathbf{B}}$ along the lines of available works in literature for general correlated MIMO channels [10], in this contribution we opt for a suboptimal precoder to keep the exposition simple and concise. The proposed precoder transmits n streams with identical powers along the directions of the eigenvectors with the n highest eigenvalues of the matrix $\mathbf{R}_c$, equivalently

$$\widetilde{\mathbf{B}} = \frac{P_{max}}{n}\begin{bmatrix}\mathbf{I}_n \\ \mathbf{0}_{(b-n)\times n}\end{bmatrix} \qquad (10)$$

where $\mathbf{0}_{(b-n)\times n}$ is a $(b-n)\times n$ matrix of zeros. Thus, (9) reduces to

$$\mathbf{r}_k = \frac{P_{max}}{n}\boldsymbol{\Gamma}_k\widetilde{\mathbf{A}}_c^H\mathbf{s}+\mathbf{z}$$

where $\widetilde{\mathbf{A}}_c$ is the submatrix obtained from $\mathbf{A}_c$ by extracting the first n rows. Then, the achievable rate for the target user k is

$$R_k = \frac{1}{n}\mathbb{E}\left[\log\det\left[\mathbf{I}_n+\frac{P_{max}}{n}\boldsymbol{\Gamma}_k\widetilde{\mathbf{A}}_c^H\widetilde{\mathbf{A}}_c\boldsymbol{\Gamma}_k^H\boldsymbol{\Sigma}_k^{-1}\right]\right]$$

where the factor n^{-1} is due to the fact that user k receives each n channel uses. Note that user k feeds back the value nR_k that should be used for encoding at the AP.

The total achievable rate in cluster $\mathcal{C}$ is given by

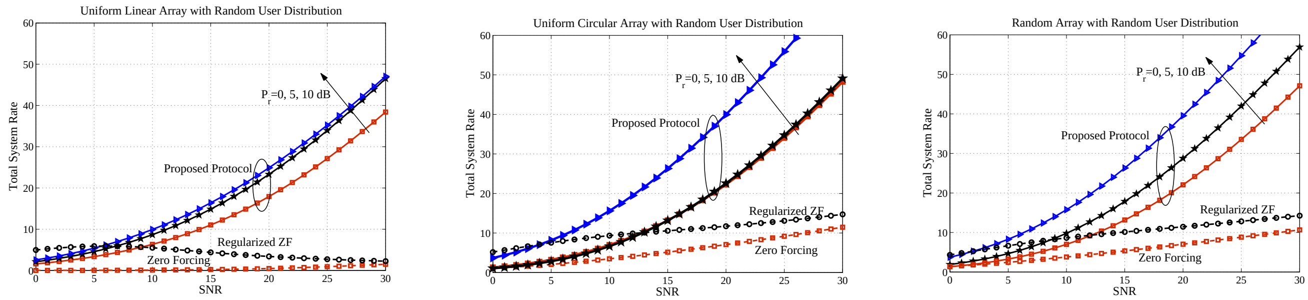
$$R_c = \sum_{k\in\mathcal{C}} R_k.$$

4. Simulation Results

In this section we assess the performance of the proposed communication system by simulations and we compare it with the benchmark JSDM system [3]. In all the simulations we adopt the following setting. We consider a system with $M = 64$ antennas and 16 users divided into 4 clusters with the same number of users, i.e. $n = 4$. At the access point, we adopt various configurations of the antenna array: Uniform Linear Array (ULA), Uniform Circular Array (UCA), and Random Array (RA). The downlink channel is fast fading with Gaussian distributed coefficients. In order to generate the covariance matrices of users' channels, we utilize a one-ring model as in [3] and the same configuration adopted in [3]: the ECSs have size $b \geq r$ with $r = 11$ where r is the number of dominant eigen-modes. In order to better understand the simulation results is relevant to recall that for the ULA the users are randomly distributed over a sector of 180 degrees while for the UCA and RA they are distributed over the full coverage area. Additionally, also the channels between pairs of users are fast fading, independent each other, and follow a complex Gaussian distribution with zero mean and unit variance.

Figures 1(a)-(c) provide a comparison between the proposed massive MIMO system with cooperation via D2D and the non-cooperative JSDM system. All the figures show the achievable total rate of the system as a function of the SNR in the downlink transmission between the access point and the user terminals. For the JSDM benchmark system, we show the performance (dashed lines) in the two cases when a zero forcing (ZF, square markers) and a regularized zero forcing (RZF, circles markers) precoding are implemented at the access point. No penalty is imposed to account for the cost of the feedback, which is essential to properly design ZF and RZF precoding. Both precoders are designed under the assumption of perfect CSI. It is worth recalling that the feedback requires a relevant portion of the useful bandwidth which can be dominant compared to the bandwidth allocated to the information when the dimensions of the ECSs are large and/or a high accuracy for the CSI is required. The performance of the proposed system is plotted in solid lines for various values of the average transmit powers at the user terminals when they act as relays, more specifically, for $P_r = 0$ dB (square markers), $P_r = 5$ dB (star markers), and $P_r = 10$ dB (triangle markers).

Since the clusters in the system with ULA antenna configuration suffer from a higher inter-cluster interference than the other systems, for this antenna configuration we obtain the poorest performance and, not surprisingly, the RZF has poorer performance at high SNR than at lower SNR. Intercluster interference is less significant for the proposed system since we adopt a reduced ECS of minimum size. For all the considered antenna configurations and the range of SNR of practical interest, the proposed system outperforms



(a) (b) (c)

Figure 1: Average Total Rate vs SNR in downlink for the proposed system (solid lines) with varying average relay powers (square marker $P_r = 0$ dB, star marker $P_r = 5$ dB, triangle marker $P_r = 10$ dB) and the benchmark JSDM system (dashed lines) with zero forcing (square marker) and regularized zero forcing (circle marker). Users randomly distributed in the coverage area and (a) ULA, (b) UCA, and (c) RA antenna configurations.

the benchmark system with ZF beamforming, even at very low levels of P_r, the averaged transmitted power at the relays. The comparison to the benchmark system with RZF beamforming is more critical: while at high SNR the proposed system has always better performance, at low SNR the same performance of the benchmark are attained at the expenses of higher transmit powers in the relaying phase. However, it is worth recalling that our plots do not account for the bandwidth required by the feedback. Additionally, this work does not explore the improvements of the proposed system that can be obtained by adopting reduced ECS of larger size and optimal precoding matrices. These further analyses are left for further studies.

Acknowledgements

This work was supported by Huawei France Research Center.

References

[1] T. Marzetta, "How much training is required for multiuser MIMO?" in *Fortieth Asilomar Conference on Signals, Systems and Computers*, Oct 2006, pp. 359–363.

[2] ——, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov 2010.

[3] J. Nam, J.-Y. Ahn, A. Adhikary, and G. Caire, "Joint Spatial Division and Multiplexing: Realizing massive MIMO gains with limited channel state information," in *Information Sciences and Systems (CISS), 2012 46th Annual Conference on*, Mar 2012, pp. 1–6.

[4] X. Rao and V. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, Jun 2014.

[5] H. Yin, L. Cottatellucci, and D. Gesbert, "Enabling massive MIMO systems in the FDD mode thanks to D2D communications," *Asilomar Conference on Signals, Systems and Computers*, Nov 2014.

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

[7] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with interference in distributed large-scale MIMO systems: A statistical approach," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 942–953, Oct 2014.

[8] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, Oct 2014.

[9] J. Chen and V. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks," *IEEE Journal on Selected Areas in Communications,*, vol. 32, no. 6, pp. 1230–1238, Jun 2014.

[10] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers ond noise: A (not so) large N analysis," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2545–2561, Oct 2003.

Samah A. M. Ghanem was born in Beirut, Lebanon, 1979. She received her BS in Electrical Engineering from An-Najah National University, Palestine, in 2001, the M.Sc. in Scientific Computing from Birzeit University, Palestine in 2009, and the Ph.D. in Telecommunications Engineering (summa cum laude) from the University of Aveiro (MAP-Tele), Portugal in 2013. She worked for the industry between 2001 to 2008, and at Ericsson LMI R&D, Ireland in 2008. She joined the Mobile Communications Dept. at Eurecom, Sophia Antipolis, France as a postdoctoral researcher in March, 2014. She was a visiting researcher at Aalborg University, Denmark in 2012 and at the University of Duisburg-Essen, Germany in 2008.

Laura Cottatellucci (S'01, M'07) is assistant professor at EURECOM since December 2006. Prior to joining EURECOM, she was research fellow at University of South Australia, Australia (Jan.-Nov. 2006) and in INRIA, Sophia Antipolis, France (Oct.-Dec. 2005). She obtained the PhD from the Technical University of Vienna, Austria (2006) and the Master degree from La Sapienza University, Rome, Italy (1995). After specialization in networking at the School for Advanced Studies Guglielmo Reiss Romoli (1996, Italy), she worked in Telecom Italia (1995-2000) as responsible of industrial projects. From April 2000 to September 2005 she worked as senior research in FTW Austria. She held visiting appointments at EURECOM, France, FAU Erlangen-Nürnberg, Germany, and NTNU, Trondheim, Norway.