

On the Energy-Efficiency of Hybrid Analog-Digital Transceivers for Single- and Multi-carrier Large Antenna Array Systems

Christos G. Tsinos *Member, IEEE*, Sina Maleki *Member, IEEE*,
Symeon Chatzinotas *Senior Member, IEEE* and Björn Ottersten *Fellow, IEEE*

Abstract—Hybrid Analog-Digital transceivers are employed with the view to reduce the hardware complexity and the energy consumption in millimeter wave/large antenna array systems by reducing the number of their Radio Frequency (RF) chains. However, the analog processing network requires power for its operation and it further introduces power losses, dependent on the number of the transceiver antennas and RF chains, that have to be compensated. Thus, the reduction in the power consumption is usually much less than it is expected and given that the hybrid solutions present in general inferior spectral efficiency than a fully digital one, it is possible for the former to be less energy efficient than the latter in several cases. Existing approaches propose hybrid solutions that maximize the spectral efficiency of the system without providing any insight on their energy requirements/efficiency. To that end, in this paper, a novel algorithmic framework is developed based on which energy efficient hybrid transceiver designs are derived and their performance is examined with respect to the number of RF chains and antennas. Solutions are proposed for fully and partially connected hybrid architectures and for both single- and multi-carrier systems under the Orthogonal Frequency Division Multiplexing (OFDM) modulation. Simulations and theoretical results provide insight on the cases where a hybrid transceiver is the most energy efficient solution or not.

Index Terms—mmWave Communications, hybrid analog and digital transceivers, MIMO, energy-efficiency

I. INTRODUCTION

The congestion in the lower parts of the electromagnetic spectrum along with the rapidly increasing data rate demands for mobile services, forces the wireless communications industry to explore systems adapted to frequencies within the so-called millimeter Wave (mmWave) band [2]–[5]. For example, mmWave band communications have already been employed with success in indoor [4], [6] and fixed outdoor (e.g. point-to-point backhaul links) [7], [8] environments. Recent advances in the hardware for mmWave transceivers have motivated the employment of the latter for outdoor cellular systems in order to provide large amounts of available spectrum areas in 5G systems.

The authors are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg 2721, (Emails: {christos.tsinos,sina.maleki,symeon.chatzinotas,bjorn.ottersten}@uni.lu).

Part of this work was presented in International Conference on Communications (ICC-2017) [1].

This work was supported by FNR, Luxembourg under the CORE projects “ECLECTIC”, “SeMiGod” and “SATSENT” and by the European Commission through the “SANSa” project (ICT-645047).

Unfortunately, the development of transceivers that function in the mmWave band is a very challenging task. Due to their nature, mmWave signals suffer from severe propagation loss, penetration loss and rain fading compared to signals in lower frequencies [9]. On the contrary, due to the short wavelength of mmWave frequencies, denser packing of more antennas in the transceiver ends is possible. Thus, a mmWave transceiver may employ large array structures for providing high beamforming gains. Furthermore, the large-scale transmit and receive arrays enable multiple stream transmission via the application of pre/post-coding techniques that result in significant improvements in the system’s spectral efficiency.

While the existing digital pre/post-coding techniques, developed in the past years for lower frequency MIMO systems, are independent of the carrier frequency, they cannot be applied in systems that employ large antenna arrays due to high demands in hardware complexity and power consumption. A fully digital transceiver requires a dedicated Radio Frequency (RF) chain per antenna which includes a number of different electronic elements (e.g., Digital-to-Analog/Analog-to-Digital converters) that may require prohibitive hardware complexity and power consumption as the antenna array size increases. Thus, a digital only mmWave transceiver seems to be currently impractical and alternative approaches must be sought.

At first, analog only beamforming approaches [7], [10], [11] were considered as a solution to the problem. The core of these techniques is a network of analog phase shifters [12], [13], driven via a single RF chain, that imposes constant modulus constraints on the beamformer. The latter is highly desirable from a hardware complexity/power consumption perspective. Unfortunately, analog only approaches cannot support multi-stream communication, and thus, they usually perform poorly compared to the fully digital approaches. In an alternative approach, the phase shifters are replaced by analog switches [14]–[16] which are simpler components. However, this comes at the cost of loss in array gain which again results in performance degradation, especially for highly correlated channels, like the ones experienced in the mmWave band.

The performance can be significantly improved if a hybrid Analog/Digital (A/D) transceiver architecture is employed. This approach is based on a two stage setup that consists of a low dimensional digital pre-coder applied in the BaseBand (BB) and an analog beamformer applied in the RF domain [17]. The RF processing part is usually implemented via a network of variable phase shifters. This architecture can

be applied to systems with limited number of RF chains and achieve performance close to the one of a fully digital approach under some conditions. Two common variations of the hybrid techniques are the fully connected one, where each of the RF chains is connected to all of the available antennas and the partially connected one, where each RF chain is connected to only some of them. The concept of hybrid pre-coding was firstly known as the so-called ‘‘antenna soft selection’’ for point-to-point MIMO systems [18], [19].

In mmWave bands, hybrid transceivers were first introduced in [17], [20]. Those approaches aim at hybrid solutions that maximize the spectral efficiency of the system when a fully connected architecture is assumed. It is further shown that the aforementioned problem is closely approximated by minimizing the Frobenius norm of the difference of the fully digital solution to the product of the analog-digital counterparts of the hybrid pre-coder. The RF pre-coder is designed based on a pre-determined codebook and thus, the resulting problem can be solved via the recent advances in sparse reconstruction signal processing literature [21], [22]. The performance of this approach is in general good when low rank channels are involved and when the hybrid transceivers have strictly more RF chains than the transmitting streams. Codebook free hybrid transceiver designs with improved performance were also developed in [23], [24]. In [23], the techniques were developed for both the fully and the partially connected architecture and for single- and multi-carrier systems that are based on the Orthogonal Frequency Division Multiplexing (OFDM) technique. Hybrid pre-coders for the partially connected architecture were also considered in [25]–[27]. Furthermore, in [28]–[30] hybrid transceivers were developed for the OFDM case and in [31], [32] for cognitive radio systems.

The above hybrid transceivers are mainly designed such that the system’s spectral efficiency is maximized. While the gains in the power consumption from reducing the number of RF chains is evident, for the hybrid architectures, they require additional power to operate the phase shifting network. Moreover, the phase shifting network along with the rest of the RF hardware (power splitters and combiners) present power losses that need to be compensated at the power amplifiers in the transmitter side and by the low noise amplifiers at the receiver’s side. Moreover, given that the spectral efficiency of a hybrid architecture is in general inferior than that of the corresponding fully digital solution, it is possible for the former to become less energy efficient than the latter after an increase in the number of its RF chains, that is its spectral efficiency per consumed Joule will be smaller. Some preliminary results regarding the study of the energy efficiency in hybrid architectures has already appeared in literature [23], [33]–[36]. Nevertheless, a framework that enables the detailed analysis and the design of a hybrid transceiver with respect to the energy efficiency performance metric, as it is extensively done in the literature of fully digital transceivers [37]–[45], has yet to appear.

In detail, the contributions of the present paper are as follows. Hybrid A/D transceivers are developed for fully and partially connected architectures and for single carrier and OFDM systems that maximize the energy efficiency of the sys-

tem. The proposed approaches derive the pre-coding and post-coding (combining) matrices by decoupling the transmitter-receiver optimization problem. Therefore, the precoder is first developed such that the ratio of the mutual information to the power consumption is maximized. Then for the derived pre-coder, a Minimum Mean Square Error (MMSE) post-coding matrix is derived. The proposed transceiver designs are codebook free and involve the solution of non-convex optimization problems. The developed algorithms are based on the so-called Alternating Direction of Multipliers Method (ADMM) [46] which is known for its good behaviour in several non-convex optimization problems [47]. Furthermore, the energy consumption of the different hybrid approaches is analyzed in detail and upper bounds on the number of RF chains beyond which a hybrid architecture is always less energy efficient than an fully digital one are established for both the fully and the partially connected architectures. The performance of the proposed energy efficient transceivers is evaluated via extensive simulations and compared to the one of a fully digital energy efficient transceiver in typical channel conditions. Both the theoretical and the numerical results provide insight on the impact of the number of RF chains and the number of antennas on the energy efficiency of the different hybrid solutions and the conditions under of which they become less energy efficient than the corresponding fully digital one.

The remainder of the paper is organized as follows. In Section II, the system model and the problem formulation are described. In Section III, the power consumption of the different architectures is analyzed. Section IV describes the fully and the partially connected hybrid transceiver designs for a single carrier system. In Section V, both the fully and the partially connected architectures are extended to the OFDM case. In section VI, the fully digital solution is briefly described and the upper bounds on the number of RF chains are derived. Section VII presents the numerical results and Section VIII concludes the work.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Let us assume a single user $R_x \times T_x$ MIMO system which is equipped with $N_t \ll T_x$ RF chains on the transmitter’s and $N_r \ll R_x$ ones on the receiver’s side, respectively. The system aims at transmitting a number of $N_s \leq \min\{N_t, N_r\}$ streams between its two ends. To that end, a $T_x \times N_s$ hybrid pre-coding matrix is applied at the transmitter, given by $\mathbf{F} = \mathbf{F}_{RF}\mathbf{F}_{BB}$ where \mathbf{F}_{RF} is the $T_x \times N_t$ analog RF precoder implemented via a phase shifting network and \mathbf{F}_{BB} is the $N_t \times N_s$ digital BB one. In a similar manner, a $R_x \times N_s$ hybrid post-coding matrix is applied at the receiver, given by $\mathbf{W} = \mathbf{W}_{RF}\mathbf{W}_{BB}$ where \mathbf{W}_{RF} is the $R_x \times N_r$ analog post-coding matrix and \mathbf{W}_{BB} is the $N_r \times N_s$ digital BB one, respectively.

The analog counterparts of the pre-coder and the post-coder, matrices \mathbf{F}_{RF} and \mathbf{W}_{RF} are mapping the signals from the RF chains to the antennas via the phase shifting network. Each phase shifter should satisfy unit modulus constraints, that is $|\mathbf{F}_{RF}(m, n)| = |\mathbf{W}_{RF}(m, n)| = 1, \forall \mathbf{F}_{RF}(m, n) \neq 0$ & $\mathbf{W}_{RF}(m, n) \neq 0$, where $\mathbf{F}_{RF}(m, n)$ and $\mathbf{W}_{RF}(m, n)$

are the elements at the m th row - n th column of matrices \mathbf{F}_{RF} and \mathbf{W}_{RF} , respectively and $|\cdot|$ is the modulus of a complex number. Moreover, the hybrid transceiver architecture can be categorized into fully connected and partially connected one, as shown in Figs. 1-2 for the transmitter side. In the fully connected architecture, the signal from each RF chain is connected via the phase shifting network to all the available antennas. To that end, the implementation of a fully connected architecture requires analog processing networks of $T_x N_t$ and $R_x N_r$ phase shifters at the transmitter and the receiver, respectively. On the other hand, in a partially connected architecture, the signal of each RF chain is connected to only $\lceil T_x/N_t \rceil$ and $\lceil R_x/N_r \rceil$ antennas at the transmitter and the receiver, respectively, where $\lceil \cdot \rceil$ function rounds a real number up to the nearest integer. Due to this type of architecture, matrix \mathbf{F}_{RF} presents a block diagonal structure, depicted as,

$$\mathbf{F}_{RF} = \begin{bmatrix} \mathbf{f}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{f}_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \mathbf{f}_{N_t} \end{bmatrix}, \quad (1)$$

where \mathbf{f}_i is a $\lceil \frac{T_x}{N_t} \rceil \times 1$ vector of unit modulus elements. The matrix \mathbf{W}_{RF} follows a similar block diagonal structure comprised by $\lceil \frac{R_x}{N_r} \rceil \times 1$ size vectors of unit modulus elements, denoted by \mathbf{g}_i , $1 \leq i \leq N_r$. Thus, the implementation of a partially connected architecture requires only a number of T_x and R_x phase shifters at the transmitter and the receiver analog processing networks, respectively. Therefore, a partially connected architecture has less hardware complexity and power consumption compared to a fully connected one, though these advantages come at the cost of performance loss.

The received signal \mathbf{y} at the receiver before and after post-coding, assuming a narrow-band block fading propagation channel¹ is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (2)$$

$$\mathbf{y}' = \mathbf{W}^H \mathbf{y} \quad (3)$$

where \mathbf{H} is the $R_x \times T_x$ channel matrix between the two transceiver ends, \mathbf{x} is the $N_s \times 1$ transmitted symbol vector under the assumption that $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \sigma_x^2 \mathbf{I}_{T_x}$, $\mathbb{E}\{\cdot\}$ is the expectation operator, σ_x^2 is the variance of the symbols transmitted by the transmitter, \mathbf{I}_{T_x} is the $T_x \times T_x$ identity matrix, and \mathbf{n} is i.i.d. complex Gaussian noise modeled as $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{R_x})$.

The mmWave channel presents limited spatial diversity due to the high free space pathloss and high antenna correlation due to the dense packing of large number of antennas. A geometrical channel model captures better these characteristics of the mmWave band. Thus, by assuming a Uniform Linear Array² (ULA) of antennas at the transceiver's ends and N_p propagation paths between them, the discrete narrow-band

¹The wide-band case is presented in Sec. V where the extension of the proposed techniques to the OFDM systems is considered.

²As it will be evident latter, the proposed techniques are applicable under any arbitrary array geometry, though we follow the ULA approach for simplicity and consistency to the existing techniques.

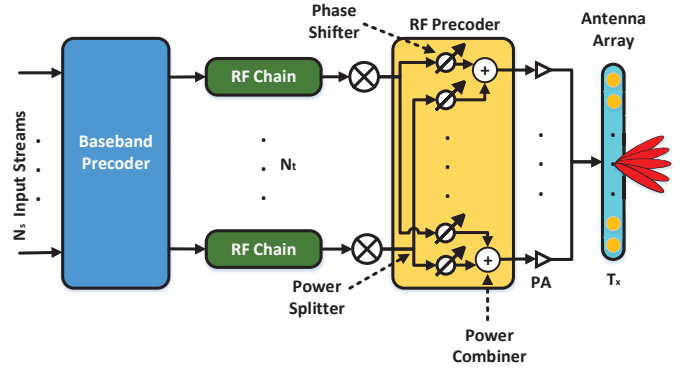


Fig. 1: A fully connected hybrid A/D transmitter where each RF chain is connected to all the antennas.

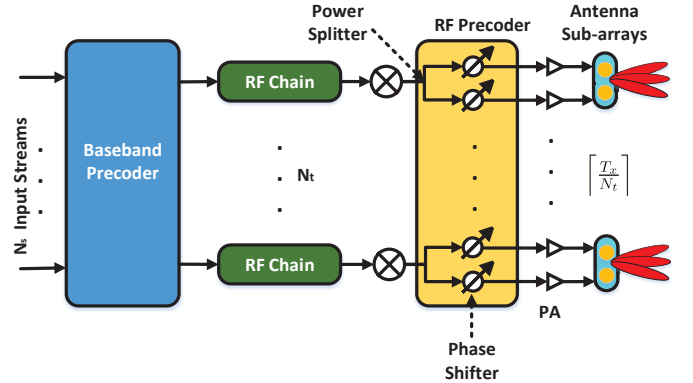


Fig. 2: A partially connected hybrid A/D transmitter where each RF chain is connected to a subset of the antennas.

channel is given by [24] [48]

$$\mathbf{H} = \sqrt{\frac{T_x R_x}{N_p}} \sum_{l=1}^{N_p} \alpha_l \mathbf{a}_r(\phi_l^r) \mathbf{a}_t(\phi_l^t)^H, \quad (4)$$

where $\alpha_l \sim \mathcal{CN}(0, 1)$ is the complex gain of the l th path, ϕ_l^r and ϕ_l^t are the azimuth angles of arrival and departure, respectively, that are uniformly distributed in $[0, 2\pi)$. Vectors $\mathbf{a}_r(\phi_l^r)$ and $\mathbf{a}_t(\phi_l^t)^H$ are the normalized receive and transmit response vectors at the azimuth angle of arrival and departure ϕ_l^r and ϕ_l^t , respectively. Furthermore, the response of an ULA of T antenna elements is given by

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{T}} \left[1, e^{jk d \sin(\phi)}, \dots, e^{j(T-1)k d \sin(\phi)} \right], \quad (5)$$

where $k = \frac{2\pi}{\lambda}$, λ is the carrier wavelength and d is the inter-element spacing. It is assumed that the channel matrix is known at both the transceiver's ends. In general, the channel estimation in hybrid architectures is a challenging task, since a baseband estimator could access only low-dimensional projections of the channel. To that end, recent advances in literature [20] exploit the relative poor scattering nature of channels in the mmWave band in order to develop efficient estimation techniques for the parametric channel model in (4) via using tools from the compressive sensing theory [21]. The design of efficient channel estimators for hybrid transceivers is still an open problem and a very active research field in

the area of communications. After the channel estimation is done, the receiver returns the required CSI to the transmitter via efficient feedback techniques [17], [49].

The aim of the present paper is to derive the optimal pre/post-coding matrices that maximize the energy efficiency of the hybrid transceiver subject to constraints on the total transmission power (P_{max}) and on the structure of the analog counterparts, \mathbf{F}_{RF} and \mathbf{W}_{RF} (i.e. unit modulus entries and block diagonal structure on the partially connected architectures). A common performance metric related to the energy efficient design is the ratio of the achieved spectral efficiency to the consumed power, given for a transceiver by [37], [41]

$$E(\mathbf{F}, \mathbf{W}) = \frac{\log_2 \det(\mathbf{I}_{L_s} + \mathbf{R}_n^{-1} \mathbf{W}^H \mathbf{H} \mathbf{F} \mathbf{F}^H \mathbf{H}^H \mathbf{W})}{P(\mathbf{F})} \quad (6)$$

where $\mathbf{R}_n = \sigma_n^2 \mathbf{W}^H \mathbf{W}$, $\det(\cdot)$ is the determinant of a matrix and $P(\mathbf{F})$ is the consumed power in that transceiver, analyzed in Sec. III.

Therefore, the energy efficient design of a transceiver requires the derivation of the optimal pre/post-coding matrices \mathbf{F} and \mathbf{W} by solving the following optimization problem

$$(\mathcal{P}_1) : \max_{\mathbf{F}, \mathbf{W}} E(\mathbf{F}, \mathbf{W}) \\ \text{s.t. } \|\mathbf{F}\|_F^2 \leq P_{max} \ \& \ \mathbf{F} \in \mathcal{S}_T \ \& \ \mathbf{W} \in \mathcal{S}_R, \quad (7)$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix, and \mathcal{S}_T - \mathcal{S}_R are the sets that impose structural constraints on the precoder and postcoder, respectively, related to the hybrid architectures.

Prior presenting the hybrid pre/post-coding solutions, the power consumption of the electronic elements of each transceiver type (digital only, fully connected and partially connected hybrid) is analyzed in detail in the following section.

III. POWER CONSUMPTION ANALYSIS OF THE DIFFERENT TRANSCEIVER ARCHITECTURES

From (6), it is evident that the power consumption of the different architectures is required for evaluating their energy efficiency. An exact computation of the dissipated power is a very difficult task and thus, our analysis is based on approximate power consumption models which are commonly adopted in relevant literature works [33], [50].

Let us start with the case of a fully digital transceiver. Based on the analysis given in [50], its power consumption $P_D(\mathbf{F})$ can be approximated by

$$P_D(\mathbf{F}) \approx \|\mathbf{F}\|_F^2 / \eta + T_x P_{RF}^T + 2P_{LO} + R_x P_{RF}^R + R_x P_{LNA} \\ = \|\mathbf{F}\|_F^2 / \eta + \mu_D. \quad (8)$$

where η is the efficiency of the Power Amplifiers (PA), $P_{RF}^T = P_{DAC} + P_{mix} + P_{fil}$ and $P_{RF}^R = P_{ADC} + P_{mix} + P_{fil}$ is the power consumed in a transmitter's and a receiver's RF chain, respectively, P_{DAC} , P_{mix} , P_{fil} , P_{ADC} and P_{LNA} is the power consumed by the Digital-to-Analog Converter (DAC), the mixer, filters, the Analog-to-Digital converter (ADC) and the Low Noise Amplifier (LNA), respectively.

For the hybrid transceivers, the power consumption on the electronic components of the RF chains is significantly

reduced due to the in general reduced number of the latter. However, there is power consumed for the operation of the phase shifting network and power losses due to the RF equipment that has to be compensated. For the fully connected case, losses are introduced in the RF domain due to the power splitters, the phase shifters and the power combiners (Fig. 1). on the transmitter side these losses were quantified based on the S-parameter representation [51] of the RF hardware component and it was shown to be equal to $P_{Loss}^{HFC} = L_S T_x L_{PS} L_C N_t$, where L_S , L_{PS} and L_C are the static power losses due to the splitters, phase shifters and combiners. In the receiver's side, the corresponding power losses are compensated by introducing $R_x(N_r + 1)$ LNAs, as described in [33]. A partial connected architecture exhibits less losses in the RF domain compared to a fully connected one. This is the case since, at the transmitter's side, the splitters divide the signal of each RF chain into only $\lceil T_x/N_t \rceil$ equal power outputs (T_x for the fully connected case) and no combiners prior feeding the signal to the antennas are used (Fig. 2). In a similar manner, at the receiver's side, the antenna signals are fed directly to the R_x phase shifters in a partially connected architecture, while in a fully connected one, splitters divide that signal into equal outputs that then are fed to the $R_x N_r$ network of phase shifters. Therefore, by following the analysis of [34], one may show that the power that needs to be compensated in the transmitter side for the partially connected architecture is given by $P_{Loss}^{HPC} = L_S \lceil T_x/N_t \rceil L_{PS}$. In a similar way, on the receiver side the corresponding power losses are compensated by employing R_x LNAs. Thus, by following the discussion given in the present paragraph we conclude that the power consumption of the fully connected P_{FCH} and the partially connected P_{PCH} architectures, is respectively approximated by

$$P_{FCH}(\mathbf{F}_{RF}, \mathbf{F}_{BB}) \approx \|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 / \eta + N_t P_{RF}^T + 2P_{LO} \\ + N_r P_{RF}^R + P_{Loss}^{HFC} + R_x(N_r + 1) P_{LNA} \\ + (T_x N_t + R_x N_r) P_{Ps} = \|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 / \eta + \mu_{HFC}, \quad (9)$$

$$P_{PCH}(\mathbf{F}_{RF}, \mathbf{F}_{BB}) \approx \|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 / \eta + N_t P_{RF}^T + 2P_{LO} \\ + N_r P_{RF}^R + P_{Loss}^{HPC} + R_x P_{LNA} + (T_x + R_x) P_{Ps} \\ = \|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 / \eta + \mu_{HPC}. \quad (10)$$

The power consumption reduction due to the employment of a partially connected architecture over a fully connected one can be seen by comparing (9) to (10). This reduction mainly stems from the fact that the first one requires only $T_x + R_x$ phase shifters while the second one requires $T_x N_t + R_x N_r$ (Sec. II). In addition, the partially connected architecture requires less power splitters and combiners compared to the fully connected one and thus, in the former, less power losses need to be compensated, as explained above.

From a first comparison of (8)-(10), one may see that there is a potential to reduce significantly the power consumption of a transceiver by resorting to a hybrid solution, at least when the latter is based on a relative small number of RF chains. A final point here that we need to discuss is regarding the computational complexity required for the derivation of the pre/post-coding matrices and how it affects the power

consumption. In general, the hybrid structures usually require high complexity iterative techniques in order to derive the solutions to the optimization problems via which the required pre/post-coders are derived. Thus, one may argue that the power consumed by these iterations could have a significant impact on the power consumption and further on the energy efficiency of the system, especially in cases where the fully digital transceiver is based on rather simple form solutions that require low computational complexity for their derivation. Thus, for a more fair comparison, the power consumed for the computation of the pre/post-coding matrices can be modeled and added in terms μ_X , $X = \{HFC, HPC\}$ as in [42]

$$\tilde{\mu}_X = \mu_X + \frac{C_X}{LD}, \quad (11)$$

where C_X is the computational complexity of the corresponding technique expressed in the number of the required floating point operations (flops), L is the coherence block length (number of symbols) and D is the transceiver's computational efficiency in flops/W. In Sec. VII, numerical results that are based on the modified $\tilde{\mu}_X$ are also presented, in order to provide insight on the impact of the computational complexity on the power consumption of the proposed techniques.

IV. ENERGY EFFICIENT HYBRID A/D TRANSCEIVERS FOR SINGLE CARRIER SYSTEMS

In this section, energy efficient hybrid transceiver designs are developed, based on the fully and the partially connected architectures. Let us start by using (6) and (9)-(10) to extend (\mathcal{P}_1) in the hybrid architectures' case, that is

$$\begin{aligned} (\mathcal{P}_2) : \quad & \max_{\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{W}_{RF}, \mathbf{W}_{BB}} E_X(\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{W}_{RF}, \mathbf{W}_{BB}) \\ \text{s.t.} \quad & \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \leq P_{max} \quad \& \\ & \mathbf{F}_{RF} \in \mathcal{X}_T \quad \& \quad \mathbf{W}_{RF} \in \mathcal{X}_R, \end{aligned} \quad (12)$$

where $X = HFC$ and $\mathcal{X}_T = \mathcal{X}_R = \mathcal{F}$ for the fully connected architecture, $X = HPC$, $\mathcal{X}_T = \mathcal{F}'_T$ and $\mathcal{X}_R = \mathcal{F}'_R$ for the partially connected one, the set of matrices \mathbf{A} with unit modulus entries $\mathbf{A}(m, n)$ is defined as $\mathcal{F} = \{\mathbf{A} \in \mathbb{C}^{M \times N} \mid |\mathbf{A}(m, n)| = 1, 1 \leq m \leq M, 1 \leq n \leq N\}$, the sets of unit modulus block diagonal matrices are defined as $\mathcal{F}'_T = \{\mathbf{A} = \text{diag}(\mathbf{f}_m), 1 \leq m \leq N_t \mid \mathbf{f}_m(l) \in \mathbb{C} \quad \& \quad |\mathbf{f}_m(l)| = 1, 1 \leq l \leq \lceil T_x/N_t \rceil\}$ and $\mathcal{F}'_R = \{\mathbf{A} = \text{diag}(\mathbf{g}_m), 1 \leq m \leq N_r \mid \mathbf{g}_m(l) \in \mathbb{C} \quad \& \quad |\mathbf{g}_m(l)| = 1, 1 \leq l \leq \lceil R_x/N_r \rceil\}$ for the transmitter and the receiver side respectively, $\text{diag}(\cdot)$ denotes a diagonal matrix, and E_X is given by

$$E_X(\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{W}_{RF}, \mathbf{W}_{BB}) = \frac{R(\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{W}_{RF}, \mathbf{W}_{BB})}{\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2/\eta + \mu_X}, \quad (13)$$

where $R(\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{W}_{RF}, \mathbf{W}_{BB}) = \log_2 \det(\mathbf{I}_{L_s} + \tilde{\mathbf{R}}_n^{-1} \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{F}_{BB}^H \mathbf{F}_{RF}^H \mathbf{H}^H \mathbf{W}_{RF} \mathbf{W}_{BB})$ and $\tilde{\mathbf{R}}_n = \sigma_n^2 \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{W}_{RF} \mathbf{W}_{BB}$.

Problem (\mathcal{P}_2) is in general intractable even for the fully digital transceiver case [52] and the common approach is to temporally decouple the designs at the transmitter and

the receiver. To that end, the optimal pre-coding matrix $\mathbf{F} = \mathbf{F}_{RF}\mathbf{F}_{BB}$ is designed by assuming Gaussian signaling and replacing in (13) the spectral efficiency (numerator) with the system's mutual information. Therefore, the corresponding optimization problem is given by

$$\begin{aligned} (\mathcal{P}_3) : \quad & \max_{\mathbf{F}_{RF}, \mathbf{F}_{BB}} \frac{\log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{F}_{BB}^H \mathbf{F}_{RF}^H \mathbf{H}^H)}{\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2/\eta + \mu_X} \\ \text{s.t.} \quad & \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \leq P_{max} \quad \& \quad \mathbf{F}_{RF} \in \mathcal{X}_T. \end{aligned} \quad (14)$$

On the receiver side, a low complexity linear design is employed via designing the post-coding matrices based on the Minimum Mean Square Error (MMSE) criterion, that is

$$\begin{aligned} (\mathcal{P}_4) : \quad & \min_{\mathbf{W}_{RB}, \mathbf{W}_{BB}} \mathbb{E} \{ \|\mathbf{x} - \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{y}\|_F^2 \} \\ \text{s.t.} \quad & \mathbf{W}_{RF} \in \mathcal{X}_R. \end{aligned} \quad (15)$$

Problems (\mathcal{P}_3) - (\mathcal{P}_4) involve non-convex cost functions and non-convex constraint sets, as well and thus are non-convex and in general hard to solve. In the following sub-sections, efficient solutions are derived by employing the so-called ADMM [47] which has been applied successfully to some non-convex cases that involve non-convex functions (e.g. Non-Negative Matrix Factorization) or non-convex sets of constraints (e.g. Least Squares Regressor Selection, Factor Model Fitting). From the numerical results presented in the simulations section (Sec. VII), the proposed ADMM based solutions achieve satisfactory performance for also the case of the non-convex optimization problems considered in the present paper.

A. Pre-coding Matrix Design

In this subsection the solution to problem (\mathcal{P}_3) is given. Let us first use the auxiliary variable $\xi = \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2$ to reformulate the previous problem as follows

$$\begin{aligned} (\mathcal{P}_5) : \quad & \max_{\mathbf{F}_{RF}, \mathbf{F}_{BB}, \xi} \frac{\log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{F}_{BB}^H \mathbf{F}_{RF}^H \mathbf{H}^H)}{\xi/\eta + \mu_X} \\ \text{s.t.} \quad & \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = \xi \quad \& \quad \mathbf{F}_{RF} \in \mathcal{X}_T \quad \& \\ & 0 \leq \xi \leq P_{max}. \end{aligned} \quad (16)$$

By inspecting (\mathcal{P}_5) , one may see that if the value of ξ is fixed, optimizing the energy efficiency is equivalent to simply maximizing the mutual information of the system, that is matrices \mathbf{F}_{RF} and \mathbf{F}_{BB} can be derived by solving the following optimization problem

$$\begin{aligned} (\mathcal{P}_6) : \quad & \max_{\mathbf{F}_{RF}, \mathbf{F}_{BB}} \log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{F}_{BB}^H \mathbf{F}_{RF}^H \mathbf{H}^H) \\ \text{s.t.} \quad & \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = \xi \quad \& \quad \mathbf{F}_{RF} \in \mathcal{X}_T. \end{aligned} \quad (17)$$

The solution to problem (\mathcal{P}_5) requires also the derivation of the optimal value of ξ . From (\mathcal{P}_6) , it is evident that the pre-coding matrices \mathbf{F}_{RF} and \mathbf{F}_{BB} are dependent on the different values of ξ . Thus, the derivation of the optimal ξ value is not a trivial task and we resort to a one-dimensional search method

for the solution (i.e. golden search [53]) over the interval $[0, P_{max}]$.

We now move to the derivation of the solution to (\mathcal{P}_6) . Of course, the solution of the latter problem can be found by employing existing literature approaches, e.g. [17], [25], though here we aim at developing solutions that achieve improved spectral efficiency with the view to improve further the energy efficiency of the transceivers. Thus, (\mathcal{P}_6) is first written in the form

$$(\mathcal{P}_7) : \min_{\mathbf{Z}, \mathbf{F}_{RF}, \mathbf{F}_{BB}} -\log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{Z} \mathbf{Z}^H \mathbf{H}^H) + \mathbb{1}_{\mathcal{T}}\{\mathbf{Z}\} + \mathbb{1}_{\mathcal{X}_T}\{\mathbf{F}_{RF}\} \quad (18)$$

$$s.t. \quad \mathbf{Z} = \mathbf{F}_{RF} \mathbf{F}_{BB}$$

where \mathbf{Z} is an auxiliary $T_x \times N_t$ matrix variable, the set \mathcal{T} for a matrix \mathbf{A} is defined as

$$\mathcal{T} = \{ \mathbf{A} \in \mathbb{C}^{T_x \times N_t} \mid \|\mathbf{A}\|_F^2 = \xi \}$$

and the indicator function $\mathbb{1}_{\mathcal{C}}\{\mathbf{A}\}$ of an arbitrary set \mathcal{C} is defined as

$$\mathbb{1}_{\mathcal{C}}\{\mathbf{A}\} = \begin{cases} \mathbf{0}, & \mathbf{A} \in \mathcal{C} \\ \infty, & \mathbf{A} \notin \mathcal{C} \end{cases} \quad (19)$$

Problem (\mathcal{P}_7) formulates the hybrid pre-coding matrix design, as a matrix factorization problem. That is, the overall precoder \mathbf{Z} is sought that maximizes the mutual information of the transceiver subject to the power constraints of set \mathcal{T} and such that it admits a decomposition into two factors, a $T_x \times N_t$ unit modulus matrix \mathbf{F}_{RF} and a $N_t \times N_s$ matrix \mathbf{F}_{BB} of complex entries.

In the following we are going to develop an iterative procedure for the solution to (\mathcal{P}_7) based on ADMM. This method, is a variant of standard augmented Lagrangian method that uses partial updates (similar to the Gauss-Seidel method for the solution of linear equations) for solving constrained optimization problems. The main characteristic of such methods is that they replace a constrained minimization problem by a series of unconstrained problems and add a penalty term to the objective function. This penalty parameter improves the robustness compared to other optimization methods for constrained problems (e.g. dual ascent method) and in particular achieves convergence without the need of specific assumptions for the objective function, i.e. strict convexity and finiteness. The interested reader may refer to [47] for further information. Let us move now to the derivation of the method. The augmented Lagrangian function of (\mathcal{P}_7) is given by

$$\mathcal{L}_T(\mathbf{Z}, \mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{A}) = -\log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{Z} \mathbf{Z}^H \mathbf{H}^H) + \mathbb{1}_{\mathcal{T}}\{\mathbf{Z}\} + \mathbb{1}_{\mathcal{X}_T}\{\mathbf{F}_{RF}\} + \langle \mathbf{A}, \mathbf{Z} - \mathbf{F}_{RF} \mathbf{F}_{BB} \rangle + \frac{\alpha}{2} \|\mathbf{Z} - \mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2, \quad (20)$$

where $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{m,n} \mathbf{A}(m,n) \mathbf{B}(m,n)$, for two matrices \mathbf{A} and \mathbf{B} , \mathbf{A} is the $T_x \times N_t$ Lagrange Multiplier matrix and α is a scalar penalty parameter.

According to the ADMM approach [47], the solution to (\mathcal{P}_7) is derived by the following steps

$$(\mathcal{P}_{7A}) : \mathbf{Z}_n = \arg \min_{\mathbf{Z}} \mathcal{L}_T(\mathbf{Z}, \mathbf{F}_{RF(n-1)}, \mathbf{F}_{BB(n-1)}, \mathbf{A}_{n-1}) \quad (21)$$

$$(\mathcal{P}_{7B}) : \mathbf{F}_{RF(n)} = \arg \min_{\mathbf{F}_{RF}} \mathcal{L}_T(\mathbf{Z}_n, \mathbf{F}_{RF}, \mathbf{F}_{BB(n-1)}, \mathbf{A}_{n-1}) \quad (22)$$

$$(\mathcal{P}_{7C}) : \mathbf{F}_{BB(n)} = \arg \min_{\mathbf{F}_{BB}} \mathcal{L}_T(\mathbf{Z}_n, \mathbf{F}_{RF(n)}, \mathbf{F}_{BB}, \mathbf{A}_{n-1}) \quad (23)$$

$$\mathbf{A}_n = \mathbf{A}_{n-1} + \alpha (\mathbf{Z}_n - \mathbf{F}_{RF(n)} \mathbf{F}_{BB(n)}). \quad (24)$$

where n denotes the iteration index.

Let us now move to the derivation of the solutions to each one of problems (\mathcal{P}_{7A}) - (\mathcal{P}_{7C}) . As it is also shown in the following, problems (\mathcal{P}_{7A}) and (\mathcal{P}_{7C}) are independent of the pre-coder structure and thus, their solutions are identical for both the fully and the partially connected architectures. Let us now start with (\mathcal{P}_{7A}) which can be written in the following form by keeping only the terms that are functions of the optimizing variable \mathbf{Z}

$$\mathbf{Z}_n = \arg \min_{\mathbf{Z}} \mathcal{L}_T^Z(\mathbf{Z}, \mathbf{F}_{RF(n-1)}, \mathbf{F}_{BB(n-1)}, \mathbf{A}_{n-1}) = \arg \min_{\mathbf{Z}} -\log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{Z} \mathbf{Z}^H \mathbf{H}^H) + \mathbb{1}_{\mathcal{T}}\{\mathbf{Z}\} + \langle \mathbf{A}_{n-1}, \mathbf{Z} - \mathbf{F}_{RF(n-1)} \mathbf{F}_{BB(n-1)} \rangle + \frac{\alpha}{2} \|\mathbf{Z} - \mathbf{F}_{RF(n-1)} \mathbf{F}_{BB(n-1)}\|_F^2. \quad (25)$$

For the latter problem, it is straightforward to see that we can set parameter α such that it is always strongly convex with respect to the optimizing variable \mathbf{Z} . The solution to this problem does not admit a closed form due to the involved cost function and thus, a projected gradient descent based approach [46] will be developed in the following. The method of projected gradient descent has been used extensively in the past in order to solve constrained optimization problems. It involves a gradient descent update step followed by a projection onto the desired set of constraints. Thus, we first start with the computation of the gradient with respect to \mathbf{Z} , given by

$$\nabla_{\mathbf{Z}} \mathcal{L}_T^Z(\mathbf{Z}, \mathbf{F}_{RF(n-1)}, \mathbf{F}_{BB(n-1)}, \mathbf{A}_{n-1}) = -\frac{2\sigma_n^{-2}}{\ln(2)} \mathbf{H}^H (\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H} \mathbf{Z} \mathbf{Z}^H \mathbf{H}^H)^{-1} \mathbf{H} \mathbf{Z} + \mathbf{A}_{n-1} + \alpha (\mathbf{Z} - \mathbf{F}_{RF(n-1)} \mathbf{F}_{BB(n-1)}). \quad (26)$$

Then, \mathbf{Z}_n is given by updating at each step the solution $\mathbf{Z}_{i,n}$ based on the previous one $\mathbf{Z}_{i-1,n}$, that is

$$\mathbf{Z}_{i,n} = \Pi_{\mathcal{T}} \{ \mathbf{Z}_{i-1,n} - \mu \nabla_{\mathbf{Z}} \mathcal{L}_T^Z(\mathbf{Z}_{i-1,n}, \mathbf{F}_{RF(n-1)}, \mathbf{F}_{BB(n-1)}, \mathbf{A}_{n-1}) \}, \quad (27)$$

where μ is a step size parameter and $\Pi_{\mathcal{T}}$ is the projection onto the set \mathcal{T} operator, derived as the solution to the following optimization problem

$$(\mathcal{P}_8) : \min_{\mathbf{A}_{\mathcal{T}}} \|\mathbf{A}_{\mathcal{T}} - \mathbf{A}\|_F^2 \quad (28)$$

$$s.t. \quad \mathbf{A}_{\mathcal{T}} \in \mathcal{T},$$

where \mathbf{A} is an arbitrary matrix and $\mathbf{A}_{\mathcal{F}}$ is its projection onto the set \mathcal{F} . By following the Lagrange multipliers method [46], it can be shown that the required projection operator admits the following closed form

$$\Pi_{\mathcal{F}}\{\mathbf{A}\} = \sqrt{\xi} \frac{\mathbf{A}}{\|\mathbf{A}\|_F}. \quad (29)$$

The iterations of (27) are running until the following termination criterion is met

$$\|\mathbf{Z}_{i,n} - \mathbf{Z}_{i-1,n}\|_F^2 < \epsilon^{gd}, \quad (30)$$

where ϵ^{gd} is a pre-defined tolerance.

We now proceed with the derivation of the solution to (\mathcal{P}_{7B}) which can be written in the following simplified form by keeping only the functions of \mathbf{F}_{RF} ,

$$\begin{aligned} \mathbf{F}_{RF(n)} = \arg \min_{\mathbf{F}_{RF}} & \mathbb{1}_{\mathcal{X}_T}\{\mathbf{F}_{RF}\} + \\ & \langle \mathbf{\Lambda}_{n-1}, \mathbf{Z}_n - \mathbf{F}_{RF} \mathbf{F}_{BB(n-1)} \rangle + \frac{\alpha}{2} \|\mathbf{Z}_n - \mathbf{F}_{RF} \mathbf{F}_{BB(n-1)}\|_F^2. \end{aligned} \quad (31)$$

The solution to this problem for the fully connected architecture does not admit a closed form and thus, it is approximated by solving the unconstrained problem and then projecting onto the set \mathcal{F} , that is

$$\begin{aligned} \tilde{\mathbf{F}}_{RF(n)} &= \frac{1}{\alpha} (\mathbf{\Lambda}_{n-1} + \alpha \mathbf{Z}_n) \mathbf{F}_{BB(n)}^H (\mathbf{F}_{BB(n)} \mathbf{F}_{BB(n)}^H)^{-1}, \\ \mathbf{F}_{RF(n)} &= \Pi_{\mathcal{F}} \left\{ \tilde{\mathbf{F}}_{RF(n)} \right\} \end{aligned} \quad (32)$$

where $\Pi_{\mathcal{F}}$ is the projection onto the set \mathcal{F} operator. The latter projection operator can be derived by solving the following optimization problem [46]

$$\begin{aligned} (\mathcal{P}'_8) : \quad & \min_{\mathbf{A}_{\mathcal{F}}} \|\mathbf{A}_{\mathcal{F}} - \mathbf{A}\|_F^2 \\ & \text{s.t. } \mathbf{A}_{\mathcal{F}} \in \mathcal{F}, \end{aligned} \quad (33)$$

where \mathbf{A} is an arbitrary matrix and $\mathbf{A}_{\mathcal{F}}$ is its projection onto the set \mathcal{F} . The solution to (\mathcal{P}'_8) is given by the phase of the complex elements of \mathbf{A} . Thus, for $\mathbf{A}_{\mathcal{F}} = \Pi_{\mathcal{F}}\{\mathbf{A}\}$ we have

$$\mathbf{A}_{\mathcal{F}}(m, n) = \begin{cases} 0, & \mathbf{A}(m, n) = 0 \\ \frac{\mathbf{A}(m, n)}{|\mathbf{A}(m, n)|}, & \mathbf{A}(m, n) \neq 0 \end{cases}, \quad (34)$$

where $\mathbf{A}_{\mathcal{F}}(m, n)$ and $\mathbf{A}(m, n)$ are the elements at the m th row - n th column of matrices $\mathbf{A}_{\mathcal{F}}$ and \mathbf{A} respectively and $|\cdot|$ is the modulus of a complex number. While, this is an inexact solution, it turns out that it behaves remarkably well, as it is verified in the simulations section (Sec.VII). This is due to the interesting property of ADMM according to which, under certain conditions, it is observed to converge even in cases where the alternating minimization steps are not carried out exactly [47]. There are theoretical results that support this statement [54], [55], though an exact analysis for the case considered here is beyond the scopes of the present paper.

We may now derive the solution of (\mathcal{P}_{7B}) for the partially connected architecture. To that end, (31) is written in the following compact form

$$\mathbf{F}_{RF(n)} = \arg \max_{\mathbf{F}_{RF}} \mathbb{1}_{\mathcal{X}_T}\{\mathbf{F}_{RF}\} +$$

$$\frac{\alpha}{2} \|\mathbf{Z}_n + \alpha^{-1} \mathbf{\Lambda}_{n-1} - \mathbf{F}_{RF} \mathbf{F}_{BB(n-1)}\|_F^2, \quad (35)$$

By exploiting the block diagonal structure of \mathbf{F}_{RF} , (35) can be further re-written as

$$\begin{aligned} \mathbf{F}_{RF(n)}(m, l) = \arg \min_{\phi_m} & \frac{\alpha}{2} \|\mathbf{Z}_n(m, :) + \alpha^{-1} \mathbf{\Lambda}_{n-1}(m, :) - \\ & e^{j\phi_m} \mathbf{F}_{BB(n-1)}(l, :)\|_F^2, \end{aligned} \quad (36)$$

where $1 \leq m \leq N_t$, $l = \lceil m \frac{N_t}{T_x} \rceil$ and $\mathbf{A}(m, :)$ is the m th row of a matrix \mathbf{A} . Eq. (36) is actually a vector approximation under phase rotation problem [23] where its solution is given by

$$\begin{aligned} \mathbf{F}_{RF(n)}(m, l) = \Pi_{\mathcal{F}} \{ & [\mathbf{Z}_n(m, :) + \alpha^{-1} \mathbf{\Lambda}_{n-1}(m, :)] \times \\ & \mathbf{F}_{BB(n-1)}(l, :)^H \}, \end{aligned} \quad (37)$$

where the projection operator $\Pi_{\mathcal{F}}$ is defined in (34).

Let us now move to the solution of (\mathcal{P}_{7C}) which can be written as

$$\begin{aligned} \mathbf{F}_{BB(n)} = \arg \min_{\mathbf{F}_{BB}} & \langle \mathbf{\Lambda}_{n-1}, \mathbf{Z}_n - \mathbf{F}_{RF(n)} \mathbf{F}_{BB} \rangle + \\ & \frac{\alpha}{2} \|\mathbf{Z}_n - \mathbf{F}_{RF(n)} \mathbf{F}_{BB}\|_F^2. \end{aligned} \quad (38)$$

By equating the gradient of (38) to zero, one may show that the solution of that problem admits the following closed form

$$\mathbf{F}_{BB(n)} = \frac{1}{\alpha} \left(\mathbf{F}_{RF(n)}^H \mathbf{F}_{RF(n)} \right)^{-1} \mathbf{F}_{RF(n)}^H (\mathbf{\Lambda}_{n-1} + \alpha \mathbf{Z}_n). \quad (39)$$

According to the termination criteria and the convergence of the ADMM sequence discussed later in this section, the primal feasibility condition of (\mathcal{P}_7) is satisfied in the sense that $\|\mathbf{Z}_n - \mathbf{F}_{RF(n)} \mathbf{F}_{BB(n)}\|_F \rightarrow 0$. Thus, it is possible for the ADMM solution to violate the power constraint of the original problem (\mathcal{P}_3) . This can be resolved by projecting the digital pre-coder solution of the ADMM sequence onto the set \mathcal{T}' defined as,

$$\mathcal{T}' = \{ \mathbf{A} \in \mathbb{C}^{T_s \times N_t} \mid \|\mathbf{F}_{RF(*)} \mathbf{A}\|_F^2 = \xi \}$$

where $\mathbf{F}_{RF(*)}$ is the solution for the analog pre-coder provided by the ADMM sequence and \mathbf{A} is again, an arbitrary matrix. The projection onto the set \mathcal{T}' operator is approximated by

$$\Pi_{\mathcal{T}'}\{\mathbf{A}\} = \sqrt{\xi} \frac{\mathbf{A}}{\|\mathbf{F}_{RF} \mathbf{A}\|_F}. \quad (40)$$

Let us now comment on the convergence of the proposed ADMM alternating minimization sequence. In the literature so far, strong convergence results for the ADMM have been derived for convex problems that involve only two blocks of variables. Moreover, strong convergence results for non-convex problems when the ADMM sequence is applied are in general unknown and also an open research problem. Problem (\mathcal{P}_7) involves three blocks of variables and on top of that, it is non-convex. In general, obtaining strong convergence results is an intractable task. Under mild conditions and following some recent results [32], [56], it can be shown that the convergence of the proposed technique can be guaranteed given that the Lagrange Multiplier sequence is bounded and parameter α is set to a large value such that the Augmented Lagrangian

Function (20) is strongly convex with respect the optimizing variables in problems (\mathcal{P}_{7A}) - (\mathcal{P}_{7C}) . The convergence analysis of the proposed technique is beyond the scopes of the present paper.

Let us now discuss the implementation aspects of the proposed approach. The core of the proposed solution is the four ADMM alternating minimization problems (\mathcal{P}_{7A}) - (\mathcal{P}_{7C}) . The matrices \mathbf{Z}_n , $\mathbf{F}_{RF(n)}$ and $\mathbf{F}_{BB(n)}$ are initialized with random values. The Lagrange multiplier matrix $\mathbf{\Lambda}_n$ is initialized with zeros. For the termination criteria of this alternating minimization procedure we propose the ones given by

$$\|\mathbf{Z}_n - \mathbf{Z}_{n-1}\|_F \leq \epsilon^z \ \& \ \|\mathbf{Z}_n - \mathbf{F}_{RF(n)}\mathbf{F}_{BB(n)}\|_F \leq \epsilon^p \quad (41)$$

where ϵ^z and ϵ^p are the corresponding tolerances. The first termination criterion guarantees the convergence of variable \mathbf{Z}_n and further the convergence of (\mathcal{P}_{7A}) to its optimal value. The second one guarantees that the primal feasibility condition of (\mathcal{P}_7) is satisfied. Due to the non-convex nature of the latter problem, it is also useful to add a termination criterion that permits the maximum number of the ADMM sequence's iterations which here is defined as N_{max} .

Another point to be discussed is the implementation aspects of the solution to (\mathcal{P}_{7A}) , given by the iterative projected gradient technique of (27). The required complexity can be greatly reduced by smooth starting the gradient descent part of the ADMM's n th iteration by setting $\mathbf{Z}_{0,n-1} = \mathbf{Z}_{n-1}$. Furthermore, the tolerance parameter ϵ^{gd} can be set to a larger value at the beginning of the ADMM sequence and decrease its value as the latter converges for better accuracy [47] [54]. A good rule of thumb is to decrease ϵ^{gd} by a power of 10 every time the quantities in (41) reach its value. Of course, the tolerance may be set to a fixed value if a more coherent implementation is desired. The complete procedure is presented in Algorithm 1.

We close this section with the derivation of the computational complexity of the proposed approach. From Algorithm 1, there are two loops and thus, the complexity is dependent on the number of iterations of these loops. For simplicity, we will assume that the inner loop associated to problem (\mathcal{P}_{7A}) is invoked m_X times and the outer one l_X^t times, where $X = HFC$ and $X = HPC$, for the fully and the partially connected architectures, respectively. By inspecting (24), (27), (32), (37) and (39), one may see that the updates of the involved variables are done via standard matrix operations. The flops required for the latter operations can be found in [57]. Based on the previous and assuming for simplicity that $N_s = N_r = N_t$, one may calculate the complexity for the fully and the partially connected cases, given by,

$$\begin{aligned} C_{HFC}^t &= (m_{HFC}/3)[R_x^3 + 3N_s T_x(11 + T_x) + 15N_s T_x + \\ & 6R_x T_x(N_s + T_x) + R_x^2(3 + 6N_s + 9T_x)] + \\ & l_{HFC}^t [(11N_s^3)/3 + 3R_x T_x + N_s^2(1 + 11T_x)] \end{aligned} \quad (42)$$

$$\begin{aligned} C_{HPC}^t &= (1/3)\{l_{HPC}^t N_s^2(21 + N_s) + 6R_x T_x(N_s + T_x) \\ & + m_{HPC}[R_x^3 + 3N_s T_x(11 + T_x) + R_x^2(3 + 6N_s + 9T_x)]\} \\ & + 3l_{HPC}^t(N_s(11 + 9N_s) + 3R_x)T_x. \end{aligned} \quad (43)$$

Algorithm 1 Hybrid Pre-coding Matrix Design

```

1: Initialize  $\xi$  with a value in  $[0, P_{max}]$ 
2: while The termination criterion of the one-dimensional search is not
   satisfied do
3:   Initialize  $\mathbf{Z}_0, \mathbf{F}_{RF(0)}, \mathbf{F}_{BB(0)}$  with random values and  $\mathbf{\Lambda}_0$  with zeros
4:   while The termination criteria of (41) are not met or  $n \leq N_{max}$  do
5:      $\mathbf{Z}_{0,n} \leftarrow \mathbf{Z}_{n-1}$ 
6:     while  $\|\mathbf{Z}_{i,n} - \mathbf{Z}_{i-1,n}\|_F^2 \leq \epsilon^{gd}$  do
7:       Update  $\mathbf{Z}_{i,n}$  from (27)
8:     end while
9:      $\mathbf{Z}_n \leftarrow \mathbf{Z}_{i,n}$ 
10:    Update  $\mathbf{F}_{RF(n)}$  from (32) or (37) (fully or partially connected),
       $\mathbf{F}_{BB(n)}$  from (39) and  $\mathbf{\Lambda}_n$  from (24)
11:   end while
12:   Update the value of  $\xi$ 
13: end while
Return:  $\mathbf{F}_{RF(n)}, \Pi_{\mathcal{T}'}\{\mathbf{F}_{BB(n)}\}$ 

```

B. Post-Coding Matrix Design

Let us now move to the post-coding matrix derivation via solving problem (\mathcal{P}_4) . In the absence of the phase only constraints on the receiver, one may find the corresponding optimal digital only post-coder by solving again the unconstrained version of problem (\mathcal{P}_4) given that the hybrid pre-coding solutions of Algorithm 1 are applied at the transmitter's side. Thus, by denoting with \mathbf{W}_D the optimal digital only MMSE post-coder, it can be shown from (2)-(3) and (15) that is given by

$$\begin{aligned} \mathbf{W}_D^H &= \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}^{-1}\mathbb{E}\{\mathbf{x}\mathbf{y}^H\}^H \\ &= (\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\mathbf{H}^H + \sigma_n^2\mathbf{I}_{R_s})^{-1}\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}. \end{aligned} \quad (44)$$

The authors in [17] used the methodology in [58] for finding linear MMSE estimators with complex structural constraints and showed that the optimal hybrid MMSE solution can be found as the weighted projection of the unconstrained optimal digital only MMSE onto the set of post-coders that admit the form $\mathbf{W}_{RF}\mathbf{W}_{BB}$ with $\mathbf{W}_{RF} \in \mathcal{X}_R$. That is, the non-convex problem (\mathcal{P}_4) , can be cast on the following equivalent form

$$\begin{aligned} (\mathcal{P}_9) : \min_{\mathbf{W}_{RB}, \mathbf{W}_{BB}} & \left\| \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}^{1/2} (\mathbf{W}_D - \mathbf{W}_{RF}\mathbf{W}_{BB}) \right\|_F^2 \\ \text{s.t.} & \mathbf{W}_{RF} \in \mathcal{X}_R. \end{aligned} \quad (45)$$

The proposed solution can be derived by developing again an ADMM based solution. To that end, (\mathcal{P}_9) is expressed in the following form

$$\begin{aligned} (\mathcal{P}_{10}) : \min_{\mathbf{G}, \mathbf{W}_{RF}, \mathbf{W}_{BB}} & \left\| \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}^{1/2} (\mathbf{W}_D - \mathbf{G}) \right\|_F^2 \\ & + \mathbb{1}_{\mathcal{X}_R}\{\mathbf{W}_{RF}\} \\ \text{s.t.} & \mathbf{G} = \mathbf{W}_{RF}\mathbf{W}_{BB}, \end{aligned} \quad (46)$$

where \mathbf{G} is an auxiliary $R_s \times N_r$ matrix variable and the indicator function of set \mathcal{X}_R is defined in (19).

The augmented Lagrangian function of (\mathcal{P}_8) is given by

$$\begin{aligned} \mathcal{L}_R(\mathbf{G}, \mathbf{W}_{RF}, \mathbf{W}_{BB}, \mathbf{\Pi}) &= \left\| \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}^{1/2} (\mathbf{W}_D - \mathbf{G}) \right\|_F^2 + \\ & \mathbb{1}_{\mathcal{X}_R}\{\mathbf{W}_{RF}\} + \langle \mathbf{\Pi}, \mathbf{G} - \mathbf{W}_{RF}\mathbf{W}_{BB} \rangle \\ & + \frac{\beta}{2} \|\mathbf{G} - \mathbf{W}_{RF}\mathbf{W}_{BB}\|_F^2, \end{aligned} \quad (47)$$

where $\mathbf{\Pi}$ is the $R_x \times N_r$ Lagrange Multiplier matrix and β is a scalar penalty parameter.

Following again the ADMM methodology, the solution to (\mathcal{P}_{10}) is given by the following alternating minimization steps,

$$(\mathcal{P}_{10A}) : \mathbf{G}_n = \arg \min_{\mathbf{G}} \mathcal{L}_R(\mathbf{G}, \mathbf{W}_{RF(n-1)}, \mathbf{W}_{BB(n-1)}, \mathbf{\Pi}_{n-1}) \quad (48)$$

$$(\mathcal{P}_{10B}) : \mathbf{W}_{RF(n)} = \arg \min_{\mathbf{W}_{RF}} \mathcal{L}_R(\mathbf{G}_n, \mathbf{W}_{RF}, \mathbf{W}_{BB(n-1)}, \mathbf{\Pi}_{n-1}) \quad (49)$$

$$(\mathcal{P}_{10C}) : \mathbf{W}_{BB(n)} = \arg \min_{\mathbf{W}_{BB}} \mathcal{L}_R(\mathbf{G}_n, \mathbf{W}_{RF(n)}, \mathbf{W}_{BB}, \mathbf{\Pi}_{n-1}) \quad (50)$$

$$\mathbf{\Pi}_n = \mathbf{\Pi}_{n-1} + \beta (\mathbf{G}_n - \mathbf{W}_{RF(n)} \mathbf{W}_{BB(n)}). \quad (51)$$

where n is again the iteration index. Note that in a similar manner to the pre-coding matrix design, problems (\mathcal{P}_{10A}) , (\mathcal{P}_{10C}) and (\mathcal{P}_{10D}) are independent of the analog post-coder structure and thus, the solutions are the same for both the fully and the partially connected architectures.

Problem (\mathcal{P}_{10A}) can be directly solved by equating the gradient of the augmented Lagrangian (47) with respect to \mathbf{G} to zero. Therefore, we have

$$\mathbf{G}_n = (\mathbb{E} \{ \mathbf{y} \mathbf{y}^H \} + \beta \mathbf{I}_{R_s})^{-1} \times (\mathbb{E} \{ \mathbf{y} \mathbf{y}^H \} \mathbf{W}_D - \mathbf{\Pi}_{n-1} + \beta \mathbf{W}_{RF(n-1)} \mathbf{W}_{BB(n-1)}) \quad (52)$$

Problem (\mathcal{P}_{10B}) is equivalent to (\mathcal{P}_{7B}) which derives the optimal analog pre-coder matrix $\mathbf{F}_{RF(n)}$. Thus, it is straightforward to see that for the fully connected architecture $\mathbf{W}_{RF(n)}$ is given by,

$$\tilde{\mathbf{W}}_{RF(n)} = (\mathbf{\Pi}_{n-1} + \beta \mathbf{G}_n) \mathbf{W}_{BB(n)}^H (\mathbf{W}_{BB(n)} \mathbf{W}_{BB(n)}^H)^{-1} \quad (53)$$

$$\mathbf{W}_{RF(n)} = \Pi_{\mathcal{F}} \left\{ \frac{1}{\beta} \tilde{\mathbf{W}}_{RF(n)} \right\}$$

and the for the partially connected one is given by,

$$\mathbf{W}_{RF(n)}(m, l) = \Pi_{\mathcal{F}} \left\{ [\mathbf{G}_n(m, :) + \beta^{-1} \mathbf{\Pi}_{n-1}(m, :)] \times \mathbf{W}_{BB(n-1)}(l, :)^H \right\}, \quad (54)$$

where $1 \leq m \leq N_r$, $l = \left\lceil m \frac{N_r}{R_x} \right\rceil$ and the projection operator $\Pi_{\mathcal{F}}$ is defined in (34).

In a similar way, the solution of (\mathcal{P}_{10C}) is equivalent to the one of (\mathcal{P}_{10C}) via which matrix $\mathbf{F}_{BB(n)}$ is updated. Therefore, it can be shown that

$$\mathbf{W}_{BB(n)} = \frac{1}{\beta} (\mathbf{W}_{RF(n)}^H \mathbf{W}_{RF(n)})^{-1} \mathbf{W}_{RF(n)}^H (\mathbf{\Pi}_{n-1} + \beta \mathbf{G}_n). \quad (55)$$

The proposed termination criteria follow from the ones of problems (\mathcal{P}_{10A}) - (\mathcal{P}_{10C}) under the same arguments. To that end, the following conditions have to be met

$$\|\mathbf{G}_n - \mathbf{G}_{n-1}\|_F \leq \epsilon^g \ \& \ \|\mathbf{W}_n - \mathbf{W}_{RF(n)} \mathbf{W}_{BB(n)}\|_F \leq \epsilon_2^p, \quad (56)$$

where ϵ^g and ϵ_2^p are the corresponding pre-defined tolerances. Regarding the convergence of the ADMM sequence, similar

Algorithm 2 Hybrid Post-coding Matrix Design

1: Initialize \mathbf{G}_0 , $\mathbf{W}_{RF(0)}$, $\mathbf{W}_{BB(0)}$ with random values and $\mathbf{\Pi}_0$ with zeros
2: **while** The termination criteria of (56) are not met or $n \leq N_{max}$ **do**
3: Update \mathbf{G}_n from (52), $\mathbf{W}_{RF(n)}$ from (53) or (54) (fully or partially connected), $\mathbf{W}_{BB(n)}$ from (55) and $\mathbf{\Pi}_n$ from (51)
4: **end while**
Return: $\mathbf{W}_{RF(n)}$, $\mathbf{W}_{BB(n)}$

comments to the ones related to (\mathcal{P}_{10A}) - (\mathcal{P}_{10C}) are holding (Sec. IV.A). The procedure is summarized in Algorithm 2.

Under the assumption that Algorithm 2 runs for l_{HFC}^r and l_{HPC}^r iterations for the fully connected and the partially connected architectures, respectively and similar to the case of Algorithm 1 one may derive its computational complexity, given for the two architectures by,

$$C_{HFC}^r = [1 + (11n_{HFC}^r)/3]N_s^3 + (2R_x^3)/3 + R_x^2(2 + 3T_x) + R_x T_x(1 + 2T_x) + N_s^2(1 + n_{HPC}^r + 11n_{HPC}^r R_x + 4T_x) + N_s \{ R_x [19n_{HPC}^r + 2(1 + n_{HPC}^r)R_x] + T_x + 2R_x T_x + T_x^2 \} \quad (57)$$

$$C_{HPC}^r = (3 + n_{HPC}^r)N_s^3/3 + (2R_x^3)/3 + R_x T_x(1 + 2T_x) + R_x^2(2 + 3T_x) + N_s^2(1 + 14n_{HPC}^r R_x + 4T_x) + N_s [2R_x(6n_{HPC}^r + R_x) + T_x + 2R_x T_x + T_x^2]. \quad (58)$$

It is now straightforward to see that the total complexity of the proposed method \mathcal{C}_X which is required for the calculation of the corresponding consumed power in (11) is the sum of the flops for the precoding and the post-coding matrices calculation, that is $\mathcal{C}_X = \mathcal{C}_X^t + \mathcal{C}_X^r$, $X = \{HFC, HPC\}$.

V. ENERGY EFFICIENT HYBRID A/D TRANSCEIVERS FOR OFDM SYSTEMS

OFDM is the most popular way to enable multi-carrier transmission through frequency selective wireless channels. Therefore, in this section, hybrid transceiver designs are developed for OFDM systems for both the fully and partially connected architectures.

Let us assume a typical OFDM system of K sub-carriers. A fully digital transmitter involves the application of a digital pre-coder in each sub-carrier, followed by an Inverse Fast Fourier Transform (IFFT) for the combination of the signals prior transmission. In the fully digital receiver's side, a FFT is applied on the received signals followed by the digital post-coder. As it is evident from the previous description, the IFFT/FFT operations are applied on the base band domain. Thus, a hybrid transceiver could apply the analog pre/post-coding counterparts after the IFFT and before the FFT in the transmitter's and the receiver's side, respectively. To that end, a direct extension of the hybrid single carrier transceivers could require a dedicated RF hardware per sub-carrier which exhibits prohibitive complexity. Therefore, current literature approaches [28], [29] assume hybrid transceivers that have dedicated BB processing per sub-carrier and a single RF processing network shared among all the sub-carriers. Thus, the design of the analog counterpart of the pre/post-coders requires the joint optimization over all the sub-carriers.

Following the above analysis, the input-output relationship for the k th sub-carrier is given by,

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{F}_k \mathbf{x}_k + \mathbf{n}_k \quad (59)$$

$$\mathbf{y}'_k = \mathbf{W}_k^H \mathbf{y}_k, \quad (60)$$

where for the k th sub-carrier, $0 \leq k \leq K-1$, \mathbf{H}_k is the $R_x \times T_x$ frequency domain channel matrix between the two transceiver ends, \mathbf{x}_k is the $N_s \times 1$ transmitted symbol vector under the assumption that $\mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\} = \sigma_x^2 \mathbf{I}_{T_x}$, \mathbf{n}_k is i.i.d. complex Gaussian noise modelled as $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{R_x})$, $\mathbf{F}_k = \mathbf{F}_{RF} \mathbf{F}_{BB(k)}$ and $\mathbf{W}_k = \mathbf{W}_{RF} \mathbf{W}_{BB(k)}$, with $\mathbf{F}_{BB(k)}$ - $\mathbf{W}_{BB(k)}$ and \mathbf{F}_{RF} - \mathbf{W}_{RF} denote the digital counterparts and the joint analog counterparts of the pre/post-coding matrices, respectively.

The frequency domain channel matrix of the k th sub-carrier, \mathbf{H}_k is given by [30], [59]

$$\mathbf{H}_k = \sqrt{\frac{T_x R_x}{N_p}} \sum_{i=0}^{N_c-1} \sum_{l=1}^{N_p} \alpha_{il} \mathbf{a}_r(\phi_{il}^r) \mathbf{a}_t(\phi_{il}^t)^H e^{-j2\pi i(k/K)}, \quad (61)$$

where N_c is the number of taps that correspond to the scattering clusters, $\alpha_{il} \sim \mathcal{CN}(0, 1)$ is the complex gain of the l th path of the i th cluster, ϕ_{il}^r and ϕ_{il}^t are the azimuth angles of arrival and departure, respectively, that are uniformly distributed in $[0, 2\pi)$ and vectors $\mathbf{a}_r(\phi_{il}^r)$ and $\mathbf{a}_t(\phi_{il}^t)^H$ are defined as in (5). The energy efficiency for an OFDM system is given by

$$E_X^{OFDM}(\mathbf{F}_{RF}, \bar{\mathbf{F}}_{BB}, \mathbf{W}_{RF}, \bar{\mathbf{W}}_{BB}) = \frac{R_{OFDM}(\mathbf{F}_{RF}, \bar{\mathbf{F}}_{BB}, \mathbf{W}_{RF}, \bar{\mathbf{W}}_{BB})}{\sum_{k=0}^{K-1} \|\mathbf{F}_{RF} \mathbf{F}_{BB(k)}\|_F^2 / \eta + \mu_X}, \quad (62)$$

where, $R_{OFDM}(\mathbf{F}_{RF}, \bar{\mathbf{F}}_{BB}, \mathbf{W}_{RF}, \bar{\mathbf{W}}_{BB}) = \frac{1}{K} \sum_{k=0}^{K-1} \log_2 \det(\mathbf{I}_{L_s} + \mathbf{R}_{nk}^{-1} \mathbf{W}_{BB(k)}^H \mathbf{W}_{RF}^H \mathbf{H}_k \mathbf{F}_{RF} \mathbf{F}_{BB(k)} \mathbf{F}_{BB(k)}^H \mathbf{F}_{RF}^H \mathbf{H}_k^H \mathbf{W}_{RF} \mathbf{W}_{BB(k)})$, $\bar{\mathbf{F}}_{BB} = [\mathbf{F}_{BB(0)}, \dots, \mathbf{F}_{BB(K-1)}]$, $\bar{\mathbf{W}}_{BB} = [\mathbf{W}_{BB(0)}, \dots, \mathbf{W}_{BB(K-1)}]$, $\bar{\mathbf{R}}_{nk} = \sigma_n^2 \mathbf{W}_{BB(k)}^H \mathbf{W}_{RF}^H \mathbf{W}_{RF} \mathbf{W}_{BB(k)}$ and $X = \{HFC, HPC\}$ for the fully and the partially connected architectures, respectively.

We are again interested in deriving the optimal pre/post-coding matrices such that the system's energy efficiency is maximized. Following the analysis of Sec. IV, the designs at the transmitter and the receiver are again temporally decoupled, and thus the pre/post-coding matrices are derived as the solutions to

$$(\mathcal{P}_{14}) : \max_{\mathbf{F}_{RF}, \bar{\mathbf{F}}_{BB}} \frac{I_{OFDM}(\mathbf{F}_{RF}, \bar{\mathbf{F}}_{BB})}{\sum_{k=0}^{K-1} \|\mathbf{F}_{RF} \mathbf{F}_{BB(k)}\|_F^2 / \eta + \mu_X} \\ s.t. \quad \sum_{k=0}^{K-1} \|\mathbf{F}_{RF} \mathbf{F}_{BB(k)}\|_F^2 \leq P_{max} \quad \& \quad \mathbf{F}_{RF} \in \mathcal{X}_T, \quad (63)$$

$$(\mathcal{P}_{15}) : \min_{\mathbf{W}_{RF}, \bar{\mathbf{W}}_{BB}} \sum_{k=0}^{K-1} \mathbb{E} \left\{ \|\mathbf{x}_k - \mathbf{W}_{BB(k)}^H \mathbf{W}_{RF} \mathbf{y}_k\|_F^2 \right\} \\ s.t. \quad \mathbf{W}_{RF} \in \mathcal{X}_R, \quad (64)$$

respectively, where $I_{OFDM}(\mathbf{F}_{RF}, \bar{\mathbf{F}}_{BB}) = \frac{1}{K} \sum_{k=0}^{K-1} \log_2 \det(\mathbf{I}_{L_s} + \sigma_n^{-2} \mathbf{H}_k \mathbf{F}_{RF} \mathbf{F}_{BB(k)} \mathbf{F}_{BB(k)}^H \mathbf{F}_{RF}^H \mathbf{H}_k^H)$. It is straightforward to see that problems (\mathcal{P}_{14}) - (\mathcal{P}_{15}) can be cast to the form of (\mathcal{P}_5) and (\mathcal{P}_9) , respectively, similar to the single carrier case. Thus, their solutions can be derived via an alternating minimization procedure (for the precoder case the line search method for the ξ variable is also required), similar to the ones of (\mathcal{P}_{7A}) - (\mathcal{P}_{7C}) and (\mathcal{P}_{10A}) - (\mathcal{P}_{10C}) , respectively. It is noteworthy that when the problems (\mathcal{P}_{7A}) , (\mathcal{P}_{7C}) , (\mathcal{P}_{10A}) , (\mathcal{P}_{10C}) are cast in the OFDM case, they can be decoupled in K independent ones, one per sub-carrier. Thus, the solution to each one of them is given again by (26)-(27), (39), (24), (52), (55) and (51), respectively.

It remains to derive matrices $\mathbf{F}_{RF(n)}$ and $\mathbf{W}_{RF(n)}$. Following the analysis for the single carrier case, it can be shown for the fully and partially connected architectures, the pre-/post-coding matrices are given by (65)-(68) which can be found at the top of the next page. Algorithms 1 and 2 can be employed to provide the pre/post-coding solutions for the OFDM case, as well, after applying on them the modifications discussed in the present section. The complexity for the OFDM case can be deduced directly from the results for the single carrier case and thus, it is omitted to avoid repetition.

VI. FULLY DIGITAL SOLUTION AND BOUNDS ON THE NUMBER OF RF CHAINS

In this section, we derive the fully digital transceiver which will serve as a benchmark for the evaluation of the proposed hybrid solutions. Furthermore, bounds on the maximum number of RF chains above which the hybrid solutions are always less energy efficient than the fully digital ones are derived.

We start with the fully digital transceiver for the single carrier case. Note that the results are directly extendible in the OFDM case, as well. By using (6), (7) and (8), (\mathcal{P}_1) can be written as

$$(\mathcal{P}_{18}) : \max_{\mathbf{F}, \mathbf{W}, \xi} \frac{\log_2 \det(\mathbf{I}_{L_s} + \mathbf{R}_n^{-1} \mathbf{W}^H \mathbf{H} \mathbf{F} \mathbf{F}^H \mathbf{H}^H \mathbf{W})}{\xi / \eta + \mu_D} \\ s.t. \quad \|\mathbf{F}\|_F^2 = \xi \quad \& \quad 0 \leq \xi \leq P_{max}, \quad (69)$$

which can be solved via combing a one-dimensional search method for the auxiliary variable ξ with the solution to $\max_{\mathbf{F}, \mathbf{W}} \log_2 \det(\mathbf{I}_{L_s} + \mathbf{R}_n^{-1} \mathbf{W}^H \mathbf{H} \mathbf{F} \mathbf{F}^H \mathbf{H}^H \mathbf{W})$, $s.t. \quad \|\mathbf{F}\|_F^2 = \xi$. The latter problem is well documented [60] and it is solved by setting $\mathbf{F} = \sqrt{\mathbf{P}} \mathbf{V}$ and $\mathbf{W} = \mathbf{U}$, where $\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H$ is the Singular Value Decomposition (SVD) of \mathbf{H} and the diagonal matrix $\sqrt{\mathbf{P}}$ is derived via the so-called "water-filling algorithm".

By comparing (\mathcal{P}_{18}) with (\mathcal{P}_5) and (\mathcal{P}_{10}) , it holds that for a given value of ξ , the spectral efficiency of the fully or the partially connected hybrid transceiver is always upper bounded by the one of the fully-digital one. Thus, at this limiting point, the energy efficiency of the hybrid architectures is greater than the one of the fully-digital one as long as $P_{HFC}(\mathbf{F}_{RF}, \mathbf{F}_{BB}) < P_D(\mathbf{F})$ and $P_{HPC}(\mathbf{F}_{RF}, \mathbf{F}_{BB}) < P_D(\mathbf{F})$, respectively. Thus, from (8)-(10), it can be deduced that for a fixed ξ value, there is a point above which an increase on the number of RF chains will result in a hybrid transceiver

$$\mathbf{F}_{RF(n)} = \Pi_{\mathcal{F}} \left\{ \frac{1}{\alpha} \left[\sum_{k=0}^{K-1} \left(\mathbf{\Lambda}_{k,n-1} \mathbf{F}_{BB(k,n-1)}^H + \alpha \mathbf{Z}_{k,n-1} \mathbf{F}_{BB(k,n-1)}^H \right) \right] \left(\sum_{k=0}^{K-1} \mathbf{F}_{BB(k,n-1)} \mathbf{F}_{BB(k,n-1)}^H \right)^{-1} \right\}, \quad (65)$$

$$\mathbf{W}_{RF(n)} = \Pi_{\mathcal{F}} \left\{ \frac{1}{\beta} \left[\sum_{k=0}^{K-1} \left(\mathbf{\Pi}_{k,n-1} \mathbf{W}_{BB(k,n)}^H + \beta \mathbf{G}_{k,n} \mathbf{W}_{BB(k,n)}^H \right) \right] \left(\sum_{k=0}^{K-1} \mathbf{W}_{BB(k,n-1)} \mathbf{W}_{BB(k,n-1)}^H \right)^{-1} \right\}, \quad (66)$$

$$\mathbf{F}_{RF(n)}(m, l) = \Pi_{\mathcal{F}} \left\{ \sum_{k=0}^{K-1} \left[\mathbf{Z}_{k,n}(m, :) + \alpha^{-1} \mathbf{\Lambda}_{k,n-1}(m, :) \right] \mathbf{F}_{BB(k,n-1)}(l, :)^H \right\}, \quad 1 \leq m \leq N_t \ \& \ l = \lceil mN_t/T_x \rceil, \quad (67)$$

$$\mathbf{W}_{RF(n)}(m, l) = \Pi_{\mathcal{F}} \left\{ \sum_{k=0}^{K-1} \left[\mathbf{G}_n(m, :) + \beta^{-1} \mathbf{\Pi}_{n-1}(m, :) \right] \mathbf{W}_{BB(n-1)}(l, :)^H \right\}, \quad 1 \leq m \leq N_r \ \& \ l = \lceil mN_r/R_x \rceil. \quad (68)$$

that is always less energy efficient than a fully-digital one. Thus, even in the ideal scenario where the hybrid solutions achieve the same spectral efficiency with the fully digital one, the hybrid solutions are not always the most efficient solution from an energy consumption perspective. One may compute those upper bounds on the number of RF chains by equating $P_D(\mathbf{F})$ to $P_{HFC}(\mathbf{F}_{RF}, \mathbf{F}_{BB})$ and $P_{HPC}(\mathbf{F}_{RF}, \mathbf{F}_{BB})$, respectively. In order to produce insightful results, let us assume that $N_t = N_r$, $T_x = R_x = N_a$ and $P_{DAC} = P_{ADC}$ (and as a consequence $P_{RF}^T = P_{RF}^R = P_{RF}$). Then, the discussed upper bounds on the RF chains for the fully (N_{RF}^{FC}) and the partially (N_{RF}^{PC}) connected architectures are given respectively by,

$$N_{RF}^{FC} = \left\lfloor \frac{2N_a P_{RF}}{2P_{RF} + L_S L_{PS} L_C N_a + N_a P_{LNA} + 2N_a P_{PS}} \right\rfloor, \quad (70)$$

$$N_{RF}^{PC} = \left\lfloor \frac{P_{RF} N_a - P_{PS} N_a + r_p}{2P_{RF}} \right\rfloor, \quad (71)$$

where $\lfloor \cdot \rfloor$ denotes the function that rounds a real number down to the closest integer and $r_p = \sqrt{N_a [(P_{RF} - P_{PS})^2 N_a - 2L_S L_{PS} P_{RF}]}$.

VII. SIMULATIONS

In this section, numerical results are presented to evaluate the energy efficiency of the proposed hybrid transceivers and compared to the one of the fully digital transceiver (Sec. VI) and to the one of existing approaches [17], [25] and [30] for the single carrier and the OFDM case, respectively. For a fair comparison, we used the framework developed in the present paper to derive the corresponding energy-efficient pre/post-coding matrices along with the optimal transmission power. That is, the solution is given by combination of the line search method to derive the optimal value of ξ with the corresponding existing approaches [17], [25], [30] that solve problems (\mathcal{P}_6) and (\mathcal{P}_9) via which the optimal pre/post-coding matrices are derived. An environment of $N_c = 5$ clusters contributing $N_p = 10$ propagation paths each one of them, is assumed and the element spacing on the ULA of each node is set to $d = \frac{\lambda}{2}$. The results of all the simulations are averaged over 10000 channel realizations. Moreover, The

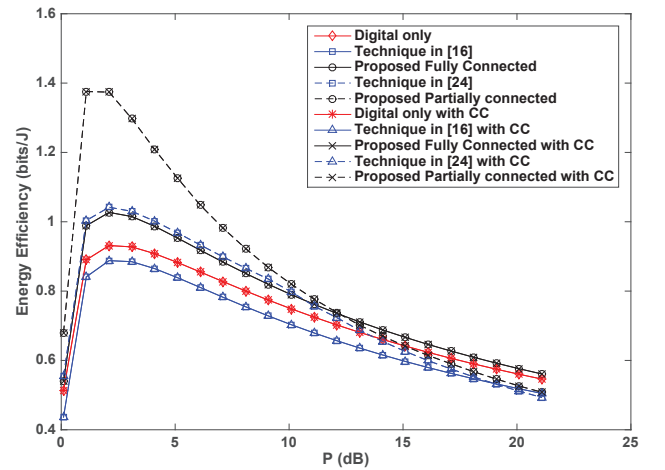


Fig. 3: Energy Efficiency of the different approaches vs Power Consumption for $T_x = R_x = 60$ and $N_t = N_r = N_s = 4$.

power consumption of the different electronic components is set to $P_{mix} = 19\text{mW}$, $P_{ADC} = P_{DAC} = 200\text{mW}$, $P_{LPF} = 14\text{mW}$, $P_{PS} = 30\text{mW}$ and $P_{LO} = 5\text{mW}$. For more information, the interested reader may check [33] and references therein. Moreover, parameters L_S , L_{PS} and L_C are computed as in [34]. Regarding parameters η , L and D values that are commonly adopted in typical communication systems are $\eta \approx 0.39$, $L = 2000$ symbols and $D = 5$ Gflops/W, respectively [42]. Let us now refer to the parameter tuning of the proposed Algorithms. These values are the same for all the experiments in this paper. For Algorithm 1, the parameters are set as $\mu = 10^{-3}$, $\alpha = 10$ for the fully connected architecture and $\alpha = 100$ for the partially connected one, $\epsilon^z = 10^{-3}$, $\epsilon^p = 10^{-4}$ and $\epsilon^{gd} = 10^{-2}$ (initial value which is updated based on the convergence of the ADMM sequence as described in Section IV.A). For Algorithm 2, the parameters are set as $\beta = 1$, for the fully connected architecture and $\beta = 10$, for the partially connected one, $\epsilon_g = 10^{-3}$ and $\epsilon_2^p = 10^{-4}$. Finally, for both algorithms $N_{max} = 500$.

In Fig. 3, the energy efficiency of the different approaches with respect the transmission power is examined for a 60×60 MIMO system with $N_s = N_t = 4$ RF chains and $N_s = 4$.

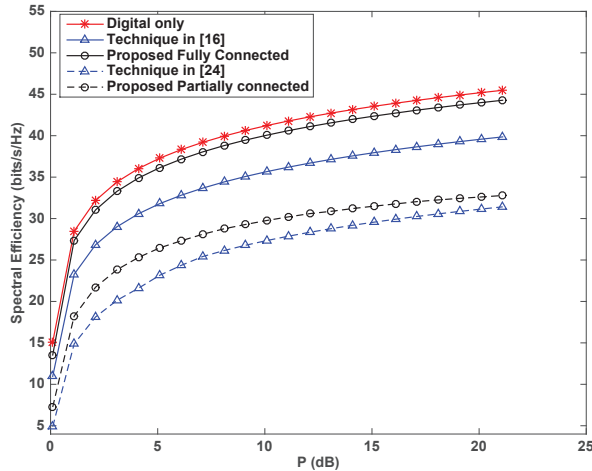


Fig. 4: Spectral Efficiency of the different approaches vs Power Consumption

The energy efficiency is calculated for the case where the consumed power is based on the hardware only consumption given by (8)-(10) and also for the case where the power consumed for the calculation of the pre/post-coding matrices is considered in (8)-(10), through the modified parameter $\tilde{\mu}_X$ in (11), as described in Sec. III. For the latter case, the depicted curves have the term “with CC (Computational Consumption)” in the corresponding label for distinction. The power consumption due to the computational complexity was derived for the proposed approach by using (42) and (57) for the fully connected and (43) and (58) for the partially connected architectures, respectively in (11). A similar approach was followed in order to compute the power consumed due to the computational complexity in all of the literature approaches that are compared to the proposed ones, based on the methods provided in the corresponding papers.

As it can be seen from a first look on the results, the power consumed due to the computational complexity of the algorithms has almost no impact on the energy efficiency of all the examined methods. This is the case, since for the typical communication scenarios examined here, the power consumed for this part remains some orders of magnitude below the one consumed on the architecture’s hardware even for the case of the proposed approaches, which exhibit relative high complexity compared to the rest examined ones. Let us compare now the different approaches. As it is depicted, the proposed solutions are more energy efficient than the existing literature ones for both the fully and the partial connected architectures. This is due to the fact that the approach in [17] is codebook based, and thus, the solution set of the RF pre-coder is restricted, while [25] is based on a diagonal baseband pre-coder that only allocates power to the transmitted streams and hence, it is not able to exploit the full capabilities of a digital pre-coder. As the power consumption increases, the energy efficiency of all the techniques decreases, since the gains on the spectral efficiency are becoming smaller compared to the power costs. Furthermore, it is observed

that for low power values, the proposed solution for the partially connected architecture is significantly more energy efficient than the fully connected one. As the power increases above that value, the performance of the partially connected architecture deteriorates significantly compared to the one of the fully connected approach and at some point, it becomes even worse than the fully digital one. Similar degradation on the performance is observed also for the approach in [25]. This is because the partially connected architectures achieve in general worse spectral efficiency than the fully-digital ones and thus, the power costs have greater impact on their energy efficiency. The same arguments explain also the relevant bad performance of the approach in [17] which presents energy efficiency always below the one of the fully digital transceiver. On the other hand, the proposed approach for the fully connected architecture is more energy efficient than the digital one, even for high power values. This is because the solution to problem (\mathcal{P}_7) results in a hybrid transceiver with spectral efficiency very close to the one of the fully digital system.

For further insight, in Fig. 4 the spectral efficiency of the proposed techniques is plotted for the examined power values following the same experimental setup with the one in Fig. 3. There, it is verified that the proposed technique for the fully connected architecture achieves very close performance to the fully digital one, and is significantly improved with respect to the one in [17]. The proposed approach for the partially connected architecture achieves also superior performance compared to the one in [25]. Furthermore, observe that in general the gap between the hybrid approaches performance and the one of the fully digital solution is increasing with an increase in power. This effect is more severe for the partially connected architectures and thus, it explains why the energy efficiency of the latter falls below the one of the fully digital and the fully connected ones for relative high power values. We close the discussion for Figs. 3-4 with the conclusion that partially connected architectures are good when the system may function (e.g. based on the requested rates that has to accommodate) on low and mid power ranges while if needs to be pushed to its limits, it is better to resort to fully connected architectures.

The impact of the number of RF chains on the energy efficiency is examined in Fig. 5. A 60×60 MIMO system is again assumed and we plot the energy efficiency for $N_t = N_r = N_s = \{2, 4, 6, 10, 12\}$. From the depicted results, the partially connected architecture and the fully digital approach become more energy efficient with the increase in the number of RF chains/streams. The fully digital architecture consumes at any case the maximum possible power in its electronic components and thus, by transmitting more streams, the spectral efficiency improvement results in a more energy efficient system. For the partially connected solutions, the energy efficiency also increases with the number of the employed RF chains. This is the case, since in partially connected architectures, the number of the employed phase shifters is independent of the number of employed RF chains and thus each new RF chain adds only the power costs for its operation, and moreover, it exhibits no further losses that have to be compensated. Contrariwise, in

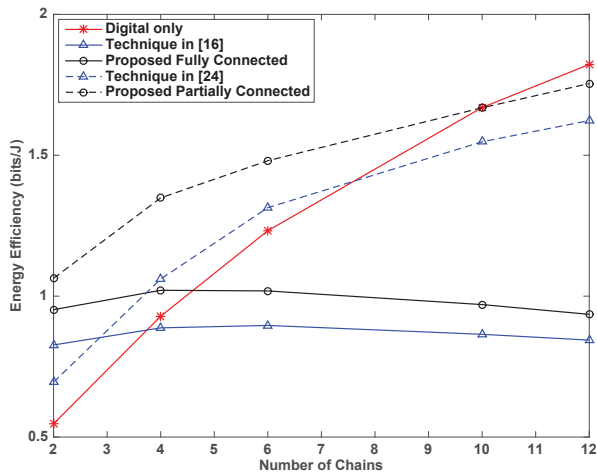


Fig. 4: Energy Efficiency of the different approaches vs Number of RF chains for $T_x = R_x = 60$.

the fully connected architectures, a new RF chain, requires T_x phase shifters at the transmitter's side and R_x phase shifters at the receiver's side. Thus, additional power must be consumed for operating the new phase shifters. Moreover, the output of the additional phase shifters is fed on the signal combiners prior the power amplifiers (Fig. 1) and this leads to further power losses (Sec. II). In a similar manner, corresponding power losses are observed also on the receiver side due to the additional signal splitting and combining requirements imposed by the introduction of a new RF chain. The latter facts explain the reduction on the energy efficiency of the fully connected architectures that eventually becomes inferior to the one of a fully digital transceiver for a much smaller increase in the number of RF chains than the one required for the partial architectures, as it is observed in Fig. 4. It is interesting that the upper bounds in (70)-(71) are $N_{RF}^{FC} = 4$ and $N_{RF}^{PC} = 37$, respectively, The bound that corresponds to the fully connected architecture is very tight while the one that corresponds to the partially connected is very loose. This is because, the former achieves spectral efficiency close to the one of the fully-digital transceiver while the latter has always lesser spectral efficiency than the fully-digital one. Thus, usually the energy efficiency drops below the one of the fully digital transceiver for fewer RF chains than the number specified by the corresponding bound for the partially connected architecture, as it is also verified in Fig. 5. Note also that, the performance of the proposed approaches is again better than the one of existing literature techniques for all the examined number of RF chains, since the former exhibit better spectral efficiency than the latter under similar power requirements.

In Fig. 6 the impact of the antenna array size is examined for $N_t = N_r = N_s = 4$. To that end we plot the energy efficiency of the different techniques for systems equipped with $T_x = R_x = \{40, 60, 80, 120\}$ antennas. As it is evident, an increase in the number of antennas results in losses on the energy efficiency, since the operation of the additional electronic equipment requires the consumption of more power.

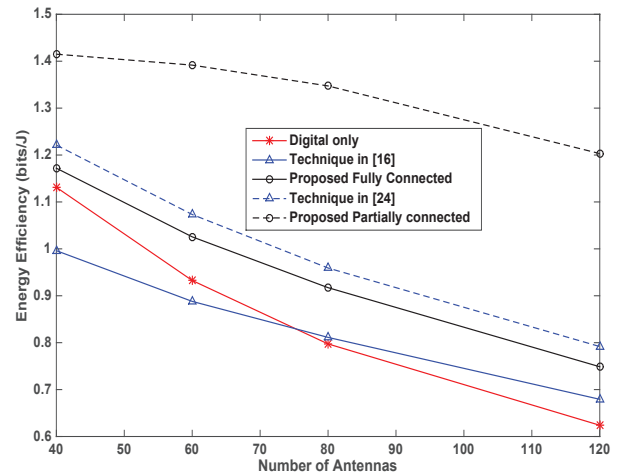


Fig. 5: Energy Efficiency of the different approaches vs Number Antennas for $N_t = N_r = N_s = 4$.

The good news is that the proposed hybrid approaches remain always more energy efficient than the fully digital one and in fact as the number of antennas increases, the gap on the performance increases. This is explained by the fact that, an additional antenna at the fully digital transceiver requires the operation of a new RF chain which requires more power than the additional PAs and LNAs that the hybrid architectures require in order to support a new antenna. Again, the proposed techniques achieve better performance than the one of the existing approaches.

Finally, in Fig. 7, the energy efficiency with respect the number of RF chains is examined for OFDM systems. The performance of the proposed techniques is depicted for $T_x = R_x = 60$ and $K = 64$. Similar conclusions to the ones of Fig. 4 can be derived. The performance of the fully digital transceiver increases with the number of the available streams. On the contrary, the energy efficiency of the fully connected architecture deteriorates significantly with an increase in the number of RF chains and falls below the one of the fully digital when the number of employed RF chains is more than four. The energy efficiency of the partially connected architecture remains always above the one of the fully digital transceiver, though it starts decreasing when the number of RF chains is more than twelve. The energy efficiency of all the techniques is in general worse compared to the single carrier case as the same power budget is allocated across $K = 64$ sub-carriers. Furthermore, the performances of the hybrid approaches drops also faster for the OFDM case since the common RF processing, applied to 64 sub-carriers is derived via a joint optimization across them. Moreover, note that the upper bounds on the number of the RF chains given in (70)-(71) hold also for the OFDM case, since their extension is quite straightforward. For comparison purposes, we plot in the same figure the technique in [30] which is the extension of the [17] to the fully connected architecture in the OFDM case. As it is observed, the latter technique is less energy efficient, than the proposed one for the fully connected architecture, due to

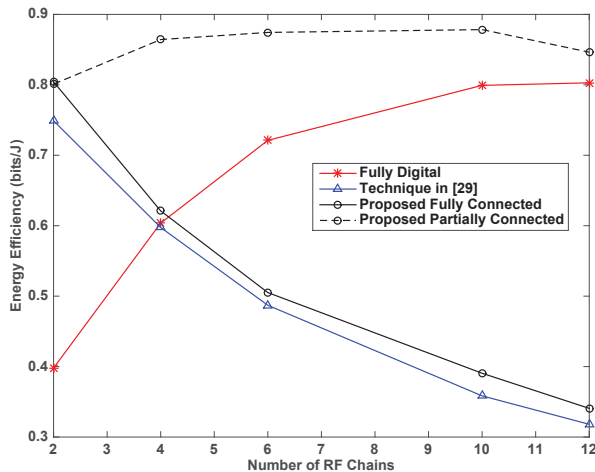


Fig. 7: Energy Efficiency of the different approaches vs Number of RF chains in OFDM systems for $T_x = R_x = 60$.

codebook requirements, as it was also discussed for the single carrier case.

VIII. CONCLUSION

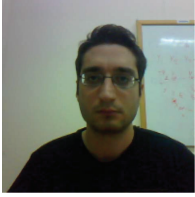
In this paper energy-efficient hybrid analog-digital transceiver designs were considered for mmWave or in general systems with large antenna arrays. A novel algorithmic framework was built via which energy efficient hybrid transceivers for fully connected and partially connected architectures were developed. At first, the proposed solutions were derived for single carrier systems and then they were extended to multi-carrier ones that employ OFDM modulation. The power consumption of the proposed transceiver was also examined in detail and upper bounds on the number of RF chains after which, a hybrid architecture always becomes less energy efficient than a fully digital one, were established. The energy efficiency of the proposed designs was examined in detail via extensive simulations, where several insights were shown regarding the performance of the hybrid energy transceivers with respect to the number of RF chains and the number of antennas. In more detail the following conclusions were reached. The fully connected hybrid architectures exhibit energy efficiency that drops quickly below the one of a fully digital solution with an increase in the number of employed RF chains/streams. On the contrary, a partially connected architecture remains more energy efficient than the fully digital solution for larger number of employed RF chains. The number of RF chains above which both of hybrid approaches are less energy efficient than the fully digital solution in general increases with an increase in the number of employed antennas at the transceiver's end. Furthermore, for low transmission power, the partially connected architecture appears to be the most energy efficient solution while as the transmission power increases (e.g. higher data rate requirements), it is better to resort to fully connected hybrid architectures. Future works include the study of the alternative approaches for

reducing the transceiver's consumption (i.e. low resolution ADCs/DACs, antenna selection) from an energy efficiency point of view.

REFERENCES

- [1] C. G. Tsinos, S. Maleki, S. Chatzinotas, and B. Ottersten, "On the energy-efficiency of hybrid analog-digital transceivers for large antenna array systems," in *International Conference on Communications, Paris, France*, May 2017.
- [2] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. on Wireless Communications*, vol. 14, no. 2, pp. 1100–1114, Feb 2015.
- [3] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [4] S. K. Yong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, pp. 1–10, 2006.
- [5] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, June 2011.
- [6] R. C. Daniels and R. W. Heath, "60 GHz wireless communications: emerging requirements and design recommendations," *IEEE Vehicular Technology Magazine*, vol. 2, no. 3, pp. 41–50, Sept 2007.
- [7] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. on Communications*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.
- [8] X. Song, C. Jans, L. Landau, D. Cvetkovski, and G. Fettweis, "A 60GHz LOS MIMO backhaul design combining spatial multiplexing and beamforming for a 100Gbps throughput," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–6.
- [9] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. on Communications*, vol. 63, no. 9, pp. 3029–3056, Sept 2015.
- [10] T. Baykas, C. S. Sum, Z. Lan, J. Wang, M. A. Rahman, H. Harada, and S. Kato, "IEEE 802.15.3c: the first IEEE wireless standard for data rates over 1 Gb/s," *IEEE Communications Magazine*, vol. 49, no. 7, pp. 114–121, July 2011.
- [11] D. J. Love and R. W. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. on Communications*, vol. 51, no. 7, pp. 1102–1110, July 2003.
- [12] C. H. Doan, S. Emami, D. A. Sobel, A. M. Niknejad, and R. W. Brodersen, "Design considerations for 60 GHz CMOS radios," *IEEE Communications Magazine*, vol. 42, no. 12, pp. 132–140, Dec 2004.
- [13] A. Hajimiri, H. Hashemi, A. Natarajan, X. Guan, and A. Komijani, "Integrated phased array systems in silicon," *Proceedings of the IEEE*, vol. 93, no. 9, pp. 1637–1655, Sept 2005.
- [14] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Trans. on Wireless Communications*, vol. 4, no. 4, pp. 1759–1772, July 2005.
- [15] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Trans. on Signal Processing*, vol. 51, no. 11, pp. 2796–2807, Nov 2003.
- [16] S. Sanayei and A. Nosratinia, "Antenna selection in mimo systems," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68–73, Oct 2004.
- [17] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, March 2014.
- [18] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. on Signal Processing*, vol. 53, no. 11, pp. 4091–4103, Nov 2005.
- [19] J. Ahmadi-Shokouh, S. H. Jamali, and S. Safavi-Naeini, "Optimal receive soft antenna selection for MIMO interference channels," *IEEE Trans. on Wireless Communications*, vol. 8, no. 12, pp. 5893–5903, December 2009.
- [20] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct 2014.

- [21] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec 2007.
- [22] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. on Information Theory*, vol. 50, no. 10, pp. 2231–2242, Oct 2004.
- [23] X. Yu, J. C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, April 2016.
- [24] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, April 2016.
- [25] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, April 2016.
- [26] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Trans. on Wireless Communications*, vol. 14, no. 5, pp. 2670–2683, May 2015.
- [27] J. A. Zhang, X. Huang, V. Dyadyuk, and Y. J. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 79–87, February 2015.
- [28] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. on Communications*, vol. 64, no. 5, pp. 1801–1818, May 2016.
- [29] —, "Gram schmidt based greedy hybrid precoding for frequency selective millimeter wave MIMO systems," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 3396–3400.
- [30] J. Lee and Y. H. Lee, "AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing," in *2014 IEEE International Conference on Communications (ICC)*, June 2014, pp. 5838–5842.
- [31] C. G. Tsinos, S. Maleki, S. Chatzinotas, and B. Ottersten, "Hybrid analog-digital transceiver designs for cognitive radio millimeter wave systems," in *50th Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 1785–1789.
- [32] C. G. Tsinos and S. Maleki and S. Chatzinotas and B. Ottersten, "Hybrid analog-digital transceiver designs for cognitive large-scale antenna array systems," *arXiv preprint arXiv:1612.02957*, 2017.
- [33] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.
- [34] A. J. G. Rodriguez, V. Venkateswaran, P. Rulikowski, and C. Masouros, "Hybrid analog-digital precoding revisited under realistic RF modeling," *IEEE Wireless Communications Letters*, vol. PP, no. 99, pp. 1–1, 2016.
- [35] P. Sudarshan, N. B. Mehta, A. F. Molisch, and J. Zhang, "Antenna selection with RF pre-processing: robustness to RF and selection non-idealities," in *Radio and Wireless Conference, 2004 IEEE*, Sept 2004, pp. 391–394.
- [36] V. Venkateswaran, F. Pivitt, and L. Guan, "Hybrid RF and digital beamformer for cellular networks: Algorithms, microwave architectures, and measurements," *IEEE Trans. on Microwave Theory and Techniques*, vol. 64, pp. 2226–2243, July 2016.
- [37] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. on Wireless Communications*, vol. 11, no. 8, pp. 2946–2957, August 2012.
- [38] R. S. Prabhu and B. Daneshrad, "Energy-efficient power loading for a MIMO-SVD system and its performance in flat fading," in *Global Telecommunications Conference, GLOBECOM, IEEE*, Dec 2010, pp. 1–5.
- [39] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals," *IEEE Trans. on Signal Processing*, vol. 59, no. 1, pp. 329–340, Jan 2011.
- [40] Z. Wang, I. Stupia, and L. Vandendorpe, "Energy efficient precoder design for MIMO-OFDM with rate-dependent circuit power," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 1897–1902.
- [41] V. S. Varma, S. Lasaulce, M. Debbah, and S. E. Elayoubi, "An energy-efficient framework for the analysis of MIMO slow fading channels," *IEEE Trans. on Signal Processing*, vol. 61, no. 10, pp. 2647–2659, May 2013.
- [42] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, June 2015.
- [43] S. He, Y. Huang, L. Yang, B. Ottersten, and W. Hong, "Energy efficient coordinated beamforming for multicell system: Duality-based algorithm design and massive mimo transition," *IEEE Trans. on Communications*, vol. 63, no. 12, pp. 4920–4935, Dec 2015.
- [44] A. Kalantari, S. Maleki, S. Chatzinotas, and B. Ottersten, "Secrecy energy efficiency optimization for MISO and SISO communication networks," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2015, pp. 21–25.
- [45] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Energy-efficient symbol-level precoding in multiuser MISO based on relaxed detection region," *IEEE Trans. on Wireless Communications*, vol. 15, no. 5, pp. 3755–3767, May 2016.
- [46] D. P. Bertsekas, "Nonlinear programming," 1999.
- [47] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [48] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 653–656, Dec 2014.
- [49] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967–2976, Aug 2005.
- [50] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089–1098, Aug 2004.
- [51] D. M. Pozar, *Microwave Engineering, 4th Edition*. John Wiley & Sons, 2011.
- [52] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, Aug 2006.
- [53] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Golden section search in one dimension," *Numerical Recipes in C: the Art of Scientific Computing*, vol. 2, 1992.
- [54] J. Eckstein and D. P. Bertsekas, "On the DouglasRachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.
- [55] E. G. Golshtein and N. Tretyakov, "Modified lagrangians in convex programming and their generalizations," in *Point-to-Set Maps and Mathematical Programming*. Springer, 1979, pp. 86–97.
- [56] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, "An alternating direction algorithm for matrix completion with nonnegative factors," *Frontiers of Mathematics in China*, vol. 7, no. 2, pp. 365–384, 2012.
- [57] S. Boyd and L. Vandenberghe, *Numerical Linear Algebra Background*. <https://web.stanford.edu/class/ee364a/lectures/>.
- [58] T. Michaeli and Y. C. Eldar, "Constrained nonlinear minimum MSE estimation," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2008, pp. 3681–3684.
- [59] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. on Vehicular Technology*, vol. 56, no. 4, pp. 1924–1934, July 2007.
- [60] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.



Christos G. Tsinos (S08-M14) received the Diploma degree in computer engineering and informatics, the MSc and the PhD degree in signal processing and communication systems and MSc in applied mathematics from the University of Patras, Greece, in 2006, 2008, 2013 and 2014 respectively. From August 2014 to June 2015 he was a Postdoctoral Researcher at University of Patras. Since July 2015 he joined as a Research Associate the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg. He

is involved or was involved in the past in a number of different R&D projects funded by national and/or EU funds. He is currently the PI of R&D Project ECLECTIC (Energy and Complexity Efficient millimeter-wave Large-Array Communications), funded under FNR CORE Framework. His current research interests include signal processing for mmWave, massive MIMO, cognitive radio and satellite communications and hyperspectral image processing, as well. Dr. Tsinos is a member of the Technical Chamber of Greece.



Sina Maleki (M13) received the B.Sc. degree from the Iran University of Science and Technology, Tehran, Iran, in 2006, and the M.Sc. and Ph.D. degrees from the Delft University of Technology, Delft, The Netherlands, in 2009 and 2013, respectively. From 2008 to 2009, he was an intern student with the Philips Research Center, Eindhoven, The Netherlands, researching on spectrum sensing for cognitive radio networks. Since 2013, he has been with the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg,

where he is working on cognitive radio for satellite communications as well as interference detection and localization in satellite communications, within the EU H2020 Project SANSA, EU FP7 Project CoRaSat, as well as Luxembourgish National Projects SATSENT and SeMIGod.



Symeon Chatzinotas (S06-M09-SM13) is currently the Deputy Head of the SIGCOM Research Group, Interdisciplinary Centre for Security, Reliability, and Trust, University of Luxembourg, Luxembourg. In the past, he has worked on numerous R&D projects for the Institute of Informatics Telecommunications, National Center for Scientific Research Demokritos, Institute of Telematics and Informatics, Center of Research and Technology Hellas, and Mobile Communications Research Group, Center of Communication Systems Research, University of Surrey,

Surrey, U.K. He has received the M.Eng. degree in telecommunications from Aristotle University of Thessaloniki, Thessaloniki, Greece, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, Surrey, U.K., in 2003, 2006, and 2009, respectively. Dr. Chatzinotas has more than 200 publications, 1600 citations and an H-Index of 22 according to Google Scholar. He is the co-recipient of the 2014 Distinguished Contributions to Satellite Communications Award, and Satellite and Space Communications Technical Committee, IEEE Communications Society, and CROWNCOM 2015 Best Paper Award. His research interests include multiuser information theory, co-operative/cognitive communications and wireless networks optimization.



Björn Ottersten (S87-M89-SM99-F04) received his M.S. degree in electrical engineering and applied physics from Linköping University, Sweden, in 1986, and his Ph.D. degree in electrical engineering from Stanford University, California, in 1989. He has held research positions at the Department of Electrical Engineering, Linköping University; the Information Systems Laboratory, Stanford University; the Katholieke Universiteit Leuven, Belgium; and the University of Luxembourg. From 1996 to 1997, he was the director of research at ArrayComm Inc,

a start-up in San Jose, California, based on his patented technology. In 1991, he was appointed a professor of signal processing with the Royal Institute of Technology (KTH), Stockholm, Sweden. From 1992 to 2004, he was the head of the Department for Signals, Sensors, and Systems, KTH, and from 2004 to 2008, he was the Dean of the School of Electrical Engineering, KTH. Currently, he is the director of the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. As Digital Champion of Luxembourg, he acts as an adviser to the European Commission. His research interests include security and trust, reliable wireless communications, and statistical signal processing. He is a Fellow of the EURASIP and has served as a member of the IEEE Signal Processing Society Board of Governors. He has served as an Associate Editor for IEEE Transactions on Signal Processing and on the Editorial Board of IEEE Signal Processing Magazine. He is currently Editor-in-Chief of the EURASIP Signal Processing Journal and a member of the Editorial Boards of the EURASIP Journal of Applied Signal Processing and Foundations and Trends in Signal Processing. He coauthored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, and 2013, and three IEEE conference papers receiving Best Paper Awards. He was the recipient of the IEEE Signal Processing Society Technical Achievement Award in 2011. He was a first recipient of the European Research Council Advanced Research Grant.