

# **A Proposed Efficient Hybrid Precoding Algorithm for Millimeter Wave Massive MIMO 5G Networks**

**A. Abdelaziz Salem1  [·](http://orcid.org/0000-0002-2399-180X) S. El‑Rabaie1 · Mona Shokair1**

Published online: 2 January 2020 © Springer Science+Business Media, LLC, part of Springer Nature 2020

# **Abstract**

Millimeter wave (mmWave) has been regarded as a candidate cellular band for future 5G networks. It exhibits an extension for current cellular bands, where its integrity with massive multiple-input-multiple output (MIMO) and precoding techniques ofers a signifcant capacity improvements. Unfortunately, analog and/or digital precoders are not energy-efficient from mmWave massive MIMO perspectives whereas, one radio frequency (RF) chain per antenna element is required. Accordingly, the hybrid precoding techniques could be introduced as cost-efective solution. It implicates a low dimensional precoding that can be executed in digital domain, followed by a large dimensional analog beam-formers to steer antenna elements. In this paper, we propose a novel low complex and efficient hybrid precoding algorithm to design the analog and digital precoders/combiners for mmWave massive MIMO transceiver. We will show that our proposed precoder approaches the fully digital (unconstrained) precoder with a negligible performance loss that makes it stand as benchmark case. Particularly, we simply design the digital precoding stage based on modifed water-flling wherein, the orthogonality criterion among transmitted data streams is guaranteed. Besides, the analog beam-formers are optimally designed via extracting steering angles from an alternative precoder which is derived from a tight and simple upper bound expression. Furthermore, the proposed algorithm will be extended to include the practical analog beam-formers that have limited phase shifter with fnite angle resolution. As a result, we will develop a quantization technique, for the analog precoder, with precision up to two bits. Furthermore, we investigate the energy efficiency (EE) performance wherein, EE degradation can be avoided even with one bit of quantization. This will be confrmed through comparing with unconstrained and unquantized precoders. Also, spectral efficiency (SE) performance displays a remarkable gain when it is fairly compared with the state-of-art.

**Keywords** Hybrid precoding · MMWave · Massive MIMO · 5G thnologies

 $\boxtimes$  A. Abdelaziz Salem ahmed.abdalaziz40@el-eng.menofa.edu.eg

<sup>&</sup>lt;sup>1</sup> Faculty of Electronic Engineering, Menoufia University, 32952 Menouf, Egypt

# **1 Introduction**

In the future 5G wireless networks, boosting SE performance is inevitably required to cover the explosive growth of spectrum demands. Bandwidth shrinkage pushes network providers to improve the system capacity via resorting to physical layer techniques wherein, SE can be improved by massive MIMO implementations and/or efective channel estimations  $[1]$  $[1]$ . Further capacity enhancements could be achieved through network densification  $[2, 3]$  $[2, 3]$  $[2, 3]$  $[2, 3]$  $[2, 3]$ , device-to-device (D2D) communications [\[4](#page-16-3)], and cloud RANs [[5](#page-16-4)]. According to the facts of spectrum scarcity and progressive growth of data rates, exploiting a new cellular band is envisioned as crucial necessity.

The mmWave bands (30 GHz - 300 GHz) have been considered for supporting high data rate with indoor and outdoor point-to-point communications [[6](#page-16-5)]. Thanks to small wavelength of mmWave, massive array of dipoles can be packed up in a small plate which offers a signifcant array gain to struggle the huge path loss. Moreover, it provides the capability to synthesize high directive beams which able to convey spatially multiplexing data streams. Further, more privileges of gain and performance can be achieved when communication systems are characterized by the integrity of mmWave, massive MIMO and hybrid precoding. On the other hand, deploying mmWave for 5G backhauling brings critical challenges. For instance, path loss proportionally increases with square of frequency and the propagation loss caused by rain attenuation [\[7](#page-16-6)].

Unfortunately, precoding over mmWave massive MIMO communications cannot be accomplished entirely at the digital domain, where it is infeasible to consider one RF chain per antenna element as in conventional MIMO systems. In fact, this limitation leads to an excessive cost of implementation complexity and power consumption. Hence, the hybrid precoding has been considered to break this limitation by reducing the number of RF chains which is responsible for joining the digital precoding stage with the analog one [[8\]](#page-16-7). The circuit of analog precoder is designed without regarding variable voltage amplifers (VGAs) [[9](#page-16-8)] where, the power loss is reduced at the cost of sacrifcing with the ability of controlling the RF magnitude. Although, the circuit design of phase shifter has been concerned in many studies [\[10–](#page-16-9)[13\]](#page-17-0), but they didn't regard designing low complex hybrid precoding algorithms for mmWave massive MIMO systems. The analog beam-formers [[10](#page-16-9), [12](#page-16-10)] are implemented for near optimal performance while achieving the full diversity. Authors of [[11](#page-16-11)] proposed an iterative limited-power baseband precoding to handle multidata streams where, both diversity and multiplexing were addressed at a limited number of the RF chains. Literature [[13](#page-17-0)] discussed the quantized phase control to introduce a precise analog to digital converter  $(A/D)$ . Although, the studies  $[10-13]$  $[10-13]$  argued issues of phase quantization and capability of handling data streams, but they didn't investigate any of performance metrics.

In the other hand, hybrid precoding techniques with mmWave massive MIMO commu-nication systems has been studied extensively [\[14–](#page-17-1)[18](#page-17-2)]. Many efforts have been introduced towards hybrid precoding structure [[8](#page-16-7), [19](#page-17-3)[–26\]](#page-17-4). Particularly, orthogonal matching pursuit (OMP) is the most vastly used algorithm wherein, the steering angles of analog beamformers were extracted from eigenvectors of mmWave channel [\[8,](#page-16-7) [19](#page-17-3), [20\]](#page-17-5). Hybrid precoding based on OMP was regarded as sparse constrained problem, due to mmWave nature, with capability of matrix reconstruction as being followed in image processing feld. However, the used constraints and approximations in OMP solution lead to unavoidable performance loss. Authors of studies [[19](#page-17-3)] and [[23](#page-17-6)] focused on reducing OMP complexity through exploiting the result of matrix inversion in each iteration of their proposed algorithm.

Most of the recent studies argued the optimality of hybrid structure from hardware design perspectives. Authors of [\[24\]](#page-17-7) studied the optimal hybrid precoder when RF chains number at least double of transmitted data streams. Nonetheless, this design didn't take into account the solution limits at wide range of RF chains and data streams. On the other hand, authors of [[27](#page-17-8)] investigated the impact of RF chains comparing with both data streams and antenna elements for point-to-point and multi-user scenarios. RF stage was designed without regarding VGAs to reduce power consumption at the analog domain[[26](#page-17-4)]. Further, Alternating optimization algorithm based hybrid prcoding was proposed in pioneering work  $[28]$  $[28]$  $[28]$  to reduce design complexity. Wherein, the spectral efficiency of hybrid precoders is directly optimized in Riemannian manifold at cost of overhead time complexity stemmed from nested loops of iterative conjugate gradient algorithm. The authors extended their work to develop a low complex precoders based on alternating optimization but, they neither considered aligning channel directions nor maintained the orthogonality among transmitted streams. Further, the limited resolution of practical phase shifters hasn't been regarded. Hence, their design aspects still require more investigations.

The majority of aforementioned studies have focused on design aspects to simplify the complexity of hybrid precoders in mmwave massive MIMO systems. This usually leads to performance loss or requires designing complex algorithms. Moreover, These studies didn't concern the integration between design simplicity and practical aspects without sacrifcing performance. As a result, we have been motivated to fnd a new algorithmic design that comprises low complexity and practical considerations, while maintaining performance approximates the unconstrained precoder.

## **1.1 Contributions**

In this section, we present a novel low complex hybrid precoding algorithm wherein, both analog and digital precoders/combiners are optimized to approach the unconstrained precoder. In more details, our contributions can be listed in the following points:

- We exploit channel information to initialize analog precoder, where a large array gain can be harvested through aligning channel modes. Then, analog beam-formers are updated by an alternative simple precoder which is derived from tight upper bound of hybrid precodering problem. This will result in low cost processing in the analog domain. Further, we enforce unit modulus constraint for each entity of analog precoding matrix to save energy consumption. As well, the digital precoder will be designed based on the legacy of modifed water-flling after some mathematical transformation where, orthogonality among transmitted data streams is guaranteed.
- With respect to the practical design aspects, we adopt phase quantization for analog beam-formers via minimizing the euclidean distance between original phases and the quantized ones. Hence, our hybrid precoding algorithm can be applied for practical phase shifters that have a limited resolution of precision. So far, simulation results will display SE comparison among quantized analog precoders (for one and two bits of quantization) and the unquantized one. Results will reveal that there is no considerable performance loss. Hence, our proposed approach not only combines low complexity and practical aspects but also, it stands as a benchmark of performance.
- Furthermore, we consider power consumption model, beside the previous SE evaluations, for the purpose of evaluating overall EE performances. Accordingly, we will debate the interplay between performance loss and capacity improvement.

– Finally, we fairly compare our proposed hybrid precoders including the quantized versions with state-of-the-art: OMP, unconstrained digital precoder and analog precoder. Results will demonstrate that our proposed design outperform OMP and approach the unconstrained precoder with marginal losses.

### **1.2 Paper Organization and Notations**

The rest of the paper is organised as follows. Section [2](#page-3-0) states the system model including the scattering channel model of mmWave, followed by problem formulation. Then, we optimize hybrid precoders and consider practical phase shifters as in Sect. [3](#page-6-0). Simulation results reveal the rigidity of proposed hybrid algorithm in Sect. [4](#page-11-0). Finally, Sect. [5](#page-16-12) will present a brief conclusion which followed by the more relevant references.

Note that, the bold face capital and small letters stand for matrix and vector, respectively.  $\mathbf{X}_{i,j}$  refers to entry on *ith* row and *jth* column of matrix **X**. The conjugate, transpose and Hermitian of matrix **X** are represented as  $\mathbf{X}^*, \mathbf{X}^T$  and  $\mathbf{X}^H$ ;  $Tr(\mathbf{X})$ , and  $\|\mathbf{X}\|_F$  denote trace and Frobenius norm; **X**<sup>†</sup> and **X**<sup>−1</sup> express the Moore-Penrose pseudo inverse and the inverse of matrix **X**, respectively.  $\mathbb{R}\{.\}$  represents the real part of complex variable and  $\mathbb{E}[.\]$  denotes the expectation of random variable.

### <span id="page-3-0"></span>**2 System Model and Problem Formulations**

In this section, we consider single cell downlink mmWave massive MIMO system. Then, hybrid precoding problem will be formulated.

#### **2.1 System Model**

Figure [1](#page-3-1) considers hybrid precoding structure wherein, mmWave channel and massive MIMO are taken into account. Base station  $(BS)$  is equipped with  $N_t$  transmitting antennas which are stacked as a uniform plane array (UPA) to transmit  $N<sub>s</sub>$  data streams. At the receiver side, we assume a single user having  $N_r$  receiving antennas. The number of RF chains, which interfacing the digital precoder with the analog one, at transmitter and receiver are designated by  $N_{RF}^t$  and  $N_{RF}^r$ ; such that  $N_s \leq N_{RF}^t \leq N_t$  and  $N_s \leq N_{RF}^r \leq N_r$ , respectively.

The transmitted signal  $N_t \times 1$  vector is mathematically written as  $\mathbf{x} = \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s}$ . The average transmitted power of Gaussian data symbol, given by  $N_s \times 1$  vector, can be expressed as  $\mathbb{E}[\mathbf{s}^H \mathbf{s}] = \frac{\mathbf{I}_{N_S}}{N_s}$ . As shown in Fig. [1,](#page-3-1) the hybrid precoding structure consists of



<span id="page-3-1"></span>**Fig. 1** Fully connected hybrid precoding transceiver system

two processing stages: an  $N_{RF}^t \times N_s$  low dimensional digital (baseband) precoder  $\mathbf{F}_{BB}$ ; followed by an  $N_t \times N_{RF}^t$  high dimensional analog precoder  $\mathbf{F}_{RF}$ . It should be noted that the total processing power of hybrid precoders is normalized to satisfy  $\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = N_s$ . The received signal after decoding processes is denoted by the following expression received signal after decoding processes is denoted by the following expression

$$
y = \sqrt{\rho} \left( \mathbf{W}_{RF} \mathbf{W}_{BB} \right)^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s} + \left( \mathbf{W}_{RF} \mathbf{W}_{BB} \right)^H \mathbf{n}.
$$
 (1)

where  $\rho$  refers to the average received power, **H** represents  $N_r \times N_t$  scattering channel matrix. Since decoding process at receiver side resembles the precoding at transmitter then, we have an  $N_{RF}^r \times N_s$  digital decoder  $\mathbf{W}_{BB}$  followed by an  $N_r \times N_{RF}^r$  analog decoder  $\mathbf{W}_{RF}$ . **n** denotes  $N_r \times 1$  noise vector which is assumed as independent and identically distributed random variable (i.i.d) with zero mean and average power  $\sigma_n^2$ , i.e.  $\mathbf{n} \sim \mathcal{CN}(0, \sigma_n^2)$ . Further, the channel state information (CSI) is assumed perfectly known for both transceiver sides. In practical scenarios, CSI is estimated frstly at receiver through training sequence and then, it is shared with transmitter using a limited feedback channel as in FDD systems [[16](#page-17-10),  $29$ ]. Based on ([1\)](#page-4-0), the SE achieved by hybrid precoding system can be written as

<span id="page-4-3"></span><span id="page-4-2"></span><span id="page-4-1"></span><span id="page-4-0"></span>
$$
SE = \log_2 \left| \mathbf{I}_{N_s} + SINR \right| \tag{2}
$$

where, *SINR* represents the signal-to-noise-plus-interference ratio which is expressed as in ([3\)](#page-4-1).

$$
SINR = \frac{\rho}{\sigma_n^2 N_s} \cdot \left| \frac{\left(\mathbf{W}_{RF} \mathbf{W}_{BB}\right)^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \left(\mathbf{F}_{RF} \mathbf{F}_{BB}\right)^H \mathbf{H}^H \left(\mathbf{W}_{RF} \mathbf{W}_{BB}\right)}{\left(\mathbf{W}_{RF} \mathbf{W}_{BB}\right)^H \left(\mathbf{W}_{RF} \mathbf{W}_{BB}\right)} \right| \tag{3}
$$

Let  $\mathbf{A} = \mathbf{W}_{RF} \mathbf{W}_{BB}$ , then we can rewrite [\(3](#page-4-1)) by [\(4](#page-4-2)).

<span id="page-4-4"></span>
$$
SINR = \frac{\rho}{\sigma_n^2 N_s} \cdot \left| \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{A}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \left( \mathbf{F}_{RF} \mathbf{F}_{BB} \right)^H \mathbf{H}^H \mathbf{A} \right|, \tag{4}
$$

Then, we substitute both of [\(4\)](#page-4-2) and the well-known mathematical defnition  $A^{\dagger} = (A^H A)^{-1} A^H$  into ([2](#page-4-3)), we can get the final expression of SE in terms of precoding matrices as in  $(5)$  $(5)$  $(5)$ .

$$
SE = \log_2 \left| \left( \mathbf{I}_{N_s} + \frac{\rho}{\sigma_n^2 N_s} \left( \mathbf{W}_{RF} \mathbf{W}_{BB} \right)^{\dagger} \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \left( \mathbf{F}_{RF} \mathbf{F}_{BB} \right)^H \mathbf{H}^H \left( \mathbf{W}_{RF} \mathbf{W}_{BB} \right) \right) \right|.
$$
 (5)

It should be noted that the hybrid structure in Fig. [1](#page-3-1) includes connecting each RF chain with each antenna element through a network of phase shifters  $N_p = N_t N_{RF}^t$ . This configuration is called *fully connected structure* which is remarked by full beam-forming gain due to large degree of freedom (DoF). Nevertheless, it has a hardware complexity that can be mitigated by reducing  $N_p$  or by replacing this structure with the *partially connected confguration* [[30](#page-17-12)–[33](#page-17-13)] at the cost of sacrifcing DoF.

### **2.2 Channel Model**

Millimeter wave channel model is characterized by clusters of scatterers (Saleh-Valenzuela model) whereas, the signal propagation is subjected to high free-space path loss. This model can be mathematically represented by the following [\[9](#page-16-8)]

$$
\mathbf{H} = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}} \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{il} \mathbf{a}_r (\phi_{il}^r, \theta_{il}^r) \mathbf{a}_t (\phi_{il}^t, \theta_{il}^t)^H.
$$
(6)

where  $N_{cl}$  states the number of clusters wherein, an  $N_{rav}$  rays can be described by path gain  $\alpha_{il}$ . We assume that  $\alpha_{il} \sim \mathcal{CN}\left(0, \sigma_{\alpha,i}^2\right)$ , such that the total average power of all clusters is indicated by  $\sum_{i=1}^{N_{cl}} \sigma_{\alpha,i}^2 = \hat{\gamma}$ . This, also, satisfies normalized average channel gain described by  $\mathbb{E}[\|\mathbf{H}\|_F^2] = N_t N_r$ . In addition,  $\mathbf{a}_t(\phi_t^t, \theta_t^t)$  and  $\mathbf{a}_r(\phi_t^t, \theta_t^r)$  represent transmit and receive array response vector for given azimuth  $[\phi_i^i, \phi_i^r]$  and elevation angles  $[\theta_i^i, \theta_i^r]$  of departure and arrival, respectively. In this paper, we suppose that massive MIMO elements are installed uniformly in *y* − *z* plane (UPA) where an  $\sqrt{N_t} \times \sqrt{N_t}$  antenna elements are equally spaced. Based on this setting, the array response of *lth* ray in *ith* cluster is given as

$$
\mathbf{a}(\phi_{il}, \theta_{il}) = \frac{1}{\sqrt{N}} \left[ 1, ..., e^{j\frac{2\pi}{\lambda}d(p\sin\phi_{il}\sin\theta_{il} + q\cos\theta_{il})}, ..., e^{j\frac{2\pi}{\lambda}d\left((\sqrt{N}-1)\sin\phi_{il}\sin\theta_{il} + (\sqrt{N}-1)\cos\theta_{il}\right)} \right]^T
$$
\n(7)

where *p* and *q* are the antenna indexes in 2D which is defined by  $0 \le p \le \sqrt{N}$  and  $0 \le q \le \sqrt{N}$ . Noting that the antenna spacing is denoted by *d* which is usually scaled by wavelength λ.

### **2.3 Problem Formulation**

Inspired by [\[8\]](#page-16-7) and [[19](#page-17-3)], the design of hybrid precoders/decoders can be decoupled into two individual sub problems: one handles analog precoding/combining problem with a unit modulus constraint of steering angles and the other solves digital precoding/combining problem constrained by an extra power constraint. Generally speaking, The hybrid precoding problem can be described by

<span id="page-5-0"></span>
$$
\begin{aligned}\n\underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}}{\text{minimize}} & \|\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_{F} \\
\text{subject to} & \mathbf{F}_{RF} \in \mathcal{Y}, \\
& \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_{F}^{2} = N_{s}.\n\end{aligned} \tag{8}
$$

where  $\mathbf{F}_{FD}$  stands for an  $N_t \times N_s$  fully digital precoder which is considered as the optimal unconstrained (virtual/digital) precoder. The frst constraint emphasizes on the feasibility of analog precoder through feasible set  $\Psi$  of steering angles, wherein unit modulus constraint of each entity has been regarded, i.e.  $|(\mathbf{F}_{RF})_{i,j}| = |(\mathbf{W}_{RF})_{i,j}| = 1$ . It has been proved<br>that the optimization problem in (8) is approximately equivalent to maximization of spec-that the optimization problem in [\(8\)](#page-5-0) is approximately equivalent to maximization of spectral efficiency in  $(5)$  $(5)$ . It is intuitive that approaching SE performance of virtual precoder can be achieved by making the hybrid precoders are sufficiently proximate the fully digital precoder. Noting that the unconstrained precoder matrices are directly deduced form the decomposition of channel matrix, i.e. the corresponding  $N<sub>s</sub>$  columns of channel eigenvectors.

The formulation in ([8](#page-5-0)) is usually named as *"matrix factorization problem"* where, we will adopt an iterative low complex algorithm to optimize the precoding matrices,  $\mathbf{F}_{RF}$  and  $\mathbf{F}_{BB}$ . In the fact, simultaneous optimization of the precoding matrices appears very complicated task due to aforementioned constraints of both analog and digital precoders. Hence, decoupling the objective in [\(8](#page-5-0)) facilitates getting the optimal solution with less complexity.

### <span id="page-6-0"></span>**3 Proposed Hybrid Precoding Algorithm**

Since the computational complexity is a crucial factor for practical designs, we have to propose a low complex and efficient hybrid precoding algorithm. In the following section, we demonstrate the main ideas of the proposed precoders, followed by the corresponding pseudo code.

#### **3.1 Digital Precoder Design Concept**

Similar to orthogonal columns of the unconstrained/fully digital precoder  $\mathbf{F}_{FD}$ , we shall enforce orthogonality constraint on the digital precoder  $\mathbf{F}_{BB}$ . This property relieves the experienced interference among the multiplexed data streams. Then, we will deploy the concept of water-flling to maintain this orthogonality depending on the efective channel  $\mathbf{H}_{e} = \mathbf{H}\mathbf{F}_{RF}$ . Let's assume the condition that satisfies orthogonal property as the following

<span id="page-6-1"></span>
$$
\mathbf{F}_{BB}^H \mathbf{F}_{BB} = \beta \mathbf{F}_D^H \beta \mathbf{F}_D = \beta^2 \mathbf{I}_{N_s} \tag{9}
$$

where  $\mathbf{F}_D$  denote unitary matrix with dimension  $N_{RF}^t \times N_s$  which is the same of  $\mathbf{F}_{BB}$ . The  $\beta$  stands for weight power coefficient or scaling factor that will be discussed in the following section. To optimize the digital precoder,the analog beam-formers are initialized with a suboptimal solution from which, an array gain and channel alignment could be earned. In particular, CSI is exploited to initialize the steering angles as  $\mathbf{F}_{RF}^{(0)} = e^{j\angle \mathbf{H}^H}$ . Once those angles have been obtained, it became a temporary solution for evaluating  $\mathbf{F}_{RR}$ . Then, hybrid beam-formers are updated as will be shown in the following sections.

#### **3.2 Upper Bound of Hybrid Precoding Problem**

In this section, we drive an upper bound of hybrid precoding problem by substituting ([9](#page-6-1)) into ([8\)](#page-5-0) and then recast the yielding results as in the following lines

<span id="page-6-2"></span>
$$
\begin{split}\n\|\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_{F}^{2}, \\
&= Tr\left( (\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB})^{H} (\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB}) \right), \\
&= Tr(\mathbf{F}_{FD}^{H}\mathbf{F}_{FD}) - Tr(\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\mathbf{F}_{BB}) \\
&- Tr(\mathbf{F}_{BB}^{H}\mathbf{F}_{RF}^{H}\mathbf{F}_{FD}) + Tr(\mathbf{F}_{BB}^{H}\mathbf{F}_{RF}^{H}\mathbf{F}_{RF}\mathbf{F}_{BB}), \\
&= Tr(\mathbf{F}_{FD}^{H}\mathbf{F}_{FD}) - \beta Tr(\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\mathbf{F}_{D}) \\
&- \beta Tr((\mathbf{F}_{RF}\mathbf{F}_{D})^{H}\mathbf{F}_{FD}) + \beta^{2} Tr((\mathbf{F}_{RF}\mathbf{F}_{D})^{H}\mathbf{F}_{D}\mathbf{F}_{RF}), \\
&= Tr(\mathbf{F}_{FD}^{H}\mathbf{F}_{FD}) - 2\beta Tr(\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\mathbf{F}_{D}) \\
&+ \beta^{2} Tr((\mathbf{F}_{RF}\mathbf{F}_{D})^{H}\mathbf{F}_{D}\mathbf{F}_{RF}), \\
&= \|\mathbf{F}_{FD}\|_{F}^{2} - 2\beta \Re \{Tr(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{RF})\} + \beta^{2} \|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2}\n\end{split}
$$

According to [\(10\)](#page-6-2), increasing the second term minimizes the objective function of [\(8](#page-5-0)). Hence, We can directly find  $\beta$  coefficient which satisfies minimization of the objective function by differentiating  $(10)$  $(10)$  $(10)$  with respect to  $\beta$  and equating the result with zero to get the following

<span id="page-7-2"></span><span id="page-7-0"></span>
$$
\beta = \frac{\Re\left\{Tr\left(\mathbf{F}_D\mathbf{F}_{FD}^H\mathbf{F}_{RF}\right)\right\}}{\|\mathbf{F}_{RF}\mathbf{F}_D\|_F^2}.
$$
\n(11)

To ensure feasibility of  $\beta$ , we substitute ([11](#page-7-0)) into [\(10\)](#page-6-2) to obtain a new form of the objective function as in the following

$$
\|\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_{F}^{2} = \|\mathbf{F}_{FD}\|_{F}^{2} - \frac{2(\Re\left\{Tr\left(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\right)\right\})^{2}}{\|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2}} + \frac{\Re\left\{Tr\left(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\right)\right\}}{\left(\|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2}\right)^{2}} \cdot \|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2},
$$
\n(12)\n
$$
\|\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_{F}^{2} = \|\mathbf{F}_{FD}\|_{F}^{2} - \frac{(\Re\left\{Tr\left(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\right)\right\})^{2}}{\|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2}}.
$$

According to the orthogonality constraint of [\(9](#page-6-1)), the upper limit of  $\|\mathbf{F}_{RF}\mathbf{F}_{D}\|$ <br>ematically deduced as  $\frac{2}{F}$  can be mathematically deduced as

<span id="page-7-1"></span>
$$
\|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2} = Tr(\mathbf{F}_{RF}\mathbf{F}_{D}\mathbf{F}_{D}^{H}\mathbf{F}_{RF}^{H}),
$$
  
let  $SVD(\mathbf{F}_{D}\mathbf{F}_{D}^{H}) = \mathbf{K} \begin{pmatrix} \mathbf{I}_{N_{s}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{K}^{H},$   

$$
\therefore \|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2} = Tr\left(\begin{pmatrix} \mathbf{I}_{N_{s}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{K}^{H}\mathbf{F}_{RF}^{H}\mathbf{F}_{RF}\mathbf{K}\right),
$$
  

$$
\leq Tr(\mathbf{F}_{RF}^{H}\mathbf{F}_{RF}).
$$
  

$$
\therefore \|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2} \leq \|\mathbf{F}_{RF}\|_{F}^{2}.
$$
 (13)

where **K** stands for eigenvectors of singular value decomposition (SVD). Noting that the result of this decomposition turns to identity if  $\mathbf{F}_D$  is always square matrix, i.e. the inequal-ity in [\(13\)](#page-7-1) is converted to strict equality if and only if  $N_s = N_{RF}^t$ . In the following lines, we utilize the result of  $(13)$  $(13)$  $(13)$  to get the upper bound of  $(12)$ 

<span id="page-7-3"></span>
$$
\|\mathbf{F}_{FD} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_F \le \|\mathbf{F}_{FD}\|_F^2 - \frac{\left(\Re\left\{Tr\left(\mathbf{F}_D\mathbf{F}_{FD}^H\mathbf{F}_{RF}\right)\right\}\right)^2}{\|\mathbf{F}_{RF}\|_F^2}
$$
(14)

By introducing some algebraic manipulation on the right-hand side of ([14](#page-7-3)), we can get

$$
\left[ \|\mathbf{F}_{FD}\|_{F}^{2} - \frac{\left(\Re\left\{Tr\left(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\right\}\right\}\right)^{2}}{\|\mathbf{F}_{RF}\|_{F}^{2}} + \frac{\|\mathbf{F}_{FD}\|_{F}^{2}}{2\|\mathbf{F}_{RF}\|_{F}^{2}} \right] - \|\mathbf{F}_{FD}\|_{F}^{2} + \frac{1}{2}\left\|X2\|\mathbf{F}_{RF}\right\|_{F}^{2}}{\left\|\mathbf{F}_{FD}\right\|_{F}^{2}} \right]
$$
\n
$$
= \|\mathbf{F}_{FD}\|_{F}^{2} - 2\left(\Re\left\{Tr\left(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{RF}\right)\right\}\right)^{2} + \|\mathbf{F}_{RF}\|_{F}^{2}
$$
\n
$$
\therefore \|\mathbf{F}_{FD}\|_{F}^{2} \approx Tr\left(\mathbf{F}_{D}\mathbf{F}_{FD}^{H}\mathbf{F}_{FD}\mathbf{F}_{FD}^{H}\right)
$$
\n
$$
\therefore \|\mathbf{F}_{FD}\|_{F}^{2} = \left\|\mathbf{F}_{FD}\mathbf{F}_{D}^{H}\right\|_{F}^{2} \forall N_{s} = N_{RF}^{t}
$$

Plugging all the previous steps into  $(14)$  $(14)$  $(14)$ , the final upper bound expression is denoted as

<span id="page-8-0"></span>
$$
\left\| \mathbf{F}_{FD} - \mathbf{F}_{RF} \mathbf{F}_{BB} \right\|_{F}^{2} \le \left\| \mathbf{F}_{FD} \mathbf{F}_{D}^{H} - \mathbf{F}_{RF} \right\|_{F}^{2}.
$$
 (15)

#### **3.3 Hybrid Precoder Design**

As we mentioned before that directly optimizing the main objective function in ([8](#page-5-0)) incurs an additional computational complexity. Accordingly, we minimize the upper bound ([15](#page-8-0)) instead of the original objective. Whereas, power transmission constraint is satisfed by normalizing the digital precodeing matrix after consecutive updates of the hybrid precoding matrices. In more details, the analog precoding subproblem is frstly solved by considering the upper bound [\(15\)](#page-8-0) as objective function with temporarily disabled transmit power constraint as described below

$$
\underset{\mathbf{F}_{RF}, \mathbf{F}_{D}}{\text{minimize}} \quad \left\| \mathbf{F}_{FD} \mathbf{F}_{D}^{H} - \mathbf{F}_{RF} \right\|_{F}^{2},
$$
\n
$$
\text{subject to} \quad \left| \left( \mathbf{F}_{RF} \right)_{i,j} \right| = 1, \quad \forall i, j,
$$
\n
$$
\mathbf{F}_{D}^{H} \mathbf{F}_{D} = \mathbf{I}_{N_{s}}.
$$
\n(16)

From optimization subproblem in ([16](#page-8-1)), we note that  $\mathbf{F}_{RF}$  is no longer multiplied with  $\mathbf{F}_{BB}$ as in the main objective. This leads to simplifed solution of analog precoder that can be described in the following closed form expression

<span id="page-8-2"></span><span id="page-8-1"></span>
$$
\angle \mathbf{F}_{RF} \approx \angle (\mathbf{F}_{FD} \mathbf{F}_{D}^{H}). \tag{17}
$$

It is cleared from Eq. ([17](#page-8-2)) that, analog precoder has been constructed from the corresponding phases of the equivalent precoder  $\mathbf{F}_{FD} \mathbf{F}_{D}^{H}$ . The mathematical expression of [\(17\)](#page-8-2) can be visualized as the projection of  $\mathbf{F}_{FD} \mathbf{F}_{D}^{H}$  on the feasible set of analog precoder  $\Psi$ , which is also equivalent to the first constraint in  $(8)$  $(8)$  $(8)$ .

On the other hand, the digital precoding subproblem can be considered as in the following optimization problem, wherein  $\mathbf{F}_{RF}$  is instantaneously fixed

<span id="page-8-3"></span>
$$
\begin{aligned}\n\text{minimize} & \quad \left\| \mathbf{F}_{FD} \mathbf{F}_{D}^{H} - \mathbf{F}_{RF} \right\|_{F}^{2} \\
\text{subject to} & \quad \mathbf{F}_{D}^{H} \mathbf{F}_{D} = \mathbf{I}_{N_{s}}.\n\end{aligned} \tag{18}
$$

The constraint in ([18](#page-8-3)) denotes power normalization with respect to  $\beta^2$ . Unfortunately, the aforementioned optimization problem is not convex in  $\mathbf{F}_D$ . Therefore, we can replace ([18](#page-8-3)) with another equivalent form as in the following

<span id="page-9-0"></span>
$$
\begin{array}{ll}\text{maximize} & \log_2 \left| \mathbf{I}_{N_s} + \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_D \mathbf{F}_D^H \left( \mathbf{H} \mathbf{F}_{RF} \right)^H \right| \\
\text{subject to} & \text{Tr} \left( \mathbf{Q} \mathbf{F}_D \mathbf{F}_D^H \right) \le P. \n\end{array} \tag{19}
$$

Where, *P* is the total transmit power which is assumed to be uniformly allocated for each data steams. For massive MIMO configurations, we can approximate  $\mathbf{Q} = \mathbf{F}_{RF}^H \mathbf{F}_{RF}$  to  $\mathbf{Q} = N_t \mathbf{I}_{N_t}$  with high probability since, the off-diagonal entries of the product  $\mathbf{F}_{RF}^H \mathbf{F}_{RF}$  are considered much less than  $N_t$  [[8](#page-16-7)]. Optimization problem and its constraint in [\(19\)](#page-9-0) are convex with respect to  $\mathbf{F}_D$ . Problem ([19](#page-9-0)) has water-filling solution. Taking into account that our problem adopts power normalization for the transmitted streams. As a result, the solution is slightly modifed to become

$$
\mathbf{F}_D = \mathbf{Q}^{-1/2} \mathbf{V}_1 \mathbf{\Lambda}_{N_s} \Rightarrow \frac{1}{\sqrt{N_t}} \mathbf{V}_1 \mathbf{I}_{N_s}
$$
(20)

where  $\Lambda_{N_s}$  denotes diagonal square matrix of power allocation. Due to power normalization criterion, we can write  $\Lambda_{N_s} \approx \sqrt{\frac{P}{N_s}}$  $\frac{P}{N_{RF}}$ **I**<sub>N<sub>s</sub></sub> = **I**<sub>N<sub>s</sub></sub>. Also, **V**<sub>1</sub> represents the first *N<sub>s</sub>* columns returned from SVD of the efective channel. This can be mathematically interpreted as

$$
SVD\left(\mathbf{H}\mathbf{F}_{RF}\mathbf{Q}^{-1/2}\right) = SVD\left(\frac{\mathbf{H}\mathbf{F}_{RF}}{\sqrt{N_t}}\right) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}_1^H
$$
\n(21)

Accordingly, we can briefy draw the design of digital precoder, at any channel realization *k*, as follows

$$
SVD\left(\frac{\mathbf{H}\mathbf{F}_{RF}^{(k)}}{\sqrt{N_t}}\right) = \mathbf{U}^{(k)}\mathbf{\Sigma}^{(k)}\left(\mathbf{V}_1^{(k)}\right)^H,
$$
  

$$
\mathbf{F}_D = \frac{1}{\sqrt{N_t}}\mathbf{V}_1^{(k)}\mathbf{I}_{N_s}.
$$
 (22)

#### **3.4 Quantized Analog Precoder**

According to ([17](#page-8-2)), each element in the  $\mathbf{F}_{RF}$  matrix can be continuously generated from infnite resolution phase shifters. But in the fact, this violates the limited capability of the phase shifters that have a fnite resolution of their produced steering angles. Hence, and for the practical purpose of implementations, we propose phase control method to quantize steering angles of analog precoder via specifed number *B* bits of precision. Then, we investigate the impact of quantization on the overall performance. Moreover, we will depict the performance gape between quantized solution and the unquantized one. Mainly, the concept of quantization depends on minimizing the Euclidean distance between the quantized and unquantized steering angles. Accordingly, the phase of an entry in the qunatized matrix is represented by

<span id="page-10-0"></span>
$$
\angle \hat{\mathbf{F}}_{RF} = \hat{\phi} = \frac{2\pi\hat{n}}{2^B} \tag{23}
$$

where  $\hat{n}$  is chosen to satisfy the following criteria

$$
\hat{n} = \underset{n \in \{0, \dots, 2^B - 1\}}{\operatorname{argmin}} \left| \phi - \frac{2\pi n}{2^B} \right|.
$$
\n(24)

where  $\phi$  refers to the unquantized phases that is obtained from ([17](#page-8-2)). Note that the digital precoding matrix is reevaluated for the quantized analog precoder ([23](#page-10-0)) to evaluate the total quantized performance.

#### **3.5 Energy Efficiency Computation**

In this part, we argue one of the most important key performance of networks. We aim to discuss the energy efficiency (EE) of the proposed hybrid precoder. EE is defined by the ratio between a benefit and cost, where the profit refers to the spectral efficiency defined by ([5\)](#page-4-4) and the cost denotes power consumption [[28](#page-17-9)] that can be modeled as in the following expression

$$
PC = P_{maj} + N_{RF}^t P_{RF} + N_t P_{PA} + N_p P_s. \tag{25}
$$

The first term, in  $(25)$  $(25)$  $(25)$  ( $P_{maj}$ ) defines the major power consumption of the main BS circuits. The second term  $(N<sup>t</sup><sub>RF</sub>, P<sub>RF</sub>)$  accounts for the total power dissipation in RF chains. Further, the third and fourth terms denote the total power loss due to power amplifer and the phase shifters.

According to equations  $(5)$  and  $(25)$  $(25)$  $(25)$ , the total EE can be described as follows

<span id="page-10-2"></span><span id="page-10-1"></span>
$$
\eta = \frac{SE}{PC}, \qquad \text{(bits/Hz/J)}.
$$

In the following lines, we shall introduce our proposed approach for hybrid precoder followed by further remarks which confrm the low complexity of the proposed solution.

# Algorithm 1 A proposed low complex and efficient algorithm for hybrid precoder

**Input:**  $\mathbf{F}_{FD}$ ,  $\mathbf{H}_{realization}$ , realization=100. Output:  $F_{RF}$ ,  $F_{BB}$ . 1:  $\mathbf{F}_{RF}^{(0)} = e^{j\angle \mathbf{H}^H}$  and initialize counter  $(k = 0)$ . 2: **for**  $k = 0$ : *realization* **do**<br>3: Fix  $\mathbf{F}_{RF}^{(k)}$  and compute  $\mathbf{F}_{D}^{(k)}$  based on (22) 4: Fix  $\mathbf{F}_{D}^{(k)}$  and update  $\mathbf{F}_{RF}^{(k+1)}$  using:  $\angle \mathbf{F}_{RF}^{(k+1)} = \angle \left(\mathbf{F}_{FD}(\mathbf{F}_{D}^{(k)})^H\right)$ 5:  $k \leftarrow k+1$ <br>6: **if** Stoppin if Stopping criterion triggers then 7: Normalize digital precoder:  $\mathbf{F}_{BB} = \frac{\sqrt{N_s} \mathbf{F}_D}{\|\mathbf{F}_{RF} \mathbf{F}_D\|_F}$ 8: end if 9: end for 10: Quantize  $\mathbf{F}_{RF}$  based on optimization problem (24) to get  $(\mathbf{F}_{RF})$ . 11: Re-evaluate the digital precoder as in line (7).

Similarly, the proposed algorithm has been deployed to evaluate decoding/combining matrices  $W_{RF}$  and  $W_{BB}$ , at receiver side, which are substituted in [\(5](#page-4-4)) to return the total SE and EE performance.

*Remark 1* Line (1) is inspired by the optimizers which require an initial suboptimal/guess point for fast convergence to fnal solutions. Therefore, this initialization reduces the overhead-time of running algorithm and brings the benefts of array (beamforming) gain. Also, *line (1)* can be interpreted as aligning data stream transmission with dominant channel modes.

*Remark 2* Updating analog precoder in line (4) is realized simply by extracting phases of alternative precoder whose dimension is  $N_t \times N_{RF}^t$ . Utilizing the upper bound instead of the original objective simplifes the solution with marginal performance loss comparing with the expected solution of the main optimization problem  $(8)$ . This loss returns to the gap between  $\|\mathbf{F}_{RF}\mathbf{F}_{D}\|_{F}^{2}$  and  $\|\mathbf{F}_{RF}\|_{F}^{2}$  as discussed in ([13](#page-7-1)). In general, this gap varies with the limits of RF chains  $N \leq N^{t} \leq 2N$ . limits of RF chains  $N_s \leq N_{RF}^t \leq 2N_s$ .

*Remark 3* We have noted that line (7) is independent of the  $\beta$  coefficient since, at transmitter, normalizing  $\mathbf{F}_{BR}$  is equivalent to normalizing  $\mathbf{F}_{D}$ . On the other hand, at receiver, the digital decoding matrix  $\mathbf{W}_{BB}$  only depends on the received signal and the associated noise. Hence, the *SINR* in ([3\)](#page-4-1) is independent of factor  $\beta$  and this reduce computational complexity of the proposed algorithm.

# <span id="page-11-0"></span>**4 Simulations Results and Discussions**

In this part of paper, we exhibit performance of the proposed algorithm. Let's assume that a number of data streams are sent through an array of antenna elements  $N_t = 144$  which have been mounted in a square plate ( $\sqrt{N_t} \times \sqrt{N_t}$ ); such that all elements are equally spaced by  $d = 0.5\lambda$ . Those streams are received by  $N_r = 36$  antennas which are similarly installed as in the transmitter side. Further, we consider mmWave (clustered) channel model with



<span id="page-12-0"></span>**Fig. 2** Spectral efficiency for different precoding algorithms at case  $N_s = N_{RF}^t = N_{RF}^r = 3$ . (Color figure online)

 $N_{cl} = 5$  clusters,  $N_{ray} = 10$  rays and the average power per cluster is  $\sigma_{\alpha,i}^2 = 1$ . The distribution of azimuth and elevation angles are assumed Laplacian with mean angular spread of  $(10^{\circ})$  for all arrival and departure angles (AoAs and AoDs) [[28](#page-17-9)]. Noting that the simulation results are averaged over 100 channel realizations.

#### **4.1 Spectral Efficiency Evaluation**

SE has been plotted against different levels of signal power  $(SNR_{dB})$  considering the worst case  $N_s = N_{RF}^t = N_{RF}^r$ ; as shown in Fig. [2](#page-12-0). The proposed algorithm (*dotted blue line*) displays a signifcant performance. The magnifed scale, at the middle-left, demonstrates very small gap that is evaluated by −5*dB* or 0.31*bits* / *s* / *HZ* of loss comparing with the unconstrained digital precoder (*black solid line*). This refects the rigidity of our algorithm and precision of the derived upper bound [\(15\)](#page-8-0) along all *SNR* range. Moreover, the proposed hybrid approach outperform the other algorithms such as OMP [[8\]](#page-16-7) and the analog beamforming. This superiority comes from not only utilizing efective channel information with modifed water-flling to keep the orthogonality criterion but also, extracting the steering angles from low complex (alternative) beamformers. Hence, the proposed approach harvests benefts of both array gain that potentially substitutes experienced losses in mmWave channel and low complexity design.

As well, the SE of quantized hybrid precoders, with one or two bits of precision, approximates the SE performance of unquantized hybrid precoders (*dotted blue line*). Particularly, the average SE performance of two bits quantized precoder (*dotted green line*) has been degraded with about −1.5*dB* comparing with the unquantized



<span id="page-13-0"></span>**Fig. 3** Spectral efficiency versus defined range of  $N_{RF}$  when  $N_s = 2$  and  $SNR = 0dB$ 

performance. Meanwhile, the proposed unquantized algorithm achieves gain more than 4.3*dB* comparing with OMP (*sold green line*). Therefore, we can frmly conclude that our algorithm (quantized and unquantized) stands as a benchmark for hybrid precoding structure.

It is worth to be mentioned that the study [[24](#page-17-7)] introduced a closed formula of hybrid precoding problem at the special case of  $N_{RF}^t \ge 2N_s$ ,  $N_{RF}^r \ge 2N_s$  at which authors' solution approximates the unconstrained digital precoder. On contrary, our proposed approach investigates a wider range of  $N_{RF}^t$  described by  $N_s \leq N_{RF}^t < 2N_s$  to check if the proposed approach able to achieve a signifcant performance or not. Fig. [3](#page-13-0) depicts fast saturation of the proposed approach (quantized and unquantized) to fxed levels of SE for the given range of RF chains. It is visualized that the average performance loss between the unqauntized hybrid algorithm and the fully digital precoder is evaluated by −5*dB*. Meanwhile, the SE degradation of the two bits quantized precoder is evaluated by about −2.2*dB* comparing with its counterpart of the one bit quantized precoder.

Furthermore, Fig. [3](#page-13-0) demonstrates that OMP outperform the one bit qunatized precoder only when  $N_{RF} > 2N_s$  due to available high DoF at reconstructing precoding matrices. On contrary, the qunatized hybrid performance using one bit of precision has a high merit at  $N_{RF} \leq 2N_s$  which makes it preferable at the worst cases of performance (i.e. at  $N_{RF} = N_s$ ) or at the cases which require more power saving. It worth noted that the same aforementioned discussions, about SE performance behavior, can be deduced with higher number of data streams, as shown in Fig. [4b](#page-14-0). Figure [4](#page-14-0)a remarks a divergence of the one bit quantized approach which slightly increases with  $N<sub>s</sub>$  since, the resolution of quantization process may not be sufficient to steer large number of data streams.



(a) Spectral efficiency at different values of  $N_s$  and  $SNR = 0$  *dB*.



<span id="page-14-0"></span>Fig. 4 Comparisons of the spectral efficiency performance



<span id="page-15-0"></span>**Fig. 5** Energy efficiency performance at  $N_s = 3$ , 8 and  $SNR = 0$ dB

### **4.2 Energy Efficiency Performance**

In this part, we will demonstrate EE behavior of the proposed algorithm based on the defni-tion of EE in [26](#page-10-2) and the corresponding parameters:  $P_{maj} = 10W$ ,  $P_{RF} = 100mW$ ,  $P_s = 10mW$ and  $P_{PA} = 100mW$  [\[9](#page-16-8)]. The simulation result is displayed in Fig. [5](#page-15-0), wherein EE is decaying with increasing the RF chains. This can be explained that phase shifters are linearly scaled with the number of used RF chains as shown by the second term of  $(25)$  $(25)$ . Therefore, it has a signifcant impact at measuring power consumption.

Based on the previous results in Fig. [3](#page-13-0) and Fig. [4](#page-14-0)b, we observe that the SE performance is fixed and the decaying rate of power consumption is faster than the increment of SE along  $N_{RF}$ range as depicted in second term of  $(25)$  $(25)$ . This explains the dramatic decaying of EE. Further, increasing the number of data streams enhances the EE performance. For instance, the EE is boosted by about 3.6*dB* at  $N_s = 8$  comparing with its counterpart at  $N_s = 3$ . Moreover, the two bits quantized EE performance is approximately equal the unquantized approach without signifcant loss for any scenario of stream number. Nevertheless, The EE of one bit quantized approach reveals quite performance loss at high number of data streams. This matches the divergence that has been explained for Fig. [4a](#page-14-0).

# <span id="page-16-12"></span>**5 Conclusions and Future Work**

We have presented an efective hybrid precoding algorithm to approach the fully digital precoder which has high complex degree of implementation in mmWave massive MIMO systems. Mainly, our proposed algorithm has been built on the legacy of tight upper bound which approximates the analog precoder to an alternative one that have low processing cost. On the other hand, the digital precoder has been optimized considering data streams orthogonality which has been maintained through modifed water flling solution. Furthermore, the proposed algorithm exhibits superior performance at any number of RF implementations where, the spectral efficiency displays a converged behavior along wide range of RF chains. Moreover, we have considered quantized analog beam-formers for practical issues wherein, the steering angles are quantized up to two bits of precision. In this case, quantized version of the proposed hybrid precoder reveals a remarkable performance without signifcant loss. Besides, tracing EE of our proposed approach returns a valuable insight for design criteria, where the hardware complexity may costs more power consumption than the expected/required growth of system capacity. As a result, network performance requires compromising between data rate demands and energy saving.

So far, it will be interesting to extend our proposal to the heterogeneous networks and further small dense cells to introduce a complete picture about the hybrid beamforming in the future networks. Also, more characterization for the hybrid precoding structure will be considered through integrating with one of new candidate modulation techniques which will be elected for future 5G networks.

# **References**

- <span id="page-16-0"></span>1. Hoydis, J., Ten Brink, S., & Debbah, M. (2013). Massive mimo in the ul/dl of cellular networks: How many antennas do we need? *IEEE Journal on selected Areas in Communications*, *31*(2), 160–171.
- <span id="page-16-1"></span>2. Li, C., Zhang, J., & Letaief, K. B. (2014). Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations. *IEEE Transactions on Wireless Communications*, *13*(5), 2505–2517.
- <span id="page-16-2"></span>3. Salem, A., El-Rabie, S., & Shokair, M. (2018). Energy efficient ultra-dense networks (udns) based on multi-objective optimization framework. *IET Networks*, *7*(6), 398–405.
- <span id="page-16-3"></span>4. Doppler, K., Rinne, M., Wijting, C., Ribeiro, C. B., & Hugl, K. (2009). Device-to-device communication as an underlay to lte-advanced networks. *IEEE Communications Magazine*, *47*(12), 42–49.
- <span id="page-16-4"></span>5. Shi, Y., Zhang, J., Letaief, K. B., Bai, B., & Chen, W. (2015). Large-scale convex optimization for ultra-dense cloud-ran. *IEEE Wireless Communications*, *22*(3), 84–91.
- <span id="page-16-5"></span>6. Torkildson, E., Madhow, U., & Rodwell, M. (2011). Indoor millimeter wave mimo: Feasibility and performance. *IEEE Transactions on Wireless Communications*, *10*(12), 4150–4160.
- <span id="page-16-6"></span>7. Akdeniz, M. R., Liu, Y., Samimi, M. K., Sun, S., Rangan, S., Rappaport, T. S., et al. (2014). Millimeter wave channel modeling and cellular capacity evaluation. *IEEE Journal on Selected Areas in Communications*, *32*(6), 1164–1179.
- <span id="page-16-7"></span>8. El Ayach, O., Rajagopal, S., Abu-Surra, S., Pi, Z., & Heath, R. W. (2014). Spatially sparse precoding in millimeter wave mimo systems. *IEEE Transactions on Wireless Communications*, *13*(3), 1499–1513.
- <span id="page-16-8"></span>9. Rappaport, T. S., Heath, R. W, Jr., Daniels, R. C., & Murdock, J. N. (2014). *Millimeter wave wireless communications*. London: Pearson Education.
- <span id="page-16-9"></span>10. Love, D. J., & Heath, R. W. (2003). Equal gain transmission in multiple-input multiple-output wireless systems. *IEEE Transactions on Communications*, *51*(7), 1102–1110.
- <span id="page-16-11"></span>11. Zhang, X., Molisch, A. F., & Kung, S. Y. (2005). Variable-phase-shift-based rf-baseband codesign for mimo antenna selection. *IEEE Transactions on Signal Processing*, *53*(11), 4091–4103.
- <span id="page-16-10"></span>12. Zheng, X., Xie, Y., Li, J., & Stoica, P. (2007). Mimo transmit beamforming under uniform elemental power constraint. *IEEE Transactions on Signal Processing*, *55*(11), 5395–5406.
- <span id="page-17-0"></span>13. Venkateswaran, V., & van der Veen, A. J. (2010). Analog beamforming in mimo communications with phase shift networks and online channel estimation. *IEEE Transactions on Signal Processing*, *58*(8), 4131–4143.
- <span id="page-17-1"></span>14. Roh, W., Seol, J. Y., Park, J., Lee, B., Lee, J., Kim, Y., et al. (2014). Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results. *IEEE Communications Magazine*, *52*(2), 106–113.
- 15. Sun, S., Rappaport, T. S., Heath, R. W., Nix, A., & Rangan, S. (2014). Mimo for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both? *IEEE Communications Magazine*, *52*(12), 110–121.
- <span id="page-17-10"></span>16. Alkhateeb, A., El Ayach, O., Leus, G., & Heath, R. W. (2014). Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, *8*(5), 831–846.
- 17. Wang, P., Li, Y., Song, L., & Vucetic, B. (2015). Multi-gigabit millimeter wave wireless communications for 5G: From fxed access to cellular networks. *IEEE Communications Magazine*, *53*(1), 168–178.
- <span id="page-17-2"></span>18. Rangan, S., Rappaport, T. S., & Erkip, E. (2014). Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*, *102*(3), 366–385.
- <span id="page-17-3"></span>19. Lee, Y. Y., Wang, C. H., & Huang, Y. H. (2015). A hybrid rf/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave mimo systems. *IEEE Transactions on Signal Processing*, *63*(2), 305–317.
- <span id="page-17-5"></span>20. Lee, J., & Lee, Y. H. (2014). Af relaying for millimeter wave communication systems with hybrid rf/ baseband mimo processing. In 2014 IEEE Conference on Communications (ICC) (pp. 5838–5842). IEEE.
- 21. Kim, M., & Lee, Y. H. (2015). Mse-based hybrid rf/baseband processing for millimeter-wave communication systems in mimo interference channels. *IEEE Transactions on Vehicular Technology*, *64*(6), 2714–2720.
- 22. Brady, J., Behdad, N., & Sayeed, A. M. (2013). Beamspace mimo for millimeter-wave communications: System architecture, modeling, analysis, and measurements. *IEEE Transactions on Antennas and Propagation*, *61*(7), 3814–3827.
- <span id="page-17-6"></span>23. Rusu, C., Méndez-Rial, R., González-Prelcicy, N., & Heath, R. W. (2015). Low complexity hybrid sparse precoding and combining in millimeter wave mimo systems. In 2015 IEEE International Conference on Communications (ICC) (pp. 1340–1345). IEEE.
- <span id="page-17-7"></span>24. Zhang, E., & Huang, C. (2014). On achieving optimal rate of digital precoder by rf-baseband codesign for mimo systems. In 2014 IEEE 80th Vehicular Technology Conference (VTC Fall) (pp. 1–5). IEEE.
- 25. Liang, L., Xu, W., & Dong, X. (2014). Low-complexity hybrid precoding in massive multiuser mimo systems. *IEEE Wireless Communications Letters*, *3*(6), 653–656.
- <span id="page-17-4"></span>26. Wang, G., & Ascheid, G. (2014). Joint pre/post-processing design for large millimeter wave hybrid spatial processing systems. In European Wireless 2014; 20th European Wireless Conference; Proceedings of. VDE (pp. 1–6).
- <span id="page-17-8"></span>27. Sohrabi, F., & Yu, W. (2016). Hybrid digital and analog beamforming design for large-scale antenna arrays. *IEEE Journal of Selected Topics in Signal Processing*, *10*(3), 501–513.
- <span id="page-17-9"></span>28. Yu, X., Shen, J. C., Zhang, J., & Letaief, K. B. (2016). Alternating minimization algorithms for hybrid precoding in millimeter wave mimo systems. *IEEE Journal of Selected Topics in Signal Processing*, *10*(3), 485–500.
- <span id="page-17-11"></span>29. Love, D. J., & Heath, R. W. (2005). Limited feedback unitary precoding for spatial multiplexing systems. *IEEE Transactions on Information Theory*, *51*(8), 2967–2976.
- <span id="page-17-12"></span>30. Singh, J., & Ramakrishna, S. (2015). On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays. *IEEE Transactions on Wireless Communications*, *14*(5), 2670–2683.
- 31. Dai, L., Gao, X., Quan, J., Han, S., & Chih-Lin, I. (2015). Near-optimal hybrid analog and digital precoding for downlink mmwave massive mimo systems. In 2015 IEEE International Conference on Communications (ICC) (pp. 1334–1339). IEEE.
- 32. Han, S., Chih-Lin, I., Xu, Z., & Rowell, C. (2015). Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Communications Magazine*, *53*(1), 186–194.
- <span id="page-17-13"></span>33. Zhang, J. A., Huang, X., Dyadyuk, V., & Guo, Y. J. (2015). Massive hybrid antenna array for millimeter-wave cellular communications. *IEEE Wireless Communications*, *22*(1), 79–87.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**A. Abdelaziz Salem** received his B.Sc. and Ms.c. degrees in electronics and electrical communications engineering from the Faculty of Electronic Engineering, Menofa University, Egypt, in 2011 and 2015, respectively. He is working as assistant lecturer at the Department of electronics and electrical communications engineering, working towards his Ph.D. degree at the Faculty of Electronic Engineering, Menofa University, Egypt. His research interests are cognitive radio networks, Game theory applications, 5G theologies and optimization theory.



**S. El‑Rabaie** (SM'92) was born in Sires Elian, Egypt, in 1953. He received the B.Sc. degree (with honors) in radio communications from Tanta University, Tanta, Egypt, in 1976, the M.Sc. degree in communication systems from Menoufa University, Menouf, Egypt, in 1981, and the Ph.D. degree in microwave Device engineering from Queen's University of Belfast, Belfast, U.K., in 1986. In his doctoral research, he constructed a Computer-Aided Design (CAD) package used in nonlinear circuit simulations based on the harmonic balance techniques. Up to February 1989, he was a Postdoctoral Fellow with the Department of Electronic Engineering, Queen's University of Belfast. He was invited as a Re-search Fellow in the College of Engineering and Technology, Northern Arizona University, Flagstaff, in 1992 and as a Visiting Professor at Ecole Polytechnique de Montreal, Montreal, QC, Canada, in 1994. He Has Authored and Co-authored of More Than 300 Papers and nineteen text Books. He was Awarded several Awards (Salah Amer Award of Electronics in 1993, The Best Researcher on (CAD) from Menoufa University in 1995). He acts as a reviewer and

member of the editorial board for several scientifc journals. He Has Shared in Translating the First Part of the Arabic Encyclopedia. Professor EL-Rabaie was the Head of the Electronic and Communication Engineering Dept., Faculty of Electronic Engineering, Menoufa University, then the Vice Dean of Postgraduate Studies and Research in the same Faculty. Prof. S. El-Rabaie is Involved now in Diferent Research Areas including CAD of Nonlinear Microwave Circuits, Nanotechnology, Digital Communication Systems, and Digital Image Processing. Now he is Reviewer of Quality Assurance and Accreditation of Egyptian Higher Education.



**Mona Shokair** received the B.E. and M.E. degrees in electronics engineering from El Menofa University, El-Menofa, Egypt, in 1993 and 1997, respectively. She received Ph.D. from Kyushu University, Japan, 2005. She was lecturer in El-Menofa University from 2006 to 2011. Since 2012, she is Associated Professor in El-Menofa University. She received VTS chapter IEEE award from Japan in 2003. Now her current research is in OFDM system, WIMAX system and cognitive radios.