Capacity Maximisation for Hybrid Digital-to-Analog Beamforming mm-Wave Systems

Osama Alluhaibi, Qasim Zeeshan Ahmed, Cunhua Pan, and Huiling Zhu
School of Engineering and Digital Arts, University of Kent, Canterbury, CT2 7NT, United Kingdom.
Email: \{oa274, q.ahmed, c.pan, h.Zhu\}@kent.ac.uk

Abstract—Millimetre waves (mm-Waves) with massive multiple input and multiple output (MIMO) have the potential to fulfill fifth generation (5G) traffic demands. In this paper, a hybrid digital-to-analog (D-A) precoding system is investigated and a particle swarm optimisation (PSO) based joint D-A precoding optimisation algorithm is proposed. This algorithm maximises the capacity of the hybrid D-A mm-Wave massive MIMO system. The proposed algorithm is compared with three known hybrid D-A precoding algorithms. The analytical and simulation results show that the proposed algorithm achieves higher capacity than the existing hybrid D-A precoding algorithms.

Index Terms—Beamforming, hybrid beamforming, optimisation, particle swarm optimisation (PSO), Millimetre wave.

I. INTRODUCTION

Mobile networks have been growing exponentially, leading to a scarcity of bandwidth. Recent studies anticipated that the global mobile data traffic will reach a 66\% annual growth rate in the next five years [1]. Recently, it is shown that millimetre-wave (mm-Wave), operating in the (30 – 300) GHz spectrum, offers a promising approach for meeting this demand by providing a larger bandwidth [1]. A reasonable short wavelength of this band enables packing a large number of antennas in the same physical space [2]. Therefore, the feasibility of implementing a massive multiple input and multiple output (MIMO) in a small aperture area is possible [3]. In a fully-digital beamforming (BF) solutions the number of radio frequency (RF) chains is equivalent to the number of transmit antennas which increase the computation complexity and power consumption of the system [3, 4]. Therefore, fully-digital BF cannot be directly applied to mm-Wave massive MIMO system, due to the fact that a large number of RF chains are required.

A simpler approach will be to use either an analog precoder system or a hybrid digital-to-analog (D-A) precoding system, where the number of RF chains is less than the number of transmitting antennas [4–11]. A fully-antenna array was used for the hybrid D-A precoding, where each RF chain was connected to all the transmit antennas [5, 6]. Nonetheless, the fully-antenna array has limitations as a) it involves higher complexity at the analog precoder [5, 9] and b) more energy is consumed since the number of phase shifters scales linearly with the number of RF chains and antennas [12]. A sub-antenna array structure for the hybrid D-A precoding was proposed, where each RF chain was connected to a specific sub-antenna array [7–11]. Therefore, in this case, the phase shifters are independent of the number of RF chains. Substantially, the sub-antenna array structure for the hybrid D-A precoding can reduce the computation complexity and power consumption of the system as compared to the fully-antenna array.

Precoding for the hybrid D-A BF system has already been proposed in [5–11]. An iterative algorithm was proposed in [4], where the analog precoder was optimised to improve the capacity of the mm-Wave system. However, the capacity achieved by [4] is much lower than the capacity of the hybrid D-A precoding system as shown in [9]. The digital precoder is fixed to an identity matrix while the analog precoder is exactly the normalised conjugate transpose of the channel as proposed in [7, 8, 10, 11]. However, in this case, the precoders have not been designed jointly. Recently, a joint analog and digital precoders have been investigated, where an iterative algorithm for the hybrid D-A precoding by utilizing the idea of a singular value decomposition (SVD) is proposed in [9]. SVD algorithm is known for its higher complexity as it requires matrix inversion [13]. The scheme in [9], optimised every RF chain successively, however, the complexity of this method is very high [13]. Therefore, in this paper, a particle swarm optimisation (PSO) algorithm is proposed to design the hybrid D-A precoding jointly. PSO is an evolutionary approach, which refines the estimates through a group of agents searching the solution space and finding the global or near an optimum solution after several iterations [14]. In this contribution, the reasons for choosing PSO is: firstly, this algorithm requires minimal tuning parameters, thereby, can be implemented in real-time applications. Secondly, PSO only requires the cost-function and does not require any differentiation, matrix inversion, resulting in reduced complexity [15]. Therefore, it can be implemented adaptively, thereby, decreasing the complexity of the system. Our simulation results, show that PSO outperforms the existing algorithms in [4, 7, 8, 10, 11]. Furthermore, the results illustrate that the proposed scheme scales easily with the increased number of RF chains and transmit antennas. Finally, the computational complexity of PSO is much lower than the SVD-based hybrid D-A precoding while the iterative analog precoder has a lower complexity.

Notation: Bold uppercase letters \( \mathbf{X} \), and lowercase letters, \( \mathbf{z} \), denote matrices and vectors, respectively. Transposition and conjugate transposition of a matrix are respectively denoted by \( (\cdot)^T \) and \( (\cdot)^H \). \(|\cdot|\) and \(|\cdot|_F\) denote the determinant and Frobenius norm of a matrix, respectively, \(|\cdot|_F\) denotes the norm of a vector. The diagonal matrix is denoted as \( \text{diag}(\cdot) \), and the operator \( \text{vec}(\cdot) \) maps the \( P \times X \) matrix to a \( PX \) vector. Finally,
\( \mathbb{C} \) denote as a complex number, while \( \sqrt{\cdot} \) is a square root of a number and \( \cup \) is denoted as union of event.

II. SYSTEM MODEL

The block diagram of the downlink mm-Wave massive MIMO system for a hybrid D-A BF is shown in Fig. 1. A digital precoder represented as \( D = \text{diag}[d_1, d_2, \cdots, d_N] \), where \( d_n \in \mathbb{C} \) for \( n = 1, 2, \cdots, N \) [9]. Due to the fact \( s_{n(i)} \) that \( D \) is a diagonal matrix, the inter-symbol interference does not exist [7]. \( N \) data symbols are precoded by \( D \) and after that, symbol \( d_n \) passes through the \( n \)-th RF chain. The digital domain signal from one RF chain is fed to \( M \) transmit antennas to perform analog precoding. The analog precoder vector is denoted by \( \bar{a}_n \in \mathbb{C}^{M \times 1} \), where all the elements of \( \bar{a}_n \) have the same amplitude \( 1/\sqrt{M} \) but different phase shifts [9]. Finally, every data symbol is transmitted by the sub-antenna array of \( M \) antennas.

A. Channel Model

Rayleigh fading or multipath Rayleigh fading has been adopted into microwave radio channel [16–29]. Mm-Wave channel will no longer follow the low frequencies conventional Rayleigh fading due to the limited number of scatters [9]. Therefore, in this paper geometric 3-dimensional (3D) Saleh-Valenzuela (SV) channel model is used as mentioned in [9, 30–32]. Channel model for the \( n \)-th RF chain is represented as

\[
\bar{h}_n = \sqrt{\frac{N M}{L}} \sum_{l=1}^{L} \left( h_{n,m}^{3D} s_{n}(\theta_l^1, \phi_l^1) s_{n}(\theta_l^2, \phi_l^2) \right),
\]

where \( \bar{h}_n \in \mathbb{C}^{1 \times M} \), \( L \) is the number of multipaths [9]. The 3D BF gain for every transmitter (Tx) antenna element \( h_{n,m}^{3D} \) is given in (2). \( \theta_l^1(\phi_l^1) \) and \( \theta_l^2(\phi_l^2) \) incorporates the zenith (azimuth) direction of departure and arrival (AOD, AOA), respectively. The steering vector \( s_{v_i} \) in (1) where \( i \in [r, t] \), is given by [5, 30]

\[
s_{v_i}(\theta_l^1, \phi_l^1) = \text{vec}[s_{v_{ix}}(\Psi)s_{v_{iy}}^T(\Psi)], \quad i \in [r, t]
\]

where

\[
s_{v_{ix}}(\Psi) = \frac{1}{\sqrt{M_x}} \left[ 1, e^{j\Psi}, \cdots, e^{j(M_x-1)\Psi} \right]^T,
\]

\[
s_{v_{iy}}(\Phi) = \frac{1}{\sqrt{M_y}} \left[ 1, e^{j\Phi}, \cdots, e^{j(M_y-1)\Phi} \right]^T.
\]

As the uniform planar antenna (UPA) structure is a preferred choice for 3D channel, we adopted a \( M = M_x \times M_y \) structure antennas, where \( M_x \) represents the \( x \)-axis while \( M_y \) represents the \( y \)-axis. The steering vectors \( s_{v_{ix}}, s_{v_{ix}} \) represent the \( x \)-axis and \( y \)-axis respectively. The values of \( \Psi \) and \( \Phi \) are calculated as

\[
h_{n,m}^{3D} = \sqrt{\frac{P_m}{L}} \sum_{l=1}^{L} \left[ F_{Rx, Z}(\bar{\phi}_l, \bar{\theta}_l) F_{Rx, A}(\bar{\phi}_l, \bar{\theta}_l) \right]^T \left[ e^{j\phi_l^{2Z}} \sqrt{\frac{1}{M_x}} e^{j\phi_l^{A2}} e^{j\phi_l^{AA}} \right] \left[ F_{Tx, Z}(\bar{\theta}_l, \bar{\phi}_l, \bar{\theta}_l) F_{Tx, A}(\bar{\phi}_l, \bar{\theta}_l, \bar{\phi}_l) \right]
\]

\[
\Psi = -2\pi \lambda^{-1} u_x \sin(\theta_l^1) \cos(\phi_l^1), \quad i \in [r, t]
\]

\[
\Phi = -2\pi \lambda^{-1} u_y \sin(\theta_l^2) \sin(\phi_l^2), \quad i \in [r, t]
\]

where \( u_x \) and \( u_y \) is the inter-element distance in the \( x \) and \( y \)-axis, respectively. In (2), \( P_m \) is the power of the \( m \)-th Tx antenna, and is calculated assuming a single slope exponential power delay profile by [30]. \( F_{Rx, Z} \) and \( F_{Rx, A} \) are the receiver (Rx) beam pattern. The azimuth (A) polarizations, \( \psi_l \) and \( \varphi_l \) are the zenith and azimuth AoA, respectively. \( F_{Tx, Z} \) and \( F_{Tx, A} \) are the Tx beam pattern for the \( n \)-th RF chain and \( \theta_l,m \) and \( \phi_l,m \) are the zenith and azimuth AoD, respectively. \( \phi_l^{2Z}, \phi_l^{A2}, \phi_l^{AA} \) are the initial random phases for zenith (ZZ), cross (ZA, AZ), and azimuth polarizations (AA) for the \( l \) tap. \( \kappa_m \) is the intra-cluster Rician \( K \)-factor associated with the \( m \)-th Tx antenna cluster [30].

B. Received Signal of Hybrid D-A BF System

The received signal for all \( N \) data symbols \( y = [y_1, y_2, \cdots, y_N]^T \), is expressed as

\[
y = H A D s + n = H G s + n,
\]

where \( H = [h_1, h_2, \cdots, h_N] \in \mathbb{C}^{N \times NM} \), \( h_n = [0_{1 \times M(n-1)}, h_n, 0_{1 \times M(N-n)}] \in \mathbb{C}^{1 \times NM} \), and \( \bar{h}_n \) is given in (1). The analog precoder \( A \) is represented as

\[
A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_N \end{bmatrix}
\]

where \( A = \text{diag}[a_1, \cdots, a_N] = [a_1, a_2, \cdots, a_N] \), \( a_n = [0_{1 \times M(n-1)}, a_n; 0_{1 \times M(N-n)}] \in \mathbb{C}^{NM \times 1} \). \( N \) data symbols are represented as \( s = [s_1, s_2, \cdots, s_N]^T \), and \( n = [n_1, \cdots, n_N] \), where \( n_x \) is the complex Gaussian random variable with zero means and a variance of \( \sigma^2, G = AD \), represents the joint hybrid precoding matrix of size \( (NM \times N) \).
The optimisation problem in (10) is a precoding, where analog and digital precoders are jointly designed and has to satisfy the following constraints:

\[ \text{C1: The Frobenius norm of } \mathbf{G} \text{ should satisfy } ||\mathbf{G}||^2_F \leq N \text{ to meet the total transmit power constraint.} \]

\[ \text{C2: As } \mathbf{D} \text{ is a diagonal matrix, and the amplitude of the analog precoding } \bar{\mathbf{a}}_n \text{ of each RF chain is fixed to } 1/\sqrt{M}. \text{ Therefore, for each non-zero elements of } \mathbf{G}, \text{ the amplitude should be equal.} \]

### III. Precoder Design for Hybrid D-A BF System

In this section, we discuss the design of the hybrid D-A precoding, where analog and digital precoders are jointly designed. As RF chains do not cause inter RF interference, \( \mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_N] \) can be designed as a block matrix where \( \mathbf{g}_n = [0_{1 \times (M-1)}; \mathbf{g}_n; 0_{1 \times (M-N-1)}] \subset \mathbb{C}^{NM \times 1} \), the matrix optimisation problem can now be solved as an N independent vectors optimisation problem. The advantages are: a) it allows us to apply our scheme for every RF chain independently and b) the result of \( \mathbf{HG} \) becomes exactly a diagonal matrix with equal elements and the upper bound of the capacity is achieved in the given model.

The capacity of the system is given as

\[ C(\mathbf{G}) = \sum_{n=1}^{N} \log_2 \left( 1 + \frac{\mathbf{h}_n^\text{T} \mathbf{g}_n \mathbf{g}_n^H \mathbf{h}_n^H}{\sigma^2} \right), \]

(11)

Undoubtedly, the design of \( \mathbf{G} \) will make the optimisation problem much easier to solve. Furthermore, each RF chain is now independently resolved and can be designed to maximise its capacity. These algorithms are independent to one another and will be initialised simultaneously. The \( n \)-th RF chain is optimised by designing the precoding vector \( \mathbf{g}_n \) as

\[ C_n(\mathbf{g}_n) = \arg \max_{\mathbf{g}_n \in \mathbb{C}^{1 \times 2}} \log_2 \left( 1 + \frac{\mathbf{h}_n^\text{T} \mathbf{g}_n \mathbf{g}_n^H \mathbf{h}_n^H}{\sigma^2} \right), \]

(12)

**A. Particle Swarm Optimisation**

PSO is a stochastic optimisation technique and details can be found in [14, 15] and the references therein. PSO algorithm is an optimisation strategy which became popular due to the fact that it is simple to implement, and quickly convergence to the desired solution [15]. It is robust against local minimas which make it appealing for real-time applications [14]. The coordinates of an agent represent the solution to the problem. Furthermore, in each iteration of PSO, velocity of each agent is adjusted towards the best location and toward the best agent. Following steps are involved to find the solution for each RF chain:

**For the \( n \)-th RF chain**

**Initialisation.** For this problem \( \mathbf{g}_n \) needs to be optimised which is a \( M \times 1 \) dimensional vector. Initialise \( P \) agents with random positions \( \mathbf{g}_{1}(0), \mathbf{g}_{2}(0), \ldots, \mathbf{g}_{P}(0) \). All positions are normalised to ensure that power of \( n \)-th RF chain is 1. The position of the agent is used to evaluate (12) and the position of the agent which maximises (12) is denoted as \( \mathbf{f}_{\text{best}} \). After that, the velocity of all the agents is randomly initialised. The \( p \)-th agent velocity is represented as \( \mathbf{v}_p \). After initialisation, the following iterative process is performed.

**Step 1.** Update the velocity \( \mathbf{v}_p \) and position \( \mathbf{g}_p \) of \( p \)-th agent

\[ \mathbf{v}_p(i + 1) = \mathbf{v}_p(i) + c_1 \mathbf{w}_1 \odot (\mathbf{g}_{\text{best}}(i) - \mathbf{g}_p(i)) \]

\[ + c_2 \mathbf{w}_2 \odot (\mathbf{f}_{\text{best}} - \mathbf{g}_p(i)), \]

(13)

\[ \mathbf{g}_p(i + 1) = \mathbf{g}_p(i) + \mathbf{v}_p(i + 1), \]

(14)

where \( \mathbf{w}_1, \mathbf{w}_2 \) are uniformly distributed random numbers. The element-wise multiplication is denoted by \( \odot \), \( c_1 \) and \( c_2 \) are positive acceleration coefficients. In the first iteration \( \mathbf{g}_{\text{best}} = \mathbf{g}_p \). After that, each agent keeps track of its own best position, which is associated with achieving the maximum value in (12). Once the position of the \( p \)-th agent is updated, its fitness \( C_n(\mathbf{g}_p(i + 1)) \) is evaluated. If the updated fitness of the agent is more than the previous best-fitness of the agent, then \( \mathbf{g}_{\text{best}}(i) = \mathbf{g}_p(i + 1) \).

**Step 2.** Finally, we compare the fitness of all the \( P \) agents and the agent which maximises (12) is compared with previous \( \mathbf{f}_{\text{best}} \) and the one which maximised (12) becomes the global best agent \( \mathbf{f}_{\text{best}} \).

**Step 3.** Repeat step-1 and step-2 until the number of iterations are complete. Now, \( \mathbf{f}_{\text{best}} = \mathbf{g}_n \).

**End for \( n \)-th RF chain**

Note that, \( \mathbf{g}_n = \bar{d}_n \mathbf{a}_n \) for all the sub-antenna array and the optimal solution \( \mathbf{g}_n \) have a similar form. After obtaining the precoder vector \( \mathbf{g}_n \) for the \( n \)-th RF chain, the same algorithm is applied to other chains. Penultimately, it is worth mentioning here that as the RF chains are independent, \( N \) independent PSO algorithms are required. After optimising the last RF chain, the optimal digital, analog, and joint hybrid precoding matrices \( \mathbf{D}, \mathbf{A} \), and \( \mathbf{G} \) are obtained.

**Output**

\[ \mathbf{G} = \text{diag}[\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_N], \]

\[ \mathbf{A} = \text{diag}[\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N], \]

\[ \mathbf{D} = \text{diag}[d_1, d_2, \ldots, d_N]. \]

Finally, as each RF chain has equivalent power of 1, therefore, the total transmit power constraint as mentioned in (C1) is satisfied

\[ ||\mathbf{G}||^2_F = ||\text{diag}[\mathbf{g}_1, \ldots, \mathbf{g}_N]||^2_F \leq N. \]

(15)

In addition, all non-zero elements of \( \mathbf{a}_n \) have fixed amplitude which makes (C2) satisfied.

### IV. Simulation Results

In this section, to validate the performance of our proposed algorithms, capacity performance per time slot versus the signal-to-noise-ratio (SNR = 1/\( \sigma^2 \)) per antenna element is compared when using different algorithms. The first algorithm employs analog precoder [4]. The second algorithm, named as
hybrid D-A precoding based on analog precoder $\mathbf{A}$, is exactly
the normalised conjugate transpose of $\mathbf{H}$, while the digital
precoder matrix $\mathbf{D} = \mathbf{I}$ is an identity matrix is implemented [7,
8, 10, 11]. The third algorithm, a hybrid D-A precoding based
on SVD method, which has been proposed in [9] is considered.
Finally, an optimal unconstrained hybrid D-A precoding
proposed in [9] is implemented based on the sub-antenna array
architecture as a benchmark. Furthermore, different RF chains
have also been implemented and multi-beampattern have been
plotted. The channels are generated according to the channel
model (1). The number of channel paths is set to $L = 3$.
The transmitter antenna array is assumed as UPA with antenna
spacing $u_q = \frac{\lambda}{2}$, $q \in \{x, y\}$. The AoAs and AoDs are taken
independent realizations of the uniform distribution within $[0, 2\pi]$.

Fig. 2 shows the learning curves of the PSO based algorithm
with a different number of agents at SNR = 5 dB. The capacity
achieved is averaged over 10,000 independent realizations of
the channel. It is be observed from Fig. 2 that by increasing the
number of agents, the algorithm converges faster and achieved
higher capacity value. However, the higher number of agents
requires more complexity as more agents have to be initialised
and more calculations have to be carried out. From Fig. 2, it can
be observed that the convergence of population size of 5 and 50
is similar. Therefore, in the sequel, the number of agents is fixed
to 10, the number of iterations is fixed to 40.

Fig. 3 and Fig. 4 plot the capacity versus SNR of the hybrid
D-A BF mm-Wave system. From Fig. 3 and Fig. 4 it is observed
that as the SNR improves, the capacity of the system increases.
The capacity achieved by analog precoder in [4] is always
lower than the capacity achieved by PSO. For example, when
SNR = 0 dB the capacity gap between the analog precoder [4]
and PSO is about 4 bits/s/Hz, while when SNR = 30 dB, the
capacity gap increases to about 14 bits/s/Hz. It can be observed
that the capacity achieved by hybrid D-A precoders proposed
in [7, 8, 10, 11] is lower than the capacity of proposed PSO.
The capacity gap between PSO and the capacity that achieved
in the hybrid D-A precoders proposed in [7, 8, 10, 11] when
SNR = 0 dB is about 3 bits/s/Hz, while when SNR = 30 dB,
the capacity gap increase to around 4 bits/s/Hz. Hybrid D-A
precoding where $\mathbf{G} = \mathbf{AD}$ is proposed, with the help of
PSO, the capacity is more than the capacity of hybrid D-A BF
system in [4, 7, 8, 10, 11]. In addition, PSO achieves the same
capacity as SVD-based hybrid D-A precoding achieved, which
is near optimal solution [9]. Likewise, the loss of PSO capacity
is due to less number of chosen agents. The capacity of PSO
can be improved by using more number of agents. Finally, it is
observed from Fig. 3 and Fig. 4 that as the number of antennas
are increasing from $M = 16$ to $M = 64$, the overall capacity
of the system is enhanced despite $N$ RF chains are fixed.
Finally, in this paper, as a 3D BF gain is considered which
means, the beampattern should be a 3D pattern. Furthermore,
when $M$ is large enough that will lead to negligible (inter RF
and inter-user)- interference. Therefore, the minimum angle for
3D beampatterns between two users that can be distinguished without interference to each other is investigated in this paper. A beampattern function is equal to array factor (AF) that has been modeled in [12], and we may rearrange AF in the following form

$$AF = \frac{\sin \left( \frac{M_2}{2} \Lambda_x \right) \sin \left( \frac{M_2}{2} \Lambda_y \right)}{\left( \frac{M_2 M_2}{4} \Lambda_x \Lambda_y \right)}$$

where $\Lambda_x = 2\pi \lambda^{-1} u_x \sin(\theta_i) \cos(\phi_i) + \beta_x, \quad \Lambda_y = 2\pi \lambda^{-1} u_y \sin(\theta_i) \sin(\phi_i) + \beta_y, \quad i \in [r, t]$, and $\beta_x, \beta_y$ is determined by user position in small cell.

Fig. 5 shows that the beam pattern generated by a transmitter with a $M = 64$ planar array. In this case, the users can be separated by $\theta^i$ and $\phi^i$ directions. The patterns are generated by using PSO precoder. It can be noticed that the beam pattern of these beams are highly directional and each user can be separated easily with the help of different angles.

![Figure 5. Beam pattern square array by using the proposed scheme for an array of size $M = 64$, $N = 2$, with different angles.](image)

V. CONCLUSIONS

In this paper, a new method based on PSO for a hybrid D-A precoding system based on a sub-antenna array architecture for a mm-Wave system has been proposed. This algorithm has maximised the capacity of the hybrid D-A BF for the mm-Wave massive MIMO system. Simulation results showed that PSO was able to achieve higher capacity than the existing hybrid D-A precoding algorithms for the mm-Wave system. In addition, our simulation result verified that the proposed PSO achieved a close performance as compared to the optimal unconstrained precoding.

ACKNOWLEDGEMENT

Osama Alluhaibi is financially supported by the Higher Education and Scientific Research of IRAQ / Kirkuk University. Dr. Zhu acknowledges support from European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 643297 (RAPID).

REFERENCES


