Physical Communication 26 (2018) 81–86

Contents lists available at ScienceDirect

Physical Communication

journal homepage: www.elsevier.com/locate/phycom

Full length article

Block-sparse hybrid precoding and limited feedback for millimeter wave massive MIMO systems

Xuefeng Liu [a,*], Weixia Zou [a,b]

[a] Key Lab. of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, Beijing 100876, PR China
[b] State Key Lab. Of Millimeter Waves, Southeast University, Nanjing 210096, PR China

ARTICLE INFO

Article history:
Received 1 May 2017
Received in revised form 24 November 2017
Accepted 3 December 2017
Available online 6 December 2017

Keywords:
Millimeter wave
MIMO
Hybrid precoding
Block-sparse reconstruction
Limited feedback

ABSTRACT

Millimeter wave (mmWave) communication using massive multiple input multiple output (MIMO) techniques has been regarded as a key enabling technology for 5G wireless system, as it can offer gigabit-per-second data rates. Due to the high cost and power consumption of mixed-signal devices in mmWave systems, the hybrid analog and digital precoding transceiver architecture has recently received considerable attention. In this paper, we fully deploy spatial structure of mmWave channel to find hybrid precoders with near-optimal performance and reduce the number of feedback bits. MmWave MIMO channel exhibits the block sparsity structure because of the limited number of scattering clusters with low angular spreads. Consequently, we formulate hybrid precoders design as a block-sparse reconstruction problem and a lower complexity algorithm for finding precoders is proposed based on the greedy sequence clustering. To assure the performance of this algorithm, a estimation method for the number of the blocks and radio frequence (RF) chains is proposed. Under this framework, for the analog precoder the number of the phases needed to be quantized is only two times minus one of that of all blocks, which remains unchanged even the number of RF chains increases. The simulation results demonstrate that the proposed algorithm can achieve the near-optimal performance and save the feedback overhead by utilizing the block sparsity information.

1. Introduction

The capacity of the existing wireless networks cannot keep up with the dramatic proliferation of mobile traffic in the near future, so it is urgent for innovative communication technologies. Millimeter wave communication has achieved gigabit-per-second data rates in indoor wireless and fixed outdoor systems [1–3]. In recent studies, MIMO system with large-scale antenna arrays operating at millimeter-wave frequencies, referred to as mmWave massive MIMO system, is regarded as a core technology of the next-generation (5G) outdoor wireless cellular systems due to significant increases in the spectral efficiency and the bandwidth [4,5].

According to the recent measurements [6], the major challenges for mmWave cellular systems are the huge propagation losses and rain attenuation. Fortunately, because of the small wavelength of mmWave signals, mmWave MIMO transmitter can be equipped with large-scale antenna arrays to provide significant array gains to achieve a reasonable signal-to-noise ratio (SNR) at the receiver. Furthermore, mmWave MIMO also can support spatial multiplexing of multiple data streams due to scattering of mmWave channel and antenna polarization [7,8].

Implementation method of the MIMO precoding is different between traditional and mmWave massive MIMO systems. For traditional MIMO systems, precoding is typically performed at baseband by controlling both the signals phase and amplitude fully digitally, which demands RF chains comparable in number to the antenna elements. Unfortunately, the high cost and power consumption of mmWave mixed-signal hardware makes the digital precoding impossible. To overcome this obstacle and achieve larger precoding gains, a hybrid precoding is proposed, which divides the precoding operations into the analog and digital domains requiring a small number of RF chains interfacing between a low-dimensional digital precoder and a high-dimensional analog precoder [9,10].

Hybrid precoding is a newly-emerged technique with the development of mmWave MIMO communication [9–12]. In [10], it is shown that near-optimal hybrid precoders can be found by minimizing the Euclidean distance between hybrid precoders and the optimal precoder. Furthermore, an algorithm for finding hybrid precoders is presented based on the concept of orthogonal matching pursuit (OMP), where the columns of analog precoding matrix are picked from certain candidate vectors, such as the array response vectors. In [11], a hybrid precoding algorithm with

* Corresponding author.
E-mail address: liuxfeng6@bupt.edu.cn (X. Liu).

https://doi.org/10.1016/j.phycom.2017.12.001
1874-4907/

low computation complexity is presented by exploiting the semi-unitary optimum precoder. In [12], effective alternating minimization algorithms are proposed by treating the hybrid precoders design as a matrix factorization problem. However, the works in [11,12] mainly revolve around how to formulate hybrid precoding problem and find the near-optimal precoders with low complexity, and take no account of how to design efficient quantization and feedback of hybrid precoders, which is essential and important for frequency division duplex (FDD) communication systems. In this paper, we fully exploit spatial block-sparse structure of mmWave channel to find the hybrid precoders with near-optimal performance and reduce the number of feedback bits.

The main contributions of the paper can be summarized as follows:

(1) We formulate the hybrid precoding as a block-sparse reconstruction problem, and present a hybrid precoding algorithm based on greedy sequence clustering with lower complexity.

(2) The estimation method for the number of the blocks and RF chains is given, which is essential for the algorithm implementation and system design in advance.

(3) The number of quantization bits greatly reduces for the analog precoder, because the columns of analog precoding matrix are well arranged according to block sparsity structure.

The rest of the paper is organized as follows. In Section 2, the mmWave channel and hybrid precoding system model are presented and the hybrid precoding problem is formulated. In Section 3, the proposed block-sparse hybrid precoding algorithm is analyzed in detail. In Section 4, the simulation results are presented and discussed. Finally, in Section 5, the conclusions are drawn.

This paper uses following notations: $\mathbf{A}$ and $\mathbf{a}$ stand for a matrix and a column, respectively; $n{:}m$ stands for the vector $[n, n+1, n+2, \ldots, m]$; The transpose and conjugate transpose of $\mathbf{A}$ are represented by $\mathbf{A}^T$ and $\mathbf{A}^H$; $\mathbf{A}^\dagger$ is the Moore–Penrose pseudo inverse of $\mathbf{A}$; $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$; $\|\mathbf{a}\|_p$ is the p-norm of $\mathbf{a}$; $diag(\mathbf{A})$ is a vector formed by the diagonal elements of $\mathbf{A}$; $\mathbf{A}(l)$ and $\mathbf{a}(l)$ denote the lth column vector of $\mathbf{A}$, lth entry of $\mathbf{a}$; $\mathbf{A}_\Omega$ denote the sub-matrix formed by collecting the columns of $\mathbf{A}$ whose indexes are in set Ω; $[\mathbf{A}, \mathbf{B}]$ denotes horizontal concatenation.

2. System model and problem formulation

Consider the single-user mmWave massive MIMO system equipped with uniform linear array as shown in Fig. 1, where N_s data streams are sent and combined by N_t transmit antennas and N_r receive antennas. For the hybrid precoding system, the transmitter and receiver are respectively equipped with N_t^{RF} and N_r^{RF} chains such that $N_s \le N_t^{RF} \le N_t$ and $N_s \le N_r^{RF} \le N_r$.

For this hardware architecture such as that in [9–12], transmitter applies an $N_t^{RF} \times N_s$ baseband precoder $\mathbf{F}_{BB}$ followed by an $N_t \times N_t^{RF}$ analog precoder $\mathbf{F}_{RF}$. The total power constraint is enforced by normalizing $\mathbf{F}_{BB}$ such that $\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = N_s$. The transmitted signal is $\mathbf{x} = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s}$, where $\mathbf{s}$ represents the symbol vector and $E[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_s}\mathbf{I}_{N_s}$. The receiver applies an $N_r \times N_r^{RF}$ analog combining matrix $\mathbf{W}_{RF}$ and an $N_r^{RF} \times N_s$ baseband combining matrix $\mathbf{W}_{BB}$, so the received signal can be written as

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

where ρ is the average received power, $\mathbf{H}$ is the $N_r \times N_t$ channel matrix, and $\mathbf{n}$ is the vector with i.i.d $\mathcal{CN}(0, \sigma_{\mathbf{n}}^2)$ noise.

Because of the scattering geometry relative to a given user varying fast, the extended Saleh–Valenzuela model is widely adopted for mmWave MIMO channel. With the channel model, the discrete-time narrowband channel $\mathbf{H}$ can be written as

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}} \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{il} \mathbf{a}_r(\phi_{il}^r) \mathbf{a}_t(\phi_{il}^t)^H \quad (2)$$

where α_{il} is the complex gain of the lth ray in the ith cluster following the complex Gaussian distribution $\mathcal{CN}(0, \sigma_{\alpha,i}^2)$ and $\sum_{i=1}^{N_{cl}} \sigma_{\alpha,i}^2 = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}}$, and N_{cl} and N_{ray} represent the number of clusters and the number of rays in each cluster. In addition $\mathbf{a}_t(\phi_{il}^t)$ and $\mathbf{a}_r(\phi_{il}^r)$ are the antenna array response vectors at the transmitter and receiver at the lth ray in the ith cluster azimuth angles of departure and arrival(AODs and AOAs), which are given by

$$\begin{aligned} \mathbf{a}_t(\phi_{il}^t) &= \frac{1}{\sqrt{N_t}}[1, e^{jkd\sin(\phi_{il}^t)}, \ldots, e^{j(N-1)kd\sin(\phi_{il}^t)}]^T \\ \mathbf{a}_r(\phi_{il}^r) &= \frac{1}{\sqrt{N_r}}[1, e^{jkd\sin(\phi_{il}^r)}, \ldots, e^{j(N-1)kd\sin(\phi_{il}^t)}]^T \end{aligned} \quad (3)$$

where $k = 2\pi/\lambda$ and d is the inter-element spacing. In this paper, the Laplacian distribution with uniformly distributed mean angles of ϕ_i over $[0, 2\pi)$ and angular spread σ_ϕ, is adopted for the random variables AODs and AOAs.

The problem to maximize the spectral efficiency for the studied system involves joint optimization over the hybrid precoders and combiners [10]. To simplify transceiver design, the problem can be decoupled by focusing on the design of the hybrid precoders [10,12]. For the hybrid precoding systems, it has been proved that maximizing the spectral efficiency is equivalent to minimizing the Euclidean distance between hybrid precoders and fully digital precoder in [10]. In [10], it has been shown that the problem of finding the near-optimal precoders to maximize the spectral efficiency can be formulated as the following sparse reconstruction problem.

$$\begin{aligned} (\mathbf{F}_{RF}, \mathbf{F}_{BB}) &= \arg\min \left\| \mathbf{F}_{opt} - \mathbf{A}\tilde{\mathbf{F}}_{BB} \right\|_F \\ \text{s.t.} \quad & \left\| \mathrm{diag}(\tilde{\mathbf{F}}_{BB}\tilde{\mathbf{F}}_{BB}^H) \right\|_0 = N_t^{RF} \\ & \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = N_s \end{aligned} \quad (4)$$

where $\mathbf{F}_{opt}$ stands for the optimal fully digital precoder, and $\mathbf{A}$ of size $N_t \times N_\phi$ is a sparse dictionary composed of array response vectors. The classical Compressed Sensing (CS) recovery problem has been extensively studied and many algorithms have been proposed, i.e. basis pursuit (BP) algorithm [13], OMP algorithm [14] and many variants of OMP algorithm and subspace pursuit (SP) algorithm [15]. In this paper, we further explore the sparse structure of mmWave MIMO channel to design a more efficient structured sparse signal recovery algorithm, with the purpose to reduce the number of feedback bits used to quantize the precoders.

3. Block-sparse hybrid precoding and feedback

3.1. Block-sparse hybrid precoding

First, in order to reveal the black-sparse structure of mmWave MIMO channel, we adopt the following approximate representation. Similar to virtual channel representation, $\mathbf{H}$ can be written approximately in more compact form

$$\mathbf{H} = \sum_i \mathbf{H}_i \approx \sum_i \mathbf{A}_{r,i}\tilde{\mathbf{H}}_i\mathbf{A}_{t,i}^H \quad (5)$$

where $\tilde{\mathbf{H}}_i$ is the angular domain channel matrix for the ith cluster. $\mathbf{A}_{r,i}$ and $\mathbf{A}_{t,i}$ are $N_r \times L_r$ and $N_t \times L_t$ matrix of array response vectors respectively, which are given by

$$\begin{aligned} \mathbf{A}_{t,i} &= [\mathbf{a}_t(g(\phi_i^t)),\ \mathbf{a}_t(g(\phi_i^t) + \frac{2\pi}{N_{\phi^t}}), \cdots \mathbf{a}_t(g(\phi_i^t) + L_t\frac{2\pi}{N_{\phi^t}})\ ] \\ \mathbf{A}_{r,i} &= [\mathbf{a}_r(g(\phi_i^r)),\ \mathbf{a}_r(g(\phi_i^r) + \frac{2\pi}{N_{\phi^r}}),\ \cdots,\ \mathbf{a}_r(g(\phi_i^r) + L_r\frac{2\pi}{N_{\phi^r}})\ ] \end{aligned} \quad (6)$$

where L_t and L_r are the sample number of AODs and AOAs respectively in ith cluster, $g(\phi_i)$ generates the first sample of AODs or