SCIENCE CHINA
Information Sciences

• **RESEARCH PAPER** •

June 2016, Vol. 59 062301:1–062301:11
doi: 10.1007/s11432-015-5392-9

Robust capacity maximization transceiver design for MIMO OFDM systems

Shaozhen GUO, Chengwen XING*, Zesong FEI & Dan LI

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081*, China*

Received April 23, 2015; accepted May 27, 2015; published online September 15, 2015

Abstract In this paper, we investigated capacity maximization problem for Multiple Input Multiple Output Orthogonal Frequency Division Multiplexing systems with imperfect channel state information (CSI). To the best of our knowledge, the considered problem is still an open problem. However, the transceiver designs for MIMO OFDM systems have been extensively studied. It seems nobody gives closed-form solutions for resource allocation for MIMO OFDM systems with statistical channel estimation errors up to date. In our work, based on practical channel estimation algorithm, the channel estimation errors are first derived and then the robust resource allocation problem has been formulated. The structure of the optimal robust precoder is first derived, based on which the optimization problem will be simplified significantly. Furthermore, based on the Lagrangian dual method, a robust power allocation algorithm is proposed. The proposed power allocation can be considered as a variant of water-filling solution named cluster water-filling solution. Finally, simulation results show that our proposed robust design outperforms the non-robust design in terms of channel capacity.

Keywords MIMO-OFDM, capacity maximization, water-filling, channel uncertainty, robust design

Citation Guo S Z, Xing C W, Fei Z S, et al. Robust capacity maximization transceiver design for MIMO OFDM systems. Sci China Inf Sci, 2016, 59(6): 062301, doi: 10.1007/s11432-015-5392-9

1 Introduction

For modern communications, higher frequency and power efficiencies are permanent goals. To realize or at least approach these goals, many new technologies and ideas are introduced into wireless communications. Orthogonal frequency division multiplexing (OFDM) and multiple input multiple output (MIMO) are widely accepted as two breakthrough technologies for wireless research [1,2]. Of course, the combination of these two famous technologies will definitely inherit their individual benefits. Inspired by the great potential performance advantage of MIMO OFDM, research of this subject has been extensively explored and is still under investigation [1,3].

Generally speaking, there is a rich body of literature work in the related areas for transceiver designs for MIMO OFDM systems and many good results have been published already. Nobody doubts the importance of MIMO OFDM systems, but somebody may think there are no problems that still need discussion in this direction. Unfortunately, the story is not so. The role of research in engineering is twofold. The first one is how to solve the new problems faced in coming new systems. The other one is to

* Corresponding author (email: xingchengwen@gmail.com)

reconsider the old problems to give new understandings and solve the well-known open problems. It is true that communication systems vary or evolve faster and faster and there are various fancy and significantly different communication systems, for example, cooperative communications, cognitive systems, cloud radio access network (C-RAN), and so on. Besides traditional coding, signal processing technologies, some new concepts such as cloud computation will also play an important role in communications. It means there will be many new challenging problems that we never encounter in the traditional systems [4–7]. Meanwhile, in the evolving procedure, some important but challenging problems are left to be solved later inevitably.

In this paper, we will discuss an open problem of capacity maximization for MIMO OFDM systems with channel estimation errors. Due to the limited length of training signals, perfect channel state information (CSI) seems only an ideal assumption and robust designs will have practical meanings. Robust designs not only suppress negative effects of channel estimation errors, but also can reveal the effects of channel estimation on whole system performance [8]. As a result, robust designs are of critical importance. In general, robust designs with imperfect CSI are much more challenging than their counterparts with perfect CSI [9].

There are many results on the robust transceiver designs for MIMO systems over flat fading channels [10–13]. The robust transceiver designs for point-to-point MIMO systems with receive correlations are discussed in [14]. On the other hand, robust linear minimum mean square error (LMMSE) transceiver designs with transmit correlations are investigated in [15]. Later, the robust transceiver designs are also proposed for more general dual-hop and multi-hop amplify-and-forward MIMO systems [16–18]. Different from the perfect CSI cases, the robust designs for flat fading channels cannot be extended directly to the cases with frequency selected fading channels. It is because however in each subcarrier the problem is exactly a MIMO transceiver design over flat fading channel, the robust model in the transceiver design is closely related with the power allocated to this subcarrier, which is unknown a priori. Compared with capacity maximization, mean square error (MSE) minimization transceiver designs for MIMO OFDM systems are a little bit easier to deal with. In [19], a robust algorithm is first proposed for the robust transceiver designs for MIMO OFDM systems. In a recent work, the optimal structure is proposed for robust MIMO OFDM systems [20]. It also reveals that for MSE the transceiver designs correspond to a variant of water-filling solutions named cluster water-filling solutions [20]. To the best of our knowledge, the robust transceiver design for MIMO OFDM systems aiming at capacity maximization is still open now.

In this paper, we take a further step to investigate the robust transceiver design for MIMO OFDM systems with channel estimation errors, which aims at capacity maximization. For channel estimation errors, we rigorously derive the estimation error model for a practical channel estimation algorithm. Based on the model, the robust transceiver design problem is formulated. To reveal its physical meaning and simplify the transceiver optimization problem, the optimal structure of the robust transceivers is derived. Then, remaining variables become some scalar variables and exploiting Lagrange dual method based on the Karush Kuhn Tucker (KKT) conditions, a new cluster water-filling solution is proposed. Finally, the simulation results demonstrate the performance advantages of our proposed solutions.

Notations: Throughout this paper, the vector is denoted by bold lower letters and the matrix is represented by bold upper letters. The notations $\boldsymbol{Z}^*$ and $\boldsymbol{Z}^\dagger$ denote the Conjugation and Hermitian transpose of the matrix $\boldsymbol{Z}$, respectively. The notation $\boldsymbol{x}^{\mathrm{T}}$ is the transpose of vector $\boldsymbol{x}$. $\mathrm{Tr}(\boldsymbol{Z})$ and $|\boldsymbol{Z}|$ are trace and determinant of the matrix $\boldsymbol{Z}$, respectively. The expectation operation is represented by symbol $\mathrm{E}\{\cdot\}$. Besides, the symbol a^+ means $\max\{0, a\}$ and the symbol $\otimes$ indicates the Kronecker product.

2 System model and channel estimation error model

In this paper, a classic point-to-point MIMO OFDM communication system is investigated, which consists of a transmitter with M_{t} transmit antennas and a receiver with M_{r} receive antennas. The transmitter transmits $N_D \leqslant \min(M_{\mathrm{t}}, M_{\mathrm{r}})$ independent streams on each subcarrier simultaneously. In a practical scenario, the transmissions can be classified into two kinds of transmissions, that is, information transmission and pilot transmission for channel estimation.

In information transmission procedure, the data vector on the kth subcarrier is denoted by $\boldsymbol{s}_k \in \mathbb{C}^{N_D \times 1}$ and without loss of generality, the covariance matrix of the transmitted signal satisfies $\mathrm{E}\{\boldsymbol{s}_k \boldsymbol{s}_k^\dagger\} = \boldsymbol{I}_{N_D}$. The received signal $\boldsymbol{y}_k$ on the kth subcarrier has the following formula:

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{F}_k \boldsymbol{s}_k + \boldsymbol{n}_k, \qquad k = 0, \ldots, K-1, \tag{1}$$

where $\boldsymbol{F}_k \in \mathbb{C}^{M_\mathrm{t} \times N_D}$ is the precoding matrix on kth subcarrier, $\boldsymbol{H}_k \in \mathbb{C}^{M_\mathrm{r} \times M_\mathrm{t}}$ denotes the channel from the transmitter to the receiver on kth subcarrier, and $\boldsymbol{n}_k \in \mathbb{C}^{M_\mathrm{r} \times 1}$ is the additive white Gaussian noise vector on kth subcarrier with $\mathrm{E}\{\boldsymbol{n}_k \boldsymbol{n}_k^\dagger\} = \sigma_n^2 \boldsymbol{I}_{M_\mathrm{r}}$. Assuming the maximum transmit power is P, then the power constraint at the transmitter can be formulated as

$$\sum_{k=0}^{K-1} \mathrm{Tr}(\boldsymbol{F}_k {\boldsymbol{F}_k}^\dagger) \leqslant P. \tag{2}$$

In the channel estimation procedure, the training sequences are transmitted, based on which the channel parameters will be estimated based on the observations at the receiver. The signal $\boldsymbol{y}$ received at the receiver in frequency domain can be compactly formulated as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \tag{3}$$

where the parameters are defined based on the following expressions:

$$\boldsymbol{y} \triangleq [{\boldsymbol{y}_0}^\mathrm{T}, {\boldsymbol{y}_1}^\mathrm{T}, \ldots, {\boldsymbol{y}_{K-1}}^\mathrm{T}]^\mathrm{T}, \tag{4}$$

$$\boldsymbol{x} \triangleq [{\boldsymbol{x}_0}^\mathrm{T}, {\boldsymbol{x}_1}^\mathrm{T}, \ldots, {\boldsymbol{x}_{K-1}}^\mathrm{T}]^\mathrm{T}, \tag{5}$$

$$\boldsymbol{H} \triangleq \mathrm{diag}[\boldsymbol{H}_0, \boldsymbol{H}_1, \ldots, \boldsymbol{H}_{K-1}], \tag{6}$$

$$\boldsymbol{n} \triangleq [{\boldsymbol{n}_0}^\mathrm{T}, {\boldsymbol{n}_1}^\mathrm{T}, \ldots, {\boldsymbol{n}_{K-1}}^\mathrm{T}]^\mathrm{T}. \tag{7}$$

Here, the transmitted signal $\boldsymbol{x}$ is pilot or training signal. Based on the frequency domain signal model given by (3), the corresponding received signal in time domain can be written as

$$\begin{aligned} \boldsymbol{r} &= (\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{r}})\boldsymbol{y} \\ &= (\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{r}})\boldsymbol{H}(\boldsymbol{\mathcal{D}} \otimes \boldsymbol{I}_{M_\mathrm{t}})(\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{t}})\boldsymbol{x} + (\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{r}})\boldsymbol{n}, \end{aligned} \tag{8}$$

where $\boldsymbol{\mathcal{D}} \in \mathbb{C}^{K \times K}$ denotes the normalized discrete Fourier transform (DFT) matrix. Define $\boldsymbol{\mathcal{H}} \triangleq (\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{r}})\boldsymbol{H}(\boldsymbol{\mathcal{D}} \otimes \boldsymbol{I}_{M_\mathrm{t}})$, $\boldsymbol{c} \triangleq (\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{t}})\boldsymbol{x}$, and $\boldsymbol{w} \triangleq (\boldsymbol{\mathcal{D}}^\dagger \otimes \boldsymbol{I}_{M_\mathrm{r}})\boldsymbol{n}$, then the matrix $\boldsymbol{\mathcal{H}}$ is a $KM_\mathrm{r} \times KM_\mathrm{t}$ block circular matrix which can be derived using the properties of DFT operation and can be written as

$$\boldsymbol{\mathcal{H}} \triangleq \begin{bmatrix} \boldsymbol{\mathcal{H}}^{(0)} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\mathcal{H}}^{(L-1)} & \boldsymbol{\mathcal{H}}^{(L-2)} & \cdots & \boldsymbol{\mathcal{H}}^{(1)} \\ \boldsymbol{\mathcal{H}}^{(1)} & \boldsymbol{\mathcal{H}}^{(0)} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\mathcal{H}}^{(L-1)} & \cdots & \boldsymbol{\mathcal{H}}^{(2)} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\mathcal{H}}^{(L-1)} & \boldsymbol{\mathcal{H}}^{(L-2)} & \boldsymbol{\mathcal{H}}^{(L-3)} & \cdots & \boldsymbol{\mathcal{H}}^{(0)} \end{bmatrix}, \tag{9}$$

where L is the length of the multipath channel, and each element $\boldsymbol{\mathcal{H}}^{(\ell)}$ denotes the lth tap of the multipath MIMO channel from the transmitter to the receiver in time domain and is given by

$$\boldsymbol{\mathcal{H}}^{(\ell)} = \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{H}_k \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}k\ell}, \qquad l = 0, 1, \ldots, L-1. \tag{10}$$

Similarly, using the characteristics of DFT operation, vector $\boldsymbol{c}$ defined in (8) can be formulated as

$$\boldsymbol{c} = \left[\underbrace{\left(\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} \boldsymbol{x}_k \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}k(0)}\right)^\mathrm{T}}_{\boldsymbol{c}_0}, \underbrace{\left(\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} \boldsymbol{x}_k \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}k(1)}\right)^\mathrm{T}}_{\boldsymbol{c}_1}, \ldots, \underbrace{\left(\frac{1}{\sqrt{K}}\sum_{k=0}^{K-1} \boldsymbol{x}_k \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}k(K-1)}\right)^\mathrm{T}}_{\boldsymbol{c}_{K-1}}\right]^\mathrm{T}, \tag{11}$$

where element $\boldsymbol{c}_i$ can be expressed as

$$\boldsymbol{c}_i = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \boldsymbol{x}_k \mathrm{e}^{\mathrm{j}\frac{2\pi}{K}ki}, \qquad i = 0, 1, \ldots, K-1. \tag{12}$$

Based on the above definitions, the received signal defined in (8) can be written as

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{\mathcal{H}}\boldsymbol{c} + \boldsymbol{w} \\ &= (\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{I}_{M_{\mathrm{r}}}) \mathrm{vec}\left(\left[\boldsymbol{\mathcal{H}}^{(0)} \cdots \boldsymbol{\mathcal{H}}^{(L-1)}\right]\right) + \boldsymbol{w}, \end{aligned} \tag{13}$$

where $\boldsymbol{C}$ is a circular matrix defined as follows:

$$\boldsymbol{C} \triangleq \begin{bmatrix} \boldsymbol{c}_0 & \boldsymbol{c}_1 & \cdots \cdots \cdots & \boldsymbol{c}_{K-1} \\ \boldsymbol{c}_{K-1} & \boldsymbol{c}_0 & \ddots \; \ddots \; \vdots & \boldsymbol{c}_{K-2} \\ \vdots & \cdots & \ddots \; \ddots \; \vdots & \vdots \\ \boldsymbol{c}_{K-L+1} & \boldsymbol{c}_{K-L+2} & \cdots \cdots \cdots & \boldsymbol{c}_{K-L} \end{bmatrix}. \tag{14}$$

For convenience, let $\boldsymbol{\theta} = \mathrm{vec}([\boldsymbol{\mathcal{H}}^{(0)} \cdots \boldsymbol{\mathcal{H}}^{(L-1)}])$, then the time domain channel estimated using the LMMSE criterion is given by [21]

$$\hat{\boldsymbol{\theta}} = \left(\sigma_n^{-2} (\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{I}_{M_{\mathrm{r}}})^{\dagger} (\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{I}_{M_{\mathrm{r}}}) + \boldsymbol{R}_{\mathrm{ch}}^{-1}\right)^{-1} \times \sigma_n^{-2} (\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{I}_{M_{\mathrm{r}}})^{\dagger} \boldsymbol{r}, \tag{15}$$

and the resultant MSE can be directly given by

$$\mathrm{E}\left\{(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^{\dagger}\right\} = \left(\boldsymbol{R}_{\mathrm{ch}}^{-1} + \sigma_n^{-2} (\boldsymbol{C}^{*} \boldsymbol{C}^{\mathrm{T}}) \otimes \boldsymbol{I}_{M_{\mathrm{r}}}\right)^{-1}, \tag{16}$$

where $\boldsymbol{R}_{\mathrm{ch}}$ is known preferentially [22] and can be calculated by $\boldsymbol{R}_{\mathrm{ch}}^{-1} = \mathrm{E}\left\{\boldsymbol{\theta}\boldsymbol{\theta}^{\dagger}\right\}$. For uncorrelated channel taps, $\boldsymbol{R}_{\mathrm{ch}} = \boldsymbol{\Xi}_{\mathrm{ch}} \otimes \boldsymbol{I}_{M_{\mathrm{r}}M_{\mathrm{t}}}$ and $\boldsymbol{\Xi}_{\mathrm{ch}} = \mathrm{diag}[\xi_{h_0}, \xi_{h_1}, \ldots, \xi_{h_{L-1}}]$, where ξ_{h_l} is the variance for lth channel tap [19].

Furthermore, using the relationship between frequency domain and time domain, the channel in frequency domain can be formulated as [21]

$$\mathrm{vec}\left([\boldsymbol{H}_0, \ldots, \boldsymbol{H}_{K-1}]\right) = \sqrt{K} \left(\boldsymbol{\mathcal{D}}_L \otimes \boldsymbol{I}_{M_{\mathrm{r}}M_{\mathrm{t}}}\right) \boldsymbol{\theta}, \tag{17}$$

where $\boldsymbol{\mathcal{D}}_L$ is the first L columns of $\boldsymbol{\mathcal{D}}$. Therefore, based on the above definitions, we have

$$\begin{aligned} &\mathrm{E}\left\{\mathrm{vec}\left([\Delta \boldsymbol{H}_0, \ldots, \Delta \boldsymbol{H}_{K-1}]\right) \mathrm{vec}^{\dagger}\left([\Delta \boldsymbol{H}_0, \ldots, \Delta \boldsymbol{H}_{K-1}]\right)\right\} \\ &\quad = \left(\boldsymbol{\mathcal{D}}_L \otimes \boldsymbol{I}_{M_{\mathrm{r}}M_{\mathrm{t}}}\right) \left(\boldsymbol{\Xi}_{\mathrm{ch}}^{-1} \otimes \boldsymbol{I}_{M_{\mathrm{t}}} + \sigma_n^{-2} (\boldsymbol{C}^{*} \boldsymbol{C}^{\mathrm{T}})\right)^{-1} \otimes \boldsymbol{I}_{M_{\mathrm{r}}} \left(\boldsymbol{\mathcal{D}}_L \otimes \boldsymbol{I}_{M_{\mathrm{r}}M_{\mathrm{t}}}\right)^{\dagger} K \\ &\quad = \left(\boldsymbol{\mathcal{D}}_L \otimes \boldsymbol{I}_{M_{\mathrm{r}}M_{\mathrm{t}}}\right) \boldsymbol{\Gamma} \otimes \boldsymbol{I}_{M_{\mathrm{r}}} \left(\boldsymbol{\mathcal{D}}_L \otimes \boldsymbol{I}_{M_{\mathrm{r}}M_{\mathrm{t}}}\right)^{\dagger} K, \end{aligned} \tag{18}$$

where $\Delta \boldsymbol{H}_k = \boldsymbol{H}_k - \hat{\boldsymbol{H}}_k$ and $\boldsymbol{\Gamma} \triangleq \left(\boldsymbol{\Xi}_{\mathrm{ch}}^{-1} \otimes \boldsymbol{I}_{M_{\mathrm{t}}} + \sigma_n^{-2} (\boldsymbol{C}^{*} \boldsymbol{C}^{\mathrm{T}})\right)^{-1}$. Taking the $M_{\mathrm{r}}M_{\mathrm{t}} \times M_{\mathrm{r}}M_{\mathrm{t}}$ block diagonal elements from (18), we have

$$\mathrm{E}\left\{\mathrm{vec}(\Delta \boldsymbol{H}_k) \mathrm{vec}^{\dagger}(\Delta \boldsymbol{H}_k)\right\} = \left(\sum_{n=0}^{L-1} \sum_{m=0}^{L-1} \left(\mathrm{e}^{-\mathrm{j}\frac{2\pi}{K}k(m-n)} \boldsymbol{\Gamma}_{m,n}\right)\right) \otimes \boldsymbol{I}_{M_{\mathrm{r}}}, \tag{19}$$

where $\boldsymbol{\Gamma}_{m,n}$ is the $M_{\mathrm{t}} \times M_{\mathrm{t}}$ matrix taken from the partition of $\boldsymbol{\Gamma}$ as

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}_{0,0} & \boldsymbol{\Gamma}_{0,1} & \cdots & \boldsymbol{\Gamma}_{0,L-1} \\ \vdots & \cdots & \ddots & \vdots \\ \boldsymbol{\Gamma}_{L-1,0} & \boldsymbol{\Gamma}_{L-1,1} & \cdots & \boldsymbol{\Gamma}_{L-1,L-1} \end{bmatrix}. \tag{20}$$

It has been proved in [16] that for arbitrary random matrix $\boldsymbol{Q}$ and $\boldsymbol{W}$, if

$$\mathrm{E}\left\{\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}^{\dagger}(\boldsymbol{W})\right\} = \boldsymbol{A} \otimes \boldsymbol{B}, \tag{21}$$

then for arbitrary matrix $\boldsymbol{R}$, we can get

$$\mathrm{E}\left\{\boldsymbol{Q}\boldsymbol{R}\boldsymbol{W}^{\dagger}\right\} = \boldsymbol{B}\mathrm{Tr}(\boldsymbol{R}\boldsymbol{A}^{\mathrm{T}}). \tag{22}$$

To obtain $\boldsymbol{\Psi}_{T,k}$, taking $\boldsymbol{R} = \boldsymbol{I}_{M_{\mathrm{t}}}$, $\boldsymbol{Q} = \boldsymbol{W} = \Delta\boldsymbol{H}_k$, $\boldsymbol{A} = \left(\sum_{n=0}^{L-1}\sum_{m=0}^{L-1}\left(\mathrm{e}^{-\mathrm{j}\frac{2\pi}{K}k(m-n)}\boldsymbol{\Gamma}_{m,n}\right)\right)$, and $\boldsymbol{B} = \boldsymbol{I}_{M_{\mathrm{r}}}$, therefore we have

$$\begin{aligned}\boldsymbol{\Psi}_{T,k} &= \mathrm{E}\left\{\boldsymbol{\Delta}\boldsymbol{H}_k\boldsymbol{\Delta}\boldsymbol{H}_k^{\dagger}\right\} \\ &= \mathrm{Tr}\left(\sum_{n=0}^{L-1}\sum_{m=0}^{L-1}\left(\mathrm{e}^{-\mathrm{j}\frac{2\pi}{K}k(m-n)}\boldsymbol{\Gamma}_{m,n}^{\mathrm{T}}\right)\right)\boldsymbol{I}_{M_{\mathrm{r}}}.\end{aligned} \tag{23}$$

Finally, based on the previous discussions on the channel estimation errors, the channel matrix has the following property:

$$\begin{aligned}\boldsymbol{H}_k &= \hat{\boldsymbol{H}}_k + \Delta\boldsymbol{H}_k, \\ \Delta\boldsymbol{H}_k &= \boldsymbol{H}_{w,k}\boldsymbol{\Psi}_{T,k}^{1/2},\end{aligned} \tag{24}$$

where $\boldsymbol{H}_{w,k}$ is a random matrix whose elements are independent and identically distributed Gaussian variables with zero mean and unit variance, and $\boldsymbol{\Psi}_{T,k}$ is the covariance matrix of channel estimation error. To obtain $\boldsymbol{\Psi}_{T,k}$, here we adopted the channel estimation method based on training sequences. Also, considering that the parameters to be estimated in time domain are much fewer than those in frequency domain, therefore, our main concern is channel estimation on time domain.

3 Capacity maximization transceiver design

For wireless designs, maximizing mutual information is undoubtedly the most widely used performance metric [17, 23]. In this section, the transceiver design and power allocation problem that aims at maximizing the system capacity is first formulated as a convex optimization problem. Then, based on the structure of the optimization problem, the transceiver design problem is decomposed into a series of independent subproblems. By solving these subproblems parallelly, the optimal transceiver structure is obtained. Furthermore, using the Lagrangian dual method, the power allocated on each subcarrier is calculated. Finally, a robust transmit power allocation algorithm to maximize the capacity of this point-to-point MIMO OFDM system is proposed.

3.1 Problem formulation

In this paper, our main goal is maximizing the mutual information of the MIMO-OFDM system. Taking the power constraint in (2) into consideration, the optimization problem of capacity maximization is formulated as

$$\begin{aligned}\max_{\boldsymbol{F}_k}\quad & \sum_{k=0}^{K-1}\log|\boldsymbol{I} + \boldsymbol{F}_k^{\dagger}\hat{\boldsymbol{H}}_k^{\dagger}\boldsymbol{K}_{\boldsymbol{F}_k}^{-1}\hat{\boldsymbol{H}}_k\boldsymbol{F}_k|, \\ \mathrm{s.t.}\quad & \boldsymbol{K}_{\boldsymbol{F}_k} = [\mathrm{Tr}(\boldsymbol{F}_k\boldsymbol{F}_k^{\dagger}\boldsymbol{\Psi}_{T,k}) + \sigma_n^2]\boldsymbol{I}, \\ & \sum_{k=0}^{K-1}\mathrm{Tr}(\boldsymbol{F}_k\boldsymbol{F}_k^{\dagger}) \leqslant P.\end{aligned} \tag{25}$$

To make the optimization problem decomposable, we introduce a series of auxiliary variables $\{P_k, k = 0, \ldots, K-1\}$. Later, we can see that P_k is actually the transmitting power on the kth subcarrier. Therefore, the optimization problem in (25) is written as

$$\begin{aligned}
\max_{\boldsymbol{F}_k, P_k} \quad & \sum_{k=0}^{K-1} \log|\boldsymbol{I} + \boldsymbol{F}_k^{\dagger} \hat{\boldsymbol{H}}_k^{\dagger} \boldsymbol{K}_{\boldsymbol{F}_k}^{-1} \hat{\boldsymbol{H}}_k \boldsymbol{F}_k|, \\
\text{s.t.} \quad & \boldsymbol{K}_{\boldsymbol{F}_k} = [\mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger} \boldsymbol{\Psi}_{T,k}) + \sigma_n^2] \boldsymbol{I}, \\
& \mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger}) \leqslant P_k, \qquad k = 0, \ldots, K-1, \\
& \sum_{k=0}^{K-1} P_k \leqslant P.
\end{aligned} \tag{26}$$

It is obvious that after reformulating the power constraint, for each fixed P_k, the original optimization problem can be decoupled into a series of independent subproblems about $\boldsymbol{F}_k$ and each subproblem can be solved parallelly [24]. After solving these subproblems, the optimal solution is easy to be derived. Therefore, in the following, the optimal structure of transceiver is first derived for fixed P_k, then based on the optimal transceiver, the eigenvalue of the transceiver and power allocation are jointly optimized by Lagrangian dual method.

3.2 The proposed solution

Based on the analysis above, we can see that for fixed P_k the optimization problem (26) can be written as

$$\begin{aligned}
\max_{\boldsymbol{F}_k} \quad & \sum_{k=0}^{K-1} \log|\boldsymbol{I} + \boldsymbol{F}_k^{\dagger} \hat{\boldsymbol{H}}_k^{\dagger} \boldsymbol{K}_{\boldsymbol{F}_k}^{-1} \hat{\boldsymbol{H}}_k \boldsymbol{F}_k|, \\
\text{s.t.} \quad & \boldsymbol{K}_{\boldsymbol{F}_k} = [\mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger} \boldsymbol{\Psi}_{T,k}) + \sigma_n^2] \boldsymbol{I}, \\
& \mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger}) \leqslant P_k, \qquad k = 0, \ldots, K-1.
\end{aligned} \tag{27}$$

Obviously, the power constraints in problem (27) have a decomposable structure in K blocks. Therefore problem (27) can be decomposed into K independent subproblems [25] and the decomposed kth subproblem can be expressed as

$$\begin{aligned}
\max_{\boldsymbol{F}_k} \quad & \log|\boldsymbol{I} + \boldsymbol{F}_k^{\dagger} \hat{\boldsymbol{H}}_k^{\dagger} \boldsymbol{K}_{\boldsymbol{F}_k}^{-1} \hat{\boldsymbol{H}}_k \boldsymbol{F}_k|, \\
\text{s.t.} \quad & \boldsymbol{K}_{\boldsymbol{F}_k} = [\mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger} \boldsymbol{\Psi}_{T,k}) + \sigma_n^2] \boldsymbol{I}, \\
& \mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger}) \leqslant P_k.
\end{aligned} \tag{28}$$

For illustrative purpose, we further define a scalar

$$\eta_{f_k} = \mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger} \boldsymbol{\Psi}_{T,k}) + \sigma_n^2. \tag{29}$$

Since the objective function is monotonically increasing about $\boldsymbol{F}_k$, the optimal solution of each subproblem always exists on the boundary [17], that is, $\mathrm{Tr}(\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger}) = P_k$. Hence, the power constraint condition in (28) can be reformulated as

$$\frac{\mathrm{Tr}[\boldsymbol{F}_k \boldsymbol{F}_k^{\dagger} (P_k \boldsymbol{\Psi}_{T,k} + \sigma_n^2 \boldsymbol{I})]}{\eta_{f,k}} = P_k. \tag{30}$$

Furthermore, if we define the following matrix:

$$\tilde{\boldsymbol{F}}_k = \frac{1}{\sqrt{\eta_{f_k}}} (P_k \boldsymbol{\Psi}_{T,k} + \sigma_n^2 \boldsymbol{I})^{1/2} \boldsymbol{F}_k, \tag{31}$$

then the problem in (28) can be rewritten as

$$\begin{aligned}\max_{\boldsymbol{F}_k} \quad & \log\left|\tilde{\boldsymbol{F}}_k^{\dagger}\left((P_k\boldsymbol{\Psi}_{T,k}+\sigma_n^2\boldsymbol{I})^{-1/2}\right)^{\dagger}\hat{\boldsymbol{H}}_k^{\dagger}\hat{\boldsymbol{H}}_k(P_k\boldsymbol{\Psi}_{T,k}+\sigma_n^2\boldsymbol{I})^{-1/2}\tilde{\boldsymbol{F}}_k+\boldsymbol{I}\right|,\\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^{\dagger})=P_k. \end{aligned} \tag{32}$$

The structure of the optimal solution for problem (32) has been given in [17] as

$$\tilde{\boldsymbol{F}}_k=\boldsymbol{U}_{\hat{\mathcal{H}}_k}\boldsymbol{\Lambda}_{\boldsymbol{F}_k}\boldsymbol{U}_{\mathrm{Arb},k}^{\dagger}, \tag{33}$$

where $\boldsymbol{U}_{\mathrm{Arb},k}$ denotes an arbitrary unitary matrix and $\boldsymbol{U}_{\hat{\mathcal{H}}_k}$ is the unitary matrix defined from the following eigenvalue decomposition:

$$\left((P_k\boldsymbol{\Psi}_{T,k}+\sigma_n^2\boldsymbol{I})^{-1/2}\right)^{\dagger}\hat{\boldsymbol{H}}_k^{\dagger}\hat{\boldsymbol{H}}_k(P_k\boldsymbol{\Psi}_{T,k}+\sigma_n^2\boldsymbol{I})^{-1/2}=\boldsymbol{U}_{\hat{\mathcal{H}}_k}\boldsymbol{\Lambda}_{\hat{\mathcal{H}}_k}\boldsymbol{U}_{\hat{\mathcal{H}}_k}^{\dagger}. \tag{34}$$

Supposing the diagonal entries of rectangular diagonal matrix $\boldsymbol{\Lambda}_{\boldsymbol{F}_k}$ are denoted as $f_{k,j}$'s, which are variables needed to be optimized later, the optimization problem can be further simplified by substituting (33) and (34) into (32):

$$\begin{aligned}\max_{f_{k,j},P_k} \quad & \sum_{k=0}^{K-1}\sum_{j=1}^{N_D}\log\left(1+f_{k,j}^2\,[\boldsymbol{\Lambda}_{\hat{\mathcal{H}}_k}]_{j,j}\right),\\ \text{s.t.} \quad & \sum_{j=1}^{N_D}f_{k,j}^2=P_k, \qquad k=0,\ldots,K-1,\\ & \sum_{k=0}^{K-1}P_k\leqslant P. \end{aligned} \tag{35}$$

Given that $[\boldsymbol{\Lambda}_{\hat{\mathcal{H}}_k}]_{j,j}$ is a very complicated function of P_k's, it is difficult to deal with the optimization problem (35) directly. Therefore, we take an alternative method to optimize the lower bound of our objective, which makes the problem more tractable. Obviously, from (34), we can get

$$[\boldsymbol{\Lambda}_{\hat{\mathcal{H}}_k}]_{j,j}\geqslant\frac{\lambda_{h_{k,j}}}{\lambda_{\max}(\boldsymbol{\Psi}_{T,k})P_k+\sigma_n^2}, \tag{36}$$

where $\lambda_{\max}(\boldsymbol{\Psi}_{T,k})$ is the maximum eigenvalue of matrix $\boldsymbol{\Psi}_{T,k}$ and $\lambda_{h_{k,j}}$ is the jth largest eigenvalue of $\hat{\boldsymbol{H}}_k^{\dagger}\hat{\boldsymbol{H}}_k$. Therefore, in the following part, we will discuss how to compute P_k's in detail and give the lower bound approximation of the original problem. The optimization problem is approximated as

$$\begin{aligned}\max_{f_{k,j},P_k} \quad & \sum_{k=0}^{K-1}\sum_{j=1}^{N_D}\log\left(1+\frac{\lambda_{h_{k,j}}f_{k,j}^2}{\lambda_{\max}(\boldsymbol{\Psi}_{T,k})P_k+\sigma_n^2}\right),\\ \text{s.t.} \quad & \sum_{j=1}^{N_D}f_{k,j}^2=P_k, \qquad k=0,\ldots,K-1,\\ & \sum_{k=0}^{K-1}P_k\leqslant P. \end{aligned} \tag{37}$$

It has been confirmed that the conditions of the optimization problem (37) are just regular conditions, which are also named as linear independence constraint qualification. Therefore, KKT conditions [26] are the necessary conditions for this optimal solution. If P_k's are given, the solution of $f_{k,j}^2$ in (37) is no doubt the water-filling solution [27]. Since the water-filling solution is unique [27], the optimal solution for $f_{k,j}^2$ with P_k given is unique as well. Based on the analysis above, we can naturally draw a conclusion that the water-filling alike solution satisfying the KKT conditions at the same time must be the optimal solution. Also, it is easy to see that P_k varies from different subcarriers, so do the water-filling levels. In the following, we will derive the solutions on different subcarriers, which satisfy the KKT conditions when adopting Lagrangian dual method. For notational simplicity, in the following derivations we replace $\lambda_{\max}(\boldsymbol{\Psi}_{T,k})$ by $\lambda_k^{\max}$.

The Lagrangian dual function of the optimization problem (37) is

$$\mathcal{L}(\{f_{k,j}\},\{P_k\}) = \sum_{k=0}^{K-1}\sum_{j=1}^{N_D}\log\left(1+\frac{\lambda_{h_{k,j}}f_{k,j}^2}{\lambda_k^{\max}P_k+\sigma_n^2}\right)-\gamma\left(\sum_{k=0}^{K-1}P_k-P\right) - \sum_{k=0}^{K-1}\left\{\mu_k\left(\sum_{j=1}^{N_D}f_{k,j}^2-P_k\right)\right\}, \tag{38}$$

where $\mu_k \geqslant 0, k=0,\ldots,K-1$, and $\gamma \geqslant 0$ are the Lagrange multipliers. Next, we decide to get the optimal solution by solving its KKT equations. The KKT conditions of the optimization problem (37) are derived to be

$$\frac{\lambda_{h_{k,j}}}{\lambda_k^{\max}P_k+\sigma_n^2}\frac{1}{\frac{\lambda_{h_{k,j}}f_{k,j}^2}{\lambda_k^{\max}P_k+\sigma_n^2}+1}f_{k,j} = \mu_k f_{k,j}\ , \qquad k=0,\ldots,K-1, \tag{39a}$$

$$\sum_j \frac{1}{\frac{\lambda_{h_{k,j}}f_{k,j}^2}{\lambda_k^{\max}P_k+\sigma_n^2}+1}\frac{\lambda_k^{\max}\lambda_{h_{k,j}}f_{k,j}^2}{(\lambda_k^{\max}P_k+\sigma_n^2)^2}-\mu_k=-\gamma, \tag{39b}$$

$$\sum_j f_{k,j}^2 = P_k,\quad \sum_k P_k = P. \tag{39c}$$

Substituting the first KKT condition (39a) into the second KKT condition (39b), we get

$$\sum_j\left\{\frac{\mu_k\lambda_k^{\max}f_{k,j}^2}{\lambda_k^{\max}P_k+\sigma_n^2}\right\}-\mu_k=-\gamma. \tag{40}$$

Using the KKT condition (39c) $\sum_j f_{k,j}^2 = P_k$, we derive

$$\mu_k\sigma_n^2/(\lambda_k^{\max}P_k+\sigma_n^2)=\gamma. \tag{41}$$

Based on (41) together with (39a), we obtain the following optimal water-filling solution.

Conclusion: The optimal solution $f_{k,j}^2$ of problem (37) on the kth subcarrier should satisfy

$$f_{k,j}^2=\left(\frac{1}{\gamma}\frac{\sigma_n^2}{\lambda_k^{\max}P_k+\sigma_n^2}-\frac{\lambda_k^{\max}P_k+\sigma_n^2}{\lambda_{h_{k,j}}}\right)^+. \tag{42}$$

Note that P_k is the power allocated to the kth subcarrier, which is also dependent on the optimized variables $f_{k,j}^2$ and may be different from one subcarrier to another. To solve for P_k, we take a sum of both sides of (42) with respect to j and get

$$P_k=\frac{1}{\lambda_k^{\max}P_k+\sigma_n^2}\frac{\sigma_n^2}{\gamma}L_k-(\lambda_k^{\max}P_k+\sigma_n^2)\sum_{j\in\phi_k}\frac{1}{\lambda_{h_{k,j}}}, \tag{43}$$

where ϕ_k is the set of subcarriers that are allocated with nonzero power and L_k is the cardinality of ϕ_k. After making some arrangement of (43), we can directly obtain the closed-form solution of P_k, which is also a function of the Lagrangian multiplier γ as follows:

$$P_k=\frac{\frac{\sigma_n^2}{\lambda_k^{\max}}+\sqrt{\left(\frac{\sigma_n^2}{\lambda_k^{\max}}\right)^2+4\left(\frac{1}{\lambda_k^{\max}}+\sum_j\frac{1}{\lambda_{h_{k,j}}}\right)\frac{\sigma_n^2}{\gamma}L_k}}{2\left(\frac{1}{\lambda_k^{\max}}+\sum_j\frac{1}{\lambda_{h_{k,j}}}\right)\lambda_k^{\max}}-\frac{\sigma_n^2}{\lambda_k^{\max}}. \tag{44}$$

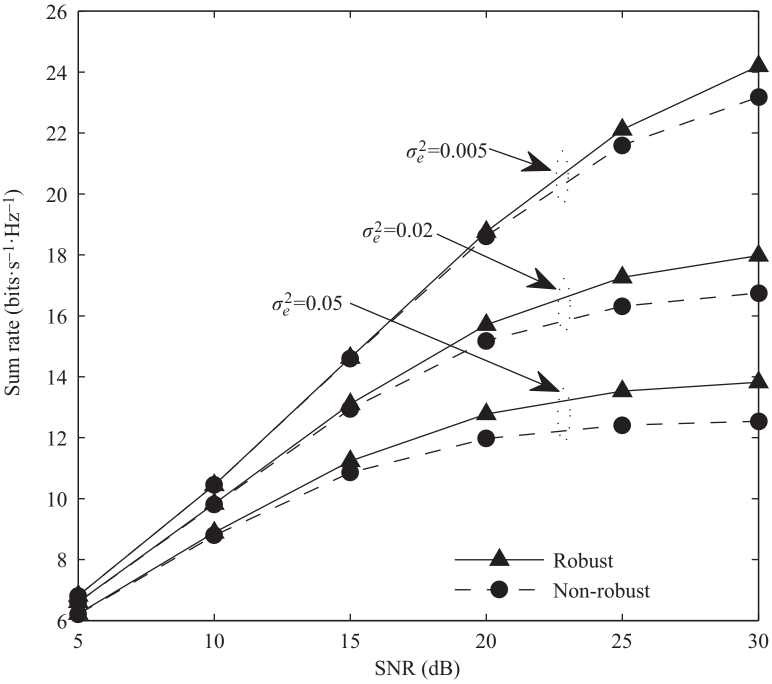

Figure 1 Sum capacity of the proposed robust algorithm and non-robust algorithm with different σ_e^2.

Now we see that γ is the only unknown variable. Based on (44) and the third KKT constraint (39c),

$$\sum_{k=0}^{K-1} P_k = P. \tag{45}$$

γ is easy to be solved by many numerical methods, for example, bisection search or golden search [28]. The detailed steps for this robust transmit power allocation algorithm are presented below:

Algorithm 1 Roboust power allocation algorithm

Require: $\phi = \phi_0 \cup \cdots \cup \phi_{K-1} = \{\text{Indexes of all the eigenchannels}\}, \frac{1}{\gamma}, P_k, \{\text{PA} = f_{k,j}^2\}$, for all k, j;
Ensure: $P_k, \{f_{k,j}^2\}$, for all k, j;
while length (find (PA < 0)) > 0 **do**
  {pos_Index} = Indexes of find (PA > 0);
  {neg_Index} = $\phi \backslash$ {pos_Index} (exclusion operation);
  Set PA (neg_Index) = 0;
  For all $(k, j) \in$ {pos_Index};
  Update $1/\gamma$ according to (44) and (45);
  Update P_k according to (44);
  Update PA (pos_Index) according to (42);
end while

4 Simulation results

In this section, the performance of the proposed robust algorithm is investigated. For the purpose of comparison, the non-robust algorithm (without considering the channel estimation error) is also simulated. A point-to-point MIMO-OFDM system where the transmitter and the receiver are equipped with the same number of antennas, $M_{\rm t} = M_{\rm r} = 4$ is considered. The length of the multipath channel is set to $L = 7$ and the number of subcarriers is $K = 256$. And the signal-to-noise ratio (SNR) is defined as $P/(K\sigma_n^2)$. The results in Figures 1 and 2 are an average in 10000 channel realizations.

On the basis of (14), we can see that $\boldsymbol{CC}^\dagger$ is a circulant block matrix. In the following simulation, we take the spatial correlation of the training sequence into account, and the time-domain training is

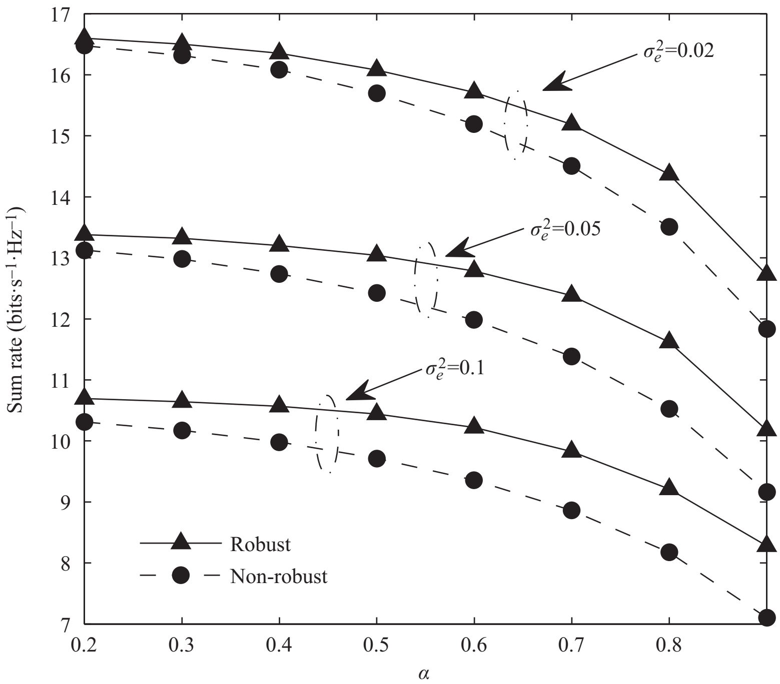

Figure 2 Sum capacity of the proposed robust algorithm and non-robust algorithm with different α.

white. Therefore, $\boldsymbol{CC}^{\dagger}$ is a block diagonal matrix. Furthermore, the extensive exponential correlation model is used to denote the spatial correlation [14,15]. In this case, the correlation matrix of the channel estimation error can be modeled as $[\boldsymbol{\Psi}_k]_{i,j} = \sigma_e^2 \alpha^{|i-j|}$, where $\sigma_e^2 = 1/\mathrm{SNR}_e$ is the variance of estimation error and SNR_e is the SNR in the earlier channel estimation phase.

First, we investigated the performance of the proposed robust algorithm when $\alpha = 0.6$. Figure 1 shows the sum capacity of the point-to-point MIMO OFDM system with different σ_e^2. It can be seen that our proposed robust design outperforms the corresponding algorithm with estimated CSI only particularly in high SNR regime. What is more, as the estimation error σ_e^2 increases, the performance gap between the robust and non-robust design becomes larger.

Figure 2 demonstrates the sum capacity of the point-to-point MIMO OFDM system for the proposed algorithm and the non-robust algorithm with different α when SNR = 20 dB. Although both algorithms have some performance degradation when α increases, the proposed algorithm achieves great performance improvement over the algorithm based on estimation CSI only. Also, with the increase of α, the gap between the proposed robust algorithm and non-robust algorithm increases.

5 Conclusion

In this paper, a robust capacity maximization transceiver design for MIMO OFDM systems was proposed. Although it seemed an old problem, to the best of our knowledge this problem was actually an open problem. Based on a practical channel estimation algorithm, the statistical properties of channel estimation errors were derived rigorously. Exploiting these statistical properties, the robust optimization problem of capacity maximization for MIMO OFDM systems was formulated first. Then, the optimal transceiver structure was analyzed in detail and the considered problem was simplified greatly. For the remaining variables, a cluster water-filling solution was proposed. After that, the simulation results showed that our proposed beamforming design outperforms that based on estimated CSI only significantly.

Acknowledgements This work was partly supported by China Mobile Research Institute (Grant No. [2014]451), National Natural Science Foundation of China (Grant No. 61421001), 111 Project of China (Grant No. B14010), and Beijing Natural Science Foundation (Grant No. 4152047).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Larsson E G, Stoica P. Space-Time Block Coding for Wireless Communications. Cambridge: Cambridge University Press, 2003
2 Tse D, Viswanath P. Fundamentals of Wireless Communication. Cambridge: Cambridge University Press. 2005
3 Bolcskei H, Gesbert D, Papadias C B, et al. Space-Time Wireless Systems: From Array Processing to MIMO Communications. Cambridge: Cambridge University Press. 2006
4 Luan T X, Gao F F, Zhang X D. Joint resource scheduling for relay-assisted broadband cognitive radio networks. IEEE Trans Wirel Commun, 2012, 11: 3090–3100
5 Zhu F C, Gao F F, Yao M L, et al. Joint information- and Jamming-beamforming for physical layer security with full duplex base station. IEEE Trans Signal Process, 2014, 62: 6391–6401
6 Zhang Z S, Long K P, Wang J P, et al. On swarm intelligence inspired self-organized networking: its bionic mechanisms, designing principles and optimization approaches. IEEE Commun Surv Tut, 2014, 16: 513–537
7 Zhang Z S, Long K P, Wang J P. Self-organization paradigms and optimization approaches for cognitive radio technologies: a survey. IEEE Trans Wirel Commun, 2013, 20: 36–42
8 Dietrich F A. Robust Signal Processing for Wireless Communications. Foundations in Signal Processing, Communications and Networking. Berlin: Springer Press, 2007
9 Zheng G, Wong K-K, Paulraj A, et al. Robust collaborative-relay beamforming. IEEE Trans Signal Process, 2009, 57: 3130–3143
10 Sampath H, Stoica P, Paulraj A. Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion. IEEE Trans Commun, 2001, 49: 2198–2206
11 Palomar D P, Cioffi J M, Lagunas M A. Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization. IEEE Trans Signal Process, 2003, 51: 2381–2399
12 Joham M, Utschick W, Nossek J A. Linear transmit processing in MIMO communications systems. IEEE Trans Signal Process, 2005, 53: 2700–2712
13 Serbetli S, Yener A. Transceiver optimization for mutiuser MIMO systems. IEEE Trans Signal Process, 2004, 52: 214–226
14 Zhang X, Palomar D P, Ottersten B. Statistically robust design of linear MIMO transceivers. IEEE Trans Signal Process, 2008, 56: 3678–3689
15 Ding M H, Blostein S D. MIMO minimum total MSE transceiver design with imperfect CSI at both ends. IEEE Trans Signal Process, 2009, 57: 1141–1150
16 Xing C W, Ma S D, Wu Y C, et al. Transceiver design for dual-hop nonregenerative MIMO-OFDM relay systems under channel uncertainties. IEEE Trans Signal Process, 2010, 58: 6325–6339
17 Xing C W, Fei Z S, Ma S D. Maximum mutual information design for amplify-and-forward multi-hop MIMO relaying systems under channel uncertainties. In: Wireless Communications and Networking Conference (WCNC), Paris, 2012. 781–786
18 Xing C W, Ma S D, Fei Z S. A general robust linear transceiver design for multi-hop amplify-and-forward MIMO relaying systems. IEEE Trans Signal Process, 2013, 61: 1196–1209
19 Rey F, Lamarca M, Vazquez G. Robust power allocation algorithms for MIMO OFDM systems with imperfect CSI. IEEE Trans Signal Process, 2005, 53: 1070–1085
20 Xing C W, Li D, Ma S D. Robust transceiver design for MIMO-OFDM systems based on cluster water-filling. IEEE Commun Lett, 2013, 17: 1451–1454
21 Zhang Z S, Zhang W, Tellambura C. Cooperative OFDM channel estimation in the presence of frequency offsets. IEEE Trans Veh Technol, 2009, 58: 3447–3459
22 Kay S M. Fundamentals of Statistical Signal Processing: Estimation Theory. Upper Saddle River: Prentice Hall PTR, 1993
23 Yuen C, Hochwald B M. Achieving near-capacity at low SNR on a multiple-antenna multiple-user channel. IEEE Trans Commun, 2009, 57: 69–74
24 Bertsekas D P, Nedic A, Ozdaglar A E. Convex Analysis and Optimization. Nashua: Athena Scientific, 2003
25 Conejo A J, Castillo E, Mínguez R, et al. Decomposition Techniques in Mathematical Programming. Berlin: Springer, 2006
26 Bertsekas D P. Constrained Optimization and Lagrange Multiplier Methods. Nashua: Athena Scientific, 1996
27 Boyd S, Vandenberghe L. Convex Optimization. Cambridge: Cambridge University Press, 2004
28 Bertsekas D P. Nonlinear Programming. Nashua: Athena Scientific, 1999