RESEARCH

Optimal linear precoding for opportunistic spectrum sharing under arbitrary input distributions assumption

Rui Zhu[1,2,3], Yifei Zhao[2], Yunzhou Li[1,2]*, Jing Wang[1,2] and Hao Hong[1,2]

Abstract

Cognitive radio network with multiple-input multiple-output is an effective method to improve not only spectrum efficiency, but also energy efficiency. In this article, a linear precoding matrix optimization algorithm, named gradient-aided mutual information optimization (GAMIO), is designed to maximize the secondary users' spectrum efficiency. Unlike the previous algorithms which were developed under a specific input assumption, the GAMIO algorithm can work without imposing any input assumption. Furthermore, a framework is also proposed to develop the energy-efficient algorithm which can work with arbitrary spectrum-efficient algorithm. In this way, an energy-efficient algorithm, which can work under arbitrary input assumption, be developed based on the GAMIO algorithm (EEGAMIO). Numerical results indicate that either the GAMIO algorithm or the EEGAMIO algorithm shows the best performance at the present time.

1 Introduction

Fixed spectrum allocation is a traditional spectrum allocation methodology for wireless communication systems. This method causes less spectrum efficiency due to the fact that most of the available spectrum is forbidden for the users except for the licensed user (primary user). Besides the spectrum efficiency, we also need an energy-efficient system in order to protect our environment and avoid the green house effect because, in recent years, the using and production of information and communication technology (ICT) contribute an increasing share to the global green house gas emissions. ICT, especially mobile telecommunication network, shows exponentially increasing energy consumption and will no doubt become a major part of energy consumption in the future. Fortunately, with the development of cognitive radio (CR) [1] network and multiple-input multiple-output (MIMO), people find new opportunities to simultaneously improve

spectrum efficiency and energy efficiency. Besides the traditional radio resource allocation research of CR network which focused on time or frequency domains, MIMO CR network can allocate the resources of space domain and offers the secondary user more freedom degrees to trade off between maximizing the spectrum/energy efficiency and minimizing the interference at the primary user.

Among all of the new technologies for MIMO CR network, linear precoding [2-6] is the most popular research topic. There are several contributions in this field: Haykin [7] gives an overview of the green wireless communication via cognitive dimension but it mainly focuses on the concept and pays less attention to the concrete application. Huang et al. [8] and Xing et al. [9] utilize game theory to maximize the sum of secondary users' utility functions under the interference-power constraint. In [10,11], resource allocation of the secondary users is studied by applying the graph-theoretic models. Zhang and Liang [12] research the sub-space algorithm performance in CR network by singular-value decomposition (SVD) theory. However, all of the aforementioned works rely on the impractical Gaussian input assumption, which often leads to substantial performance degradation in real application. In order to overcome this shortage, a parameterized iterative algorithm [13], which can work

*Correspondence: liyunzhou@tsinghua.edu.cn
[1] Department of Electronic Engineering, Tsinghua University, Beijing, China
[2] Research Institute of Information Technology, State Key Laboratory on Microwave and Digital Communications, Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing, China
Full list of author information is available at the end of the article

under equip-probable discrete input assumption, was proposed. Though the parameterized iterative algorithm cannot directly be used in CR network because of the interference constraint, an algorithm, named branch and bound-aided mutual information optimization (BAMIO), was proposed in [14,15] to optimize the precoding matrix in CR network. However, these algorithms can work only under the equip-probable finite alphabet inputs assumption. Thanks to the contribution of [16-18], a fundamental relation between the mutual information and the minimum mean square error (MMSE) in Gaussian channel was unveiled. Furthermore, an algorithm named mercury water filling (M/WF) [19] was proposed to maximize the mutual information over parallel channel under arbitrary inputs assumption. But M/WF can work only when the channel matrix is diagonal and cannot be used in the CR network. This constraint limited its application. All of the above-mentioned drawbacks motivate the research of this article. Our goal is to address a linear precoding algorithm of spectrum/energy efficiency which can be used in CR network without imposing any input assumption.

The main contributions of this article are summarized as follows: First, this article formulates the spectrum and energy-efficient problem for CR network. Second, in the case of spectrum efficiency, a new algorithm based on the gradient-aided mutual information optimization (GAMIO) is proposed. Different from the previous works, the GAMIO algorithm does not limit any input assumption. It is more realistic than the previous algorithm which developed under the Gaussian or any other specific input assumption. Third, in the case of energy efficiency, this article proposes a framework to develop the energy-efficient algorithm which can work with any spectrum-efficient algorithm. In this way, it is easy to develop an energy-efficient algorithm (EEGAMIO) which can inherit the advantage of GAMIO algorithm and work under arbitrary input assumption.

The remainder of this article is organized as follows: In Section 2, the MIMO CR network system model is proposed. Based on this model, we formulate the spectrum- and energy-efficient optimization problems. Two typical channel capacity expressions are also proposed to examine the algorithm performance in simulation. Linear precoding algorithm is given in Section 3 to solve the proposed optimization problem. The performance evaluation is shown in Section 4. Finally, conclusions are drawn in Section 5.

Notation: Bold face uppercase letters denote matrices, bold face lower case letters denote vectors. The superscript $(.)^{\dagger}$ and $(.)^{*}$ stand for the conjugate transpose operator and the conjugate operator, respectively. The operator **diag**($\mathbf{a}$) denotes a diagonal matrix with elements given by $\mathbf{a}$. Trace($\mathbf{A}$) denotes the trace operation of matrix $\mathbf{A}$. The operator $\mathbf{E}(.)$ denotes the statistical expectation

and $\mathbf{E_a}(.)$ denotes the statistical expectation with respect to $\mathbf{a}$. The operator $\|.\|$ and $|.|$ denote the l_2 norm and the matrices determinant, respectively.

2 System model and problem formulation

As shown in Figure 1, a MIMO CR network with K primary receivers and a single pair of secondary user is considered. All of the primary users and secondary users share the same spectrum for transmission. We consider the scenario where each user is equipped with multiple antennas. The channel state information (CSI) from the secondary transmitter to every receiver is perfectly known at the secondary transmitter. Under such assumptions, the secondary transmitter is able to adapt linear precoding matrix to optimize spectrum/energy efficiency.

In practice, the CSI between primary transmitter and secondary receiver can be obtained at the secondary transmitter by sensing the primary transmitter-emitted signal. On the other hand, the CSI between the secondary transmitter and secondary receiver can be obtained by periodical training when the time-division-duplexing model is employed. Of course, it is difficult to get the perfect CSI for the secondary transmitter. Hence, the achieved spectrum/energy efficiency is an upper bounder.

Based on the model which we mentioned above, we further assume that a Gaussian channel composed N_t transmitted antennas and N_r received antennas is considered. Then, the MIMO signal can be formulated as
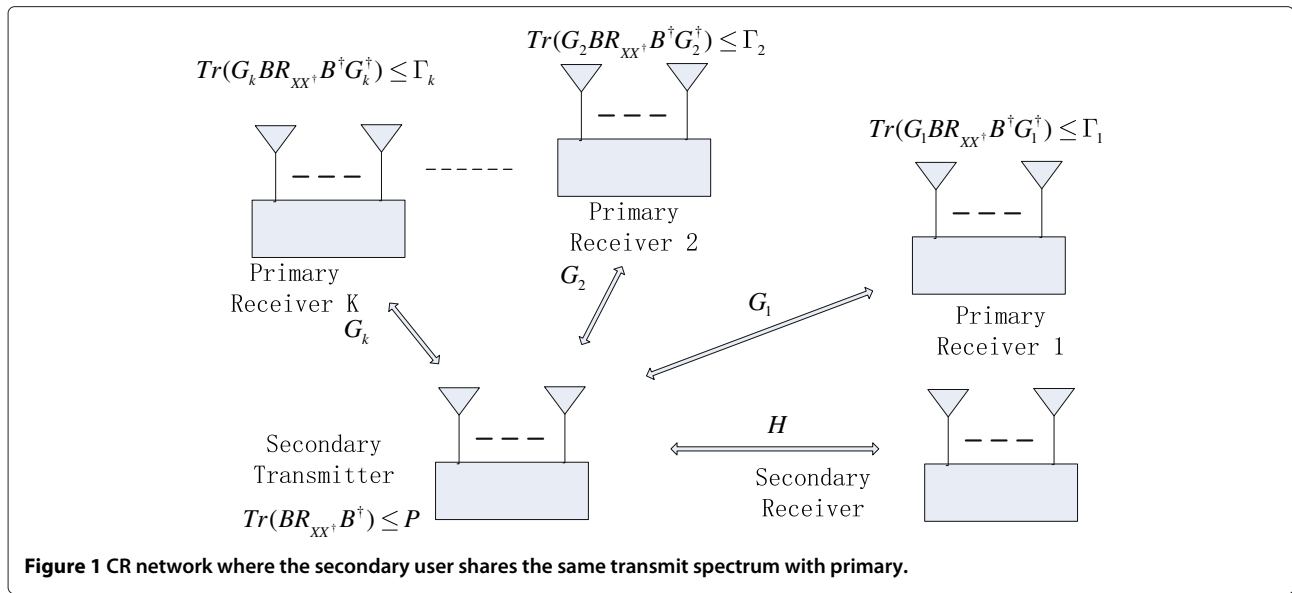
$$\mathbf{y}' = \mathbf{H}'\mathbf{Bx} + \mathbf{z}'. \tag{1}$$

In (1), $\mathbf{y}'$ and $\mathbf{x}$ denote the received and transmitted signal vectors, respectively, and $\mathbf{H}' \in C^{N_r \times N_t}$ denotes the secondary users channel (in this article, we always assume that the channel matrix is full rank). $\mathbf{z}'$ is the Gaussian additive noise vector. It is assumed that $\mathbf{z}' \sim CN(0, \sigma^2)$. $\mathbf{B}$ is the linear precoding matrix which we want to optimize. We denote the transmit covariance matrix of secondary users signal as $\mathbf{R_{xx^{\dagger}}}$, $\mathbf{R_{xx^{\dagger}}} = \mathbf{E}\left[\mathbf{xx}^{\dagger}\right]$, where the expectation $\mathbf{E}[.]$ can be referenced from the codebook. In order to simplify the representation, we normalize $\mathbf{y}'$ and $\mathbf{H}'$ with the noise vector $\mathbf{z}'$. So the equivalent signal can be represented as

$$\mathbf{y} = \mathbf{HBx} + \mathbf{z}. \tag{2}$$

The variables $\mathbf{y}$, $\mathbf{H}$, and $\mathbf{z}$ are $\mathbf{y} = \mathbf{R}_{\mathbf{z}'\mathbf{z}'}^{-0.5}\mathbf{y}'$, $\mathbf{H} = \mathbf{R}_{\mathbf{z}'\mathbf{z}'}^{-0.5}\mathbf{H}'$, and $\mathbf{z} = \mathbf{R}_{\mathbf{z}'\mathbf{z}'}^{-0.5}\mathbf{z}'$, respectively. $\mathbf{R}_{\mathbf{z}'\mathbf{z}'}$ is the covariance matrix of $\mathbf{z}'$. It is clear that $\mathbf{z} \sim CN(0, 1)$. And the covariance matrix of $\mathbf{z}$ is an identity matrix $\mathbf{I}$. The total transmitted power for secondary transmitter is denoted as P_t. So the power constraint must be held as

$$\text{Trace}(\mathbf{BR_{xx^{\dagger}}B^{\dagger}}) \leq P_t. \tag{3}$$

Figure 1 CR network where the secondary user shares the same transmit spectrum with primary.

Just like we mentioned above, we assume that there are K primary receivers in the CR network and each primary receiver is equipped with N_r antennas. For every primary receiver, there must be an interference power constraint over all received antennas. The interference power constraint can be represented as

$$\text{Trace}(\mathbf{G}_k \mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{G}_k^\dagger) \leq \Gamma_k, \ k = 1, \ldots, K. \tag{4}$$

$\mathbf{G}_k$ represents the channel matrix between secondary transmitter and kth primary receiver. From [7], it is clear that the obtained upper bound for the capacity loss of primary user is a function of the interference power Γ_k. So the capacity loss of any primary user due to the secondary transmission can arbitrarily be small by controlling the interference power Γ_k.

Now, it is time to formulate the efficient problem. The first goal is to design the linear precoding matrix of secondary user to maximize its spectrum efficiency under its own transmit power constraint together with a set of interference power constraints at primary users. The problem can be represented as (Pr1)

$$\begin{aligned} \text{maximize:} \quad & C = \mathcal{I}(\mathbf{x}, \mathbf{y}) \\ & \text{Trace}(\mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger) \leq P_t \\ \text{subject } to : \quad & \text{Trace}(\mathbf{G}_k \mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{G}_k^\dagger) \leq \Gamma_k, \\ & k = 1, \ldots, K. \end{aligned} \tag{5}$$

We notice that many researchers usually develop their algorithm based on the Gaussian input assumption even though the input signal is not always a Gaussian signal, because it has a very simple and elegant expression (6) and can be solved effectively by convex optimization algorithm [20,21].

$$\mathcal{I}(\mathbf{x}, \mathbf{y}) = \log_2 \left| \mathbf{I} + \mathbf{H} \mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{H}^\dagger \right| \tag{6}$$

Actually, the channel capacity has different expressions in different situations. For example, when we consider the equip-probable discrete signaling constellations, such as M-ary PSK, PAM, and QAM, the channel capacity is more likely to be represented as [12,20,21]

$$\begin{aligned} \mathcal{I}(X, Y) = N_t \log_2 M - \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \\ \times \mathbf{E} \left\{ \log_2 \sum_{k=1}^{M^{N_t}} e^{-(\|\mathbf{HB}(\mathbf{x}_m - \mathbf{x}_k)\|^2 - \|\mathbf{z}\|^2)} \right\}. \end{aligned} \tag{7}$$

where M is the number of points in the signal constellation. Of course, there are many other channel capacity expressions in practical scenarios (e.g., TCM). Due to the limited space, we do not further enumerate them. It will obviously result in performance loss if we ignore the real input signal distribution. Thus, a spectrum-efficient optimization algorithm which can adapt to various input signals is needed. In Section 4, we will examine GAMIO algorithm under the Gaussian input assumption and 2-symbol equip-probable discrete input assumption. The simulation results show a consistent conclusion.

The second goal is to design the linear precoding matrix of secondary user to maximize its energy efficiency under not only its own transmit power constraint, but also a set of interference power constraints at primary users. The problem can be represented as (Pr2)

$$\begin{aligned} \text{maximize:} \quad & C/P = \mathcal{I}(\mathbf{x}, \mathbf{y}) / \text{Trace}(\mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger) \\ & \text{Trace}(\mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger) \leq P_t \\ \text{subject to} : \quad & \mathcal{I}(\mathbf{x}, \mathbf{y}) \geq C_t \\ & \text{Trace}(\mathbf{G}_k \mathbf{B} \mathbf{R}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{G}_k^\dagger) \leq \Gamma_k, k = 1, \ldots, K. \end{aligned} \tag{8}$$

where C_t represents the minimal channel capacity for the quality of service. The energy-efficient problem has a close relationship with the spectrum-efficient problem. The algorithm for Pr1 can be modified to deal with Pr2 directly.

3 Efficient optimization algorithm under arbitrary input assumption

In this section, we present the linear precoding optimization algorithm. First, we give the gradient of mutual information with respect to precoding matrix **B**.

3.1 Gradient of mutual information with respect to the precoding matrix B

To start with, we give the gradient of mutual information with respect to the linear precoding matrix **B**, which was initially proposed in [17]. Since the channel is AWGN, the conditional output of **y** can be represented as

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}) = \frac{1}{\pi^n} e^{(-\|\mathbf{y}-\mathbf{HBx}\|^2)}. \tag{9}$$

where n represents the dimension of **x**. The mutual information is

$$\mathcal{I}(\mathbf{x};\mathbf{y}) = \mathbf{E}\left[\log\frac{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})}{p_{\mathbf{y}}(\mathbf{y})}\right] = -n\log(\pi e) - \mathbf{E}\left[\log p_{\mathbf{y}}(\mathbf{y})\right]$$

$$= -n\log(\pi e) - \int p_{\mathbf{y}}(\mathbf{y})\log p_{\mathbf{y}}(\mathbf{y})d\mathbf{y}. \tag{10}$$

Then the derivation of (10) is

$$\frac{\partial\mathcal{I}(\mathbf{x};\mathbf{y})}{\partial(\mathbf{HB})^*} = -\int (1+\log p_{\mathbf{y}}(\mathbf{y}))\frac{\partial p_{\mathbf{y}}(\mathbf{y})}{\partial(\mathbf{HB})^*}d\mathbf{y}$$

$$= -\int (1+\log p_{\mathbf{y}}(\mathbf{y}))\mathbf{E}_{\mathbf{x}}\left[\frac{\partial p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})}{\partial(\mathbf{HB})^*}\right]d\mathbf{y}. \tag{11}$$

From (9), we know that the derivation of $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})$ is

$$\frac{\partial p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})}{\partial(\mathbf{HB})^*} = -p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})\frac{\partial}{\partial(\mathbf{HB})^*}((\mathbf{y}-\mathbf{HBx})^{\dagger}(\mathbf{y}-\mathbf{HBx}))$$

$$= p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})(\mathbf{y}-\mathbf{HBx})\mathbf{x}^{\dagger} = -\nabla_{\mathbf{y}}p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})\mathbf{x}^{\dagger}. \tag{12}$$

Therefore, (11) can be rewritten as

$$\frac{\partial\mathcal{I}(\mathbf{x};\mathbf{y})}{\partial(\mathbf{HB})^*} = \int (1+\log p_{\mathbf{y}}(\mathbf{y}))\mathbf{E}_{\mathbf{x}}\left[\nabla_{\mathbf{y}}p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})\mathbf{x}^{\dagger}\right]d\mathbf{y}$$

$$= \mathbf{E}_{\mathbf{x}}\left[(\int (1+\log p_{\mathbf{y}}(\mathbf{y}))\nabla_{\mathbf{y}}p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})d\mathbf{y})\mathbf{x}^{\dagger}\right]. \tag{13}$$

Notice that

$$\int (1+\log_e p_{\mathbf{y}(\mathbf{y})})\frac{\partial p_{\mathbf{y}|\mathbf{x}(y)}}{\partial\mathbf{y}}d\mathbf{y} = \left[(1+\log_e p_{\mathbf{y}(\mathbf{y})})p_{\mathbf{y}|\mathbf{x}(y)}\right]|_{\mathbf{y}=-\infty}^{\mathbf{y}=\infty}$$

$$-\int \frac{1}{p_{\mathbf{y}(\mathbf{y})}}\frac{\partial p_{\mathbf{y}(\mathbf{y})}}{\partial\mathbf{y}}p_{\mathbf{y}|\mathbf{x}(y)}d\mathbf{y}, \tag{14}$$

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})(1+\log p_{\mathbf{y}}(\mathbf{y})) \to 0, \ \|\mathbf{y}\| \to \infty. \tag{15}$$

Then (13) can be represented as

$$\frac{\partial\mathcal{I}(\mathbf{x};\mathbf{y})}{\partial(\mathbf{HB})^*} = \mathbf{E}\left[\left(-\int \frac{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})}{p_{\mathbf{y}}(\mathbf{y})}\nabla_{\mathbf{y}}p_{\mathbf{y}}(\mathbf{y})d\mathbf{y}\right)\mathbf{x}^{\dagger}\right]. \tag{16}$$

We also notice that $p_{\mathbf{x}|\mathbf{y}} << p_{\mathbf{x}}$ for almost every **y**. Based on this condition and $\frac{\partial p_{\mathbf{x}|\mathbf{y}}}{\partial p_{\mathbf{x}}}(\mathbf{x},\mathbf{y}) = \frac{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})}{p_{\mathbf{y}}(\mathbf{y})}$ (Radon-Nikodym derivative), we can rewrite (16) as

$$\frac{\partial\mathcal{I}(\mathbf{x};\mathbf{y})}{\partial(\mathbf{HB})^*} = -\int \nabla_{\mathbf{y}}p_{\mathbf{y}}(\mathbf{y})\mathbf{E}_{\mathbf{x}}\left[\frac{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})}{p_{\mathbf{y}}(\mathbf{y})}\mathbf{x}^{\dagger}\right]d\mathbf{y}$$

$$= -\int \nabla_{\mathbf{y}}p_{\mathbf{y}}(\mathbf{y})\mathbf{E}\left[\mathbf{x}^{\dagger}|y\right]d\mathbf{y}. \tag{17}$$

Now, using

$$\nabla_{\mathbf{y}}p_{\mathbf{y}}(\mathbf{y}) = \nabla_{\mathbf{y}}\mathbf{E}_{\mathbf{x}}\left[p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})\right] = \mathbf{E}_{\mathbf{x}}\left[\nabla_{\mathbf{y}}p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})\right]$$

$$= -\mathbf{E}_{\mathbf{x}}\left[p_{\mathbf{y}|\mathbf{x}}(\mathbf{y})(\mathbf{y}-\mathbf{HBx})\right]$$

$$= -\mathbf{E}\left[p_{\mathbf{y}}(\mathbf{y})(\mathbf{y}-\mathbf{HBx})|\mathbf{y}\right]$$

$$= -p_{\mathbf{y}}(\mathbf{y})(\mathbf{y}-\mathbf{HBE}\left[\mathbf{x}|\mathbf{y}\right]), \tag{18}$$

we can get

$$\frac{\partial\mathcal{I}(\mathbf{x};\mathbf{y})}{\partial(\mathbf{HB})^*} = \int \left[p_{\mathbf{y}}(\mathbf{y})(\mathbf{y}-\mathbf{HBE}\left[\mathbf{x}|\mathbf{y}\right])\right]\mathbf{E}\left[\mathbf{x}^{\dagger}|\mathbf{y}\right]d\mathbf{y}. \tag{19}$$

From the definition of conditional expectation $\mathbf{E}\left[x|y\right] = \int xp(x|y)dx$, we have

$$\int \left[yp_{\mathbf{y}}(\mathbf{y})\right]\mathbf{E}\left[\mathbf{x}^{\dagger}|\mathbf{y}\right]d\mathbf{y} = \int \left[yp_{\mathbf{y}}(\mathbf{y})\right]\int \mathbf{x}^{\dagger}p(\mathbf{x}^{\dagger}|\mathbf{y})d\mathbf{x}^{\dagger}d\mathbf{y}$$

$$= \int\int \mathbf{x}^{\dagger}yp(\mathbf{x}^{\dagger},\mathbf{y})d\mathbf{x}^{\dagger}d\mathbf{y}. \tag{20}$$

Because $\mathbf{y} = \mathbf{HBx} + \mathbf{z}$, (20) can be represented as

$$\iint \mathbf{x}^{\dagger}(\mathbf{HBx}+\mathbf{z})p(\mathbf{x}^{\dagger},\mathbf{y})d\mathbf{x}^{\dagger}d\mathbf{y} = \mathbf{HBE}\left[\mathbf{x}^{\dagger}\mathbf{x}\right]+\mathbf{E}\left[\mathbf{x}^{\dagger}\mathbf{z}\right]$$

$$= \mathbf{HBE}\left[\mathbf{x}^{\dagger}\mathbf{x}\right]. \tag{21}$$

Similarly, we also have

$$\int p_{\mathbf{Y}}(\mathbf{y})\mathbf{HBE}\left[\mathbf{x}|\mathbf{y}\right]\mathbf{E}\left[\mathbf{x}^{\dagger}|\mathbf{y}\right]d\mathbf{y} = \mathbf{HBE}\left[\mathbf{E}\left[\mathbf{x}|\mathbf{y}\right]\mathbf{E}\left[\mathbf{x}^{\dagger}|\mathbf{y}\right]\right]$$

$$= \mathbf{HBE}\left[\mathbf{x}|\mathbf{y}\right]\mathbf{E}\left[\mathbf{x}^{\dagger}|\mathbf{y}\right].$$

(22)

From (19), (21), and (22), we can get

$$\frac{\partial \mathcal{I}(\mathbf{x};\mathbf{y})}{\partial(\mathbf{HB})^{*}} = \mathbf{HB}(\mathbf{E}\left[\mathbf{xx}^{\dagger}\right] - \mathbf{E}[\mathbf{E}[\mathbf{x}|\mathbf{y}]\mathbf{E}[\mathbf{x}^{\dagger}|\mathbf{y}]]). \quad (23)$$

We also notice that

$$\mathbf{E}\left[(\mathbf{x}-\mathbf{E}[\mathbf{x}|\mathbf{y}])(\mathbf{x}-\mathbf{E}[\mathbf{x}|\mathbf{y}])^{\dagger}\right] = \mathbf{E}\left[\mathbf{xx}^{\dagger}\right] - \mathbf{E}[\mathbf{E}[\mathbf{x}|\mathbf{y}]$$

$$\times \mathbf{E}[\mathbf{x}^{\dagger}|\mathbf{y}]].$$

(24)

We denote (24) as **T** which represents the MMSE matrix. From (23), (24), and the chain rule of derivation, we know that

$$\frac{\partial \mathcal{I}(\mathbf{x};\mathbf{y})}{\partial \mathbf{B}} = \mathbf{H}^{\dagger}\mathbf{HBT}. \quad (25)$$

Of course, MMSE matrix have different expressions under different input assumptions. Under the Gaussian input assumption, we have [14]

$$\mathbf{T} = \mathbf{E}\left[(\mathbf{x}-\mathbf{E}[\mathbf{x}|\mathbf{y}])(\mathbf{x}-\mathbf{E}[\mathbf{x}|\mathbf{y}])^{\dagger}\right]$$

$$= \mathbf{R}_{\mathbf{xx}^{\dagger}}\mathbf{H}^{\dagger}(\mathbf{I}+\mathbf{HR}_{\mathbf{xx}^{\dagger}}\mathbf{H}^{\dagger})^{-1}.$$

(26)

If we assume that the signal follows the finite equipprobable discrete distribution, the MMSE matrix is represented as

$$\mathbf{T} = \mathbf{E}\left[(\mathbf{x}-\mathbf{E}[\mathbf{x}|\mathbf{y}])(\mathbf{x}-\mathbf{E}[\mathbf{x}|\mathbf{y}])^{\dagger}\right] = \frac{1}{M^{N_{t}}(2\pi)^{\frac{M}{2}}}$$

$$\int \sum_{i=1}^{M^{N_{t}}}(\mathbf{x}_{i}-\mathbf{E}[\mathbf{x}|\mathbf{y}])(\mathbf{x}_{i}-\mathbf{E}[\mathbf{x}|\mathbf{y}])^{\dagger}e^{-\frac{(\mathbf{y}-\mathbf{x}_{i})(\mathbf{y}-\mathbf{x}_{i})^{\dagger}}{2}}d\mathbf{y}.$$

(27)

3.2 Optimization algorithm of Pr1

In this section, we propose an iterative algorithm to solve Pr1. The linear precoding matrix **B** is updated by the cutting-plane method. The idea is proposed in [22,23]. First, the algorithm localizes a set of candidate **B** in a closed set. Second, half of the closed set will be eliminated from the candidate set by choosing appropriate direction. The iteration will not stop until the candidate set is small enough. Usually, a common choice of the update direction

is the gradient of objective function which we want to optimized. So, from (25), We know it is

$$d_{1} = \frac{\partial \mathcal{I}(\mathbf{x};\mathbf{y})}{\partial \mathbf{B}} = \mathbf{H}^{\dagger}\mathbf{HBT}. \quad (28)$$

Unfortunately, we cannot promise that the whole candidate set satisfies the constraint condition. So, we have to examine them and choose the right direction in order to eliminate the *bad candidates* which does not satisfy the constraint. The feasible update direction is (29) which is the derivation of (3) and (4) [24]. If the constraint (3) or (4) cannot be satisfied, the GAMIO algorithm should take (29) instead of (28) for good candidates which satisfy the constraint condition.

$$\begin{cases} d_{2} = \mathbf{BR}_{\mathbf{xx}^{\dagger}} \\ d_{3,k} = \mathbf{G}_{k}^{\dagger}\mathbf{G}_{k}\mathbf{BR}_{\mathbf{xx}^{\dagger}}, \ k = 1,\dots,K \end{cases} \quad (29)$$

Beside the update direction, how to choose the initial candidate set is also an important element of GAMIO algorithm. A reasonable choice is the minimal size ellipsoid contained the optimal precoding matrix $\mathbf{B}^{*}$. So, we should take the initial set **A** and center $\mathbf{B_0}$ as (30) based on the transmitted power constraint (3).

$$\mathbf{A} = \mathbf{diag}([P_{t},P_{t},\dots,P_{t}]),$$

$$\mathbf{B_0} = \mathbf{diag}\left(\left[P_{t}^{\frac{1}{2}},P_{t}^{\frac{1}{2}},\dots,P_{t}^{\frac{1}{2}}\right]\right).$$

(30)

According to the property of gradient, the GAMIO algorithm will converge to the global optimal value if the Pr1 is a convex/concave problem. Of course, not all of input signal distribution promise Pr1 is a convex/concave problem. The GAMIO algorithm performance will greatly depend on the initial value if the Pr1 is not a convex/concave problem. Usually, it will converge to a local optimal value. By the way, it should be noted that the Gaussian input assumption leads Pr1 to a concave problem. So, the GAMIO algorithm can achieve the global optimal value. It has the same performance as the interior-point algorithm which is considered as the best algorithm under the Gaussian input assumption.

From [25], we know that the GAMIO, being a kind of cutting-plane algorithm, is a polynomial-time algorithm. The iteration complexity is decided by the threshold ϵ_{1} and initial set **A**. Without loss of generality, we assume that the channel capacity is Lipschitz. Then the iteration complexity is $O(2n^{2}\ln(\frac{LR}{\epsilon_{1}}))$. L represent the Lipschitz condition constant. R is decided by the initial set **A** and represents the volume of ellipsoid. The most complex part of GAMIO algorithm is to estimate the MMSE matrix based on the current precoding matrix. In order to reduce the time-consuming operations, we take the trick which employed in [15] and update the MMSE matrix recursively. Let s be the number of operations in the MMSE matrix evaluation. Then the computational complexity of

GAMIO algorithm is $O(2sn^2 \ln(\frac{LR}{\varepsilon}))$. Of course, s is different under different input assumption and can be adjusted according to the algorithm precision. The GAMIO algorithm is concluded as follows.

GAMIO algorithm under arbitrary input assumption

1. Initial the candidate set:

$$\mathbf{A} = \mathbf{diag}([P_t, P_t, \ldots, P_t]),$$

$$\mathbf{B}_0 = \mathbf{diag}([P_t^{\frac{1}{2}}, P_t^{\frac{1}{2}}, \ldots, P_t^{\frac{1}{2}}]),$$

$$\begin{cases} d_1 = \mathbf{H}^\dagger \mathbf{HBT} \\ d_2 = \mathbf{BR}_{\mathbf{xx}^\dagger} \\ d_{3,k} = \mathbf{G}_k^\dagger \mathbf{G}_k \mathbf{BR}_{\mathbf{xx}^\dagger}, \ k = 1, \ldots, K; \end{cases}$$

2. Repeat:

    (a) If $\mathrm{Trace}(\mathbf{BR}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger) > P_t$,

    $$d = d_2.$$

    Else if
    $\mathrm{Trace}(\mathbf{G}_k \mathbf{BR}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{G}_k^\dagger) > \Gamma_k, \ k = 1, \ldots, K,$

    $$d = d_{3,k}, \ k = 1, \ldots, K.$$

    Else

    $$d = -d_1;$$

    (b) Update as

    $$t = (\mathrm{Trace}(d^\dagger \mathbf{A}d))^{\frac{1}{2}}$$
    $$\mathbf{B}_0 = \mathbf{B}_0 - \frac{\mathbf{A}}{N+1}\frac{d}{t}$$
    $$\mathbf{A} = \frac{N^2}{N^2-1}(\mathbf{A} - \frac{2}{N+1}\mathbf{A}\frac{d}{t}(\frac{d}{t})^T\mathbf{A}^T);$$

3. Until $t < \varepsilon_1$, end.

3.3 Optimization algorithm of Pr2

Just as we mentioned before, Pr2 closely coupled with Pr1. In general, Pr2 is not a concave problem. So, we cannot solve it by traditional optimization method directly. Thus, we try to find a sub-optimal solution instead. We notice that the Pr2 decrease as its denominator value increases, while it will increase with the raise of its numerator value. So, we can minimize the denominator value and maximize the numerator value separately to optimize Pr2. It can be done by combining arbitrary spectrum efficient optimization algorithm with bi-section method. So, the Pr2 can be separated into two parts. The first part can be represented as (Pr3)

$$\begin{aligned} \text{maximize:} \quad & C = \mathcal{I}(\mathbf{x}, \mathbf{y}) \\ & \mathrm{Trace}(\mathbf{BR}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger) \leq P_i^* \\ \text{subject to} \ : \quad & \mathrm{Trace}(\mathbf{G}_k \mathbf{BR}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{G}_k^\dagger) \leq \Gamma_k \\ & k = 1, \ldots, K \end{aligned} \quad . \quad (31)$$

The only difference between Pr3 and Pr1 is the constraint power P_i^* which represents current power con-

straint during ith iteration. Obviously, $P_t \geq P_i^*$. If we have $\mathcal{I}(\mathbf{x}^*, \mathbf{y}^*) \geq C_t$, then

$$P_{\mathrm{Max}} = P_i^*, \ P_{i+1}^* = \frac{P_{\mathrm{Min}} + P_{\mathrm{Max}}}{2}. \quad (32)$$

Otherwise, if $I(\mathbf{x}^*, \mathbf{y}^*) < C_t$, then

$$P_{\mathrm{Min}} = P_i^*, \ P_{i+1}^* = \frac{P_{\mathrm{Min}} + P_{\mathrm{Max}}}{2}. \quad (33)$$

where $\mathcal{I}(\mathbf{x}^*, \mathbf{y}^*)$ represents the optimal value of Pr3. The initial values of P_{Min} and P_{Max} are 0 and P_t, respectively. The iteration will not stop until P_{Min} and P_{Max} are close enough. The computational complexity of this new iterative approach is tightly related with the spectrum-efficient algorithm for Pr3. Without any doubt, although arbitrary spectrum-efficient algorithm can work in this frame to optimize the energy efficiency, the energy-efficient algorithm which is developed based on the GAMIO algorithm (EEGAMIO) is a better choice than any other algorithm to deal with various input signal. From [23], we know that the iteration complexity of bi-section method is

$$\log_2 \frac{P_{\mathrm{Max}} - P_{\mathrm{Min}}}{\varepsilon_2} = \log_2 \frac{P_t}{\varepsilon_2} \quad (34)$$

where ε_2 represents the threshold of bi-section method. Based on the conclusion of Section 3.2, we can get the iteration and computational complexity of EEGAMIO energy-efficient algorithm are $O(2n^2 \ln(\frac{LR}{\varepsilon_1})\log_2 \frac{P_t}{\varepsilon_2})$ and $O(2sn^2 \ln(\frac{LR}{\varepsilon_1})\log_2 \frac{P_t}{\varepsilon_2})$ respectively. Table 1 summarize the complexity of the proposed algorithms and the framework of the energy-efficient algorithm is concluded as follows.

The framework of energy-efficient algorithm

1. Initial the searching set:

$$P_{\mathrm{min}} = 0, \ P_{\mathrm{max}} = P_t, \ P_0^* = \frac{P_{\mathrm{min}} + P_{\mathrm{max}}}{2};$$

2. Repeat:

    (a) Solve the problem

    $$\begin{aligned} \text{max} : \quad & \mathcal{I}(\mathbf{x}, \mathbf{y}) \\ & \mathrm{Trace}(\mathbf{BR}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger) \leq P_i^* \\ s.t : \quad & \mathrm{Trace}(\mathbf{G}_k \mathbf{BR}_{\mathbf{xx}^\dagger} \mathbf{B}^\dagger \mathbf{G}_k^\dagger) \leq \Gamma_k, \ k = 1 \ldots K; \end{aligned}$$

Table 1 The complexity of the proposed algorithms

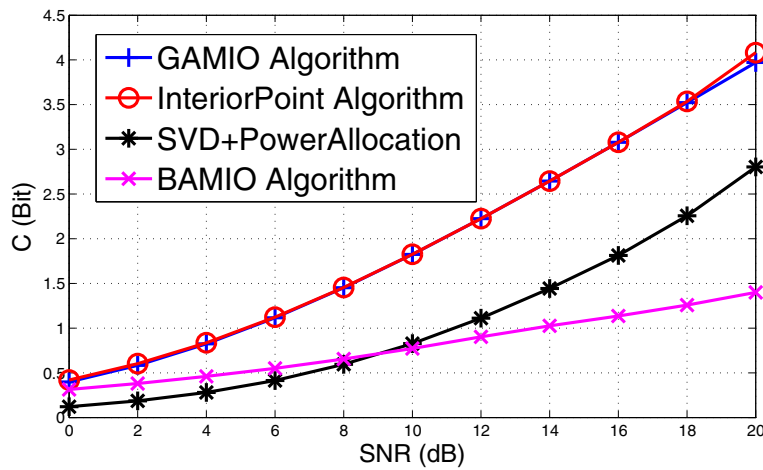| Case	Iteration complexity	Computational complexity
Spectrum efficiency	$O(2n^2 \ln(LR/\varepsilon_1))$	$O(2sn^2 \ln(LR/\varepsilon_1))$
Energy efficiency	$O(2n^2 \ln(LR/\varepsilon_1)$ $\times \log_2(P_t/\varepsilon_2))$	$O(2sn^2 \ln(LR/\varepsilon_1)$ $\times \log_2(P_t/\varepsilon_2))$

Figure 2 Channel capacity versus SNR under the Gaussian input assumption.

(b) If $\mathcal{I}(\mathbf{x}^*, \mathbf{y}^*) \geq C_t$,

$$P_{\max} = P_i^*, \quad P_{i+1}^* = \frac{P_{\min} + P_{\max}}{2};$$

Else if $\mathcal{I}(\mathbf{x}^*, \mathbf{y}^*) < C_t$,

$$P_{\min} = P_i^*, \quad P_{i+1}^* = \frac{P_{\min} + P_{\max}}{2}.$$

3. Until $P_{\max} - P_{\min} \leq \varepsilon_2$, end.

4   Simulation results

In this section, we demonstrate some simulation results. The simulation is separated into two parts. For simplicity, we assume that there is only one primary user in the CR network.

4.1   Spectrum efficiency

In first part, we test the spectrum efficiency of GAMIO algorithm. We take (6) and (7) as the channel capacity for the Gaussian input assumption and the finite equip-probable discrete input assumption, respectively. The channel matrix from secondary transmitter to primary receiver is $G = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$ and the channel matrix between secondary user is $H = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. The average signal-noise-radio (SNR) of CR MIMO system is given by SNR $= \frac{P_t}{\sigma^2}$. The total transmit power P_t is 2 W. The interference-power constraint Γ is 1 W.

In Figures 2 and 3, we test the achievable channel capacity of secondary user under different situations. In Figure 2, we assume that the input signal follows the
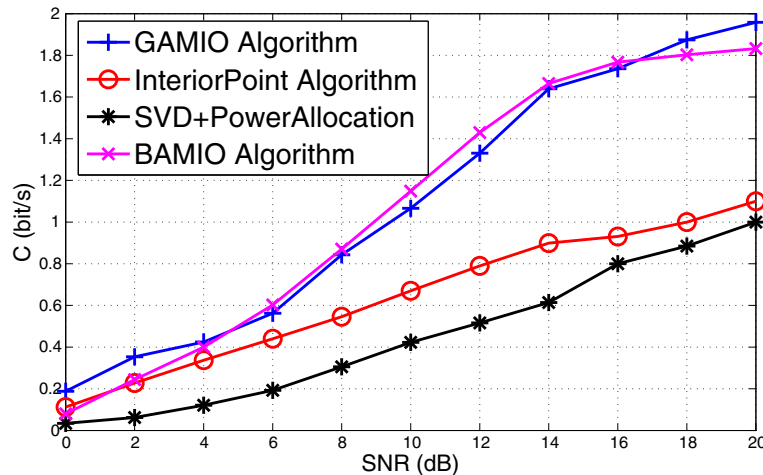
Figure 3 Channel capacity versus SNR under the 2-symbol equip-probable discrete input assumption.
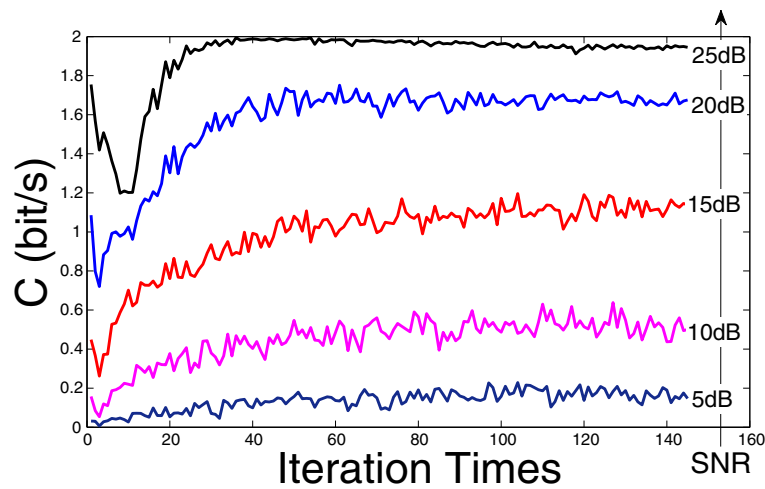
Figure 4 Channel capacity versus iteration times under the 2-symbol equip-probable discrete input assumption.

Gaussian distribution. So, the channel capacity can be expressed as (6). The line denoted with circle signs represents the performance of interior-point algorithm. It has the best performance. The line with '+' represents the performance of GAMIO algorithm. It has the same performance as interior-point algorithm because the channel capacity (6) is a concave problem under the Gaussian input assumption. The line with '*' represents the performance of sub-space algorithm which is a well-known suboptimal algorithm [12]. The line with '×' represents the performance of BAMIO algorithm and much lower than the others because the BAMIO algorithm is developed under equip-probable discrete input assumption.

In Figure 3, we assume that the input signal follows the 2-symbol equip-probable discrete distribution (e.g., BPSK). So the channel capacity can be expression as (7) with $M = 2$. Just as what Figure 2 shows, the first line with '+' represents the performance of GAMIO algorithm and the second line with '×' represents the performance of BAMIO algorithm. They almost have the same performance. Similarly, the third line denoted with circle signs represents the performance of interior-point algorithm and the fourth line with '*' represents the performance of sub-space algorithm. Figure 3 shows that the performance of GAMIO and BAMIO algorithm is much better than interior-point algorithm and sub-space algorithm. It almost reaches the best capacity, 2 bit/s, when the SNR is 20 dB. On the other hand, the capacity of interior-point and sub-space algorithm lower than 1 bit/s because they intend to allocate too much energy to the best channel and degrade the channel capacity.

Figures 2 and 3 show that, unlike the traditional spectrum-efficient algorithm which developed under specific input assumption, the GAMIO algorithm considers the specific input signal distribution and achieves the

best performance under various input assumption. This result is consistent with our analysis and previous research results [8,10].

Figure 4 gives the convergence performance results of GAMIO algorithm with different SNR. As shown in Figure 4, the algorithm converges to the optimal value before the iteration ends. On the other hand, we can also find that the convergence rate improves with SNR increasing.

In order to unveil the meaning of different channel capacity expression, we present a BPSK MIMO system as Figure 5 shows. Figure 6 gives the BER of this system with different SNR, Figure 6 indicates that the BER of GAMIO and BAMIO algorithms is obviously lower than the interior-point algorithm and SVD algorithm. This result is consistent with Figure 3. And Figure 2 cannot reflect this property. Figures 2, 3, and 6 show that the input signal distribution affect the *real* performance
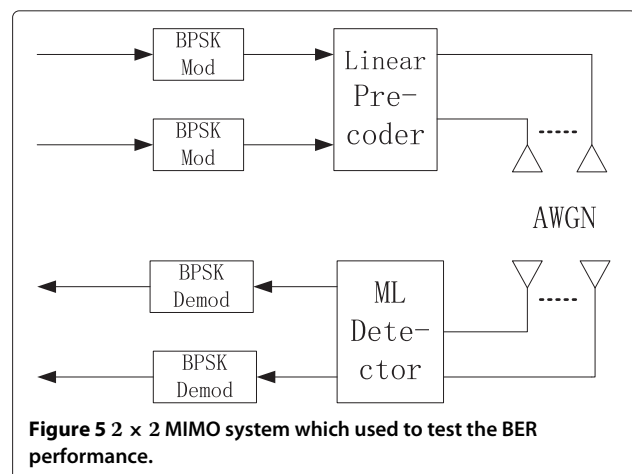

Figure 5 2×2 MIMO system which used to test the BER performance.
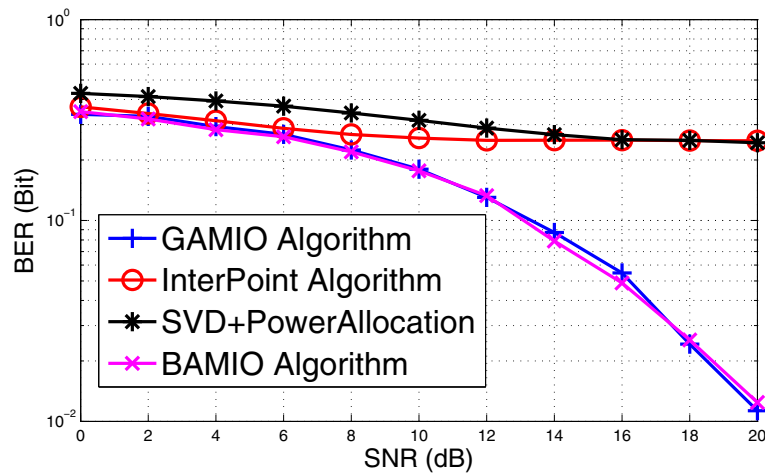
Figure 6 BER versus SNR when we use the precoding matrix which proposed by SVD+PowerAllocation, interior-Point, and GAMIO algorithm, respectively.

greatly in practical scenarios. So, the GAMIO algorithm is more useful than any other spectrum-efficient algorithm, which can work only under specific input assumption, for various practical system design.

4.2 Energy efficiency

In the second part, we simulate the energy efficiency problem which is defined as Pr2. Just like the first part,

we also take into account two kinds of channel capacity expression. The first one is based on the Gaussian input assumption and the other one based on 2-symbol finite equip-probable discrete input assumption respectively. All of the assumptions are the same as the first part except adding a minimum channel capacity constraint, $C_t = 1$ bit/s.

In Figure 7, we test the achievable energy efficiency under different situations. Under the Gaussian input
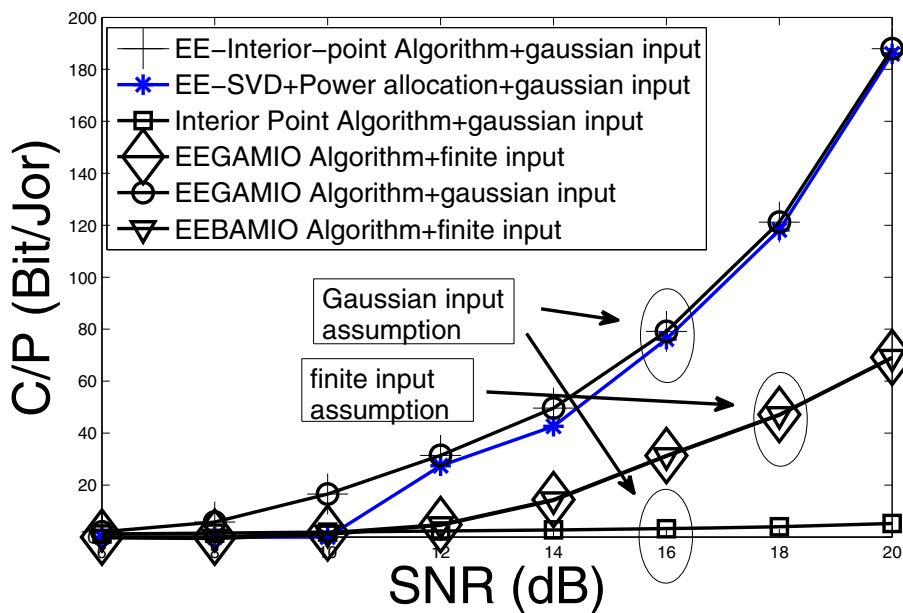
Figure 7 Different energy-efficient algorithm performance versus SNR under the Gaussian input assumption and 2-symbol equip-probable discrete input assumption, respectively.
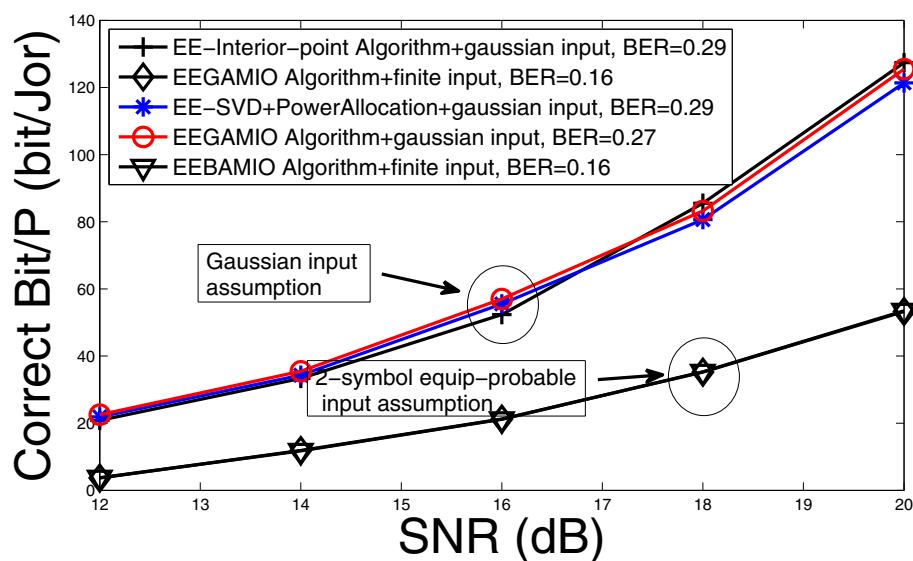
Figure 8 Energy efficiency versus SNR in the MIMO system as Figure 5.

assumption, we develop energy-efficient algorithm based on the interior-point algorithm (EE-interior-point), subspace algorithm (EE-SVD), and GAMIO algorithm (EEGAMIO), respectively. The results are noted with plus, star, and circle, respectively. As Figure 7 shown, the performances of EE-interior-point algorithm and EEGA-MIO algorithm have no significant difference under Gaussian input assumption. EE-SVD algorithm cannot satisfy the capacity constraint when the SNR is less than 12 dB. But it is almost the same as the other algorithms when SNR is beyond 12 dB. Similarly, we also develop energy-efficient algorithm based on the BAMIO algorithm and GAMIO algorithm under the 2-symbol equip-probable discrete input assumption. The performance line is annotated by diamond and inverted-triangle, respectively. There are no performance results for EE-interior-point algorithm and EE-SVD algorithm because they cannot satisfy the capacity constraint under the 2-symbol equip-probable discrete input assumption. Figure 7 shows that, under either Gaussian or 2-symbol equip-probable discrete input assumption, the EEGAMIO algorithm can reach the best performance.

The last line denoted with square sign represents the energy efficiency of traditional interior-point algorithm under the Gaussian input assumption. It is much lower than all of the others. This result shows that the proposed framework of energy-efficiency is effective.

Figure 8 shows the energy efficiency of MIMO system which Figure 5 figured. The lines which denoted with plus, star, circle, diamond, and inverted-triangle sign are same as Figure 7. As Figure 8 shown, although the Gaussian input assumption can keep high energy efficiency, it brings more loss in the BER performance. So the Gaussian input

assumption is more difficult to satisfy the channel capacity constraint than the 2-symbol equip-probable discrete input assumption in the MIMO system as Figure 5. It also reveal that the input signal distribution affect the energy efficiency greatly.

5 Conclusion

In this article, linear precoding matrix optimization for the secondary user in a CR network is considered. We set both spectrum and energy efficiency as our goals. Unlike previous work, we do not limit the input signal following specific distribution. Instead, we propose a spectrum-efficient algorithm (GAMIO), which can work well under not only the Gaussian input assumption but also any other kind of input assumption. With this advantage, the GAMIO algorithm is more practical than existing algorithms, such as interior-point algorithm, SVD algorithm, or BAMIO algorithm. In the case of energy efficiency, we propose a framework to optimize the energy efficiency. With this framework, we can get different energy-efficient algorithms based on different spectrum-efficient algorithms. In order to deal with various signal, the EEGAMIO algorithm is a smart choice. The simulation results show that these algorithms can significantly increase the energy efficiency.

Author details
[1]Department of Electronic Engineering, Tsinghua University, Beijing, China. [2]Research Institute of Information Technology, State Key Laboratory on Microwave and Digital Communications, Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing, China. [3]The Telecommunication Engineering Institute, Air Force Engineering University, Shaanxi, China.

References
1. J Mitola, GQ Maguire, Cognitive radios: making software radios more personal. Personal Commun. **6**(4), 13–18 (1999)
2. A Scaglione, P Staica, S Barbarossa, Optimal designs for space-time linear precoders and decoders. IEEE Trans. Signal Process. **50**(5), 1051–1064 (2002)
3. H Sampath, A Paulraj, Linear precoding for space-time coded systems with known fading correlations. IEEE Commun. Lett. **6**(6), 234–241 (2004)
4. S Zhou, GB Giannakis, Optimal transmitter eigen beamforming and space-time block coding based on channel correlations. IEEE Trans. Commun. **50**(5), 1673–1690 (2003)
5. JP Kermoal, A stachastic MIMO radio channel model with expermental validation. IEEE J. Sel. Areas Commun. **20**(6), 1211—1226 (2002)
6. MHM Costa, Writing on dirty paper. IEEE Trans. Inf. Theory. **29**(3), 439–441 (1983)
7. S Haykin, Cognitive radio: brain-empowered wireless communications. J. Sel. Areas Commun. **23**(2), 201–220 (2005)
8. J Huang, R Berry, ML Honig, Auction-based spectrum sharing. ACM/Springer Mob. Netw. Appl. J. (MONET). **11**(3), 405–418 (2006)
9. Y Xing, CN Mathur, MA Haleem, R Chandramouli, KP Subbalakshmi, Dynamic spectrum access with QoS and interference temperature constraints. Trans. Mob. Comput. **6**(4), 423–433 (2007)
10. H Zheng, C Peng, in *Proceedings of the IEEE International Conference on Communication*, vol. 5. Collaboration and fairness in opportunistic spectrum access (IEEE Press, USA, 2005), pp. 3132–3136
11. W Wang, X Liu, in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 1. List-coloring based channel allocation for open-spectrum wireless networks (IEEE Press, USA, 2005), pp. 690–694
12. R Zhang, YC Liang, Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks. IEEE J. Sel. Topic Signal Process. **2**, 88–102 (2008)
13. C Xiao, YR Zheng, Z Ding, Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels. IEEE Trans. Signal Process. **59**(7), 3301–3314 (2011)
14. W Zeng, C Xiao, J Lu, Globally optimal precoder design with finite-alphabet input for cognitive radio networks. IEEE J. Sel. Topics Commun. **30**(10), 1861–1874 (2012)
15. W Zeng, C Xiao, M Wang, J Lu, Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI. IEEE Trans. Signal Process. **60**(6), 3134–3148 (2012)
16. D Guo, SS Shitz, S Verdu, information, Mutual minimum mean-square error in Gaussian channels. IEEE Trans. Inf. Theory. **51**(4), 1261–1282 (2005)
17. DP Palomar, S Verdu, Gradient of mutual information in linear vector Gaussian channels. IEEE Trans. Inf. Theory. **52**, 141–154 (2006)
18. SM Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory (Prentice-Hall)*. (Englewood Cliffs, NJ, 1993)
19. A Lozano, AM Tulino, S Verdu, Optimum power allocation for parallel Gaussian channels with arbitrary input distributions. IEEE Trans. Inf. Theory. **52**(7), 3033–3051 (2006)
20. TE Duncan, On the calculation of mutual information. SIAM. J. Appl. Math. **19**, 215–220 (1970)
21. G Jengren, M Skoglund, B Ottersen, Combining beamforming and orthogonal space-time block coding. IEEE Trans. Inf. Theory. **48**(3), 611–617 (2002)
22. TM Cover, JA Thomas, *Elements of Information Theory*. (Wiley, New York, 1991)
23. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press Cambridge, UK, 2004)
24. JR Magnus, H Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. (Wiley, New York, 1999)
25. S Boyd, EE392o Course Notes Stanford University (2004). http://www.stanford.edu/class/ee392o/elp.pdf