Thèse de Doctorat

Lei ZHAO

Mémoire présenté en vue de l'obtention
du grade de Docteur de l'Université de Nantes
Sous le label de l'Université Nantes Angers Le Mans

Discipline : Electronique
Spécialité : Communications numériques
Laboratoire : IETR UMR 6164

Soutenance le 1er octobre 2014

École doctorale Sciences et Technologies de l'Information et Mathématiques (STIM)

BEAMFORMING AND POWER ALLOCATION TECHNIQUES FOR MULTIUSER MIMO BROADCAST CHANNELS

JURY

Président :	**M. Salah BOURENNANE**, Professeur, Institut Fresnel, Ecole Centrale Marseille
Rapporteurs :	**M. Gilles BUREL**, Professeur, Université de Bretagne Occidentale, Brest
	M. Jean-Pierre CANCES, Professeur, Université de Limoges
Examinateurs :	**M. Guillaume ANDRIEUX**, Maître de Conférences, IUT La Roche s/Yon
Directeurs de Thèse :	**M. Pascal CHARGE**, Professeur, Ecole polytechnique de l'université de Nantes
	M. Yide WANG, Professeur, Ecole polytechnique de l'université de Nantes

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Chargé and co-supervisor Prof. Wang, for their professional guidance and unconditional support throughout the entire three years of my PhD study. I thank them for the countless hours they spent discussing my research work, giving professional suggestions, and correcting the manuscripts. I also thank them for giving me a down-to-earth attitude and spirit in my future career.

Of course, I have to thank my friends and every colleague in the lab, for making these three years in France so enjoyable.

In addition, I would also like to thank China Scholarship Council (CSC) for the financial support.

I also want to express my gratitude to my parents, for their selfless support, I cannot imagine coming so far without their love and understanding.

Finally, I thank my girlfriend Han Zheng for her sweet and thoughtful love, which has substantially supported me for my PhD study.

Résumé de la thèse en français

Pour augmenter les débits de transmission, de nouvelles techniques permettant d'améliorer l'efficacité spectrale des systèmes de télécommunication sont proposées. Parmi les techniques prometteuses, Multiple-Input Multiple-Output (MIMO) est très étudiée [1], [2]. La technique MIMO permet en effet de transmettre plusieurs symboles de données différents de manière simultanée et sur la même bande de fréquence. Cette technique a d'abord été étudiée pour des liaisons point-à-point, c'est-à-dire entre un émetteur et un récepteur, chacun équipé de plusieurs antennes. Les travaux [3], [4] et [5] indiquent notamment que l'efficacité spectrale obtenue pour des systèmes équipés de plusieurs antennes peut s'avérer importante lorsque le canal offre beaucoup de diversité. Dans [6] et [7] les auteurs soulignent que pour un canal à bruit additif gaussien indépendant et identiquement distribué, la capacité des systèmes MIMO peut croître linéairement avec le nombre d'antennes de transmission ou de réception.

Depuis quelques années, une extension de MIMO nommée multi-utilisateurs MIMO (MU-MIMO) attire de plus en plus d'attention par rapport au système MIMO classique. MU-MIMO permet en effet d'établir des communications simultanées avec plusieurs utilisateurs répartis dans l'espace, chacun équipé d'antennes multiples. Ce système permet donc de servir simultanément plusieurs utilisateurs avec une grande efficacité spectrale en exploitant la diversité spatiale. Ce service est possible au prix d'un traitement du signal plus intensif [8], [9]. Le schéma de base est celui d'une station de base en lien avec plusieurs utilisateurs simultanément dans la même bande de fréquence, et qui exploite les différentes signatures spatiales induites par la dispersion géographique des utilisateurs. Cette technique est également connue sous le nom de Space-Division Multiple Access (SDMA) [10]. Il est montré qu'en appliquant la technique DPC (Dirty-Paper-Coding), le système MU-MIMO a la capacité d'annuler les interférences connues non causalement au niveau de l'émetteur. A l'aide de techniques de formation de voies, il est possible de supprimer les autres interférences. La formation de voies a été utilisée dans les systèmes MIMO [11] de manière optionnelle pour améliorer le rapport signal-sur-bruit (RSB) au niveau du récepteur [12]. Afin de profiter pleinement de l'amélioration des performances apportée par le multiplexage spatial dans les systèmes MU-MIMO, les techniques de formation de voies sont essentielles pour éliminer ou minimiser les interférences multi-utilisateurs (MUI).

Le chapitre 2 introduit les canaux MIMO, MU-MIMO, et les canaux d'accès multiple (MAC). La liaison descendante, de la station de base avec des antennes multiples vers plusieurs utilisateurs, chacun étant équipé d'une ou de plusieurs antennes, est définie comme un canal de diffusion MU-MIMO. L'objet de l'étude est de déterminer la formation de voies optimale en termes de capacité de ce système. S'agissant ici d'un problème d'optimisation non convexe et non-concave [13], la dualité entre le canal descendant MU-MIMO de diffusion et le canal MAC de liaison montante est exploitée. Cependant, malgré l'exploitation de cette propriété, la complexité reste élevée pour résoudre le problème. Afin de réduire cette complexité, des méthodes sous-optimales telles que ZF-DPC [14], SA-DPC [15] et SBD-DPC [16] ont été proposées. Pour ce type de méthodes qualifiées de non linéaires, les signaux à transmettre sont généralement encodés d'une manière séquentielle. On distingue alors à chaque étape, une partie de l'interférence, connue au niveau de l'émetteur, et étant causée par les signaux précédemment encodés. Cette interférence peut être éliminée simplement par la technique DPC. Ceci offre plus de degrés de liberté pour la détermination des vecteurs de formation de voies à l'émission, puisqu'il reste à annuler uniquement la partie restante de l'interférence. Cette stratégie conduit généralement à des performances importantes en termes de débit pour ces méthodes non linéaires. La méthode ZF-DPC ne permet de considérer qu'une seule antenne de réception par utilisateur, tandis que la méthode SBD-DPC permet de considérer plusieurs antennes de réception par utilisateur. La méthode SA-DPC proposée dans [15] détermine la formation de voies en émission et en réception pour les flux de données d'une de manière séquentielle. Cette méthode fonctionne même si le nombre total d'antennes en réception est supérieur au nombre d'antennes à l'émission. Un aperçu de ces techniques d'optimisation non-linéaires nécessitant l'algorithme DPC, est donné dans le chapitre 3.

Bien que la capacité globale fournie par ces dernières méthodes s'approche de la capacité optimale, elles nécessitent la mise en œuvre de l'algorithme DPC qui augmente la complexité. Par conséquent, des solutions de traitement linéaires, telles que la méthode ZF [17], la méthode BD [18], la méthode RBD [4], la méthode CB [19] et la méthode ZF-SA [20] ont été proposées. Ces méthodes permettent aussi d'annuler complètement l'interférence par la formation de voies. Les méthodes ZF et BD consistent, pour chaque utilisateur, à trouver un vecteur de formation de voies à l'émission orthogonal à l'espace formé par les autres utilisateurs. Les méthodes ZF-SA et CB permettent de considérer les cas où le nombre total d'antennes en réception est plus grand que le nombre d'antennes à l'émission. La condition de ZF est relaxée dans la méthode MMSE [21], et les vecteurs de formation voies à l'émission sont calculés par les valeurs propres généralisées. Toutes ces techniques sont décrites dans le chapitre 3. La mise en œuvre de ces méthodes reste raisonnable en termes de complexité. Cependant pour toutes ces techniques, l'optimisation se fait sous une contrainte de puissance cumulée.

Dans la pratique, chaque antenne émettrice possède un amplificateur de puissance dont la linéarité est nécessairement limitée; en particulier dans le cas du système OFDM qui présente de forts PAPR [22]. Ainsi, il est plus réaliste de chercher à optimiser le débit avec une contrainte de puissance non plus totale mais relative à chaque antenne d'émission. Des techniques de formation voies avec une contrainte de puissance par antenne ont été étudiées dans [23], [24], [25], [26] et [27]. Les travaux [23] et [24] se placent dans le cas où chaque utilisateur est équipé d'une seule antenne et considèrent un pré-codage de ZF. Dans [25] cette contrainte ZF est relaxée. Dans [26] les auteurs analysent le cas où chaque utilisateur est équipé de plusieurs antennes, et l'algorithme DPC est utilisé en considérant un ordre d'attribution prédéfini des utilisateurs. Cette méthode [26] attribue à chaque utilisateur un certain nombre de flux de données en fonction du rang de sa matrice de canal. Ce dernier point n'est cependant pas optimal car les canaux offrant un débit faible ou même négligeable pour un utilisateur pourraient introduire néanmoins de fortes contraintes pour les autres utilisateurs [15]. Dans [27], la solution optimale sous contrainte de puissance par antenne est obtenue par l'exploitation de la dualité des canaux montants et descendants. Les techniques mises en œuvre pour résoudre ce problème d'optimisation convergent malheureusement assez lentement [25]. L'état de l'art des techniques de formation de voies sous la contrainte de puissance par antenne dans les canaux de diffusion MU-MIMO est également donné dans le chapitre 3. Dans la première partie du chapitre 4, nous proposons une approche alternative à la méthode SA-DPC sous la contrainte de la puissance totale. La manière de déterminer les vecteurs de formation de voies à l'émission dans la méthode SA-DPC, ne permet pas d'introduire facilement la contrainte de puissance par antenne. Dans la méthode proposée, le vecteur de la formation de voies à l'émission est obtenu en déterminant d'abord le sous-espace vectoriel auquel il doit appartenir, puis en cherchant dans cet espace le vecteur qui permet d'optimiser le débit global. La méthode proposée se comporte de manière identique à la méthode SA-DPC lorsque la contrainte de la puissance totale est imposée. Par contre, le procédé proposé peut être facilement modifié pour prendre en compte la contrainte plus réaliste de puissance par antenne. Dans la deuxième partie du chapitre 4, une méthode de formation de voies sous la contrainte de puissance par antenne est proposée. Puisque la solution optimale du problème initial est difficile à obtenir, le problème est scindé en deux sous-problèmes classiques, l'un est un problème d'optimisation SDP (semidefinite-programming), et l'autre se résout par la technique MRC (maximal-ratio-combining). La comparaison avec les méthodes de la littérature montre le bénéfice apporté par cette nouvelle méthode. De plus, la méthode proposée fonctionne même si le nombre total d'antennes de réception est plus grand que le nombre d'antennes d'émission.

Lorsque le RSB est faible, l'interférence apparaît comme négligeable par rapport au bruit gaussien additif. Dans ce cas, il n'est donc pas indispensable de vouloir supprimer l'interférence complètement. Ainsi la contrainte d'annulation de l'interférence peut être relaxée, mais le problème de l'optimisation de l'allocation de puissance s'avère alors être un problème NP [21]. Dans la littérature, plusieurs approches sous-optimales qui permettent d'optimiser conjointement les vecteurs de formation de voies à l'émission et l'allocation de puissance ont été développées. Dans [28], les auteurs ont analysé la situation où chaque utilisateur dispose d'une seule antenne de réception. Cette technique a été généralisée dans [21] où chaque utilisateur est équipé de plusieurs antennes de réception pour assurer plusieurs flux de données. Cependant, la méthode d'allocation de puissance proposée dans [21] se trouve être un problème d'optimisation GP (geometric-programming) itératif, qui présente une complexité de calcul assez élevée. Dans [29] l'optimisation porte sur le rapport SINR (signal-to-interference-plus-noise ratio) moyen de chaque utilisateur. Cependant ce critère n'est pas toujours optimal, car la dégradation du TEB (taux d'erreurs binaires) apparaît principalement lorsque le SINR d'un sous-canal est faible même si le SINR moyen de l'utilisateur est élevé. Dans [30], la relation est établie entre le SINR par utilisateur et les débits pondérés (weighted sum rate) dans le cas d'une seule antenne de réception par utilisateur. L'optimisation du débit est aussi étudiée pour le cas multicellulaire MIMO dans [31] et [32], et le résultat optimal est trouvé au prix d'une complexité de calcul exponentielle. Dans [33], [34] et [35], des solutions optimales locales ont été proposées avec une complexité raisonnable.

Motivée par [21], au chapitre 5, une nouvelle méthode d'allocation de puissance dans le contexte MU-MIMO est proposée. Nous adoptons la technique de formation de voies MMSE [21], qui est une stratégie efficace pour la résolution d'un tel problème d'optimisation [29], [36]. En outre, contrairement à la technique d'allocation de puissance GP dans [21], la méthode d'allocation de puissance proposée attribue la puissance d'émission totale de manière itérative selon le principe du water-filling, attribuant ainsi plus de puissance aux canaux ayant les plus forts gains. Cette stratégie réduit considérablement la complexité de calcul par rapport à la méthode GP. En outre, la méthode proposée attribue la puissance d'émission totale de manière itérative, en prenant en compte à chaque itération la puissance allouée lors des itérations précédentes. Cette technique permet d'atteindre des débits proches de la capacité du canal. L'algorithme proposé est intéressant sur le plan de sa mise en œuvre en pratique. Les résultats numériques permettent de valider la technique proposée.

Le chapitre 6 donne la conclusion et quelques perspectives à ces travaux. L'objectif des méthodes étudiées est l'augmentation du débit global, mais il serait important de prendre en considération le RSB à chaque sous-canal. En effet, un faible RSB peut entrainer un fort TEB, qui peut même s'avérer non envisageable en pratique.

Dans ce mémoire, afin de maximiser le débit global, certains utilisateurs pour lesquels le sous-canal n'est pas favorable, sont négligés. Or bien souvent en pratique, le système doit garantir un débit minimum à chaque utilisateur. Il est donc nécessaire aussi de se pencher sur ce problème de qualité de service minimum pour chaque utilisateur.

Parmi les méthodes que nous avons proposées, certaines nécessitent la mise en œuvre de l'algorithme DPC, ce qui entraîne bien souvent une complexité importante des équipements émetteurs et récepteurs. Les recherches futures pourraient porter sur la réduction de cette complexité.

Les canaux non sélectifs en fréquence sont considérés dans cette thèse, il serait intéressant d'étudier le cas où les canaux sont sélectifs en fréquence.

Dans cette étude, nous avons supposé que les canaux sont parfaitement connus. La sensibilité des méthodes proposées par rapport à la connaissance imparfaite des canaux mérite d'être étudiée.

Contents

List of Figures

List of Acronyms

MIMO Multiple-Input Multiple-Output

BC Broadcast Channels

MAC Multiple Access Channels

Non-linear beamforming techniques

- **ZF-DPC** Zero Forcing-Dirty Paper Coding
- **SZF-DPC** Successive Zero Forcing-Dirty Paper Coding
- **SA-DPC** Successive Allocation-Dirty Paper Coding

Linear beamforming techniques

- **ZF** Zero Forcing
- **BD** Block Diagonalization
- **CB** Coordinated Beamforming
- **ZF-SA** Zero Forcing-Successive Allocation
- **MMSE** Minimum Mean Square Error

Beamforming techniques under per-antenna power constraint

- **Per-OPT** Per-Optimal
- **PBD-DPC** Per Block Diagonalization-Dirty Paper Coding

1

Introduction

1.1 Background

With the emerging of the fourth generation (4G) wireless systems, commercial wireless communications such as mobile multimedia, mobile online game, and high quality video etc. come into our daily life. Meanwhile, the requirement for radio spectrum also increases strongly with this fast development of wireless communication industry and business. The radio spectrum resource being limited, it becomes more and more expensive [37]. In order to meet the requirement of extremely high data rates, a large number of new techniques that can improve the spectrum efficiency and data rates are being studied, and tremendous research efforts are also undertaken to develop advanced coding, modulation, signal processing and multiple-access schemes for improving the quality and spectral efficiency of wireless links (e.g. FDMA, CDMA, OFDM, etc.). Multiple-Input Multiple-Output (MIMO) which transmits several different data symbols at the same time and on the same frequency, is one key technique among them because of its ability to enhance the channel capacity of cellular systems at no extra cost of spectrum [1], [2]. MIMO technique is first investigated in point to point scenario, that is, the transmitter equipped with multiple transmit antennas and the receiver equipped with multiple receive antennas. The work in [3], [4] and [5] predicts that remarkable spectral efficiencies for wireless systems with multiple antennas can be obtained when the channel exhibits rich scattering. The work in [6] and [7] points

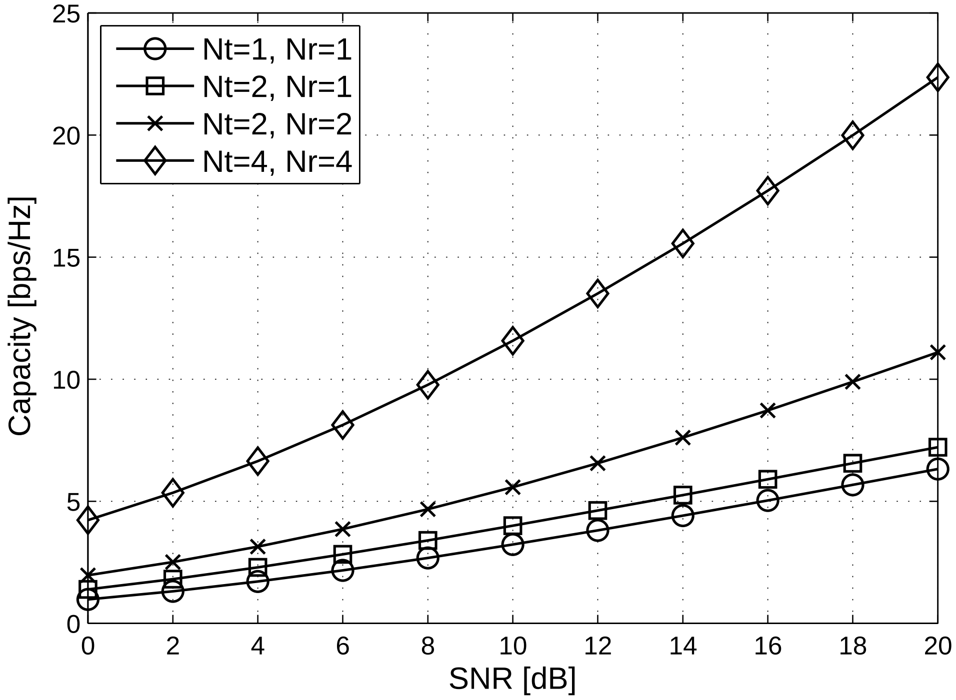

Figure 1.1: Illustration of capacities with different antennas configurations.

out that for an independent and identically distributed Gaussian noise channel, the capacity of MIMO systems can grow linearly with the number of transmit or receive antennas. Figure 1.1 shows the different capacities with different antennas configurations over fading channel. When SNR is 10dB, the capacity for single transmit and receive antenna system is 3 bps/Hz, approximately. A two transmit antennas and one receive antenna system would achieve 4 bps/Hz. A four transmit antennas and four receive antennas system can reach 12 bps/Hz. In addition, the existing 802.11n [38] and 802.16e [39] standards also employ MIMO systems.

The benefits offered by MIMO systems are built on two underlying gains (i.e., spatial diversity and spatial multiplexing), which come with the increased cost of radio frequency hardware. Compared with the conventional Single-Input Single-Output (SISO) systems, MIMO systems have more degrees of freedom regarding the signal transmission. Generally, there are three major transmission models in MIMO system: Diversity [40], Multiplexing [41], and Diversity mixed with Multiplexing [42].

1, MIMO Diversity

Wireless channels severely suffer from fading phenomena, which causes unreliability in data decoding. Fundamentally, the spatial diversity scheme sends multiple copies of the signal through multiple transmit antennas, so that the probability that all the signal components fade simultaneously is reduced. Therefore, the reliability of the data reception is enhanced and improved [43].

Receive diversity can be used in Single-Input Multiple-Output (SIMO) channels. The receive antennas receive the signal with independent fading. Then the receiver combines these signals so that the resulted signal exhibits considerably reduced fading [44]. The receive diversity order is characterized by the number of independently fading branches, and the maximum receive diversity order is equal to the number of receive antennas in SIMO channels. The transmit diversity is applicable to Multiple-Input Single-Output (MISO) channels [8], [11]. The transmit diversity order corresponds to the number of independently fading paths that a symbol passes through. Therefore, the maximum transmit diversity order of MISO system is equal to the number of transmit antennas. For a general MIMO system with N_r receive antennas and N_t transmit antennas, the maximum diversity order that can be achieved is

$$D = N_r \times N_t \tag{1.1}$$

where the channel between each transmit-receive antenna pair is assumed to fade independently.

2, MIMO Multiplexing

In spatial multiplexing, a high rate signal is split into multiple lower rate streams and each stream is transmitted from a different transmit antenna in the same frequency channel. [4] has shown that in the high SNR region, the capacity of a channel with independent and identically distributed (i.i.d.) Rayleigh fading between each transmit-receive antenna pair is given by

$$C(SNR) = \min\{N_r, N_t\} \log(SNR) + O(1) \tag{1.2}$$

where SNR is the signal to noise ratio. The spatial multiplexing transmission offers a linear increase with the number of receive or transmit antennas in the transmission rate for the same bandwidth and with no additional power expenditure [10]. Compared with the spatial diversity transmission, the spatial multiplexing transmission aims to maximize the system capacity. One typical spatial multiplexing transmission model is Bell Laboratory Layered Space-Time (BLAST) system [45]. The maximum multiplexing gain of BLAST system is

$$r = \min\{N_r, N_t\} \tag{1.3}$$

The spatial multiplexing configuration can also be applied in a multiuser system.

Recently, an extension of MIMO named Multiuser MIMO (MU-MIMO) system attracts more attention compared with MIMO system since spatially distributed users with multiple antennas can be served at the same time by using spatial diversity, at the cost of some more signal processing [8], [9]. The base station communicates with the multiple users simultaneously in the same frequency channel by exploiting differences in spatial signatures induced by spatially dispersed users, this technique is also known as space-division multiple access (SDMA) [10]. It is shown that similar capacity scaling to MIMO systems can be achieved by dirty paper coding (DPC) technique in MU-MIMO systems. Some advantage of MU-MIMO systems can be obtained with the aid of beamforming techniques. By beamforming we mean all methods applied at the transmitter that facilitate detection at the receiver [46]. Although beamforming is not a new concept and has been used in MIMO systems as well [11], it is optional and used only to improve the SNR at the receiver [12]. However, in order to fully exploit the increased performance of spatial multiplexing in MU-MIMO systems, beamforming techniques are essential to eliminate or minimize multiuser interference (MUI). Note that normally, beamforming techniques are performed with the help of the known downlink channel state information (CSI) [1] at the base station.

The downlink, from the base station with multiple antennas to multiple users with one or more antennas per user, is denoted as MU-MIMO broadcast channels. Considering the optimal beamforming technique in terms of capacity for this system, duality between the downlink MU-MIMO broadcast channels and the uplink multiple access channels (MAC), where several transmitter send different symbols to one common receiver, has to be used. The optimal beamforming vectors are then obtained with an iterative and numerically complex process. To reduce the computational complexity, some suboptimal near capacity methods are proposed. For example, zero forcing beamforming, coordinated beamforming, and DPC combined with user scheduling and zero forcing beamforming methods. Currently, there are still some challenges and problems for MU-MIMO beamforming techniques [48].

1. The assumption that full CSI available at the transmit side is valid in Time Division Duplex (TDD) systems because the uplink and downlink share the same frequency band. For Frequency Division Duplex (FDD) systems, however, the CSI needs to be estimated at the receiver and fed back to the transmitter. With beamforming techniques employed at the transmit side, the required computational effort for each receiver can be reduced, and eventually the receiver structure can be simplified [47].

1.2 Motivation and methodology

It is proven that DPC can achieve the sum capacity of the MU-MIMO broadcast channels [49], [13]. However, optimizing the transmit covariance matrix in DPC directly is difficult because it is a non-concave optimization problem [13]. To avoid the complex processing, suboptimal solutions such as ZF-DPC method [14], SA-DPC method [15] and SBD-DPC method [16] are proposed, which divide the interference into two parts, one part is removed by DPC and the other part is suppressed by beamforming techniques. ZF-DPC method only supports single receive antenna per user, and SBD-DPC method extends it to the multiple receive antenna per user case. SA-DPC method proposed in [15] finds the transmit beamforming and receive combining vectors of one data stream at each step for the user who can bring the largest throughput increase, the total number of receive antennas is further extended and can be larger than the number of transmit antennas. We denote these DPC involved beamforming techniques as non-linear methods. Even if the sum rates provided by these suboptimal non-linear methods are close to the sum capacity, obviously, the implementation of DPC increases the complexity of the transmitter and receiver design.

Therefore, linear processing solutions, such as ZF method [17], BD method [18], RBD method [4], CB method [19] and ZF-SA method [20] that eliminate the interference completely by beamforming technique are considered. ZF and BD methods search for the transmit beamforming vectors of each user in the null space of the space spanned by other users. ZF-SA and CB methods adopt the receive combining techniques and extend the total number of receive antennas to be larger than that of transmit antennas. MMSE method proposed in [21] relaxes the zero-forcing condition, the transmit beamforming vectors are found by generalized eigenvalue technique. It can be seen that the sum rates offered by these methods are close to the sum capacity, and they are easy to implement, but these transmit beamforming vectors are designed under the assumption of a total power constraint.

In practice, the power amplifier of each antenna is limited individually by its linearity. Especially in an orthogonal frequency division multiplexing (OFDM) system, the peak-to-average power ratio (PAPR) is high [22]. Thus, a power constraint imposed on each transmit antenna is more realistic. Per-antenna power constraint beamforming techniques are studied in [23], [24], [25], [26] and [27]. [23] and [24] investigate the scenario where each user is equipped with a single antenna under the constraint of zero-forcing precoding. In [25], this constraint is relaxed and DPC technique is used to further improve the performance. In [26], authors analyze the case where each user is equipped with multiple antennas via block diagonalization, and DPC is used under the assumption of a preset user order. However, the method in [26] assigns to a certain user a number of data streams depending on the rank of the relative channel matrix. This is suboptimal if, for instance, some of the subchannels are weak. In that case, the contribution of these subchannels to the sum rate might be negligible while they may impose severe constraints on the subchannels of subsequent users [15]. In [27], the optimal solution under per-antenna power constraint is exploited through the duality of MU-MIMO broadcast channels and corresponding uplink multiple access channels. Lagrange duality method and ellipsoid method are used to find the optimal value, which converge unfortunately quite slowly [25].

In this thesis, we propose a successive allocation of data streams to users. Motivated by [15], one data stream is assigned at each step, the corresponding transmit beamforming and receive combining vectors are designed to maximize the global throughput. Moreover, a more practical per-antenna power constraint is imposed to the transmit antennas compared with [15], in which only a total transmit power constraint is considered. It is shown that the optimal solution in the original problem is difficult to obtain. In the proposed method, this problem is first divided into two classical optimization problems, which can be solved with existing standard algorithms. Then, we alternatively solve each subproblem under the assumption that the other is fixed similarly to [21]. The convergence can be achieved within a small number of iterations. At each step, the data stream is allocated to the user who brings the largest increase of the global throughput. The non-causally known interference is pre-subtracted through DPC technique before transmission, and the remaining interference is eliminated by the transmit beamforming and receive combining vectors. Note that the receive combining technique being adopted in the proposed method, the number of total receive antennas can be larger than that of transmit antennas.

In this thesis, we also propose an efficient power allocation method in multiuser MIMO broadcast channels. Since the original problem is non-deterministic polynomial-time (NP) hard when the interference is not removed completely, the optimal solution has extremely high computational complexity. Inspired by the classical water-filling algorithm, which assigns more power to the subchannels with large channel gains, we iteratively use water-filling algorithm to perform the power allocation. Simulation results show that the performance is close to the optimal value, and the complexity is substantially reduced.

1.3 Contributions

The main contributions of this thesis are summarized as follows.

- An alternative approach to SA-DPC method is proposed. SA-DPC method calculates the transmit beamforming vector directly. In the proposed method, we first find the subspace where the transmit beamforming vector should lie in, then the one that maximizes the global throughput is selected. It is shown that the proposed method can be easily adapted to the scenario where the per-antenna power constraint is imposed.
- A new greedy data stream allocation method in multiuser MIMO broadcast channels under the per-antenna power constraint is proposed. Since the spatial diversity in multiuser MIMO broadcast channels is fully exploited, compared with PBD-DPC method in [26], a better sum rate performance is achieved by the proposed method. In the proposed method, receive combining technique is adopted, and the data streams are assigned to users successively. The number of total receive antennas may be larger than that of transmit antennas.
- An efficient power allocation method is proposed. Compared with the optimal solution, the proposed method has low computational complexity, and the performance is close to the optimum value.

1.4 Publications

- L. Zhao, Y. Wang, and P. Chargé. Zero-Forcing DPC Beamforming Design for Multiuser MIMO Broadcast Channels. Submitted to Signal Processing. Under review.
- L. Zhao, Y. Wang, and P. Chargé. Low Complexity Power Allocation in MU-MIMO Broadcast Channels. Submitted to International Journal of Electronics and Communications. Under review.
- L. Zhao, Y. Wang, and P. Chargé. Efficient Power Allocation Strategy in Multiuser MIMO Broadcast Channels. PIMRC 2013.
- L. Zhao, Y. Wang, and P. Chargé. Efficient Iterative Water-filling Power Allocation Method in MU-MIMO Broadcast Channels. MCC 2013.
- L. Zhao, Y. Wang, and P. Chargé. Joint Beamforming Design and Power Allocation for Multiuser MIMO Broadcast Channels. SIFWICT 2013.
- L. Zhao, P. Chargé, and Y. Wang. A Novel Zero-Forcing Transmit Data Scheme for Multiuser MIMO Broadcast Channels. SIFWICT 2013.

1.5 Outline of the thesis

The remainder of this thesis is organized into five chapters as listed below.

Chapter 2 introduces the structures of MIMO channel, MU-MIMO channel, and multiple access channels (MAC) channel. Brief capacity calculations of these channels are also given. MIMO channel is decomposed into several parallel SISO channels, and the well-known water-filling algorithm is used to perform the power allocation. For MU-MIMO channel, DPC technique is used to help to find the channel capacity.

Chapter 3 overviews the state of the art of beamforming techniques in the literature. Firstly, the non-linear methods taking advantage of DPC technique are discussed, then the linear beamforming methods are presented. After that, we consider the practical issues and address beamforming techniques under per-antenna power constraint. At last, the performance of these methods in terms of global throughput is compared, and we also give the advantages and disadvantages of each method.

Chapter 4 addresses the proposed beamforming methods. The first part introduces the proposed beamforming method under total power constraint, and the second part presents the proposed beamforming method under per-antenna power constraint. We also show the simulation results of each method, and give the performance comparisons with other methods.

Chapter 5 introduces the proposed power allocation method when the interference is not completely removed. The original problem is a NP hard problem, and the optimal solution has very high computational complexity. Motivated by the classical water-filling algorithm, we propose a suboptimal method with a very low complexity. The performance is also quite close to the optimal value.

Chapter 6 gives the conclusions of this thesis and the possible directions in the future works.

2

Channel model

A signal propagating through a wireless channel arrives at the destination along a number of different paths, collectively referred to as multipath. These paths arise from scattering, reflection and diffraction of the radiated energy by objects in the environment or refraction in the medium. The different propagation mechanisms influence path loss and fading models differently.

The signal power changes due to three effects : mean propagation (path) lose, macroscopic fading and microscopic fading. The mean propagation loss in macrocellular environment comes from inverse square law power loss, absorption by water and foliage and the effect of ground reflection. Mean propagation loss is range dependent. Macroscopic fading results from a blocking effect by buildings and natural features and is also known as long term fading or shadowing. Microscopic fading results from the constructive and destructive combination of multipaths and is also known as short term fading or fast fading. Multipath propagation results in the spreading of signal in different dimensions. There are delay spread, Doppler (or) frequency spread (Time-varying multipath channel) and angle spread. These spreads have significant effects on the signal. Mean path loss, macroscopic fading, microscopic fading, delay spread, Doppler spread and angle spread are the main channel effects. The details have been covered by a number of excellent papers and books [50], [51], [52], [53], [54]. They are beyond the scope of this dissertation. Our goal here is to optimize the capacity by using some well investigated propagation models.

2.1 MIMO channel model

We assume a complex baseband representation for the signal and channel unless otherwise specified. Consider a MIMO system with N_t transmit antennas and N_r receive antennas (Figure 2.1). The MIMO channel is given by the $N_r \times N_t$ matrix $\boldsymbol{H}(\tau, t)$ with

$$\boldsymbol{H}(\tau,t) = \begin{bmatrix} h_{1,1}(\tau,t) & h_{1,2}(\tau,t) & \cdots & h_{1,N_t}(\tau,t) \\ h_{2,1}(\tau,t) & h_{2,2}(\tau,t) & \cdots & h_{2,N_t}(\tau,t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r,1}(\tau,t) & h_{N_r,2}(\tau,t) & \cdots & h_{N_r,N_t}(\tau,t) \end{bmatrix} \tag{2.1}$$

where $h_{i,j}(\tau, t)$ is function of time, delay and amplitude gain between the jth transmit antenna and the ith receive antenna. In [10], the elements of $\boldsymbol{H}$ are shown as independent zero mean circularly symmetric complex Gaussian random variables (Rayleigh random variables), with suitable choices of the scatterer location, antenna element patterns, and scattering model. Some properties of $\boldsymbol{H}$ are summarized below:

$$\mathbb{E}\{h_{i,j}(\tau,t)\} = 0 \tag{2.2}$$

$$\mathbb{E}\{|h_{i,j}(\tau,t)|^2\} = 1 \tag{2.3}$$

$$\mathbb{E}\{h_{i,j}(\tau,t)h_{m,n}(\tau,t)^*\} = 0 \quad \text{if } i \neq m \text{ or } j \neq n \tag{2.4}$$

If the transmitted signal vector is $\boldsymbol{s}(t) \in \mathbb{C}^{N_t \times 1}$, then the received signal vector is obtained as

$$\boldsymbol{y}(t) = \boldsymbol{H}(\tau,t)\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{2.5}$$

where $\boldsymbol{n}(t) \in \mathbb{C}^{N_t \times 1}$ is the Gaussian noise with independent and identically distributed (i.i.d.) entries of zero mean and variance σ^2. In the flat fading channel, since the output at any instant of time is independent of inputs at previous times, the received signal can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n} \tag{2.6}$$

2.2 MIMO capacity

We focus on the MIMO capacity in the frequency flat channel, the capacity in the frequency selectivity channel is not in the scope of this dissertation. First the channel

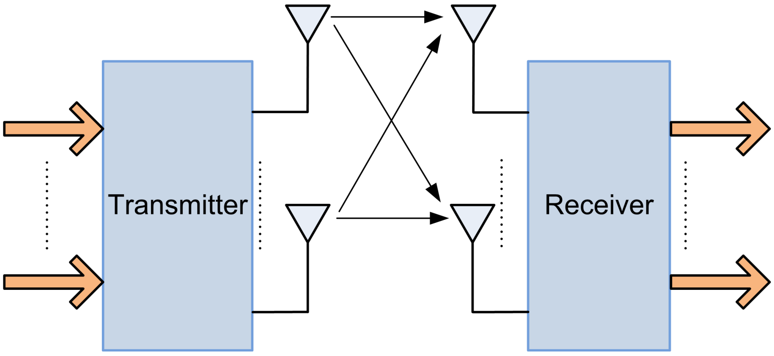

Figure 2.1: Diagram of MIMO system.

$\boldsymbol{H}$ is supposed to be known at both the transmitter and receiver, the capacity of the MIMO channel is defined as [10]

$$C = \max_{f(\boldsymbol{s})} I(\boldsymbol{s}; \boldsymbol{y}) \tag{2.7}$$

where $f(\boldsymbol{s})$ is the probability distribution of the vector $\boldsymbol{s}$, and $I(\boldsymbol{s}; \boldsymbol{y})$ is the mutual information between vector $\boldsymbol{s}$ and $\boldsymbol{y}$. Note that

$$I(\boldsymbol{s}; \boldsymbol{y}) = H(\boldsymbol{y}) - H(\boldsymbol{y}|\boldsymbol{s}) \tag{2.8}$$

where $H(\boldsymbol{y})$ is the differential entropy of the vector $\boldsymbol{y}$, and $H(\boldsymbol{y}|\boldsymbol{s})$ is the conditional differential entropy of the vector $\boldsymbol{y}$, given the knowledge of the vector $\boldsymbol{s}$. Since the vector $\boldsymbol{s}$ and $\boldsymbol{n}$ are independent. i.e.,

$$H(\boldsymbol{y}|\boldsymbol{s}) = H(\boldsymbol{s} + \boldsymbol{n}|\boldsymbol{s}) = H(\boldsymbol{n}) \tag{2.9}$$

Then we have

$$I(\boldsymbol{s}; \boldsymbol{y}) = H(\boldsymbol{y}) - H(\boldsymbol{n}) \tag{2.10}$$

Maximizing the mutual information $I(\boldsymbol{s}; \boldsymbol{y})$ reduces to maximizing $H(\boldsymbol{y})$. Note that the covariance matrix of $\boldsymbol{y}$, $\boldsymbol{R}_{\boldsymbol{yy}} = \mathbb{E}[\boldsymbol{y}\boldsymbol{y}^H]$ satisfies

$$\boldsymbol{R}_{\boldsymbol{yy}} = \boldsymbol{H}\boldsymbol{R}_{\boldsymbol{ss}}\boldsymbol{H}^H + \sigma^2 \boldsymbol{I}_{N_r} \tag{2.11}$$

where $\boldsymbol{R}_{\boldsymbol{ss}}$ is the covariance matrix of $\boldsymbol{s}$. We know that amongst all vectors $\boldsymbol{y}$ with a given covariance matrix $\boldsymbol{R}_{\boldsymbol{yy}}$, the differential entropy $H(\boldsymbol{y})$ is maximized when $\boldsymbol{y}$ is a zero mean circularly symmetric complex Gaussian vector [55]. This in turn implies that $\boldsymbol{s}$ must be a zero mean circularly symmetric complex Gaussian vector, and its distribution is completely characterized by $\boldsymbol{R_{ss}}$. The differential entropies of vector $\boldsymbol{y}$ and $\boldsymbol{n}$ are given by [10]

$$H(\boldsymbol{y}) = \log_2|\pi e \boldsymbol{R_{yy}}| \quad \text{bps/Hz} \tag{2.12}$$

$$H(\boldsymbol{n}) = \log_2|\pi e \sigma^2 \boldsymbol{I}_{N_r}| \quad \text{bps/Hz} \tag{2.13}$$

therefore, $I(\boldsymbol{s};\boldsymbol{y})$ reduces to [5]

$$I(\boldsymbol{s};\boldsymbol{y}) = \log_2|\boldsymbol{I}_{N_r} + \frac{1}{\sigma^2}\boldsymbol{H}\boldsymbol{R_{ss}}\boldsymbol{H}^H| \tag{2.14}$$

and it follows from (2.7) that the capacity of the MIMO channel is given by

$$C = \max_{\text{trace}(\boldsymbol{R_{ss}})=P_T} \log_2|\boldsymbol{I}_{N_r} + \frac{1}{\sigma^2}\boldsymbol{H}\boldsymbol{R_{ss}}\boldsymbol{H}^H| \tag{2.15}$$

where P_T is the total transmit power. The capacity is often referred to as the error-free spectral efficiency or the data rate per unit bandwidth that can be sustained reliably over the MIMO link.

In the above section, the properties of $\boldsymbol{H}$ are presented. Now, we study the capacity of a MIMO channel taking advantage of the properties of $\boldsymbol{H}$. We assume that the CSI is perfectly known to both the receiver and the transmitter. The transmitter can benefit from this information in order to improve the MIMO channel capacity under the constraint of a fixed transmission power P_T. By using the knowledge of the CSI, the transmit power can be allocated in an optimal way on the transmit antennas. The idea behind this method, called water-filling, is to distribute more power to strong channels and less power to weak channels.

Consider a MIMO channel $\boldsymbol{H}$ with rank of r, through which a normalized signal vector $\boldsymbol{x}$ of dimension r is transmitted. Before transmission, the signal vector $\boldsymbol{x}$ is multiplied by the allocated power and the transmit beamforming matrix, i.e.,

$$\boldsymbol{s} = \boldsymbol{V}'\sqrt{\boldsymbol{P}}\boldsymbol{x} \tag{2.16}$$

where the unitary matrix $\boldsymbol{V}'$ is obtained from the singular value decomposition (SVD) of $\boldsymbol{H}$ (i.e., $\boldsymbol{H} = \boldsymbol{U}'\Sigma\boldsymbol{V}'^H$). $\sqrt{\boldsymbol{P}}$ is a diagonal matrix indicating the allocated power for the signal vector $\boldsymbol{x}$. At the receiver, the received signal vector $\boldsymbol{y}$ is multiplied by

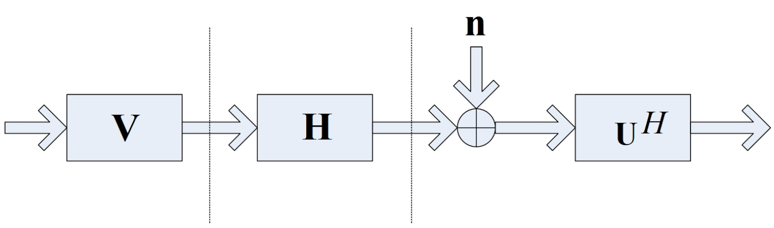

Figure 2.2: Diagram of MIMO transmission.

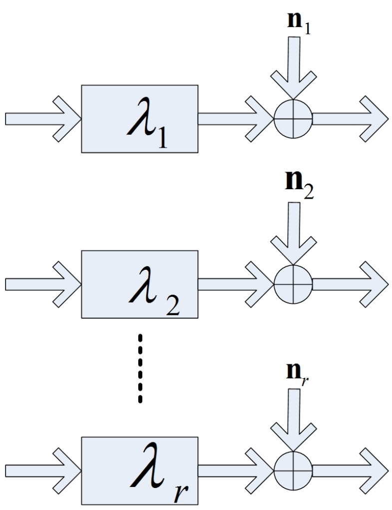

Figure 2.3: Diagram of the decomposed parallel SISO channels.

the matrix $\boldsymbol{U}'^H$. The effective input-output relation for this system is given by

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{U}'^H \boldsymbol{H} \boldsymbol{V}' \sqrt{\boldsymbol{P}} \boldsymbol{x} + \boldsymbol{U}'^H \boldsymbol{n} \\ &= \boldsymbol{\Sigma} \sqrt{\boldsymbol{P}} \boldsymbol{x} + \tilde{\boldsymbol{n}} \end{aligned} \tag{2.17}$$

where $\tilde{\boldsymbol{n}}$ is the $r \times 1$ transformed noise vector with covariance matrix $\mathbb{E}\{\tilde{\boldsymbol{n}}\tilde{\boldsymbol{n}}^H\} = \sigma^2 \boldsymbol{I}_r$. Notice that $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values of the channel matrix $\boldsymbol{H}$. (2.17) shows that $\boldsymbol{H}$ can be explicitly decomposed (see Fig 2.3) into r parallel Single Input Single Output (SISO) channels satisfying

$$y_i = \lambda_i \sqrt{p_i} x_i + \tilde{n}_i, \quad i = 1, 2, \cdots, r. \tag{2.18}$$

The capacity of the MIMO channel is the sum of individual parallel SISO channel capacities and is given by

$$C = \sum_{i=1}^{r} \log_2(1 + \frac{\lambda_i^2 p_i}{\sigma^2}) \tag{2.19}$$

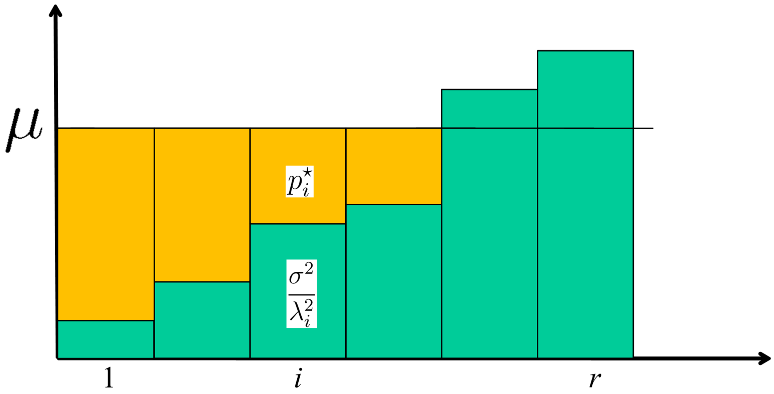

Figure 2.4: Scheme of the Water-filling algorithm.

p_i reflects the transmit power in the ith subchannel and satisfies $\sum p_i = P_T$.

Since the transmitter can access the spatial subchannels, it can allocate variable power across the sub-channels to maximize the mutual information. The mutual information maximization problem now becomes

$$C = \max_{\sum p_i = P_T} \sum_{i=1}^{r} \log_2(1 + \frac{\lambda_i^2 p_i}{\sigma^2}) \tag{2.20}$$

The objective function for the maximization is concave with respect to the variables p_i $(i = 1, \cdots, r)$ and can be maximized using Lagrangian method. The optimal power allocation policy $p_i^\star$, satisfies

$$p_i^\star = (\mu - \frac{\sigma^2}{\lambda_i^2})_+, \quad i = 1, \cdots, r \tag{2.21}$$

$$\sum_{i=1}^{r} p_i^\star = P_T. \tag{2.22}$$

where μ is a constant determined by the total transmission power and $(a)_+$ implies

$$(a)_+ = \begin{cases} a & \text{if} \quad a \geq 0 \\ 0 & \text{if} \quad a < 0 \end{cases} \tag{2.23}$$

This optimal power allocation solution is often referred as water-filling algorithm [56], which is pictorially described as Figure 2.4.

If the channel has no preferred direction and is completely unknown to the transmitter, the vector $\boldsymbol{s}$ may be chosen to be statistically non-preferential, i.e. $\boldsymbol{R}_{\boldsymbol{ss}} = \frac{P_T}{N_t}\boldsymbol{I}_{N_t}$. This implies that the signals are independent and equi-powered at the transmit anten-

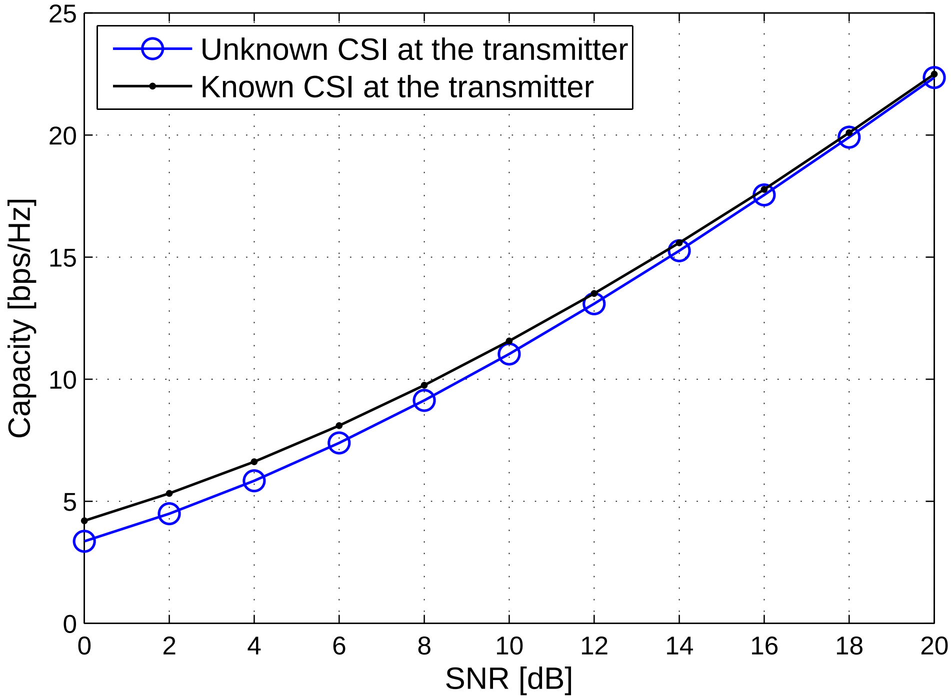

Figure 2.5: Comparison between the capacities of unknown CSI and known CSI at the transmitter, $N_t = 4$, $N_r = 4$.

nas. The capacity of the MIMO channel in the absence of channel knowledge at the transmitter is given by [10]

$$C = \log_2 |\boldsymbol{I}_{N_r} + \frac{P_T}{N_t \sigma^2} \boldsymbol{H}\boldsymbol{H}^H| \tag{2.24}$$

The capacity of the MIMO channel when channel is known to the transmitter is necessarily greater than that when the channel is unknown to the transmitter. This point can also be observed by the simulation results in Figure 2.5.

2.3 MU-MIMO channel model

When a base station with multiple antennas supports multiple users with one or more antennas per user, we refer to this class of systems as multiuser MIMO (MU-MIMO). The downlink (forward link) from the base station to the users is a vector broadcast channel and the uplink (reverse link) is a vector multiple access channel. We focus on the downlink MU-MIMO broadcast channels, where a base station is equipped with N_t transmit antennas and serves K users, each user has $N_{r,k}$ receive

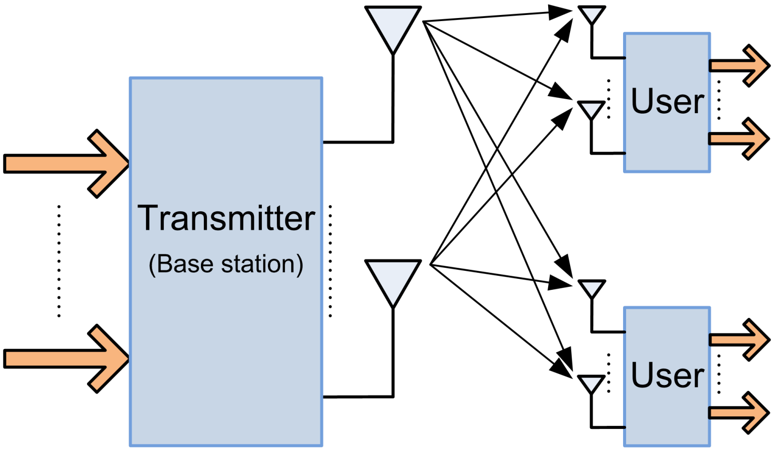

Figure 2.6: Diagram of MU-MIMO system.

antennas. The received signal $\boldsymbol{y}_k \in \mathbb{C}^{L_k \times 1}$ by the kth user is

$$\boldsymbol{y}_k = \boldsymbol{H}_k \sum_{j=1}^{K} \boldsymbol{V}_j \boldsymbol{x}_j + \boldsymbol{n}_k \tag{2.25}$$

where L_k is the number of data streams of the kth user; $\boldsymbol{H}_k \in \mathbb{C}^{N_{r,k} \times N_t}$ denotes the channel between the transmitter and the kth user; $\boldsymbol{x}_j \in \mathbb{C}^{L_j \times 1}$ is the transmit data vector for the jth user; $\boldsymbol{V}_j \in \mathbb{C}^{N_t \times L_j}$ denotes the transmit beamforming matrix, and the allocated transmit power is included. Therefore, compared with (2.16), the transmitted signal can be represented as

$$\boldsymbol{s} = \sum_{j=1}^{K} \boldsymbol{V}_j \boldsymbol{x}_j \tag{2.26}$$

Note that in downlink MU-MIMO broadcast channels, the transmit information for each user is emitted simultaneously, and each user can receive the information of all the users. Therefore, the transmit beamforming technique is essential for users to eliminate the interference, and enhance the desired information meanwhile.

2.4 MU-MIMO capacity

Before presenting the capacity calculation, we introduce dirty paper coding (DPC) first, which plays an important role for MU-MIMO capacity calculation.

In [57], the case where an additive white Gaussian noise channel corrupted by an interference known at the transmitter but unknown at the receiver is modeled as

$$Y = X + S + Z \tag{2.27}$$

where X and Y are the desired and received signals, respectively, S is the non-causally known interference, and Z is the unknown Gaussian noise. [57] shows that the capacity of this channel under the transmit power constraint is the same as if S did not exist. This technique is also referred to DPC technique.

Successive encoding of transmit information was proved to be optimum in terms of sum capacity in downlink MU-MIMO broadcast channels [14]. Given a preset user order in MU-MIMO broadcast channels, for the first encoded user, similarly to (2.25), the received signal can be written as

$$\boldsymbol{y}_1 = \boldsymbol{H}_1\boldsymbol{V}_1\boldsymbol{x}_1 + \boldsymbol{H}_1\sum_{j\geq 2}^{K}\boldsymbol{V}_j\boldsymbol{x}_j + \boldsymbol{n}_1 \tag{2.28}$$

At the time of encoding the first user, signals from the following users are unknown, we receive it (i.e., $\boldsymbol{H}_1\sum_{j\geq 2}^{K}\boldsymbol{V}_j\boldsymbol{x}_j$) with Gaussian noise together at the receiver.

For the kth ($k \geq 2$) user encoding, the received signal can be written as

$$\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{V}_k\boldsymbol{x}_k + \boldsymbol{H}_k\sum_{j<k}^{k-1}\boldsymbol{V}_j\boldsymbol{x}_j + \boldsymbol{H}_k\sum_{j>k}^{K}\boldsymbol{V}_j\boldsymbol{x}_j + \boldsymbol{n}_k \tag{2.29}$$

Note that at the time of encoding the kth user, the second term at the right hand of (2.29) (i.e., $\boldsymbol{H}_k\sum_{j<k}^{k-1}\boldsymbol{V}_j\boldsymbol{x}_j$) is known perfectly at the transmitter, which can be regarded as non-existing by DPC technique. In this case, the data rate of the kth user is given by

$$R_k = \log_2\frac{|\sigma^2\boldsymbol{I} + \sum_{j=k}^{K}\boldsymbol{H}_k\boldsymbol{V}_j\boldsymbol{V}_j^H\boldsymbol{H}_k^H|}{|\sigma^2\boldsymbol{I} + \sum_{j=k+1}^{K}\boldsymbol{H}_k\boldsymbol{V}_j\boldsymbol{V}_j^H\boldsymbol{H}_k^H|} \tag{2.30}$$

Define the transmit covariance matrix as $\boldsymbol{Q}_k = \mathbb{E}[\boldsymbol{V}_k\boldsymbol{x}_k{\boldsymbol{x}_k}^H\boldsymbol{V}_k^H]$. Note that the transmit data vector $\boldsymbol{x}_k$ satisfies $[\boldsymbol{x}_k\boldsymbol{x}_k^H] = \boldsymbol{I}$. Therefore, we have $\boldsymbol{Q}_k = \boldsymbol{V}_k\boldsymbol{V}_k^H$. The sum capacity of MU-MIMO broadcast channels can be written as

$$\begin{aligned} \max_{\{\boldsymbol{Q}_k\}_{k=1}^K} & \sum_{k=1}^{K} \log_2 \frac{|\sigma^2 \boldsymbol{I} + \sum_{j=k}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|}{|\sigma^2 \boldsymbol{I} + \sum_{j=k+1}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|} \\ \text{subject to} & \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k) \leq P_T \\ & \boldsymbol{Q}_k \geq 0 \end{aligned} \tag{2.31}$$

(2.31) is neither a convex nor a concave problem. Direct optimization will generally involve an exhaustive search over the entire space of covariance matrices that satisfy the power constraint and over the set of encoding orders, which is obviously very costly. Alternative methods to solve this problem have been proposed in [6], [49], [27], [58], [59], [60], [61], that exploit the relationship between the capacity region of the broadcast channels and that of its dual multiple access channels, which will be presented in the next chapter.

2.5 MAC model and capacity

Given MU-MIMO broadcast channels as described in (2.25), the system model for the dual multiple access channels (MAC) (Figure 2.7) is

$$\boldsymbol{t} = \sum_{k=1}^{K} \boldsymbol{H}_k^H \boldsymbol{U}_k \boldsymbol{x}_k^M + \boldsymbol{n}^M \tag{2.32}$$

where $\boldsymbol{t} \in \mathbb{C}^{N_t \times 1}$ is the received signal at the base station; $\boldsymbol{H}_k^H \in \mathbb{C}^{N_t \times N_{r,k}}$ denotes the channel between the kth user and the base station, notice that the channel matrix of MAC is the transpose conjugate of its dual broadcast channels; $\boldsymbol{U}_k$ is the transmit beamforming matrix of the kth user. Define $\boldsymbol{Q}_k^M = \boldsymbol{U}_k \boldsymbol{U}_k^H$ as the transmit covariance matrix of the kth user, and $\boldsymbol{n}^M \in \mathbb{C}^{N_t \times 1}$ is the Gaussian noise. Under a sum of transmit power constraint, i.e.

$$\sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k^M) \leq P_T \tag{2.33}$$

it has been shown in [27] that the set of achievable rates in MAC by successively decoding users, which is optimum in terms of capacity, is equal to the set of achievable rates in the dual broadcast channels by performing a successive encoding of users. Moreover, given a set of covariance matrices and a particular decoding order, a method

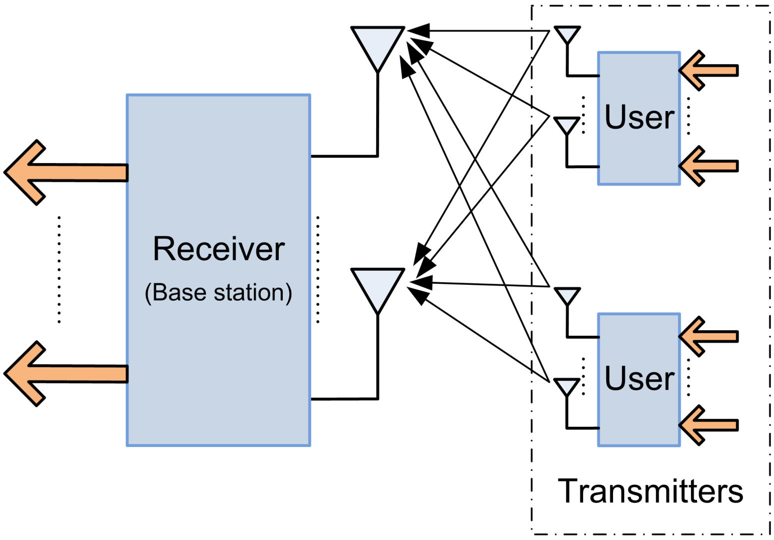

Figure 2.7: Diagram of MAC system.

has been found to compute the covariance matrices that achieve the same rates in the broadcast channel by encoding users in reverse order, i.e., the user decoded first in the multiple access channel is encoded last in the broadcast channel. Note that MAC with constraint (2.33) is merely a mathematical tool that allows the computation of optimum operational points in broadcast channel. Obviously, a common power constraint shared by non-cooperating users lacks practical relevance.

As a consequence of these results, the maximization of (2.31) can be indirectly performed by first maximizing the sum of achievable rates in the dual MAC and then computing the covariance matrices that achieve that sum rate in the broadcast channels. Fortunately, the sum capacity optimization in MAC, given by

$$\begin{aligned} \max_{\{\boldsymbol{Q}_k^M\}_{k=1}^K} \sum_{k=1}^{K} \log_2 & \frac{|\sigma^2 \boldsymbol{I} + \sum_{j=1}^{k} \boldsymbol{H}_j^H \boldsymbol{Q}_j^M \boldsymbol{H}_j|}{|\sigma^2 \boldsymbol{I} + \sum_{j=1}^{k-1} \boldsymbol{H}_j^H \boldsymbol{Q}_j^M \boldsymbol{H}_j|} \\ \text{subject to} \quad & \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k^M) \leq P_T \\ & \boldsymbol{Q}_k^M \geq 0 \end{aligned} \tag{2.34}$$

is a concave problem and therefore, can be maximized by using convex optimization techniques. Details will be given in the next chapter. In [27] it has been shown that the maximum value of (2.34) achieves the sum capacity of its dual MU-MIMO broadcast channels.

3

Beamforming techniques in MU-MIMO broadcast channels

In the past decade, a great deal of research has been directed toward the development of transmit beamforming techniques for the downlink MU-MIMO broadcast channels (BC).

It is shown that the optimal transmit strategy given by information theory is DPC, which achieves the capacity region. Unfortunately, DPC does not directly lead to a realizable transmission strategy because of the coupled structure of the transmitted signals. The BC optimization problems are usually non-convex and thus cannot be solved directly. The key technique used to overcome this difficulty is to transform the non-convex BC problem into a convex MAC problem via so called BC-MAC duality relationship. However, the computational complexity of the sum capacity optimization is still significant. Consequently, there has been substantial interest in developing transmission strategies that approach the performance of optimal solution and are easier to realize in practice.

Non-linear zero-forcing DPC techniques separate the interference into two parts, one part which is non-causally known at the base station can be removed by DPC technique and the rest part is eliminated by transmit beamforming vectors. Since these methods have quite low computational complexities, and their performances are close to the optimal solution, substantial research attentions have been observed in recent years.

In addition, linear beamforming techniques that avoid the non-linear DPC-like processing is also promising since the structures of the transmitter and receiver are much more simple and easier to implement, such as zero-forcing (ZF), block diagonalization (BD), and coordinated beamforming (CB).

3.1 Optimal solution

As discussed in Chapter 2, owing to the special structure of the BC, the associated capacity region computation and beamforming optimization problems are typically non-convex, and thus cannot be solved directly. One feasible approach is to consider the respective dual MAC problems, which are easier to deal due to their convexity properties. In the literature, two different BC-MAC dualities are studied substantially. One is subject to a total transmit power constraint (denoted as conventional BC-MAC duality), and another one is based on minimax duality.

Conventional BC-MAC duality discussed in [6], [49], [13], [58], [62], [63], [28], [29] states that, under a single transmit total power constraint, the capacity region of the BC is identical to that of its dual MAC under the same total power constraint. The channel matrix associated with the dual MAC is the conjugate transposed channel matrix of the corresponding BC, and the noise covariance matrices of both the BC and its dual MAC are identity matrices.

The conventional BC-MAC duality is first observed by [62], and is applied to solve the sum power minimization problem for the BC with signal-to-interference-plus-noise ratio (SINR) constraint. Several methods are developed independently to prove the conventional BC-MAC duality. The proof in [62] is based on the equivalence between the optimal solutions of the power minimization problems for the BC and MAC with SINR constraint. [49] proves the conventional BC-MAC duality by presenting the explicit transformation between the transmit covariance matrix of the BC and that of the MAC, and applies this duality to solve the sum-capacity problem. The conventional BC-MAC duality is widely applied to solve a number of BC problems. [28] and [29] solve the SINR balance problem for the BC, by maximizing the minimal SINR problem among all the users under the total power constraint, and by transforming this problem into its dual MAC problem. The conventional BC-MAC duality is also used in [58] to show that DPC achieves the sum capacity. Moreover, the entire capacity region for the BC can be obtained using the conventional BC-MAC duality.

Minimax duality [59], [60], [64], under a single sum power constraint or a set of linear power constraints, states that any boundary point of a BC capacity region can be obtained by solving a minimax optimization problem in its dual MAC. The channel matrix of the dual MAC is the conjugate transposed channel matrix of the corresponding BC, and the noise covariance matrix of the dual MAC is unknown for the minimization step of the minimax optimization problem.

Minimax duality proposed in [59], unifies the conventional BC-MAC duality. However, only the sum capacity is considered in [59]. Furthermore, [64] extends minimax duality to solve the capacity region computation problem and beamforming optimization problem for the BC with a per-antenna power constraint. Using a minimax optimization approach, the sum capacity of the MIMO-BC is also studied in [60].

In [27], authors propose a general BC-MAC duality that combines the conventional BC-MAC duality and minimax duality. It can be applied to solve the capacity region computation of the BC under the total power constraint. In addition, the optimal rate region under per-antenna power constraint can also be obtained with multiple linear transmit covariance constraints. In the following, we focus on the case where a total power constraint is imposed to exploit the capacity region of the BC. The general BC-MAC duality compares the SINR of each data stream for both the primal BC and its dual MAC. At the BS of the auxiliary MAC, successive interference cancellation (SIC) is deployed to decode the information of each user [65]. For this dual MAC, the decoding order among the users as well as the data streams of each user is the reverse of the encoding order in the primal BC. Let $SINR_{i,j}$ and $SINR_{i,j}^M$ denote the SINR of the (i, j)th data stream in BC and its dual MAC, respectively. According to DPC principle, encoded data streams have non-causal information about earlier encoded data streams, and thus the interference due to the earlier encoded data streams can be completely removed. Therefore, we have

$$SINR_{i,j} = \frac{p_{i,j}\left|\boldsymbol{u}_{i,j}^H \boldsymbol{H}_i \boldsymbol{v}_{i,j}\right|^2}{\sum_{k=i+1}^{K}\sum_{l=1}^{N} p_{k,l}\left|\boldsymbol{u}_{i,j}^H \boldsymbol{H}_i \boldsymbol{v}_{k,l}\right|^2 + \sum_{l=j+1}^{N} p_{i,l}\left|\boldsymbol{u}_{i,j}^H \boldsymbol{H}_i \boldsymbol{v}_{i,l}\right|^2 + \sigma^2} \tag{3.1}$$

and

$$SINR_{i,j}^M = \frac{q_{i,j}\left|\boldsymbol{v}_{i,j}^H \boldsymbol{H}_i^H \boldsymbol{u}_{i,j}\right|^2}{\boldsymbol{v}_{i,j}^H\left(\sum_{k=1}^{i-1}\sum_{l=1}^{N} q_{k,l}\boldsymbol{H}_k{}^H \boldsymbol{u}_{k,l}\boldsymbol{u}_{k,l}{}^H \boldsymbol{H}_k + \sum_{l=1}^{j-1} q_{i,l}\boldsymbol{H}_i{}^H \boldsymbol{u}_{i,l}\boldsymbol{u}_{i,l}{}^H \boldsymbol{H}_i + \boldsymbol{I}\right)\boldsymbol{v}_{i,j}} \tag{3.2}$$

where $\boldsymbol{v}_{i,j}$ denotes both the transmit beamforming vector in the primal BC and the receive combining vector in the dual MAC of the jth data stream of the ith user; $\boldsymbol{u}_{i,j}$ is both the transmit beamforming vector in the dual MAC and receive combining vector in the primal BC of the jth data stream of the ith user; $p_{i,j}$ and $q_{i,j}$ denote the transmit power of the jth data stream of the ith user in the primal BC and the dual MAC, respectively; N is the number of data streams for each user.

The general BC-MAC duality states that : For a given primal MIMO BC with fixed transmit beamforming and receive combining vectors $\boldsymbol{v}_{i,j}$ and $\boldsymbol{u}_{i,j}$, and fixed transmit power $p_{i,j}$ for each data stream, which satisfy trace$(\sum_{i,j} p_{i,j}\boldsymbol{v}_{i,j}\boldsymbol{v}_{i,j}^H) = P_T$, we can always find a set of transmit power $q_{i,j}$ for its dual MAC with fixed transmit beamforming and receive combining vectors $\boldsymbol{u}_{i,j}$ and $\boldsymbol{v}_{i,j}$, which satisfy $\sum_{i,j}\sigma^2 q_{i,j} = P_T$, such that the achievable SINR tuple of the primal BC is the same as that of its dual MAC, i.e.,

$$SINR_{i,j} = SINR_{i,j}^M \tag{3.3}$$

The proof is in Appendix A.

In the method proposed in [27], with a given user order, the received signal at the kth user is

$$\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{V}_k\boldsymbol{x}_k + \boldsymbol{H}_k\sum_{l>k}^{K}\boldsymbol{V}_l\boldsymbol{x}_l + \boldsymbol{n}_k \tag{3.4}$$

then the capacity optimization problem is written as

$$\begin{aligned} \max_{\{\boldsymbol{Q}_k\}_{k=1}^K} & \sum_{k=1}^{K}\log_2\frac{|\sigma^2\boldsymbol{I}+\sum_{j=k}^{K}\boldsymbol{H}_k\boldsymbol{Q}_j\boldsymbol{H}_k^H|}{|\sigma^2\boldsymbol{I}+\sum_{j=k+1}^{K}\boldsymbol{H}_k\boldsymbol{Q}_j\boldsymbol{H}_k^H|} \\ \text{subject to} & \sum_{k=1}^{K}\text{trace}(\boldsymbol{Q}_k)\le P_T \\ & \boldsymbol{Q}_k \ge 0 \end{aligned} \tag{3.5}$$

where $\boldsymbol{Q}_j = \boldsymbol{V}_j\boldsymbol{V}_j^H$ is the transmit covariance matrix. Unfortunately, as discussed in Chapter 2, (3.5) is neither a convex nor a concave problem, direct optimization is difficult. According to the general BC-MAC duality, (3.5) can be transformed into a dual MAC problem as

$$\begin{aligned} \max_{\{\boldsymbol{Q}_k^M\}_{k=1}^K} & \sum_{k=1}^{K}\log_2\frac{|\boldsymbol{I}+\sum_{j=1}^{k}\boldsymbol{H}_j\boldsymbol{Q}_j^M\boldsymbol{H}_j^H|}{|\boldsymbol{I}+\sum_{j=1}^{k-1}\boldsymbol{H}_j\boldsymbol{Q}_j^M\boldsymbol{H}_j^H|} \\ \text{subject to} & \sum_{k=1}^{K}\text{trace}(\sigma^2\boldsymbol{Q}_k^M)\le P_T \\ & \boldsymbol{Q}_k^M \ge 0 \end{aligned} \tag{3.6}$$

where $\boldsymbol{Q}_k^M$ indicates the transmit covariance matrix in the dual MAC. The general BC-MAC duality proved that (3.5) and (3.6) can achieve the same optimal value. From the proof of general BC-MAC duality in Appendix A, we can see that the SINR relationship between the BC and its dual MAC relies crucially on the reciprocity relationship [5]. In [27], a MAC-BC covariance transformation algorithm between $\boldsymbol{Q}_k^M$ and $\boldsymbol{Q}_k$ is proposed to find $\boldsymbol{Q}_k$ when $\boldsymbol{Q}_k^M$ is obtained (Algorithm 1). Note that the objective function of (3.6) can be reordered as [66]

$$\begin{aligned}
&\sum_{k=1}^{K} \log_2 \frac{|\boldsymbol{I} + \sum_{j=1}^{k} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H|}{|\boldsymbol{I} + \sum_{j=1}^{k-1} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H|} \\
&= \sum_{k=1}^{K} \log_2 |\boldsymbol{I} + \sum_{j=1}^{k} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H| - \sum_{k=1}^{K} \log_2 |\boldsymbol{I} + \sum_{j=1}^{k-1} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H| \\
&= \sum_{k=1}^{K-1} \log_2 |\boldsymbol{I} + \sum_{j=1}^{k} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H| + \log_2 |\boldsymbol{I} + \sum_{j=1}^{K} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H| \\
&\quad - \sum_{k=1}^{K-1} \log_2 |\boldsymbol{I} + \sum_{j=1}^{k} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H| \\
&= \log_2 |\boldsymbol{I} + \sum_{j=1}^{K} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H|
\end{aligned} \tag{3.7}$$

Therefore, the problem (3.6) can be simplified as

$$\begin{aligned}
&\max_{\{\boldsymbol{Q}_k^M\}_{k=1}^K} \log_2 |\boldsymbol{I} + \sum_{k=1}^{K} \boldsymbol{H}_k \boldsymbol{Q}_k^M \boldsymbol{H}_k^H| \\
&\text{subject to} \quad \sum_{k=1}^{K} \text{trace}(\sigma^2 \boldsymbol{Q}_k^M) \leq P_T \\
&\qquad\qquad\quad \boldsymbol{Q}_k^M \geq 0
\end{aligned} \tag{3.8}$$

This is a convex problem and the standard techniques (e.g. CVX, Yalmip) can be used to get the optimal transmit covariance matrix $\boldsymbol{Q}_k^{M\star}$. Then the BC transmit covariance matrix $\boldsymbol{Q}_k$ is found by the MAC-BC covariance transformation algorithm. Note that the decoding order of the BC is the reverse of the encoding order of its dual MAC.

The optimal solution can be found by this method. But in practice, suboptimal methods, that are much easier to implement with no significant performance degradation, also attract substantial attentions. Next, we introduce several famous suboptimal

Algorithm 1 MAC-BC covariance transformation algorithm

Input $\boldsymbol{Q}_k^M$, $1 \leq k \leq K$
for $k = 1 : K$ **do**
  Compute the eigen value decomposition of $\boldsymbol{Q}_k^{M\star} = \boldsymbol{U}_k \boldsymbol{\Lambda}_k {\boldsymbol{U}_k}^H$
  for $j = 1 : N$ **do**
    Define $\boldsymbol{u}_{k,l}$ as the lth column of $\boldsymbol{U}_k$
    Define $q_{k,l}$ as the lth diagonal element of $\boldsymbol{\Lambda}_k$
    Obtain $\boldsymbol{v}_{k,j}$ by MMSE receiver
    $\boldsymbol{v}_{k,j} = (\sum_{i=1}^{k-1} \sum_{l=1}^{N} q_{i,l} {\boldsymbol{H}_i}^H \boldsymbol{u}_{i,l} {\boldsymbol{u}_{i,l}}^H \boldsymbol{H}_i + \sum_{l=1}^{j-1} q_{k,l} {\boldsymbol{H}_k}^H \boldsymbol{u}_{k,l} {\boldsymbol{u}_{k,l}}^H \boldsymbol{H}_k + \boldsymbol{I})^{-1} {\boldsymbol{H}_k}^H \boldsymbol{u}_{k,j}$
    Normalize $\boldsymbol{v}_{k,j}$
    Compute $p_{k,j}$ using $\text{SINR}_{k,j} = \text{SINR}_{k,j}^M$ [27]
  end for
  Compute $\boldsymbol{Q}_k = \sum_{l=1}^{N} p_{k,l} \boldsymbol{v}_{k,l} \boldsymbol{v}_{k,l}^H$
end for

methods. The DPC involved method is denoted as non-linear method and otherwise it is denoted as linear method.

3.2 Non-linear beamforming techniques

The first non-linear beamforming technique is ZF-DPC method. In this method, the total number of receive antennas can not be larger than that of transmit antennas, and each user is supposed to have one receive antenna. Then, this method is extended to SZF-DPC method, in which each user can have multiple receive antennas, but the total number of receive antennas can not be larger than that of transmit antennas. Then, SZF-DPC method is extended to SA-DPC method, in which the total number of receive antennas may be larger than that of transmit antennas. In Figure 3.1, the evolution of these non-linear beamforming techniques is given.

3.2.1 Tomlinson Harashima Precoder

Non-linear methods separate the interference into two parts, one part is removed by zero-forcing technique, and the rest part which is non-causally known is eliminated by DPC techniques. A practical implementation of DPC technique is the pre-equalization Tomlinson Harashima Precoder (THP) technique, which is proposed in [67] and [68], aiming to pre-subtract the non-causally known interference at the transmitter. THP is initially proposed for single-input single-output channels in the presence of inter-

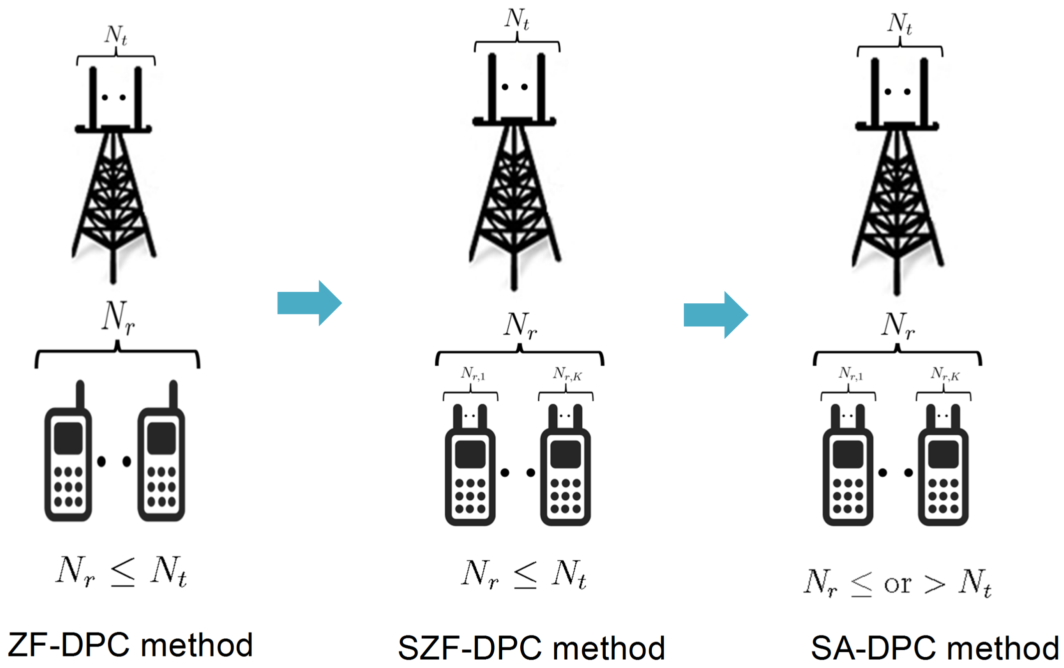

Figure 3.1: Evolution of non-linear beamforming techniques

symbol interference (ISI), and it is extended to BC in [69]. Several different criteria are proposed using THP including zero-forcing (ZF) [69], minimum mean square error (MMSE) [70]. In [71], a robust THP is investigated with imperfect channel state information (CSI) at the transmitter. The main idea of THP is that, the non-causally known interference produced by the previous precoded symbols can be pre-canceled before transmission at the transmitter, and the modulo operation can be adopted to ensure that transmit power does not exceed the power constraint.

Now, we introduce the pre-equalization THP technique that pre-subtracts the non-causally known interference at the base station. Suppose a base station with N_t transmit antennas transmits information to K users, each user is equipped with one single receive antenna, the received signal by the kth user is

$$y_k = \boldsymbol{h}_k^H \boldsymbol{v}_k x_k + \boldsymbol{h}_k^H \sum_{j=1}^{k-1} \boldsymbol{v}_j x_j + \boldsymbol{h}_k^H \sum_{j=k+1}^{K} \boldsymbol{v}_j x_j + n_k \quad (1 \leq k \leq K) \tag{3.9}$$

where $\boldsymbol{v}_k$ is the transmit beamforming vector for the kth user. The transmit symbol x_k can be written as

$$x_k = a_k - \sum_{j=1}^{k-1} b_{k,j} x_j \quad (1 \leq k \leq K) \tag{3.10}$$

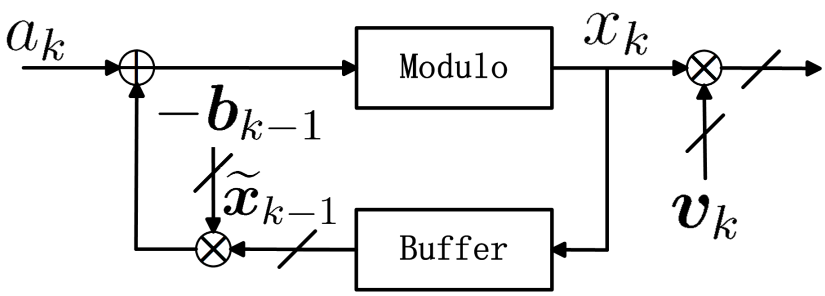

Figure 3.2: Block diagram of the transmitter

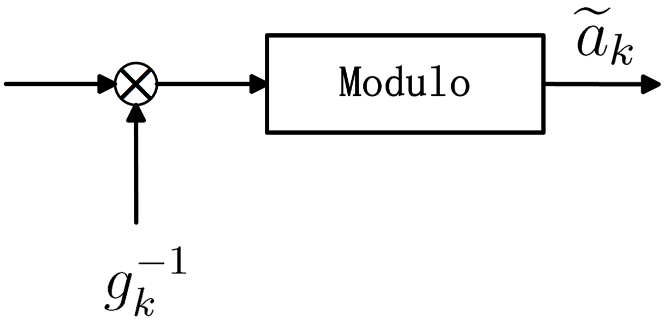

Figure 3.3: Block diagram of the receiver

where a_k is the information symbol and $b_{k,j}$ is defined as

$$b_{k,j} = \frac{\boldsymbol{h}_k^H \boldsymbol{v}_j \sqrt{p_j}}{\boldsymbol{h}_k^H \boldsymbol{v}_k \sqrt{p_k}} \tag{3.11}$$

Using (3.11) and (3.10) in (3.9), we have

$$y_k = \boldsymbol{h}_k^H \boldsymbol{v}_k \sqrt{p_k} a_k + \boldsymbol{h}_k^H \sum_{j=k+1}^{K} \boldsymbol{v}_j x_j + n_k \quad (1 \leq k \leq K) \tag{3.12}$$

Therefore, the non-causally known interference $\boldsymbol{h}_k^H \sum_{j=1}^{k-1} \boldsymbol{v}_j x_j$ can be pre-subtracted completely from the base station. In next sections, we will show how to suppress the residual interference $\boldsymbol{h}_k^H \sum_{j=k+1}^{K} \boldsymbol{v}_j x_j$.

Define $\boldsymbol{b}_k = [b_{k,1}, \cdots, b_{k,k-1}, 0, \cdots, 0]^T$ and $\widetilde{\boldsymbol{x}}_k = [x_1, \cdots, x_{k-1}, 0, \cdots, 0]^T$ as a measure of pre-subtracting the known interference term in (3.9), which is shown in Figure 3.2. The transmit symbol x_k are successively generated from the information symbols a_k in (3.10).

If the information symbol a_k is uniformly distributed in an M-ary QAM constellation, then $\mathbb{E}\{\boldsymbol{a}\boldsymbol{a}^H\} = \boldsymbol{I}_K$ with $\boldsymbol{a} = [a_1, \cdots, a_K]^T$. The modulo device is used to ensure that the transmit power does not exceed the power constraint, and $\mathbb{E}\{\boldsymbol{x}\boldsymbol{x}^H\} = \boldsymbol{I}_K$ is guaranteed [69].

At the kth user, the received signal is reformulated as

$$y_k = \boldsymbol{h}_k^H \boldsymbol{v}_k \sqrt{p_k}(a_k + q_k) + n_k \quad (1 \leq k \leq K) \tag{3.13}$$

where $q_k \in \{2\sqrt{M} \cdot (q_I + iq_Q) | q_I, q_Q \in \mathbb{Z}\}$ is introduced by the modulo operator at the transmitter. y_k is divided by $g_k = \boldsymbol{h}_k^H \boldsymbol{v}_k \sqrt{p_k}$ before passing through the modulo operator to satisfy the same constellation boundaries as for the transmitter. After following the modulo device, q_k is removed and the original data symbol is estimated as $\tilde{a}_k = a_k + n_k / g_k$ as shown in Figure 3.3.

3.2.2 ZF-DPC method

ZF-DPC method [14] is based on LQ decomposition of the channels where each user with only one single receive antenna is encoded successively. For each user in a given order, the interference caused by the previously encoded users is considered as the non-causally known interference, and it can be eliminated by DPC technique. The interference caused by the subsequent encoded users is removed by beamforming technique at the transmitter. As we discussed in Section 3.2.1, after pre-subtracting the non-causally known interference $\boldsymbol{h}_k^H \sum_{j=1}^{k-1} \boldsymbol{v}_j x_j$, we have

$$y_k = \boldsymbol{h}_k^H \boldsymbol{v}_k \sqrt{p_k} a_k + \boldsymbol{h}_k^H \sum_{j=k+1}^{K} \boldsymbol{v}_j x_j + n_k \quad (1 \leq k \leq K) \tag{3.14}$$

The term $\boldsymbol{h}_k^H \sum_{j=k+1}^{K} \boldsymbol{v}_j x_j$ can be removed by appropriate transmit beamforming vectors, i.e., we force

$$\boldsymbol{h}_k^H \sum_{j=k+1}^{K} \boldsymbol{v}_j x_j = 0 \tag{3.15}$$

Suppose $\boldsymbol{v}_k = \boldsymbol{v}'_k \sqrt{p_k}$, where $\boldsymbol{v}'_k$ and p_k is the normalized transmit beamforming vector and the assigned power for the kth user, respectively. If the number of users K is less than or equal to the number of transmit antennas N_t (i.e., $K \leq N_t$), and the channel matrix $\boldsymbol{H}$ is described as

$$\boldsymbol{H} = \Big[\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K\Big]^H \in \mathbb{C}^{K \times N_t} \tag{3.16}$$

then the transmit beamforming vectors can be obtained by performing LQ decomposition of the channel matrix $\boldsymbol{H} = \boldsymbol{L}\boldsymbol{Q}$. $\boldsymbol{v}'_k$ $(\forall k)$ is chosen as the kth column of the matrix $\boldsymbol{Q}^H \in \mathbb{C}^{N_t \times N_t}$, and after performing DPC technique, (3.14) is reformulated as

$$y_k = l_{k,k}\sqrt{p_k}x_k + n_k \quad (1 \leq k \leq K) \tag{3.17}$$

where $l_{k,k}$ is the kth diagonal element of the lower triangular matrix $\boldsymbol{L}_{K \times N_t}$. Similarly to section 2.2, water-filling algorithm is used to find the optimal p_k. In the following, we ignore the power allocation unless it is specified. Note that water-filling algorithm is always the optimal solution if interference is completely removed.

The pseudo code of ZF-DPC method is given in Algorithm 2.

ZF-DPC method only supports one single receive antenna for each user, and the user permutation also has a great effect on the throughput performance. It is shown that the optimal user order needs exhaustive search [14]. In Figure 3.4, the achievable sum rate provided by ZF-DPC method and the sum capacity (denoted as DPC method) is compared. It can be seen that the gap between the ZF-DPC method and DPC method becomes small with the increase of SNR. The reason is that, the transmit beamforming vectors are designed to remove one part of the interference while the Gaussian noise, which plays a dominant role in low SNR region, is ignored.

Algorithm 2 Pseudo code of ZF-DPC method

Suppose a preset user order
Build the matrix $\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^H$ as (3.16)
Do LQ decomposition of $\boldsymbol{H}$ as $\boldsymbol{H} = \boldsymbol{L}\boldsymbol{Q}$
$\boldsymbol{v}_k$ $(\forall k)$ is chosen as the kth column of $\boldsymbol{Q}^H$
Perform water-filling power allocation
Use DPC pre-subtract the non-known interference

3.2.3 SZF-DPC method

Successive ZF-DPC method (denoted as SZF-DPC method) [16] is considered as a generalization of ZF-DPC method, it extends ZF-DPC method to the case where each user has multiple receive antennas. Suppose a base station with N_t transmit antennas transmits information to K users, the kth user has $N_{r,k}$ receive antenna. The received signal of the kth user $\boldsymbol{y}_k \in {}^{N_{r,k} \times 1}$ is written as

$$\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{V}_k\boldsymbol{x}_k + \boldsymbol{H}_k\sum_{j=1}^{k-1}\boldsymbol{V}_j\boldsymbol{x}_j + \boldsymbol{H}_k\sum_{j=k+1}^{K}\boldsymbol{V}_j\boldsymbol{x}_j + \boldsymbol{n}_k \quad (1 \leq k \leq K) \tag{3.18}$$

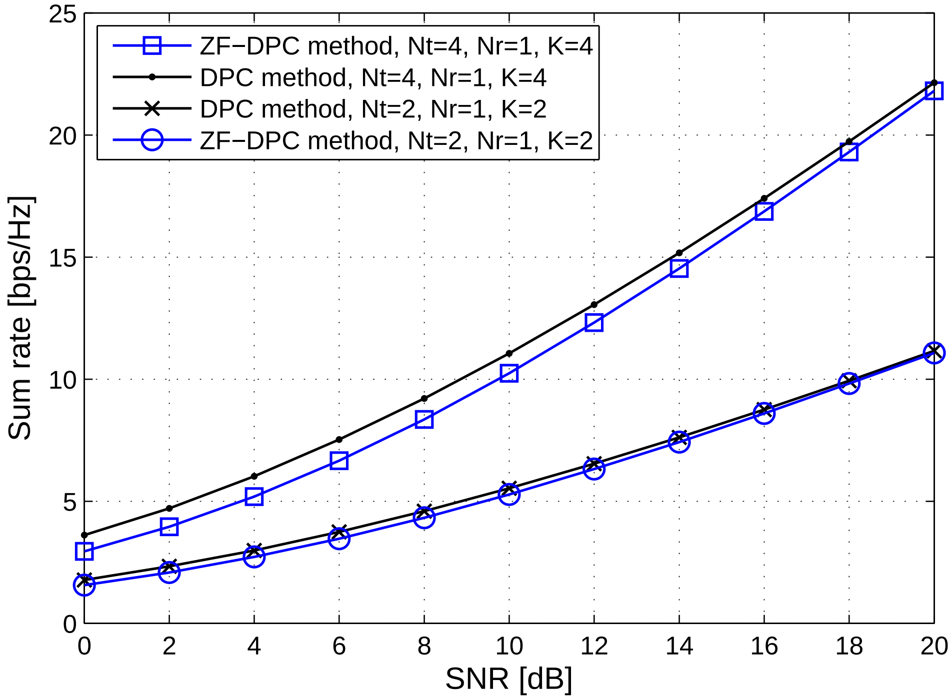

Figure 3.4: Comparison between the achievable sum rate of ZF-DPC method and the sum capacity.

where $\boldsymbol{H}_k \in \mathbb{C}^{N_{r,k}\times N_t}$ denotes the channel between the transmitter and the kth user; $\boldsymbol{x}_k \in \mathbb{C}^{N_{r,k}\times 1}$ is the transmit data vector for the kth user; $\boldsymbol{V}_k \in \mathbb{C}^{N_t\times N_{r,k}}$ denotes the transmit beamforming matrix. Similarly to ZF-DPC method, SZF-DPC method is also a combination of DPC and zero-forcing technique. The non-causally known interference term $\boldsymbol{H}_k \sum_{j=1}^{k-1} \boldsymbol{V}_j\boldsymbol{x}_j$ is removed according to DPC principle, and the residual interference term $\boldsymbol{H}_k \sum_{j=k+1}^{K} \boldsymbol{V}_j\boldsymbol{x}_j$ is eliminated completely by transmit beamforming design, i.e.,

$$\boldsymbol{H}_k \sum_{j=k+1}^{K} \boldsymbol{V}_j\boldsymbol{x}_j = \boldsymbol{0} \tag{3.19}$$

Define the matrix $\hat{\boldsymbol{H}}_k$ as

$$\hat{\boldsymbol{H}}_k = \begin{bmatrix}\boldsymbol{H}_1^H & \boldsymbol{H}_2^H & \cdots & \boldsymbol{H}_{k-1}^H\end{bmatrix}^H \in \mathbb{C}^{(\sum_{j=1}^{k-1} N_{r,j})\times N_t} \tag{3.20}$$

If (3.19) is defined, the transmit beamforming matrix $\boldsymbol{V}_k$ must lie in the null space of the space spanned by $\hat{\boldsymbol{H}}_k$. In the literature, several matrix decomposition methods are proposed to find the transmit beamforming matrix $\boldsymbol{V}_k$, such as SVD method in [16], and SGO method in [72]. Here, we use SVD method to demonstrate how the transmit beamforming matrix $\boldsymbol{V}_k$ can be obtained. The SVD of $\hat{\boldsymbol{H}}_k$ is

$$\hat{\boldsymbol{H}}_k = \hat{\boldsymbol{U}}_k\hat{\boldsymbol{\Lambda}}_k[\hat{\boldsymbol{V}}_k^{(1)}\hat{\boldsymbol{V}}_k^{(0)}]^H \tag{3.21}$$

where $\hat{\boldsymbol{U}}_k$ and $\hat{\boldsymbol{\Lambda}}_k$ denote the left singular matrix and the matrix of ordered singular values of $\hat{\boldsymbol{H}}_k$, respectively. $\hat{\boldsymbol{V}}_k^{(1)}$ and $\tilde{\boldsymbol{V}}_k^{(0)}$ denote the right singular matrices each consisting of the singular vectors corresponding to non-zero singular values and zero singular values, respectively. Therefore, to satisfy (3.19), we can choose the transmit beamforming matrix $\boldsymbol{V}_k$ from the first $N_{r,k}$ columns of $\hat{\boldsymbol{V}}_k^{(0)}$, i.e.,

$$\boldsymbol{V}_k = (\hat{\boldsymbol{V}}_k^{(0)})_{1:N_{r,k}} \tag{3.22}$$

where $N_{r,k}$ is the number of transmit data streams for the kth user, notice that it is also the number of receive antennas of the kth user. Then, the downlink MU-MIMO broadcast channels can be considered as parallel and interference free, and water-filling power allocation is performed as the optimal solution. Algorithm 3 gives the pseudo code of SZF-DPC method.

SZF-DPC method also requires exhaustive search over all the user permutations to find the optimal solution. Moreover, the total number of receive antennas is restricted

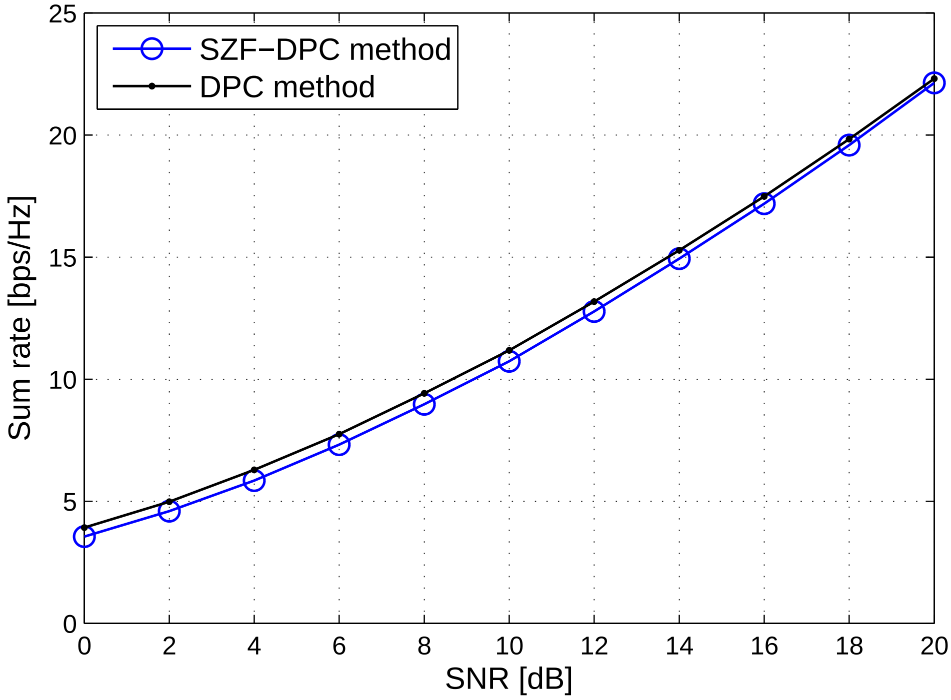

Figure 3.5: Comparison between the achievable sum rate of SZF-DPC method and the sum capacity, $N_t = 4$, $N_{r,k} = 2$, $\forall k$, and $K = 2$.

by the number of transmit antennas, i.e. $N_t \geq \sum_{k=1}^{K} N_{r,k}$. The transmit beamforming matrix can be found for each served user only in the case where the dimension of the null space of the space spanned by $\hat{\boldsymbol{H}}_k$ is larger than zero. In Figure 3.5, the simulation results also show that the gap between SZF-DPC method and DPC method becomes small with high SNR.

Algorithm 3 Pseudo code of the SZF-DPC method

Suppose a preset user order
for $k = 1 : K$ **do**
  Build the matrix $\hat{\boldsymbol{H}}_k = [\boldsymbol{H}_1{}^H, \boldsymbol{H}_2{}^H, \cdots, \boldsymbol{H}_{k-1}{}^H]^H$ as (3.20)
  Do SVD decomposition of $\hat{\boldsymbol{H}}_k$ and find $\boldsymbol{V}_k$ as (3.22)
end for
Perform water-filling power allocation
Use DPC pre-subtract the non-known interference $\boldsymbol{h}_1^H \boldsymbol{v}_1$

3.2.4 SA-DPC method

Successive allocation DPC method (denoted as SA-DPC method) proposed in [15] assigns only one data stream at each step to the user who brings the largest throughput increase. At the ith $(1 \leq i \leq L)$ step, to avoid the ith data stream interfering the previously allocated $i-1$ data streams, the channel matrix of each candidate is projected into the subspace

$$\boldsymbol{H}_k^i = \boldsymbol{H}_k \boldsymbol{T}_i \quad \forall k \tag{3.23}$$

where the projection matrix $\boldsymbol{T}_i = \boldsymbol{T}_{i-1} - \boldsymbol{v}'_{i-1} {\boldsymbol{v}'_{i-1}}^H$ represents the projector of the subspace defined by the intersection of the null space of the already established subchannels with $\boldsymbol{T}_1 = \boldsymbol{I}_{N_t}$, and $\boldsymbol{v}'_i$ is denoted as the normalized transmit beamforming vector of the ith data stream. It is clear that the largest channel gain is obtained by selecting the normalized transmit beamforming and receive combining vectors as the right and left singular vectors associated to the largest singular value of the projected channel matrix $\boldsymbol{H}_k^i$, respectively. Then the ith data stream is assigned to the user who can offer the largest channel gain among all the candidates.

Since the transmit beamforming vector of the ith data stream lies in the null space of the space spanned by the previously allocated $i-1$ data streams, the interference caused by the ith data stream to the previously allocated $i-1$ ones is zero. Furthermore, the interference caused by the previously allocated $i-1$ data streams to the ith one is non-causally known, which can be removed before transmission by DPC principle. Algorithm 4 gives the pseudo code of SA-DPC method.

Figure 3.6 shows the achievable sum rate of SA-DPC method with two different system configurations. It can be seen that the sum rate provided by SA-DPC method is quite close to the sum capacity under total transmit power constraint. In the next chapter, we will present SA-DPC method in a different way, which performs identically as SA-DPC method, but when the practical per-antenna power constraint is considered. This new way to demonstrate SA-DPC method can be easily extended to cope with this more practical case.

3.3 Summary

In Figure 3.7, we summarize the interference removing techniques in non-linear methods. We can see that the interference from the previous users is removed by DPC technique, and only the interference that comes from the following users is canceled by beamforming techniques.

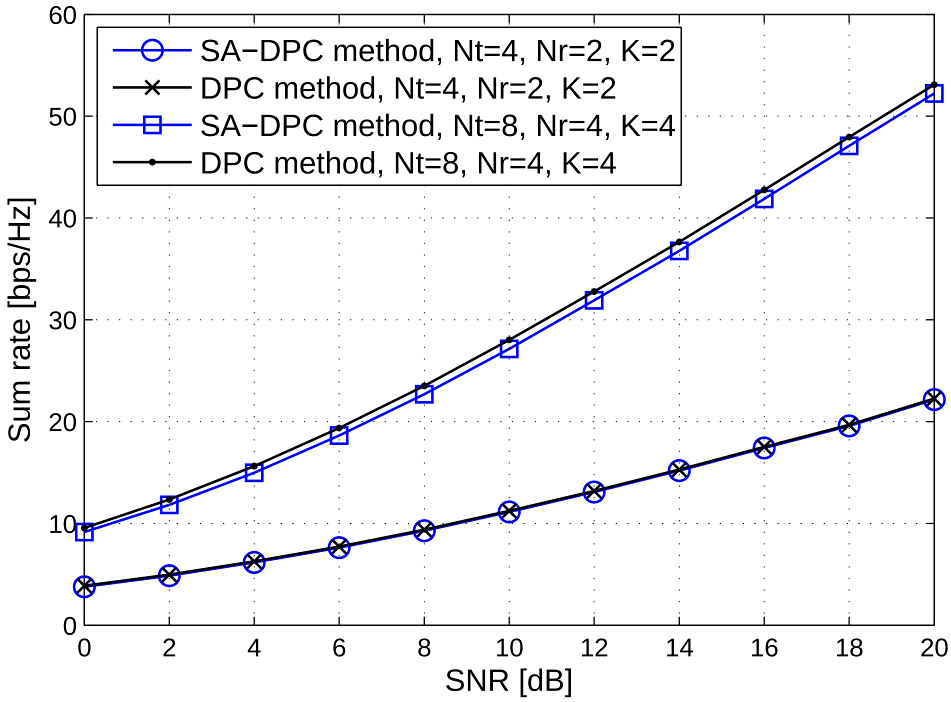

Figure 3.6: Comparison between the achievable sum rate of SA-DPC method and the sum capacity.

Algorithm 4 Pseudo code of the SA-DPC method

Initialization: $\boldsymbol{T}_1 = \boldsymbol{I}_{N_t}$
for $i = 1 : N_t$ **do**
  for $k = 1 : K$ **do**
    Build the matrix $\boldsymbol{H}_k{}^i = \boldsymbol{H}_k \boldsymbol{T}_i$ as (3.23)
    Do SVD decomposition of $\boldsymbol{H}_k{}^i$ and find the temporary transmit beamforming and receive combining vectors
    Perform water-filling power allocation
    Calculate the temporary sum Denoted as R_k
  end for
  Calculate the largest sum rate $C_i = \max(R_k)$
  if $C_{i-1} \geq C_i$
    break
  end
end for

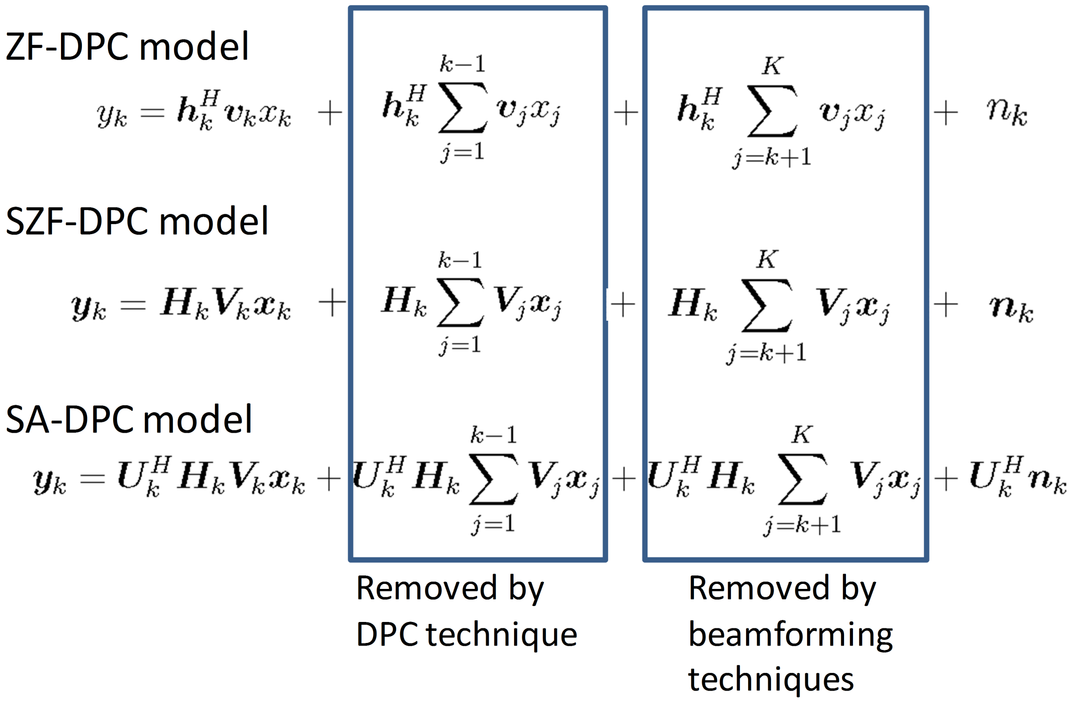

Figure 3.7: Interference removing techniques in non-linear methods

3.4 Linear beamforming techniques

Although THP is a practical implementation of DPC technique, the structure of the transmitter and receiver is complex in this way. Another popular direction aims to suppress the interference only via transmit beamforming vectors, we denote it as linear beamforming techniques.

The first linear beamforming technique is ZF method, In this method, the total number of receive antennas can not be larger than that of transmit antennas, and each user is supposed to have one receive antenna. Then, it is extended to multiple receive antennas for each user by Block Diagonalization (BD) method and Coordinated beamforming (CB) method. BD method works for situations with multiple antennas and multiple data streams intended for each user, but the total number of receive antennas is no more than that of transmit antennas. CB method extends the constraint of each user's receiver antennas to unconstraint, but the data stream is limited to one for each user. In ZF-SA method, the total number of receive antennas may be larger than that of transmit antennas, and each user can have multiple receive antennas. In MMSE method, it is the same as ZF-SA method, but the sum rates optimization is performed in a different way. The evolution of these linear beamforming techniques is given in Figure 3.8.

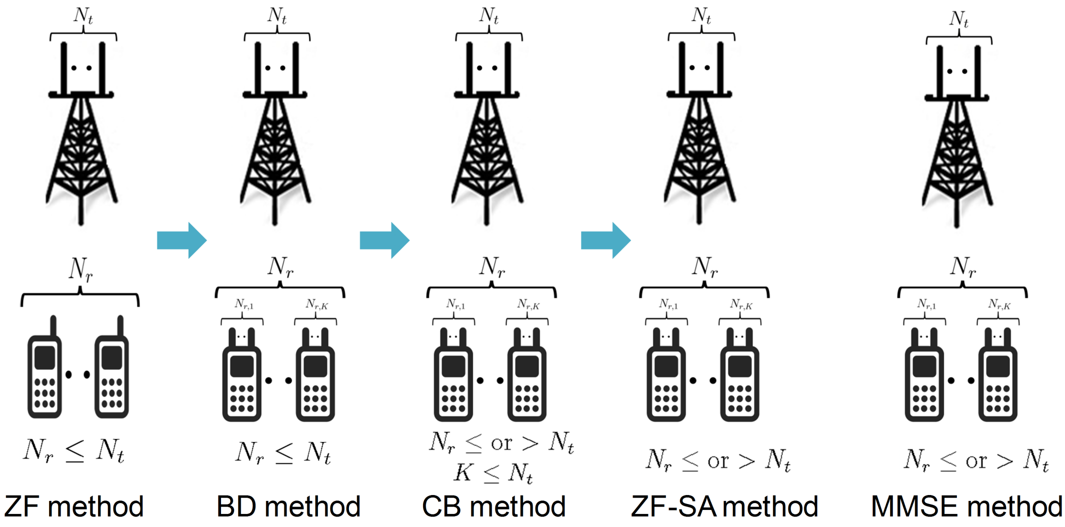

Figure 3.8: Evolution of linear beamforming techniques

3.4.1 ZF method

Zero-forcing (ZF) or channel inversion decomposes the channel into several parallel scalar channels with only additive noise, and the interference is removed completely by transmit beamforming techniques. Suppose a base station with N_t transmit antennas transmits information to K users, each user is equipped with one single receive antenna. The received signal by the kth user is

$$y_k = \boldsymbol{h}_k^H \boldsymbol{v}_k x_k + \boldsymbol{h}_k^H \sum_{j=1, j\neq k}^{K} \boldsymbol{v}_j x_j + n_k \quad (1 \leq k \leq K) \tag{3.24}$$

ZF method transmits the signals towards the intended user with nulls steered in the direction of the other users, i.e., $\boldsymbol{h}_j^H \boldsymbol{v}_k = 0 \ \forall j \neq k$. The users will receive only the desired signal without any interference because of the perfect nulling. In this case, the received data at the kth user can be written [17]

$$y_k = \boldsymbol{h}_k^H \boldsymbol{v}_k x_k + n_k \tag{3.25}$$

The corresponding vector equation is

$$\boldsymbol{y} = \boldsymbol{H}^H \boldsymbol{V} \boldsymbol{x} + \boldsymbol{n} \tag{3.26}$$

Therefore, if the normalized transmit beamforming vector of the kth user is selected as

$$\boldsymbol{v}'_k = \frac{\boldsymbol{h}_k^{(\dagger)}}{\sqrt{\| \boldsymbol{h}_k{}^{(\dagger)} \|_F^2}} \tag{3.27}$$

where $\boldsymbol{h}_k^{(\dagger)}$ is the kth column of the pseudo inverse of $\boldsymbol{H}$, denoted as $\boldsymbol{H}^{(\dagger)}$. Then it is shown that the interference can be canceled completely. In this case, the SNR of the kth user is

$$\text{SNR}_k = \frac{|\boldsymbol{h}_k^H \boldsymbol{v}'_k|^2 p_k}{\sigma^2} \tag{3.28}$$

The optimal p_k is given by water-filling algorithm. Finally, the achievable sum rate of ZF is found like

$$C = \sum_{k=1}^{K} \log_2(1 + \frac{|\boldsymbol{h}_k^H \boldsymbol{v}'_k|^2 p_k}{\sigma^2}) \tag{3.29}$$

Algorithm 5 gives the pseudo code of ZF method. It can be seen that ZF beamforming technique is easy to implement, but this technique suffers from the noise enhancement problem. For example, if two or more user channels $\boldsymbol{h}_i$ are close to each other, corresponding users will receive very little power (power reduction).

Algorithm 5 Pseudo code of the ZF method

Suppose a preset user order
Build the matrix $\boldsymbol{H}^H = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]$
Do pseudo inverse of $\boldsymbol{H}$ denoted as $\boldsymbol{H}^{(\dagger)}$
$\boldsymbol{v}_k$ $(\forall k)$ is chosen as the kth column of $\boldsymbol{H}^{(\dagger)}$
Perform water-filling power allocation

Figure 3.9 illustrates the noise enhancement problem (also known as the power efficiency problem). The transmit beamforming vector of the first user $\boldsymbol{v}_1$ is orthogonal to $\boldsymbol{h}_2$ but delivers very little power along $\boldsymbol{h}_1$, which is the desired user direction. The problem becomes worse as $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ become closer. Figure 3.10 illustrates the achievable sum rate of ZF method with two different configurations, we can see that the gap between DPC method and this simple ZF method is large, this is not only because of the noise enhancement problem, the spatial multiplexing is not fully exploited neither.

3.4.2 BD method

To cope with the limitations of ZF method, a related strategy named BD method for situations with multiple antennas and multiple data streams intended for each receiver is proposed [18]. The multiuser MIMO model with BD method employs N_t transmit

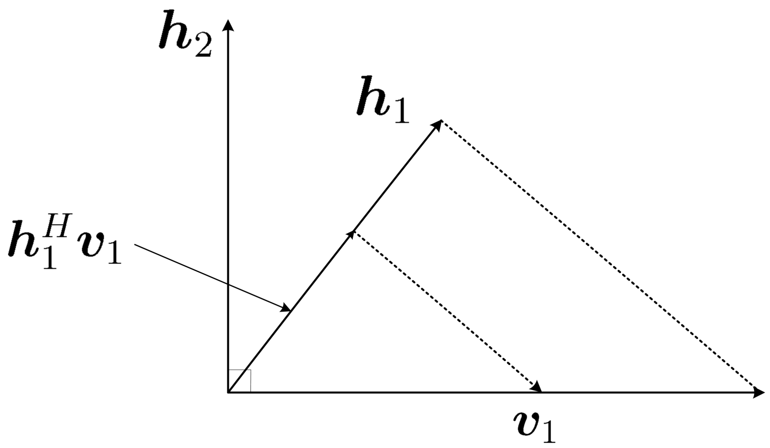

Figure 3.9: Scheme of illustrating the noise enhancement problem problem. $\boldsymbol{v}_1$ has gain $<< 1$ along $\boldsymbol{h}_1$

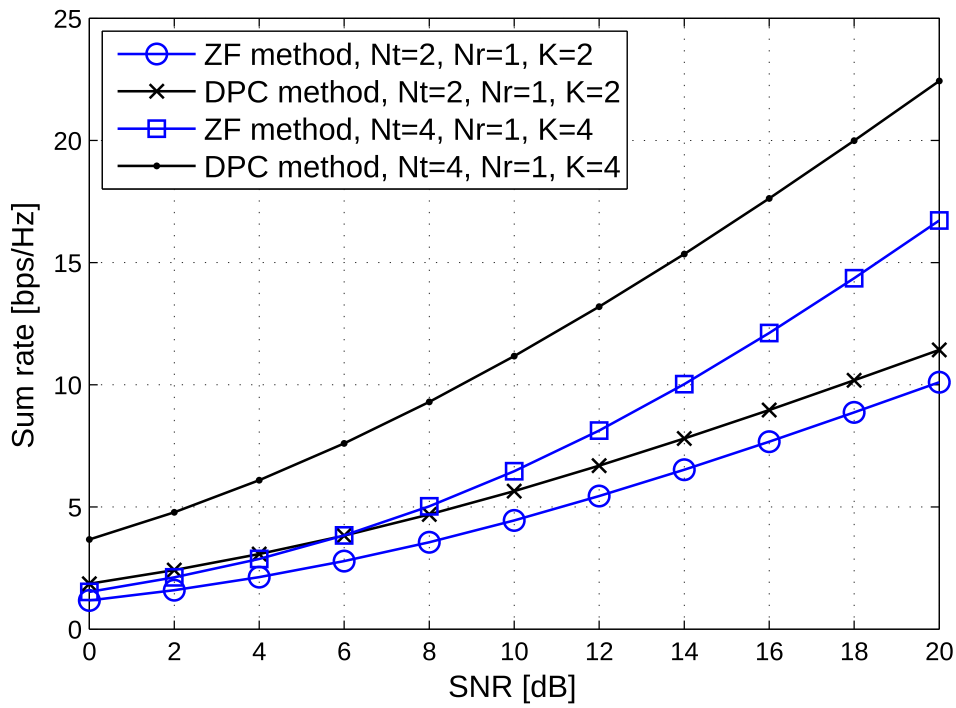

Figure 3.10: Comparison between the achievable sum rate of ZF method and sum capacity.

antennas with K users, each equipped with $N_{r,k}$ receive antennas and receiving their own data streams. The received signal at the kth receiver $\boldsymbol{y}_k \in \mathbb{C}^{N_{r,k}\times 1}$ is

$$\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{V}_k\boldsymbol{x}_k + \boldsymbol{H}_k\sum_{l=1,l\neq k}^{K}\boldsymbol{V}_l\boldsymbol{x}_l + \boldsymbol{n}_k \tag{3.30}$$

where $\boldsymbol{H}_k \in \mathbb{C}^{N_{r,k}\times N_t}$ denotes the channel between the transmitter and the kth user; $\boldsymbol{V}_k \in \mathbb{C}^{N_t\times N_{r,k}}$ denotes the transmit beamforming matrix for the kth user, which is a cascade of two precoding matrices $\boldsymbol{B}_k$ and $\boldsymbol{D}_k$ i.e.,

$$\boldsymbol{V}_k = \boldsymbol{B}_k\boldsymbol{D}_k \tag{3.31}$$

where $\boldsymbol{B}_k \in \mathbb{C}^{N_t\times N_{r,k}}$ removes the interuser interference and $\boldsymbol{D}_k \in \mathbb{C}^{N_{r,k}\times N_{r,k}}$ is used for parallelizing and power allocation. $\boldsymbol{B}_k$ is chosen such that the subspace spanned by its columns lies in the null space of $\boldsymbol{H}_l(\forall l \neq k)$ that is, $\boldsymbol{H}_l\boldsymbol{B}_k = 0$ for $l = 1, ..., k-1, k+1, ..., K$. If we define $\tilde{\boldsymbol{H}}_k$ as

$$\tilde{\boldsymbol{H}}_k = [\boldsymbol{H}_1^T \cdots \boldsymbol{H}_{k-1}^T\boldsymbol{H}_{k+1}^T \cdots \boldsymbol{H}_K^T]^T \tag{3.32}$$

then $\boldsymbol{B}_k$ can be obtained from the null space of $\tilde{\boldsymbol{H}}_k$. Note that the SVD of $\tilde{\boldsymbol{H}}_k$ is

$$\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{U}}_k\tilde{\boldsymbol{\Lambda}}_k[\tilde{\boldsymbol{V}}_k^{(1)}\tilde{\boldsymbol{V}}_k^{(0)}]^H \tag{3.33}$$

where $\tilde{\boldsymbol{U}}_k$ and $\tilde{\boldsymbol{\Lambda}}_k$ denote the left singular matrix and the matrix of ordered singular values of $\tilde{\boldsymbol{H}}_k$, respectively. $\tilde{\boldsymbol{V}}_k^{(1)}$ and $\tilde{\boldsymbol{V}}_k^{(0)}$ denote the right singular matrices each consisting of the singular vectors corresponding to non-zero singular values and zero singular values, respectively. To cancel the interference of the other users, we choose the precoder $\boldsymbol{B}_k$ as the first $N_{r,k}$ columns of $\tilde{\boldsymbol{V}}_k^{(0)}$, i.e.,

$$\boldsymbol{B}_k = (\tilde{\boldsymbol{V}}_k^{(0)})_{1:N_{r,k}} \tag{3.34}$$

where $N_{r,k}$ is the number of transmit data streams for the kth user. Notice that it is also the number of receive antennas of the kth user. The received signal $\boldsymbol{y}_k$ is given by

$$\boldsymbol{y}_k = \boldsymbol{H}_{eff,k}\boldsymbol{D}_k\boldsymbol{x}_k + \boldsymbol{n}_k \tag{3.35}$$

where $\boldsymbol{H}_{eff,k} = \boldsymbol{H}_k\boldsymbol{B}_k \in \mathbb{C}^{N_{r,k}\times N_{r,k}}$ denotes the effective channel of the kth user and the size of $\boldsymbol{x}_k$ is $N_{r,k}\times 1$. Since the kth user receives its own data stream without any interference from other users, to achieve the highest sum rate, Water-filling algorithm is adopted to allocate the transmit power to the data streams. Define the SVD of $\boldsymbol{H}_{eff,k}$

$$\boldsymbol{H}_{eff,k} = \boldsymbol{U}_k[\boldsymbol{\Lambda}_k 0][\boldsymbol{V}_k^{(1)}\boldsymbol{V}_k^{(0)}]^H \tag{3.36}$$

where $\boldsymbol{V}_k^{(1)}$ denotes the set of the right singular vectors corresponding to non-zero singular values and $\boldsymbol{U}_k$ is the left singular matrix. Taking $\boldsymbol{D}_k = \boldsymbol{V}_k^{(1)}\boldsymbol{P}_k^{\frac{1}{2}}$, we have

$$\boldsymbol{V}_k = (\tilde{\boldsymbol{V}}_k^{(0)})_{1:N_{r,k}}\boldsymbol{V}_k^{(1)}\boldsymbol{P}_k^{\frac{1}{2}} \tag{3.37}$$

where $\boldsymbol{P}_k$ denotes a diagonal matrix whose elements scale the power transmitted into each of the column of $\boldsymbol{V}_k^{(1)}$. At the receiver, $\boldsymbol{U}_k$ is used as the receive combining matrix to decode the received signal. The maximum achievable sum rate of the BD method is given by

$$C = \max \log|\boldsymbol{I} + \frac{\boldsymbol{\Lambda}^2\boldsymbol{P}}{\sigma_n^2}| \tag{3.38}$$

where $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\Lambda}_1, \cdots, \boldsymbol{\Lambda}_K)$, $\boldsymbol{P} = \text{diag}(\boldsymbol{P}_1, \cdots, \boldsymbol{P}_K)$.

The pseudo code of BD method is presented in Algorithm 6, and Figure 3.11 gives the simulation results of the sum rates. It is shown that even the receive antenna number of each user is extended to multiple, the noise enhancing problem still exists. Therefore, the gap between the achievable sum rate of BD method and DPC method is still large.

To decode the received signal, knowledge of the decoding matrix $\boldsymbol{U}_k$ is required at each receiver. The decoding matrix $\boldsymbol{U}_k$ depends, on $\boldsymbol{H}_{eff,k}$, but $\boldsymbol{H}_{eff,k}$ also consists of the original channel matrix $\boldsymbol{H}_k$ and the nulling matrix $\tilde{\boldsymbol{V}}_k^{(0)}$. Because the nulling matrix is calculated by using partial information about the CSI of other users, the receiver needs to either calculate the decoding matrix directly from the estimated channel of $\boldsymbol{H}_{eff,k}$ or the transmitter can send specific information to calculate $\boldsymbol{U}_k$ at the kth receiver, which is call a coordination information [19].

3.4.3 CB method

Another approach, unlike ZF, allows a larger number of receiver antennas than the number of data streams for each receiver. We refer this method as CB method. CB method also enforces a zero interference property like channel inversion but requires an iterative optimization procedure to find the transmit beamforming and receive combining vectors [19].

Algorithm 6 Pseudo code of BD method

for $k = 1 : K$ **do**
  Build $\tilde{\boldsymbol{H}}_k$ (3.32)
  Obtain $\tilde{\boldsymbol{V}}_k^{(0)}$ by using the SVD operation (3.33)
  Calculate $\boldsymbol{H}_{eff,k} = \boldsymbol{H}_k \tilde{\boldsymbol{V}}_k^{(0)}$
  Obtain the first $N_{r,k}$ singular values $\lambda_{k,1}, \cdots, \lambda_{k,N_{r,k}}$ and the corresponding singular vector $\boldsymbol{V}_k^{(1)}$ by using the SVD operation on $\boldsymbol{H}_{eff,k}$
  Do water-filling algorithm to determine the optimal power distribution matrices $\boldsymbol{P}_k$
  Calculate $\boldsymbol{V}_k = (\tilde{\boldsymbol{V}}_k^{(0)})_{1:N_{r,k}} \boldsymbol{V}_k^{(1)} \boldsymbol{P}_k^{\frac{1}{2}}$ (3.37)
end for

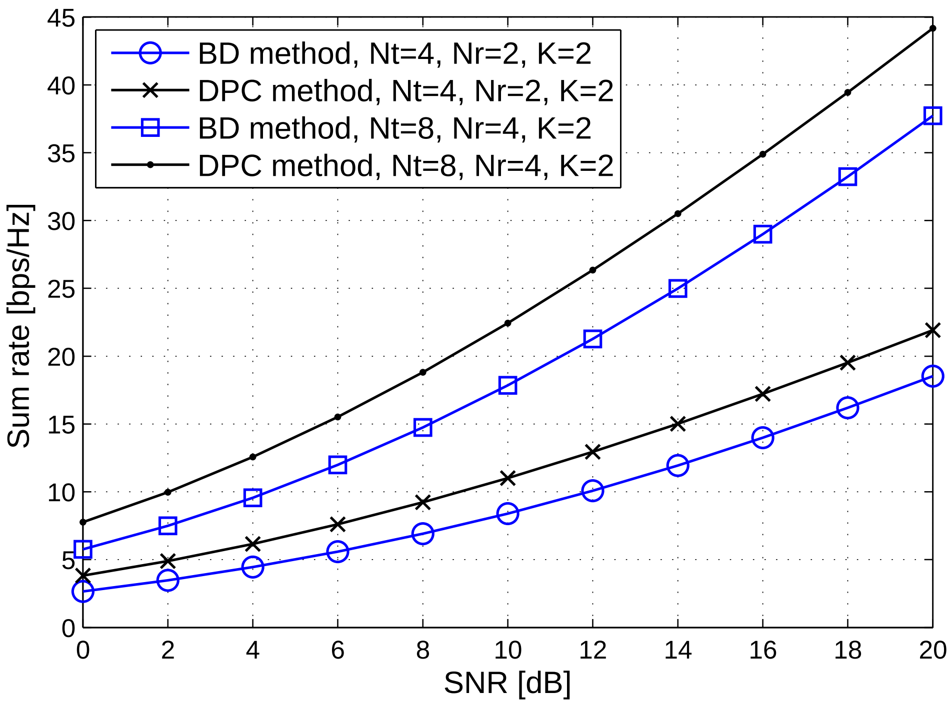

Figure 3.11: Comparison between the achievable sum rate of BD method and the sum capacity.

The system model for CB method is designed with N_t antennas at the transmitter and N_r receiver antennas for each user. The number of users must be less or equal to the number of transmit antennas. The received signal at the kth user after receive combining vector is

$$y_k = \boldsymbol{u}_k^H \boldsymbol{H}_k \boldsymbol{v}_k x_k + \boldsymbol{u}_k^H \boldsymbol{H}_k \sum_{l=1, l\neq k}^{K} \boldsymbol{v}_l x_l + \boldsymbol{u}_k^H \boldsymbol{n}_k \tag{3.39}$$

where $\boldsymbol{v}_k$ and $\boldsymbol{u}_k$ denotes the transmit beamforming and receive combining vectors for the kth user, respectively. In CB method, the base station chooses $\boldsymbol{v}_k$ in the null space of the space spanned by $\boldsymbol{u}_l^H \boldsymbol{H}_l (\forall l \neq k)$, that is, $\boldsymbol{u}_l^H \boldsymbol{H}_l \boldsymbol{v}_k = 0$ for $l = 1, \cdots, k-1, k+1, \cdots, K$. Then $\boldsymbol{v}_k$ will cause zero interference to other users. Specifically, CB method forms an equivalent channel matrix for the kth user

$$\tilde{\boldsymbol{H}}_k = [\tilde{\boldsymbol{h}}_1 \cdots \tilde{\boldsymbol{h}}_{k-1} \tilde{\boldsymbol{h}}_{k+1} \cdots \tilde{\boldsymbol{h}}_K]^H \tag{3.40}$$

where $\tilde{\boldsymbol{h}}_i^H = \boldsymbol{u}_i^H \boldsymbol{H}_i$, and then finding a transmit beamforming vector $\boldsymbol{v}_k$ that satisfies $\tilde{\boldsymbol{H}}_k \boldsymbol{v}_k = 0$.

Assuming that $K = N_t$ and that the channels are sufficiently rich, $\tilde{\boldsymbol{H}}_k$ will be full-rank and of dimension $(K-1) \times K$, the null-space has dimension one and there is only one zero singular value. Define the SVD of $\tilde{\boldsymbol{H}}_k$ as

$$\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{U}}_k \tilde{\boldsymbol{\Lambda}}_k [\tilde{\boldsymbol{V}}_k^{(1)} \tilde{\boldsymbol{v}}_k^{(0)}]^H \tag{3.41}$$

where $\tilde{\boldsymbol{U}}_k$ and $\tilde{\boldsymbol{\Lambda}}_k$ denote the left singular matrix and the matrix of singular values of $\tilde{\boldsymbol{H}}_k$, respectively, and $\tilde{\boldsymbol{V}}_k^{(1)}$ and $\tilde{\boldsymbol{v}}_k^{(0)}$ are the right singular matrix and vector each corresponding to non-zero singular values and zero singular value, respectively. The transmit beamforming vector of the kth user should lie in the space spanned by $\tilde{\boldsymbol{v}}_k^{(0)}$, consequently, we take $\boldsymbol{v}_k = \tilde{\boldsymbol{v}}_k^{(0)}$. If maximum ratio combining is used at the receiver, the receive combining vector is taken as $\boldsymbol{u}_k = \boldsymbol{H}_k \boldsymbol{v}_k$, then $\boldsymbol{v}_k$ (or $\boldsymbol{u}_k$) is optimized iteratively under the assumption that $\boldsymbol{u}_k$ (or $\boldsymbol{v}_k$) is fixed.

Algorithm 7 gives the pseudo code of CB method, and the simulation results are given in Figure 3.12. We can see that with the number of transmit antennas increasing, the sum rate gap between CB method and DPC method becomes larger. This is mainly because only one data stream is assigned to each user, and the spatial multiplexing is not fully exploited.

Algorithm 7 Pseudo code of CB method

Initialize the receive combining vector $\boldsymbol{u}_k$ $\forall k$ as random unitary vector
$i = 0$
Repeat $i \leftarrow i + 1$
for $k = 1 : K$ **do**
  $\tilde{\boldsymbol{H}}_k = [\tilde{\boldsymbol{h}}_1 \cdots \tilde{\boldsymbol{h}}_{k-1}\tilde{\boldsymbol{h}}_{k+1} \cdots \tilde{\boldsymbol{h}}_K]^H$
  $\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{U}}_k \tilde{\boldsymbol{\Lambda}}_k [\tilde{\boldsymbol{V}}_k^{(1)} \tilde{\boldsymbol{v}}_k^{(0)}]^H$
  $\boldsymbol{v}_k(i) = \boldsymbol{v}_k^{(0)}$
  $\boldsymbol{u}_k(i) = \boldsymbol{H}_k \boldsymbol{v}_k(i)$
end for
Until $\| \boldsymbol{v}_k(i) - \boldsymbol{v}_k(i-1) \| < \epsilon$
$\boldsymbol{v}_k = \boldsymbol{v}_k(i)$
$\boldsymbol{u}_k = \boldsymbol{u}_k(i)$

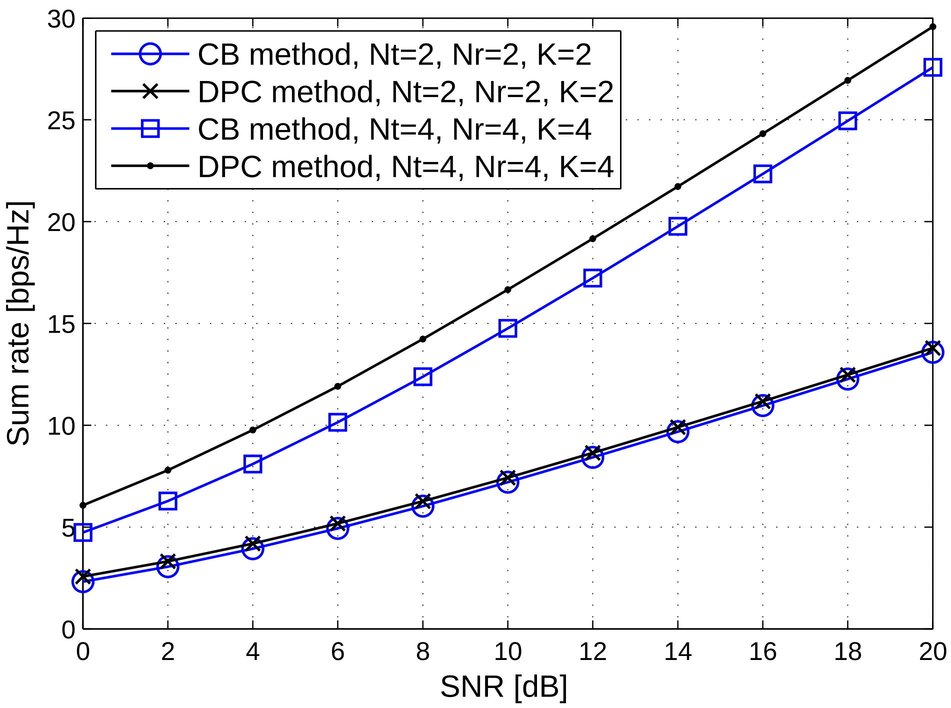

Figure 3.12: Comparison between the achievable sum rate of CB method and the sum capacity.

3.4.4 ZF-SA method

ZF-SA method proposed in [20] also uses receive combining technique, and extends MU-MIMO broadcast channels to a more general case. This method allocates data streams successively to the users, at each step one data stream is assigned to the user who brings the largest increase of the global throughput. For presentation convenience and without loss of generality, we consider each data stream separately. The lth $(1 \leq l \leq L)$ data stream after MU-MIMO broadcast channels and the receive combining vector is [64]

$$y_l = \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l x_l + \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \sum_{\substack{j=1 \\ j \neq l}}^{K} \boldsymbol{v}_j x_j + \boldsymbol{u}_l^H \boldsymbol{n}_{\pi(l)} \tag{3.42}$$

where $\pi(l)$ indicates that the lth data stream is allocated to the $\pi(l)$th user. The first data stream is assigned to the user who has the largest data rate. For the lth $(2 \leq l \leq L)$ data stream allocation, we assume that the receive beamforming vectors of the previously allocated data streams are fixed, the original sum rate optimization problem is approximated to a concave generalized eigenvalue problem, and each possible receive beamforming vector $\boldsymbol{u}_{(l)}(k)$ for the lth data stream is calculated over all the K users [73], i.e.,

$$C_l(k) \geq \log_2 \left(1 + \frac{P_T}{\left\| \begin{bmatrix} \overline{\boldsymbol{H}}_{l-1} \\ \boldsymbol{u}_l(k)^H \boldsymbol{H}_k \end{bmatrix}^{(\dagger)} \right\|^2} \right) \tag{3.43}$$

where $C_l(k)$ is the temporary sum rate if the lth data stream is allocated to the kth user; and $\overline{\boldsymbol{H}}_{l-1}$ is defined as

$$\overline{\boldsymbol{H}}_{l-1} = \begin{bmatrix} \boldsymbol{u}_1^H \boldsymbol{H}_{\pi(1)} \\ \vdots \\ \boldsymbol{u}_{l-1}^H \boldsymbol{H}_{\pi(l-1)} \end{bmatrix} \tag{3.44}$$

In order to suppress the interference, the transmit beamforming vectors are selected from the pseudo inverse of $\overline{\boldsymbol{H}}_{l-1}$. By using the lower bound (3.43), the sum rate optimization can be rewritten as

$$\begin{aligned}
&\min_{\boldsymbol{u}_l(k)} \left\| \begin{bmatrix} \overline{\boldsymbol{H}}_{l-1} \\ \boldsymbol{u}_l(k)^H \boldsymbol{H}_k \end{bmatrix}^{(\dagger)} \right\|^2 \\
&\text{subject to} \quad \boldsymbol{u}_l(k)^H \boldsymbol{u}_l(k) = 1 \\
&\qquad\qquad \boldsymbol{u}_j{}^H \boldsymbol{u}_l(k) = 0, \forall j < l \text{ for } \pi(j) = k
\end{aligned} \tag{3.45}$$

Using the successive update of the pseudo inverse with LQ decomposition of matrix $\overline{\boldsymbol{H}}_{l-1} = \boldsymbol{L}_{l-1}\boldsymbol{Q}_{l-1}^H$, we have [74]

$$\begin{aligned}
\left\| \begin{bmatrix} \overline{\boldsymbol{H}}_{l-1} \\ \boldsymbol{u}_l(k)^H \boldsymbol{H}_k \end{bmatrix}^{(\dagger)} \right\|^2 &= \text{trace}((\boldsymbol{L}_{l-1}^{-1})^H \boldsymbol{L}_{l-1}^{-1}) \\
&+ \frac{1 + \boldsymbol{u}_l(k)^H \boldsymbol{H}_k \boldsymbol{Q}_{l-1} \boldsymbol{L}_{l-1}^{-1} (\boldsymbol{L}_{l-1}^{-1})^H \boldsymbol{Q}_{l-1}^H \boldsymbol{H}_k^H \boldsymbol{u}_l(k)}{\boldsymbol{u}_l(k)^H \boldsymbol{H}_k (\boldsymbol{I} - \boldsymbol{Q}_{l-1}\boldsymbol{Q}_{l-1}^H) \boldsymbol{H}_k^H \boldsymbol{u}_l(k)}
\end{aligned} \tag{3.46}$$

As the matrix $\boldsymbol{L}_{l-1}$ is independent of the index k and the receive combining vector $\boldsymbol{u}_l(k)$, the optimization of the receive combining vector for the kth user reduces to

$$\begin{aligned}
&\max_{\boldsymbol{u}_l(k)} \frac{\boldsymbol{u}_l(k)^H \boldsymbol{H}_k (\boldsymbol{I} - \boldsymbol{Q}_{l-1}\boldsymbol{Q}_{l-1}^H) \boldsymbol{H}_k^H \boldsymbol{u}_l(k)}{\boldsymbol{u}_l(k)^H (\boldsymbol{I} + \boldsymbol{H}_k \boldsymbol{Q}_{l-1} \boldsymbol{L}_{l-1}^{-1} (\boldsymbol{L}_{l-1}^{-1})^H \boldsymbol{Q}_{l-1}^H \boldsymbol{H}_k^H) \boldsymbol{u}_l(k)} \\
&\text{subject to} \quad \boldsymbol{u}_l(k)^H \boldsymbol{u}_l(k) = 1 \\
&\qquad\qquad \boldsymbol{u}_j{}^H \boldsymbol{u}_l(k) = 0, \forall j < l \text{ for } \pi(j) = k
\end{aligned} \tag{3.47}$$

The objective function in (3.47) is maximized by choosing $\boldsymbol{u}_l(k)$ to be the generalized eigenvector belonging to the principal generalized eigenvalue of the matrix pair $\boldsymbol{H}_k(\boldsymbol{I} - \boldsymbol{Q}_{l-1}\boldsymbol{Q}_{l-1}^H)\boldsymbol{H}_k^H$ and $\boldsymbol{I} + \boldsymbol{H}_k\boldsymbol{Q}_{l-1}\boldsymbol{L}_{l-1}^{-1}(\boldsymbol{L}_{l-1}^{-1})^H\boldsymbol{Q}_{l-1}^H\boldsymbol{H}_k^H$. In order to suppress the interference, each transmit beamforming vector $\boldsymbol{v}_j$ $(1 \leq j \leq l)$ should be orthogonal to row vectors $\boldsymbol{u}_i^H \boldsymbol{H}_{\pi(i)}$ $(\forall i \neq j)$. Therefore, the transmit beamforming vectors of the already allocated data streams can be obtained via the pseudo inverse of

$$\overline{\boldsymbol{H}}_l(k) = \begin{bmatrix} \overline{\boldsymbol{H}}_{l-1} \\ \boldsymbol{u}_l(k)^H \boldsymbol{H}_k \end{bmatrix} \tag{3.48}$$

Finally, the lth data stream is assigned to the user who contributes the largest increase of the total throughput with the previously selected users. The pseudo code of ZF-SA method is given in Algorithm 8.

ZF-SA method can remove the entire interference without any constraint on the number of receive antennas. As far as we know, ZF-SA method has the best performance in terms of sum rate in the literature (Figure 3.13). However, compared with other methods (e.g. BD method and CB method) under zero-forcing constraint, the computational complexity of ZF-SA is significant.

Algorithm 8 Pseudo code of ZF-SA method

```
Initialization: l = 0
while l ≤ N_t do
    l ← l + 1
    for k = 1 : K do
        Solve (3.47) to find the receive combining vector
        Do the pseudo inverse of H̄_l(k) to find the transmit beamforming vectors
        Perform water-filling power allocation
        Calculate the temporary sum, denoted as R_k
    end for
    Calculate the largest sum rate [C_l, π(l)] = max(R_k)
    if C_{l-1} ≥ C_l then
        π(l) = 0
        break
    end if
end while
```

3.4.5 MMSE method

Note that the aforementioned methods (ZF, BD, CB and ZF-SA) aim to cancel all the interference, meanwhile, the global throughput is optimized. But sometimes interference is not vital to remove, especially at low SNR region. In this case, Gaussian noise plays a more important role. In this subsection, MMSE method is introduced, which seeks to maximize the SINR of each data stream instead of removing the interference. Consider the downlink transmission, the received signal $\boldsymbol{y}_k \in \mathbb{C}^{N_{r,k}\times 1}$ at the kth user after receive combining matrix can be written as [21], [29]

$$\boldsymbol{y}_k = \boldsymbol{U}_k^H \boldsymbol{H}_k \boldsymbol{V}_k \sqrt{\boldsymbol{P}_k} \boldsymbol{x}_k + \boldsymbol{U}_k^H \boldsymbol{H}_k \Big(\sum_{i=1, i\neq k}^{K} \boldsymbol{V}_i \sqrt{\boldsymbol{P}_i} \boldsymbol{x}_i \Big) + \boldsymbol{U}_k^H \boldsymbol{n}_k \tag{3.49}$$

where $\boldsymbol{H}_k \in \mathbb{C}^{N_{r,k}\times N_t}$ denotes the channel between the transmitter and the kth user; $\boldsymbol{V}_k \in \mathbb{C}^{N_t \times N_{r,k}}$ and $\boldsymbol{U}_k \in \mathbb{C}^{N_{r,k}\times N_{r,k}}$ are the normalized transmit beamforming and the receive combining matrices of the kth user, respectively; the diagonal matrix $\boldsymbol{P}_k^M$

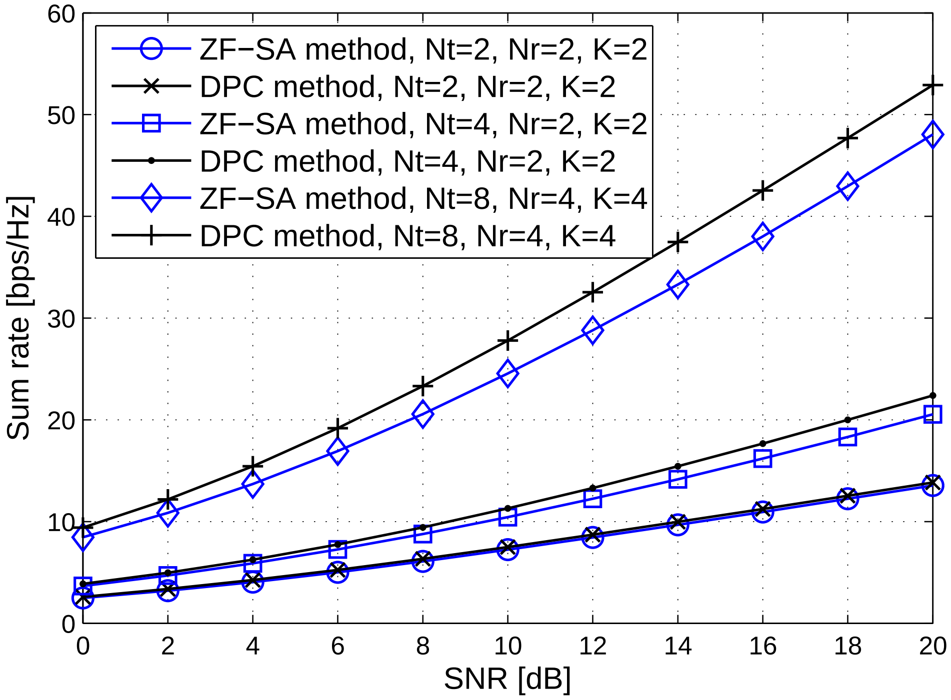

Figure 3.13: Comparison between the achievable sum rate of ZF-SA method and the sum capacity.

denotes the corresponding transmit power. For presentation convenience and without loss of generality, we consider each data stream separately. The lth $(1 \leq l \leq L)$ data stream after MU-MIMO broadcast channels and the receive combining vector is [64]

$$y_l = \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l \sqrt{p_l} x_l + \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \sum_{\substack{j=1 \\ j \neq l}}^{L} \boldsymbol{v}_j \sqrt{p_j} x_j + \boldsymbol{u}_l^H \boldsymbol{n}_{\pi(l)} \tag{3.50}$$

where $\pi(l)$ and p_l indicate that the lth data stream is allocated to the $\pi(l)$th user and the allocated power for the $\pi(l)$th user, respectively. Note that the interference is not canceled completely, water-filling power allocation method does not work here. Similarly to [21] and [30], the sum rate optimization problem is given by

$$\begin{aligned} \max_{\{\boldsymbol{u}_l, \boldsymbol{v}_l, p_l\}_{l=1}^L} & \sum_{l=1}^{L} \log_2 \Big(1 + \frac{p_l |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l|^2}{\sum_{j=1, j \neq l}^{L} p_j |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_j|^2 + \sigma^2}\Big) \\ \text{subject to} \quad & \| \boldsymbol{v}_l \| = 1, \ \| \boldsymbol{u}_l \| = 1, \ 1 \leq l \leq L \\ & \sum_{l=1}^{L} p_l \leq P_T, \ p_l \geq 0, \ 1 \leq l \leq L \end{aligned} \tag{3.51}$$

It can be observed from (3.51) that for the lth data stream, the transmit beamforming vector $\boldsymbol{v}_l$ and power p_l are difficult to optimize, since $\boldsymbol{v}_l$ and p_l are coupled with $\boldsymbol{v}_j$ and p_j respectively, these are non-convex problems. However, the optimization problem with respect to receive combining vector $\boldsymbol{u}_l$ can be solved efficiently if the transmit beamforming vectors and transmit power allocation are fixed. In this case, the optimization problem with respect to the receive combining vector $\boldsymbol{u}_l$ is

$$\begin{aligned} \max_{\{\boldsymbol{u}_l\}_{l=1}^L} & \sum_{l=1}^{L} \log_2 \Big(1 + \frac{\boldsymbol{u}_l^H \boldsymbol{S}(l) \boldsymbol{u}_l}{\boldsymbol{u}_l^H \boldsymbol{T}(l) \boldsymbol{u}_l}\Big) \\ & \text{subject to} \ \| \boldsymbol{u}_l \| = 1, \ 1 \leq l \leq L. \end{aligned} \tag{3.52}$$

where $\boldsymbol{S}(l) = p_l \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l \boldsymbol{v}_l^H \boldsymbol{H}_{\pi(l)}^H$, and $\boldsymbol{T}(l) = \sum_{\substack{j=1 \\ j \neq l}}^{L} p_j \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_j \boldsymbol{v}_j^H \boldsymbol{H}_{\pi(l)}^H + \sigma^2 \boldsymbol{I}_{\pi(l)}$. In (3.52), the log function is monotone increasing, the optimization problem is equivalent to maximizing each SINR term separately, which is a standard generalized eigenvalue problem, that is

$$\text{SINR}_l = \frac{\boldsymbol{u}_l^H \boldsymbol{S}(l) \boldsymbol{u}_l}{\boldsymbol{u}_l^H \boldsymbol{T}(l) \boldsymbol{u}_l} \tag{3.53}$$

Clearly, the SINR value of the lth data stream SINR_l is maximized by choosing $\boldsymbol{u}_l$ as the generalized eigenvector of $(\boldsymbol{S}(l),\ \boldsymbol{T}(l))$ corresponding to the dominant generalized eigenvalue, i.e.,

$$\hat{\boldsymbol{u}}_l = \left(\sum_{\substack{j=1\\ j\neq l}}^{L} p_j \boldsymbol{H}_{\pi(l)}\boldsymbol{v}_j\boldsymbol{v}_j^H\boldsymbol{H}_{\pi(l)}^H + \sigma^2\boldsymbol{I}_{\pi(l)}\right)^{-1} \boldsymbol{H}_{\pi(l)}\boldsymbol{v}_l \tag{3.54}$$

with $\boldsymbol{u}_l = \hat{\boldsymbol{u}}_l / \|\hat{\boldsymbol{u}}_l\|_F$. This beamforming technique is also known as the minimum mean square error (MMSE) receiver [21], [30].

Next we calculate the transmit beamforming matrix in the dual MAC. According to BC-MAC duality, the received signal after receiving combining matrix in dual MAC is

$$\boldsymbol{y}_k^M = \boldsymbol{V}_k^H\boldsymbol{H}_k^H\boldsymbol{U}_k\sqrt{\boldsymbol{P}_k^M}\boldsymbol{x}_k^M + \boldsymbol{V}_k^H(\sum_{j=1,j\neq k}^{K}\boldsymbol{H}_j^H\boldsymbol{U}_j\sqrt{\boldsymbol{P}_j^M}\boldsymbol{x}_j^{\mathrm{M}}) + \boldsymbol{V}_k^H\boldsymbol{n}^M \tag{3.55}$$

where ${\boldsymbol{H}_k}^H \in \mathbb{C}^{N_t\times N_{r,k}}$ denotes the channel between the kth user and the base station; $\boldsymbol{U}_k \in \mathbb{C}^{N_t\times N_{r,k}}$ and $\boldsymbol{V}_k \in \mathbb{C}^{N_{r,k}\times N_{r,k}}$ are the normalized transmit beamforming and the receive combining matrices, respectively; the diagonal matrix $\boldsymbol{P}_k^M$ is the corresponding power allocation in the dual MAC similarly to $\boldsymbol{P}_k$. We consider the transmit beamforming vector design while receive combining vectors and allocated power are fixed. As it can be observed in (3.51), the transmit beamforming vector design problem is non-convex and the solution is difficult to find. However, it can be transformed into a generalized eigenvalue problem taking advantage of the uplink-downlink duality. According to (3.49) and (3.55), the transmit beamforming matrix $\boldsymbol{V}_k$ of the kth user in the downlink corresponds to the receive combining matrix $\boldsymbol{V}_k$ in the dual MAC, then the sum rate optimization of the dual MAC is

$$\begin{aligned} &\max_{\{\boldsymbol{u}_l,\boldsymbol{v}_l,{p_l}^M\}_{l=1}^L} \sum_{l=1}^{L}\log_2(1+\frac{{p_l}^M|\boldsymbol{v}_l^H\boldsymbol{H}_{\pi(l)}^H\boldsymbol{u}_l|^2}{\sum_{j=1,j\neq l}^{L}{p_j}^M|\boldsymbol{v}_l^H\boldsymbol{H}_{\pi(j)}^H\boldsymbol{u}_j|^2+\sigma^2}) \\ &\text{subject to} \quad \|\ \boldsymbol{u}_l\ \|=1,\ \|\ \boldsymbol{v}_l\ \|=1,\ 1\leq l\leq L \\ &\qquad \sum_{l=1}^{L}{p_l}^M \leq P_T,\ q_l\geq 0,\ 1\leq l\leq L \end{aligned} \tag{3.56}$$

In the dual MAC, similarly to (3.51), we suppose that the transmit beamforming vectors and transmit power allocation are fixed. Then (3.56) is reformulated as

$$\max_{\{\boldsymbol{v}_l\}_{l=1}^L} \sum_{l=1}^{L} \log_2(1 + \frac{\boldsymbol{v}_l^H \boldsymbol{S}^M(l) \boldsymbol{v}_l}{\boldsymbol{v}_l^H \boldsymbol{T}^M(l) \boldsymbol{v}_l}) \quad \text{subject to } \| \boldsymbol{v}_l \| = 1,\ 1 \leq l \leq L \tag{3.57}$$

where $\boldsymbol{S}^M(l) = q_l \boldsymbol{H}_{\pi(l)}^H \boldsymbol{u}_l \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)}$, and $\boldsymbol{T}^M(l) = \sum_{j=1,j\neq l}^{L} q_j \boldsymbol{H}_{\pi(j)}^H \boldsymbol{u}_j \boldsymbol{u}_j^H \boldsymbol{H}_{\pi(j)} + \sigma^2 \boldsymbol{I}_{N_t}$. Similarly to (3.52), the optimal transmit beamforming vector $\boldsymbol{v}_l$ is the generalized eigenvector of $(\boldsymbol{S}^M(l),\ \boldsymbol{T}^M(l))$ corresponding to the dominant generalized eigenvalue, i.e.,

$$\hat{\boldsymbol{v}}_l = \left(\sum_{j=1,j\neq l}^{L} q_j \boldsymbol{H}_{\pi(j)}^H \boldsymbol{u}_j \boldsymbol{u}_j^H \boldsymbol{H}_{\pi(j)} + \sigma^2 \boldsymbol{I}_{N_t} \right)^{-1} \boldsymbol{H}_{\pi(l)}^H \boldsymbol{u}_l \tag{3.58}$$

with normalization $\boldsymbol{v}_l = \hat{\boldsymbol{v}}_l / \|\hat{\boldsymbol{v}}_l\|_F$.

If the transmit and receive beamforming vectors obtained as described above are supposed to be fixed, the power allocation optimization problem can be written as

$$\max_{\{p_l\}} \sum_{l=1}^{L} \log_2(1 + \frac{p_l G_{ll}}{\sum_{j=1,j\neq l}^{L} p_j G_{lj} + \sigma^2}) \quad \text{subject to } \sum_{l=1}^{L} p_l \leq P_T,\ p_l \geq 0, 1 \leq l \leq L \tag{3.59}$$

where $G_{lj} = |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_j|^2$ and $G_{ll} = |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l|^2$ represent the coefficient of the interference to the lth subchannel caused by the jth subchannel and the channel gain of the lth subchannel, respectively. This optimization problem is a NP hard problem and the optimal solution needs solving geometric programming (GP) iteratively [21], [75]. The pretty high computational complexity makes it hard to implement in practice. In chapter 5, we will introduce the proposed suboptimal method to solve (3.59), which has a much low computational complexity with a negligible performance degradation.

Algorithm 9 gives the pseudo code of MMSE method, and Figure 3.14 shows the sum rate performance. We can see that the achievable sum rate provided by MMSE method is very close to the sum capacity in low SNR region, while with SNR increase, the performance degrades significantly. The reason is that, in high SNR region, the interuser interference is stronger than Gaussian noise, while MMSE method does not take it into account, which results in performance degradation.

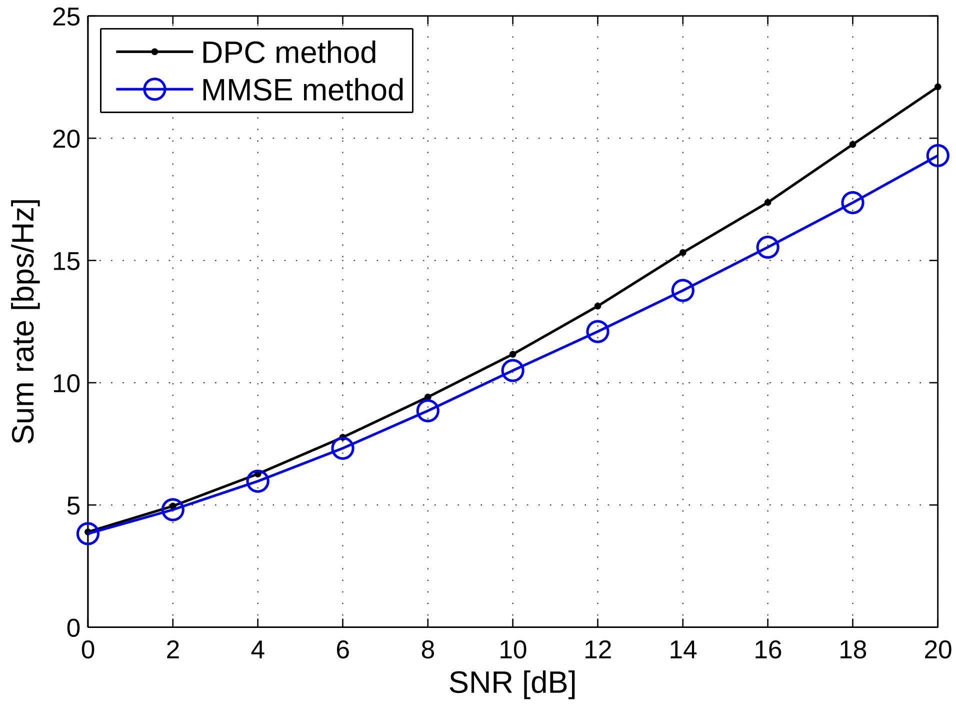

Figure 3.14: Comparison between the achievable sum rate of MMSE method and the sum capacity,

---

Algorithm 9 Pseudo code of MMSE method

Initialization: $[\boldsymbol{U}_k, \boldsymbol{\Sigma}_k, \boldsymbol{V}_k] = \text{svd}(\boldsymbol{H}_k), \ \forall k$

Repeat $i \leftarrow i + 1$

1, Downlink power allocation (3.59).

Suppose $\boldsymbol{V}_k$ and $\boldsymbol{U}_k$ are fixed, $\forall k$.

2, Downlink receive beamforming design (3.52).

Suppose power allocation and $\boldsymbol{V}_k$ are fixed, $\forall k$.

3, Uplink power allocation (3.59).

Suppose $\boldsymbol{V}_k$ and $\boldsymbol{U}_k$ are fixed, $\forall k$.

4, Uplink receive beamforming design (3.57).

Suppose power allocation and $\boldsymbol{U}_k$ are fixed, $\forall k$.

Until the sum rate convergence, i.e. $|C_i^{DL} - C_{i-1}^{DL}| \leq \epsilon$, or $i > i_{\max}$.

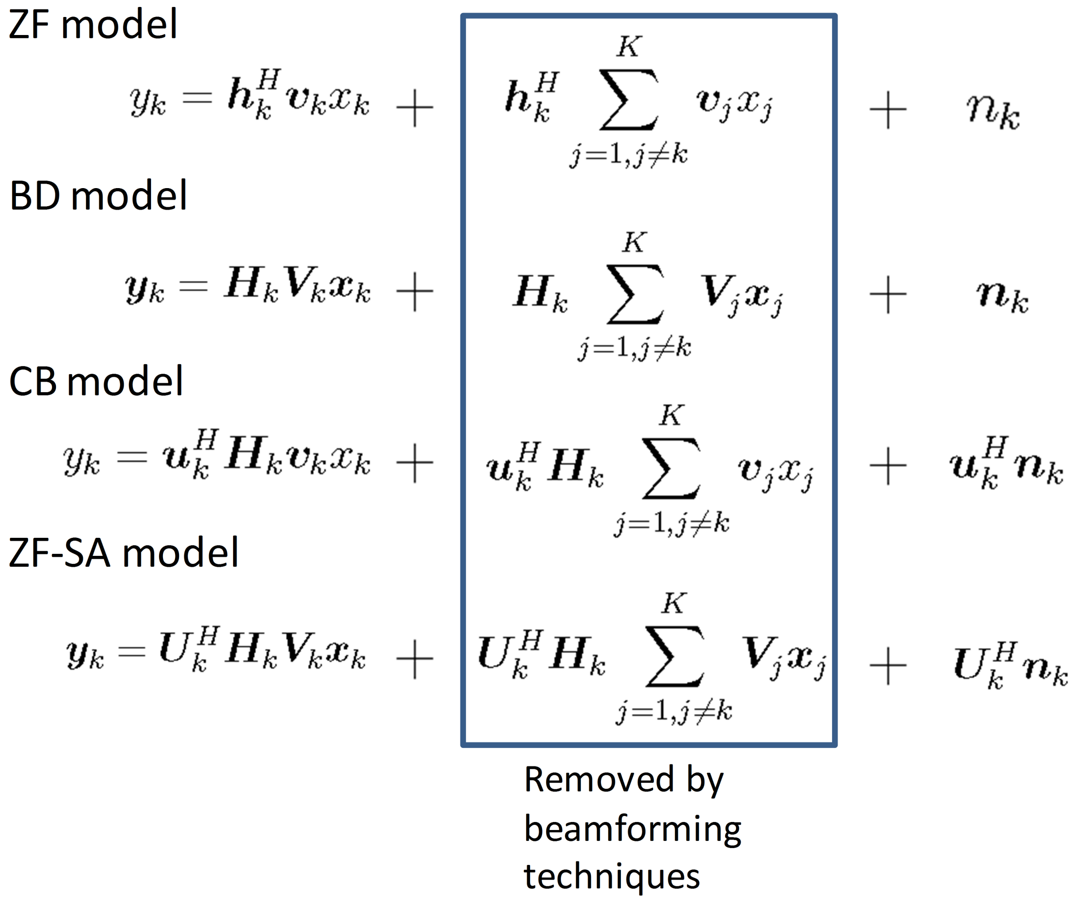

Figure 3.15: Interference removing techniques in linear methods

3.5 Summary

In Figure 3.15, we summarize the interference removing techniques in linear methods. Compared with non-linear methods, we can see that all the interference from other users is canceled by beamforming techniques.

3.6 Performance comparisons

In this section, we briefly give a summary of transmit beamforming techniques in the literature, and some simulation results show the performance comparison of different methods.

First, ZF method is the most simple method, one single receive antenna for each user is supported, and the transmit beamforming vector is obtained by the pseudo inverse of the build channel matrix. Since the goal of transmit beamforming vectors are

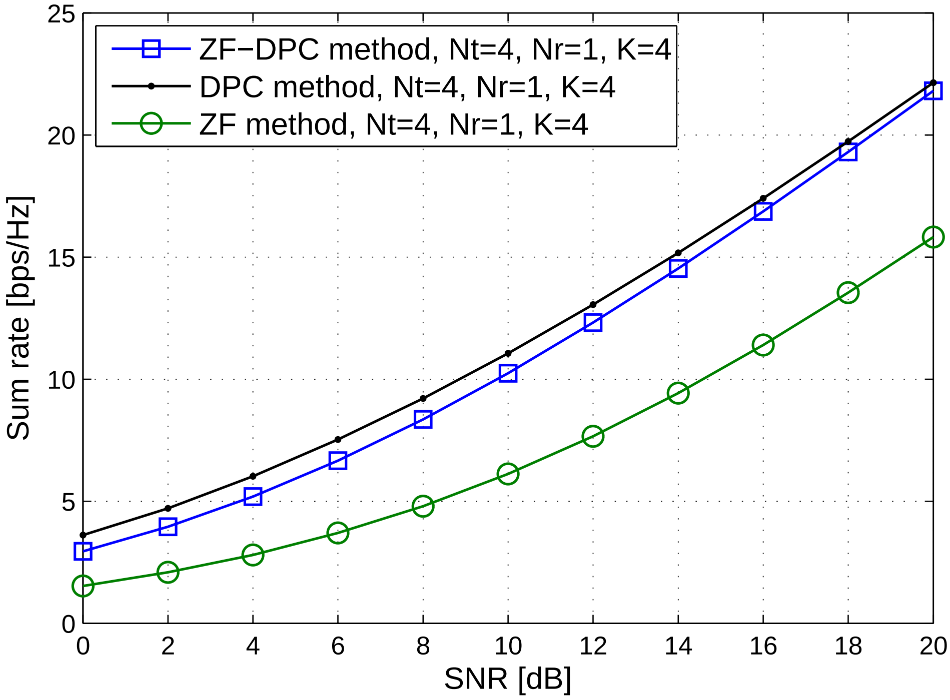

Figure 3.16: Sum rate comparison of ZF-DPC method, ZF method and the sum capacity.

to remove the entire interference, the number of degrees of freedom of transmit beamforming design is very limited. ZF-DPC extends this situation to the case where only one part of the interference is removed by transmit beamforming vectors, and the rest of the interference which is known non-causally to the transmitter can be eliminated by DPC technique. Thus, the number of degrees of freedom for choosing transmit beamforming vectors becomes larger, and the performance is improved. In Figure 3.16, the sum rate provided by ZF method and ZF-DPC method is given.

BD method and SZF-DPC method are extensions of ZF method and ZF-DPC method, respectively, both of them support the multiple receive antenna case. The beamforming vectors in BD method are used to remove the entire interference while in SZF-DPC method, only one part of interference is removed by transmit beamforming vectors. Similarly to Figure 3.16, in Figure 3.17, the simulation results show that SZF-DPC method outperforms BD method in terms of sum rate because of the extended number of degrees of freedom available for choosing transmit beamforming vector. Both ZF-SA method and SA-DPC method allocate the data streams successively. At each step, only one data stream is assigned to the user who can bring the largest sum rate increase. In addition, the total receive antenna number can be larger than transmit

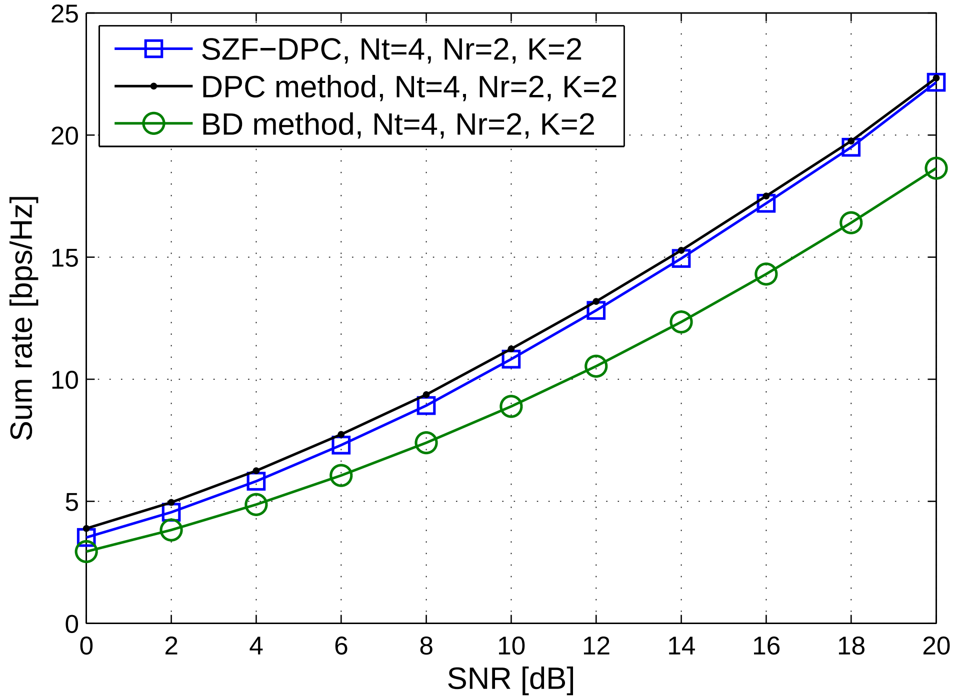

Figure 3.17: Sum rate comparison of SZF-DPC method, BD method and the sum capacity.

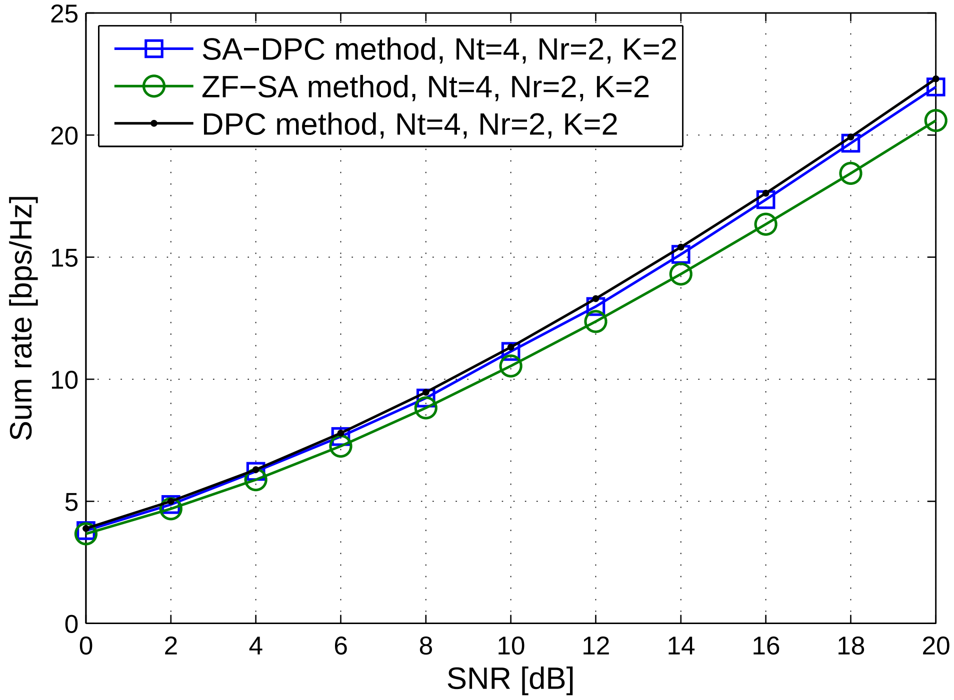

Figure 3.18: Sum rate comparison of SA-DPC method, ZF-SA method and the sum capacity.

antenna number. ZF-SA works for the zero-forcing case while SA-DPC works with DPC technique. In the Figure 3.18, we can also find that SA-DPC method has better performance in terms of sum rate compared with ZF-SA method. This comes from the larger number of degrees of freedom available for transmit beamforming design in SA-DPC method.

In Figure 3.19, the simulation results of all the methods in the literature with the configuration of $N_t = 4$ transmit antenna and $K = 2$ users with $N_r = 2$ receive antennas each are given. Generally, DPC based methods have better performance than zero-forcing methods, since the transmit beamforming vector has larger number of degrees of freedom if DPC technique is used. But to implement this DPC technique, the complexity of the structures of transmitter and receiver is also increased. We can also find that MMSE method has a pretty good performance at very low SNR region. The reason is that, the interference at this region is negligible. MMSE method considers it together with Gaussian noise, while in other methods, the transmit beamforming vectors are designed only to remove this interference.

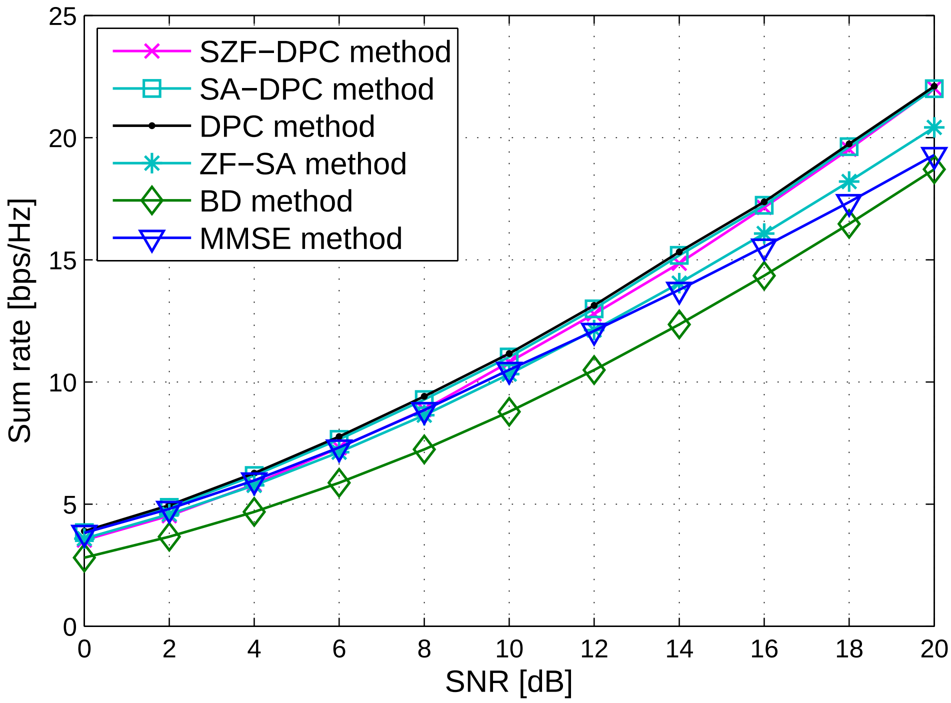

Figure 3.19: Sum rate comparison of SZF-DPC method, SA-DPC method, ZF-SA method, BD method, MMSE method and the sum capacity.

3.7 Beamforming techniques under per-antenna power constraint

As discussed in the introduction, in practice, the power amplifier of each antenna is limited individually by its linearity. A power constraint imposed on each transmit antenna is more realistic. In this section, we first introduce the optimal solution of transmit beamforming design under per-antenna power constraint, then a suboptimal solution is presented.

3.7.1 Per-OPT method

Similarly to the optimal solution under total power constraint in Section 3.1, Per-OPT method proposed in [27] also transforms the original non-convex sum rate optimization problem into a convex one. Taking advantage of the duality between BC and the corresponding MAC, using the existing standard optimization packages (e.g. CVX) and Lagrange duality method alternatively, the original problem converges to the optimal value eventually.

For convenience, we discuss the case in which the base station is equipped with $N_t = 2$ transmit antennas, and each antenna has a power constraint. Suppose K users are served, and each user has N_r receive antennas. If the per-antenna power constraint is considered and DPC technique is used, the sum rate optimization problem can be written as

$$\begin{aligned} \max_{\{\boldsymbol{Q}_k\}_{k=1}^K} & \sum_{k=1}^{K} \log_2 \frac{|\sigma^2 \boldsymbol{I} + \sum_{j=k}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|}{|\sigma^2 \boldsymbol{I} + \sum_{j=k+1}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|} \\ \text{subject to} & \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_1) \leq P_1 \\ & \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_2) \leq P_2 \end{aligned} \tag{3.60}$$

where $\boldsymbol{Q}_k = \boldsymbol{V}_k \boldsymbol{V}_k^H \in \mathbb{C}^{2\times 2}$ is the transmit covariance matrix of the kth user, and $\boldsymbol{A}_i$ $(1 \leq i \leq 2)$ is a diagonal matrix with only a 1 in the ith diagonal element and zeros elsewhere.

In order to use the duality between MU-MIMO broadcast channels and MAC channels, the auxiliary function $g(\lambda_1, \lambda_2)$ is introduced and (3.60) is reformed as

$$\begin{aligned} g(\lambda_1, \lambda_2) := \max_{\{\boldsymbol{Q}_k\}_{k=1}^K} & \sum_{k=1}^{K} \log_2 \frac{|\sigma^2 \boldsymbol{I} + \sum_{j=k}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|}{|\sigma^2 \boldsymbol{I} + \sum_{j=k+1}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|} \\ \text{subject to} & \\ \lambda_1 \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_1) & + \lambda_2 \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_2) \leq \lambda_1 P_1 + \lambda_2 P_2 \end{aligned} \tag{3.61}$$

In [27] it is proven that the optimal value of (3.61) for any given pair of λ_1 and λ_2 ($\lambda_1 \geq 0$, $\lambda_2 \geq 0$) is an upper bound on the optimal value of (3.60) and the upper bound is tight. Therefore, we can solve (3.61) and the following minimization problem

$$\min_{\lambda_1 \geq 0, \lambda_2 \geq 0} g(\lambda_1, \lambda_2) \tag{3.62}$$

alternatively, until the optimal value is achieved.

Note that according to the general BC-MAC duality, (3.61) is equivalent to the following dual MAC problem

$$\begin{aligned} \max_{\{\boldsymbol{Q}_k^M\}_{k=1}^K} & \sum_{k=1}^{K} \log_2 \frac{|\lambda_1 \boldsymbol{A}_1 + \lambda_2 \boldsymbol{A}_2 + \sum_{j=1}^{k} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H|}{|\lambda_1 \boldsymbol{A}_1 + \lambda_2 \boldsymbol{A}_2 + \sum_{j=1}^{k-1} \boldsymbol{H}_j \boldsymbol{Q}_j^M \boldsymbol{H}_j^H|} \\ \text{subject to} & \sum_{k=1}^{K} \text{trace}(\sigma^2 \boldsymbol{Q}_k^M) \leq \lambda_1 P_1 + \lambda_2 P_2 \\ & \boldsymbol{Q}_k^M \geq 0 \end{aligned} \tag{3.63}$$

where $\boldsymbol{Q}_k^M$ indicates the transmit covariance matrix of the kth user in MAC. This is a convex problem and it can be solved via standard techniques. After obtaining the optimal $\boldsymbol{Q}_k^M$, the optimal solution of (3.61) can then be found via the MAC-BC covariance transformation algorithm discussed in Section 3.1.

Next, we introduce the ellipsoid method to solve (3.62). The Lagrangian function of (3.61) can be written as

$$\begin{aligned} L(\{\boldsymbol{Q}_k\}_{k=1}^K, \{\lambda_1, \lambda_2\}) = \max_{\{\boldsymbol{Q}_k\}_{k=1}^K} & \sum_{k=1}^{K} \log_2 \frac{|\sigma^2 \boldsymbol{I} + \sum_{j=k}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|}{|\sigma^2 \boldsymbol{I} + \sum_{j=k+1}^{K} \boldsymbol{H}_k \boldsymbol{Q}_j \boldsymbol{H}_k^H|} \\ & + \lambda_1 P_1 + \lambda_2 P_2 - \lambda_1 \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_1) - \lambda_2 \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_2) \end{aligned} \tag{3.64}$$

If $\boldsymbol{\lambda} = [\lambda_1, \lambda_2]^T$ is defined with random initialized positive value, and $\boldsymbol{s} = [s_1, s_2]^T$ is the subgradient of the Lagrangian function (3.64) at a set of fixed $\{\lambda_1, \lambda_2\}$, i.e.,

$$\begin{aligned} s_n &= \frac{\partial L(\{\boldsymbol{Q}_k\}_{k=1}^K, \{\lambda_1, \lambda_2\})}{\partial \lambda_n} \\ &= P_n - \sum_{k=1}^{K} \text{trace}(\boldsymbol{Q}_k \boldsymbol{A}_n), \quad 1 \leq n \leq 2 \end{aligned} \tag{3.65}$$

then $\boldsymbol{\lambda}$ is updated as follows:

$$\boldsymbol{\lambda}^+ = \boldsymbol{\lambda} - \frac{1}{(N_t+1)\sqrt{\boldsymbol{s}^T \boldsymbol{M} \boldsymbol{s}}} \boldsymbol{M}\boldsymbol{s} \tag{3.66}$$

and $\boldsymbol{M}$ is upgraded as

$$\boldsymbol{M}^+ = \frac{N_t^2}{N_t^2 - 1}\left(\boldsymbol{M} - \frac{2}{(N_t+1)\boldsymbol{s}^T\boldsymbol{M}\boldsymbol{s}} \boldsymbol{M}\boldsymbol{s}\boldsymbol{s}^T\boldsymbol{M}\right) \tag{3.67}$$

Note that $\boldsymbol{M}$ is initialized as unit matrix and $N_t = 2$ according to the assumption. (3.66) and (3.67) are optimized alternatively with each other fixed, and it is proven that the optimal value will be achieved eventually when it converges [76]. Algorithm 10 gives the pseudo code of Per-OPT method.

In Figure 3.20, simulation results are given to show the achievable sum rate of Per-OPT method. We can see that the sum rate is a little lower than the sum capacity due to the practical per antenna power constraint. With the increase of SNR, the gap becomes smaller since the equal power allocation is optimal in the high SNR region.

Algorithm 10 Pseudo code of Per-OPT method

Initialize $\boldsymbol{\lambda}$ as random vector
$i = 0$
Repeat $i \leftarrow i + 1$
Solve (3.63) by toolbox such as CVX
Update $\boldsymbol{\lambda}$ (3.66), denoted as $\boldsymbol{\lambda}(i)$
Until $\| \boldsymbol{\lambda}(i) - \boldsymbol{\lambda}(i-1) \| < \epsilon$

3.7.2 PBD-DPC method

Instead of optimizing the sum rate in the dual MAC as Per-OPT method, PBD-DPC method proposed in [26] takes advantage of SZF-DPC method, and tries to optimize the sum rate in BC directly under the per-antenna power constraint. The sum rate

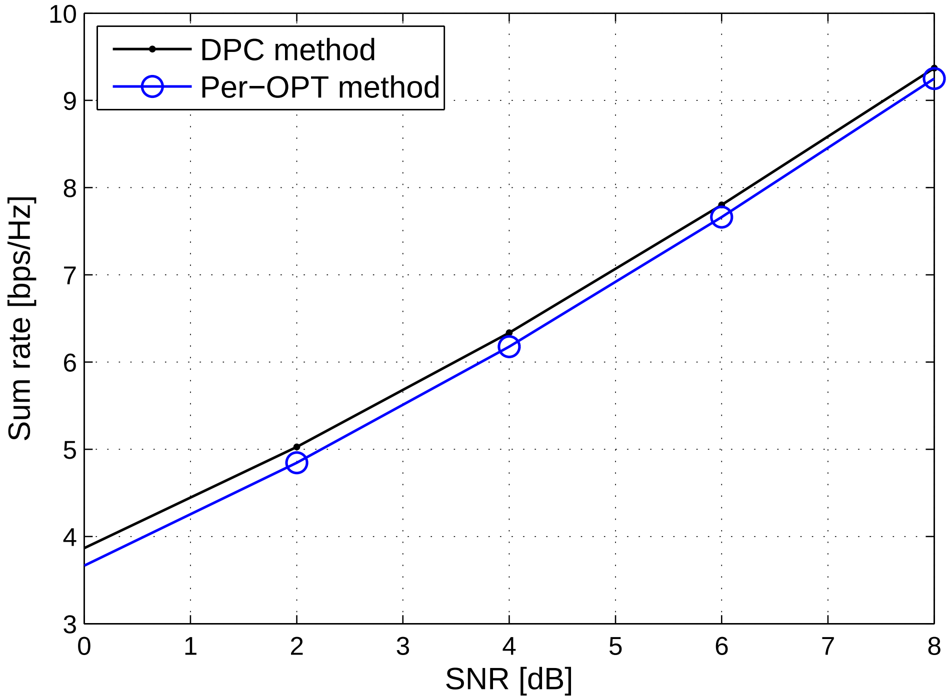

Figure 3.20: Sum rate comparison of DPC method, Per-OPT method, $N_t = 4$, $N_{r,k} = 2$, $\forall k$, and $K = 2$.

optimization of MU-MIMO broadcast channels is formulated to

$$\begin{aligned} \max_{\{\boldsymbol{V}_k\}_{k=1}^K} & \sum_{k=1}^K \log_2 \frac{|\sigma^2 \boldsymbol{I} + \sum_{j=k}^K \boldsymbol{H}_k \boldsymbol{V}_j \boldsymbol{V}_j^H \boldsymbol{H}_k^H|}{|\sigma^2 \boldsymbol{I} + \sum_{j=k+1}^K \boldsymbol{H}_k \boldsymbol{V}_j \boldsymbol{V}_j^H \boldsymbol{H}_k^H|} \\ \text{subject to} \quad & [\sum_{k=1}^K \boldsymbol{V}_k \boldsymbol{V}_k^H]_{(n,n)} \leq P_n, \ 1 \leq n \leq N_t \end{aligned} \tag{3.68}$$

Similarly to SZF-DPC method, if the transmit beamforming matrices $\boldsymbol{V}_k$ $(1 \leq k \leq K)$ are designed successively, and $\boldsymbol{V}_k$ $(\forall k)$ is denoted by $\boldsymbol{V}_k = \boldsymbol{B}_k \boldsymbol{D}_k$, then to remove the residual interference, $\boldsymbol{B}_k$ should lie in the null space of the space spanned by

$$\bar{\boldsymbol{H}}_k = \left[\boldsymbol{H}_1^T \ \boldsymbol{H}_2^T \ \cdots \ \boldsymbol{H}_{k-1}^T\right]^T \tag{3.69}$$

$\boldsymbol{D}_k$ is designed to maximize the throughput under the per-antenna power constraint. The residual interference which is non-causally known to the transmitter can be removed by DPC technique. Since the interference is suppressed completely in this way, (3.68) can be rewritten as

$$\begin{aligned} \max_{\{\boldsymbol{\Omega}_k\}_{k=1}^K} & \sum_{k=1}^K \log_2 |\boldsymbol{I} + \frac{1}{\sigma^2} \boldsymbol{H}_k \boldsymbol{B}_k \boldsymbol{\Omega}_k \boldsymbol{B}_k^H \boldsymbol{H}_k^H| \\ \text{subject to} \quad & [\sum_{k=1}^K \boldsymbol{B}_k \boldsymbol{\Omega}_k \boldsymbol{B}_k^H]_{(n,n)} \leq P_n, \ 1 \leq n \leq N_t \\ & \text{rank}(\boldsymbol{\Omega}_k) \leq N_{r,k}, \ 1 \leq k \leq K \end{aligned} \tag{3.70}$$

where $\boldsymbol{\Omega}_k = \boldsymbol{D}_k \boldsymbol{D}_k^H$. Problem (3.70) is a convex program when the constraint $\text{rank}(\boldsymbol{\Omega}_k) \leq N_{r,k}$ is omitted, and numerical optimization tools can be used to solve it. Here we introduce how to solve this problem by Lagrange duality method, which is also used in Per-OPT method. Note that other more efficient methods can also be found in [26].

Consider the Lagrangian function of (3.70) is given as

$$\begin{aligned} L(\{\boldsymbol{\Omega}_k\}, \{\lambda_n\}) = & \sum_{k=1}^K \log_2 |\boldsymbol{I} + \frac{1}{\sigma^2} \boldsymbol{H}_k \boldsymbol{B}_k \boldsymbol{\Omega}_k \boldsymbol{B}_k^H \boldsymbol{H}_k^H| \\ & - \sum_{n=1}^{N_t} \lambda_n ([\sum_{k=1}^K \boldsymbol{B}_k \boldsymbol{\Omega}_k \boldsymbol{B}_k^H]_{(n,n)} - P_n) \end{aligned} \tag{3.71}$$

Since strong duality holds for (3.70), its optimal solution can be found via the following

Lagrange duality method

$$\min_{\lambda_n \geq 0} \max_{\boldsymbol{\Omega}_k \geq 0} L(\{\boldsymbol{\Omega}_k\}, \{\lambda_n\}) \tag{3.72}$$

Similarly to (3.63), the two-stage iterative algorithm works as follows. For fixed λ_n, the set of covariance matrix that maximizes $L(\{\boldsymbol{\Omega}_k\}, \{\lambda_n\})$ can be obtained by water-filling algorithm. Next, for a set of given $\{\boldsymbol{\Omega}_k\}$, the ellipsoid method can be used to update $\{\lambda_n\}$. However, the ellipsoid method converges slowly to the optimum in this form. In [26], it is proven that the optimal solution always satisfies the constraint $\text{rank}(\boldsymbol{\Omega}_k) \leq N_{r,k}$. Algorithm 11 gives the pseudo code of PBD-DPC method. It can be

Algorithm 11 Pseudo code of PBD-DPC algorithm

```
Suppose a preset user order
for k = 1 : K do
    Build the matrix H̄_k = [H_1^T H_2^T ··· H_{k-1}^T]^T as (3.69)
    Do SVD decomposition of Ĥ_k and find B_k
end for
Solve (3.70) to get Ω_k
Do eigenvalue decomposition to get D_k
Use DPC pre-subtract the non-known interference
```

observed that PBD-DPC method finds the transmit beamforming matrix for each user without considering its individual channel gain, which means that, when the channel gain of the kth user is weak, it will contribute a negligible throughput to the sum rate but impose severe constraint on the subsequent users. Figure 3.21 illustrates the sum rate provided by Per-OPT method, PBD-DPC method, and the sum capacity of the channels. We can see that a large gap exists between the optimal Per-OPT method and PBD-DPC method.

3.8 Conclusion

In this chapter, the state of the art of beamforming techniques in MU-MIMO broadcast channels is overviewed. First, we present the optimal solution in terms of sum capacity. Then, the non-linear and linear suboptimal methods that have low complexities are introduced. After that, beamforming techniques under per-antenna power constraint are presented. At last, we give some simulation results and the performance comparisons between different methods. Based on the observations of the beamforming techniques in the literature, we propose two new beamforming methods in MU-MIMO broadcast channels, which will be introduced in the next chapter.

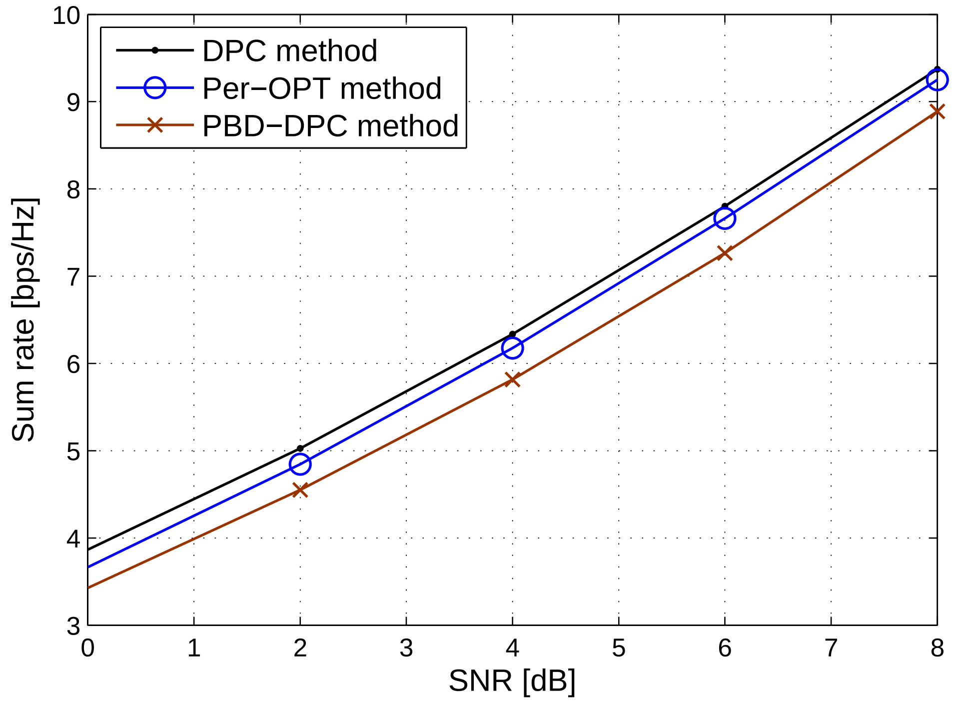

Figure 3.21: Sum rate comparison of DPC method, Per-OPT method, PBD-DPC method, $N_t = 4$, $N_{r,k} = 2$, $\forall k$, and $K = 2$.

4

Proposed beamforming methods

In the first part of this chapter, we propose an alternative approach to SA-DPC method under the total power constraint. In SA-DPC method, the transmit beamforming vector is selected as the right singular vector corresponding to the largest singular value of the projected channel matrix, the per-antenna power constraint is hard to be imposed in this way. In the proposed method, instead of obtaining the transmit beamforming vector directly, we first find the subspace where the transmit beamforming vector should lie in, then the transmit beamforming vector is selected in the above subspace to optimize the global throughput. The proposed method performs identically as SA-DPC method when the total power constraint is imposed, even though they are derived from different ways. Moreover, it is shown that the proposed method can be easily modified to the more realistic per-antenna power constraint. In the second part of this chapter, a beamforming method under the per-antenna power constraint is proposed. Since the optimal solution to the original problem is difficult to obtain, in the proposed method, this problem is divided into two classical optimization problems, which can be solved with existing standard algorithms. We alternatively solve each subproblem under the assumption that another one is fixed, the convergence can be achieved within a small number of iterations. Similarly to SA-DPC method and the first proposed method, one data stream is allocated to the user who brings the largest global throughput increase at each step. The non-causally known interference is pre-subtracted through DPC technique before transmission, and the remaining interference is eliminated by the transmit beamforming and receive combining vectors.

4.1 Under total transmit power constraint

Consider downlink MU-MIMO broadcast channels with K users, where a base station is equipped with N_t transmit antennas and transmits $\sum_k L_k = L$ data streams to the users, each user has $N_{r,k}$ receive antennas and receives L_k data streams. The channel state information (CSI) is supposed to be perfectly known at the base station. Since the channel gains vary with different users [20], [29], an optimal L_k $(0 \leq L_k \leq N_{r,k})$ for the kth $(\forall k)$ user should be found to maximize the global throughput. In this section, for presentation convenience and without loss of generality, we consider each data stream separately. The lth $(1 \leq l \leq L)$ data stream after MU-MIMO broadcast channels and the receive combining vector is [64]

$$y_l = \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l x_l + \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \sum_{j=1}^{l-1} \boldsymbol{v}_j x_j + \boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \sum_{j=l+1}^{L} \boldsymbol{v}_j x_j + \boldsymbol{u}_l^H \boldsymbol{n}_{\pi(l)} \tag{4.1}$$

where $\pi(l)$ indicates that the lth data stream is allocated to the $\pi(l)$th user. If the data streams are allocated in a successive way and DPC technique is used, the non-causally known interference in (4.1) (i.e., $\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \sum_{j=1}^{l-1} \boldsymbol{v}_j x_j$) coming from the previously allocated data stream can be pre-canceled before transmission. In this case, the optimization of the sum rate under the total transmit power constraint can be rewritten as

$$\begin{aligned} \max_{\{\boldsymbol{u}_l, \boldsymbol{v}_l\}_{l=1}^L} & \sum_{l=1}^{L} \log_2\Big(1 + \frac{|\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l|^2}{\sum_{j=l+1}^{L} |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_j|^2 + \sigma^2}\Big) \\ \text{subject to} \quad & \boldsymbol{u}_l^H \boldsymbol{u}_l = 1, \ 1 \leq l \leq L \\ & \text{trace}\Big(\sum_{l=1}^{L} \boldsymbol{v}_l \boldsymbol{v}_l^H\Big) \leq P_T \end{aligned} \tag{4.2}$$

The maximum value of (4.2) is difficult to obtain. Instead of maximizing (4.2) directly, we try to cancel the residual interference $\sum_{j=l+1}^{L} |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_j|^2$ in (4.2) completely, meanwhile, the sum rate is maximized.

It can be seen that the optimal user order and data stream allocation need to search all the possibilities [14]. The computational complexity of this pure global exhaustive search is pretty high when the number of users is large. In the proposed method, data streams are allocated successively to the K users. At each step, we search firstly over all users, then one data stream is assigned to the user who brings the largest throughput increase. This greedy strategy reduces the computational complexity since not all the possibilities are explored. To simplify the presentation, we focus on beamforming design by ignoring power allocation. Note that with zero-forcing constraint, water-filling power allocation principle is the optimal solution.

Specifically, the lth $(1 \leq l \leq L)$ data stream is supposed to be allocated to the $\pi(l)$th user $(1 \leq \pi(l) \leq K)$ at the lth step. The non-causally known interference produced by the previously allocated $l-1$ data streams $\sum_{j=1}^{l-1} |\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_j|^2$ can be removed completely through DPC technique. Additionally, we force that the lth data stream does not interfere with the previously allocated $l-1$ data streams (i.e., $\boldsymbol{u}_j^H \boldsymbol{H}_{\pi(j)} \boldsymbol{v}_l = 0$ $(1 \leq j \leq l-1)$). Under these two constraints, the rate optimization of the lth data stream is performed with

$$\begin{aligned} &\max_{\{\boldsymbol{u}_l, \boldsymbol{v}_l\}} \log_2(1 + \frac{|\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{v}_l|^2}{\sigma^2}) \\ &\text{subject to} \quad \|\boldsymbol{u}_l\| = 1, \ \|\boldsymbol{v}_l\| = 1 \\ &\qquad\qquad \boldsymbol{u}_j^H \boldsymbol{H}_{\pi(j)} \boldsymbol{v}_l = 0, \ 1 \leq j \leq l-1 \end{aligned} \tag{4.3}$$

In (4.3), the constraints $\boldsymbol{u}_j^H \boldsymbol{H}_{\pi(j)} \boldsymbol{v}_l = 0$ $(1 \leq j \leq l-1)$ imply that $\boldsymbol{v}_l$ must lie in the null space of the space spanned by

$$\boldsymbol{N}_l = \begin{bmatrix} \boldsymbol{u}_1^H \boldsymbol{H}_{\pi(1)} \\ \vdots \\ \boldsymbol{u}_{l-1}^H \boldsymbol{H}_{\pi(l-1)} \end{bmatrix} \in \mathbb{C}^{(l-1) \times N_t} \tag{4.4}$$

As a result of LQ decomposition of $\boldsymbol{N}_l$, it comes that

$$\boldsymbol{N}_l = \begin{bmatrix} \boldsymbol{L}_{(l-1)\times(l-1)} & \boldsymbol{0}_{(l-1)\times(N_t-l+1)} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{(1)}^H \\ \boldsymbol{A}_l^H \end{bmatrix} \tag{4.5}$$

where $\boldsymbol{L}_{(l-1)\times(l-1)}$ is a lower triangular matrix, and the columns of $\boldsymbol{A}_l \in \mathbb{C}^{N_t \times (N_t-l+1)}$ form an orthonormal basis of the null space of $\boldsymbol{N}_l$, i.e., $\boldsymbol{N}_l \boldsymbol{A}_l = \boldsymbol{0}$, and $\boldsymbol{A}_l^H \boldsymbol{A}_l = \boldsymbol{I}$. Thus, the columns of $\boldsymbol{A}_l$ span the space where the transmit beamforming vector of the lth data stream $\boldsymbol{v}_l$ must lie in (i.e., $\boldsymbol{v}_l = \boldsymbol{A}_l \boldsymbol{f}_l$). The optimization problem (4.3) can then be rewritten as

$$\begin{aligned} &\max_{\{\boldsymbol{u}_l, \boldsymbol{f}_l\}} \log_2(1 + \frac{|\boldsymbol{u}_l^H \boldsymbol{H}_{\pi(l)} \boldsymbol{A}_l \boldsymbol{f}_l|^2}{\sigma^2}) \\ &\text{subject to} \quad \|\boldsymbol{u}_l\| = 1, \ \|\boldsymbol{f}_l\| = 1 \end{aligned} \tag{4.6}$$

The data rate of the lth data stream is maximized by choosing $\boldsymbol{u}_l$ and $\boldsymbol{f}_l$ as the left and right singular vectors corresponding to the dominant singular value of matrix $\boldsymbol{H}_{\pi(l)}\boldsymbol{A}_l$, respectively. Then the transmit beamforming vector is calculated as $\boldsymbol{v}_l = \boldsymbol{A}_l\boldsymbol{f}_l$. Since $\boldsymbol{A}_l$ and $\boldsymbol{f}_l$ are unitary matrix and unitary vector, respectively, we can see that $\boldsymbol{v}_l$ and $\boldsymbol{u}_l$ are also unitary vectors. Finally, the water-filling algorithm is performed to allocate the transmit power over the allocated data streams.

The lth data stream is allocated to the user (among all the candidates) who brings the largest global throughput increase. The algorithm stops when there is no global throughput increase. Note that in order to obtain the null space of $\boldsymbol{N}_l$, the number of total allocated transmit data streams L should be less or equal to that of transmit antennas (i.e. $\sum_{k=1}^{K} L_k = L \leq N_t$).

We can see that the interference term $\boldsymbol{u}_l^H\boldsymbol{H}_{\pi(l)}\sum_{j=l+1}^{L}\boldsymbol{v}_j x_j$ in (4.1) can be suppressed completely by the above beamforming method. Note that the non-causally known interference $\boldsymbol{u}_l^H\boldsymbol{H}_{\pi(l)}\sum_{j=1}^{l-1}\boldsymbol{v}_j x_j$ is pre-canceled at the base station, the sum rate of the proposed beamforming method can be calculated as

$$\sum_{l=1}^{L}\log(1+\frac{|\boldsymbol{u}_l^H\boldsymbol{H}_{\pi(l)}\boldsymbol{v}_l|^2}{\sigma^2}) \tag{4.7}$$

Compared with SA-DPC method, which finds the transmit beamforming vector as the right singular vector associated to the largest singular value of the projected channel matrix, in the proposed method, the subspace that the transmit beamforming vector should lie in is found first, then the one that maximizes the channel gain is selected. It is another way to find exactly the same performance as SA-DPC method. Algorithm 12 gives the pseudo code of the proposed beamforming method under total power constraint (Prop-T method). In Figure 4.1, we can see that the sum rate of Prop-T method is very close to DPC method, which is the maximum sum capacity.

Considering the computational complexity, an approximate measurement is the number of floating point operations (flops) [77], [78]. A real addition, multiplication or division operation is counted as one flop. A complex addition and multiplication have two flops and six flops, respectively [79]. In the proposed method (Prop-T method), LQ decomposition is performed once at the lth $(0 \leq l \leq L)$ step, which needs $4(l-1)^2(3N_t-l+1)$ flops [78], then SVD is used to find the dominant singular vectors, the number of flops is $4N_r^2N_t + 8N_r + N_t^2 + 9N_t^3$ [79]. While in SA-DPC method, the projection matrix is calculated at each step, which needs $8N_t^3 - 10N_t^2$ flops, then SVD is also performed to find the dominant singular vectors. Figure 4.2

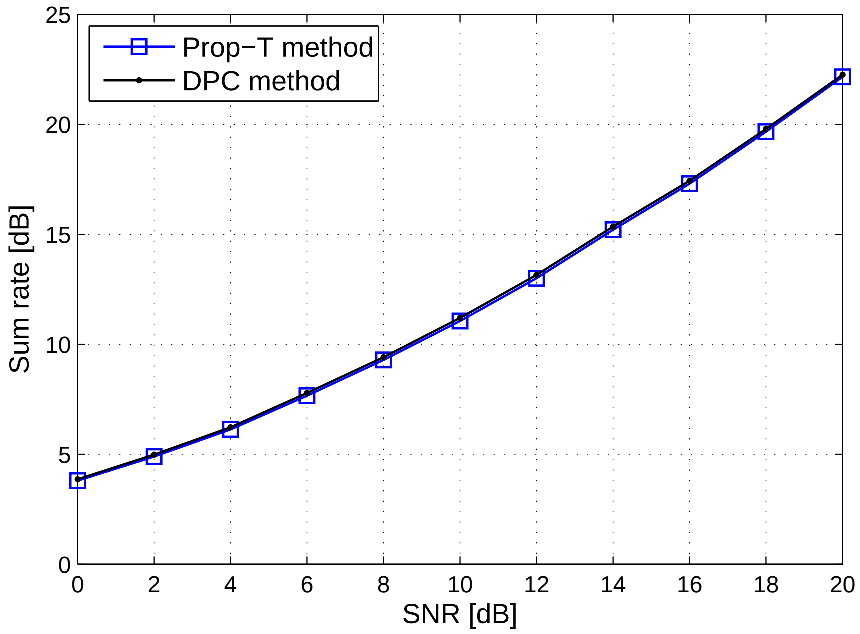

Figure 4.1: Sum rate comparison of DPC method and Prop-T method method, $N_t = 4$, $N_{r,k} = 2$, $\forall k$, and $K = 2$.

Algorithm 12 Pseudo code of Prop-T method.

Initialization: $\boldsymbol{A}_1 = \boldsymbol{I}_{N_t},\ l = 0$
while $l \le N_t$ **do**
  $l \leftarrow l + 1$
  for $k = 1 : K$ **do**
    Maximize (4.6) denoted as R_k
  end for
  Calculate the largest sum rate $[C_l, \pi(l)] = \max(R_k)$
  if $C_{l-1} \ge C_l$ **then**
    $\pi(l) = 0$
    break
  end if
end while

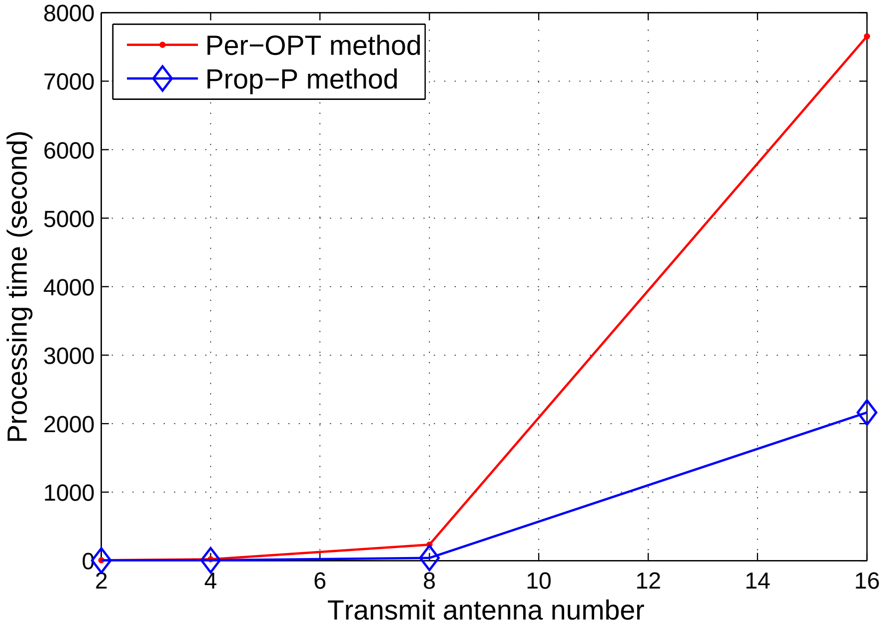

Figure 4.2: Complexity comparison of SA-DPC and Prop-T methods. $N_t = 4$, and $N_{r,k} = 2,\ \forall k$.

illustrates the number of needed flops of SA-DPC and Prop-T methods in terms of transmit antenna number. We can see that the complexity of Prop-T method is relatively higher than that of SA-DPC method, but Prop-T method can be easily modified to take the per-antenna power constraint into account by designing the vector $\boldsymbol{f}_i$, which will be presented in the following section.

4.2 Under per-antenna power constraint

Similarly to the development in Section 4.1, if the per-antenna constraint is imposed, the optimization problem at the lth step can be formulated as

$$\begin{aligned} \max_{\{\boldsymbol{u}_j, \boldsymbol{f}_j\}_{j=1}^{l}} & \sum_{j=1}^{l} \log_2\left(1 + \frac{\boldsymbol{u}_j^H \boldsymbol{H}_{\pi(j)} \boldsymbol{A}_j \boldsymbol{f}_j \boldsymbol{f}_j^H \boldsymbol{A}_j^H \boldsymbol{H}_{\pi(j)}^H \boldsymbol{u}_j}{\sigma^2}\right) \\ \text{subject to} \quad & \boldsymbol{u}_j^H \boldsymbol{u}_j = 1,\ 1 \le j \le l \\ & \Big[\sum_{j=1}^{l} \boldsymbol{A}_j \boldsymbol{f}_j \boldsymbol{f}_j^H \boldsymbol{A}_j^H\Big]_{(n,n)} \le P_n,\ 1 \le n \le N_t \end{aligned} \tag{4.8}$$

Compared with (4.6), we consider the sum of all the allocated data streams instead of only the lth one under the per-antenna power constraint.

The objective function in (4.8) is difficult to maximize under these constraints. Even with fixed receive combining vector $\boldsymbol{u}_j$, the objective function is still neither convex nor concave with respect to $\boldsymbol{f}_j$. However, we can observe that the maximum value of the objective function can be achieved by maximal-ratio combining (MRC) when $\boldsymbol{f}_j$ is assumed to be fixed, i.e.,

$$\boldsymbol{u}_j^+ = \frac{\boldsymbol{H}_{\pi(j)}\boldsymbol{v}_j}{\|\boldsymbol{H}_{\pi(j)}\boldsymbol{v}_j\|_F} = \frac{\boldsymbol{H}_{\pi(j)}\boldsymbol{A}_j\boldsymbol{f}_j}{\|\boldsymbol{H}_{\pi(j)}\boldsymbol{A}_j\boldsymbol{f}_j\|_F}, \quad 1 \le j \le l \tag{4.9}$$

where $\boldsymbol{u}_j{}^+$ indicates the upgraded $\boldsymbol{u}_j$ after each iteration. In this paper, an iterative process is proposed to solve (4.8). The variable $\boldsymbol{u}_j$ (or $\boldsymbol{f}_j$) is optimized by assuming the other $\boldsymbol{f}_j$ (or $\boldsymbol{u}_j$) fixed.

Inspired by [25], we define $\boldsymbol{F}_j = \boldsymbol{f}_j\boldsymbol{f}_j^H$, then (4.8) can be reformulated as

$$\begin{aligned}
&\max_{\{\boldsymbol{u}_j,\boldsymbol{F}_j\}_{j=1}^l} \sum_{j=1}^{l} \log_2\Big(1 + \frac{\boldsymbol{u}_j^H\boldsymbol{H}_{\pi(j)}\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H\boldsymbol{H}_{\pi(j)}^H\boldsymbol{u}_j}{\sigma^2}\Big) \\
&\text{subject to} \quad \boldsymbol{u}_j^H\boldsymbol{u}_j = 1, \ \text{rank}(\boldsymbol{F}_j) \le 1, \ \text{and } \boldsymbol{F}_j \ge 0, \ 1 \le j \le l \\
&\qquad [\sum_{j=1}^{l} \boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H]_{(n,n)} \le P_n, \ 1 \le n \le N_t
\end{aligned} \tag{4.10}$$

Note that the optimal solution of (4.10) is still difficult to find. However, if the receive combining vector $\boldsymbol{u}_j$ is assumed to be fixed, the constraint $\text{rank}(\boldsymbol{F}_j) \le 1$ is supposed to be omitted, and denoting $\boldsymbol{h}_j^H = \boldsymbol{u}_j^H\boldsymbol{H}_{\pi(j)}$, it comes then that (4.10) can be reformulated to

$$\begin{aligned}
&\max_{\{\boldsymbol{F}_j\}_{j=1}^l} \sum_{j=1}^{l} \log_2\Big(1 + \frac{\boldsymbol{h}_j^H\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H\boldsymbol{h}_j}{\sigma^2}\Big) \\
&\text{subject to} \quad [\sum_{j=1}^{l} \boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H]_{(n,n)} \le P_n, \ 1 \le n \le N_t \\
&\qquad \boldsymbol{F}_j \ge 0, \ 1 \le j \le l
\end{aligned} \tag{4.11}$$

This is a convex optimization problem, which can be solved by optimization toolbox such as CVX [80] or YALMIP [81]. The more efficient interior-point method [25] can also be applied, which uses the barrier method in [82] to solve this problem. In [25], it is also shown that the optimal $\boldsymbol{F}_j^\star$ in (4.11) always satisfies the rank-1 constraint (i.e., $\text{rank}(\boldsymbol{F}_j^\star) \le 1$).

The closed-form of $\boldsymbol{F}_j^\star$ using Lagrange duality method similarly to [24] can be written as

$$\boldsymbol{F}_j^\star = \frac{\lambda_j}{\delta_j^2}(\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-1}\boldsymbol{A}_j^H \boldsymbol{h}_j \boldsymbol{h}_j^H \boldsymbol{A}_j (\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-1}, \quad 1 \le j \le l \tag{4.12}$$

where $\boldsymbol{B}_\mu = \sum_{n=1}^{N_t} \mu_n \boldsymbol{B}_n$, and $\boldsymbol{B}_{n(N_t \times N_t)} = \text{diag}(0, \cdots 0, 1, 0, \cdots, 0)$ is a diagonal matrix with only the nth diagonal element being 1. Details are given in the appendix B. It can be observed from (4.12) that $\boldsymbol{F}_j^\star$ indeed satisfies the rank-1 constraint. From $\boldsymbol{F}_j = \boldsymbol{f}_j \boldsymbol{f}_j^H$, $\boldsymbol{f}_j^\star$ can be written as

$$\boldsymbol{f}_j^\star = \frac{\sqrt{\lambda_j}}{\delta_j}(\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-1}\boldsymbol{A}_j^H \boldsymbol{h}_j, \quad 1 \le j \le l \tag{4.13}$$

and the corresponding transmit beamforming vector is

$$\boldsymbol{v}_j = \boldsymbol{A}_j \boldsymbol{f}_j^\star = \frac{\sqrt{\lambda_j}}{\delta_j}\boldsymbol{A}_j(\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-1}\boldsymbol{A}_j^H \boldsymbol{h}_j, \quad 1 \le j \le l \tag{4.14}$$

Then $\boldsymbol{u}_j$ is upgraded according to (4.9). We alternatively calculate (4.9) and (4.14) until the increase of maximum value of the objective function is below a preset threshold.

Finally, the lth data stream is allocated to the user (among all the candidates), who brings the largest throughput increase with the previously selected users. The algorithm stops when there is no global throughput increase.

Similarly to section 4.1, in order to obtain the null space of $\boldsymbol{N}_l$, the number of total allocated transmit data streams L should be less or equal to that of transmit antennas (i.e. $\sum_{k=1}^{K} L_k = L \le N_t$). As defined in Section 2, L_k is the number of allocated data streams for the kth user, so we have $L_k \le N_{r,k}$. In addition, since receive combining technique is adopted, the number of total receive antennas can be larger than that of transmit antennas, in this case, we can have $\sum_{k=1}^{K} N_{r,k} \ge N_t$.

Note that the proposed method under total power constraint is a suboptimal method to solve the optimization problem (4.2). Moreover, if the stronger per-antenna power constraint is imposed, the sum rate is lower than that obtained under the total power constraint. Let C_{DPC}, C_T and C_P denote the maximum value of (4.2), (4.7) and (4.8), respectively, it follows that

$$C_P \le C_T \le C_{DPC} \tag{4.15}$$

The algorithm of the whole process is summarized in Algorithm 13.

Algorithm 13 Proposed beamforming and data allocation method under per-antenna power constraint.

Initialization: $\boldsymbol{A}_1 = \boldsymbol{I}_{N_t},\ l = 0$
while $l \leq N_t$ **do**
  $l \leftarrow l + 1$
  for $k = 1 : K$ **do**
    $i = 0$, $r_0 = 0$
    Repeat $i \leftarrow i + 1$
      Find the optimal $\boldsymbol{f}_1^\star, \cdots, \boldsymbol{f}_l^\star$ (4.13)
      Update $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_l$ (4.9)
      Update $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_l$ (4.5)
      Calculate the temporary sum rate as r_i
    Until $|r_i - r_{i-1}| < \epsilon$
    Define $R_k = r_i$
  end for
  Calculate the largest sum rate $[C_l, \pi(l)] = \max(R_k)$
  if $C_{l-1} \geq C_l$ **then**
    $\pi(l) = 0$
    break
  end if
end while

4.3 Simulation results

In this section, the performance of the discussed methods is evaluated by simulations. For simplification, each user is assumed to have the same number of receive antennas, and the entries of channel matrix are i.i.d. Gaussian random variables with zero-mean, independent real and imaginary parts with equal variance (i.e., uncorrelated Rayleigh-fading channels). We first consider MU-MIMO broadcast channels with $N_t = 8$ transmit antennas at the base station, and $K = 4$ users under a natural order, each user is equipped with two receive antennas (i.e., $N_{r,k} = 2, \forall k$). The power of each transmit antenna is limited to $P_n = P_T / N_t$, with P_T the total transmit power. The threshold ϵ is preset as 10^{-3}. Note that the same transmit beamforming and receive combining vectors are obtained by both Prop-T method and SA-DPC method, their performance is therefore identical. In Figure 4.3, we show the sum capacity provided by DPC method, the average sum rates provided by Per-OPT method [27], PBD-DPC method [26], Prop-T method, and the proposed method under per-antenna power constraint (denoted as Prop-P method), respectively. We can see that the performance of Prop-P method is slightly lower than that of Prop-T method. This comes from the more severe per-antenna power constraint. On the other hand, Prop-P method outperforms

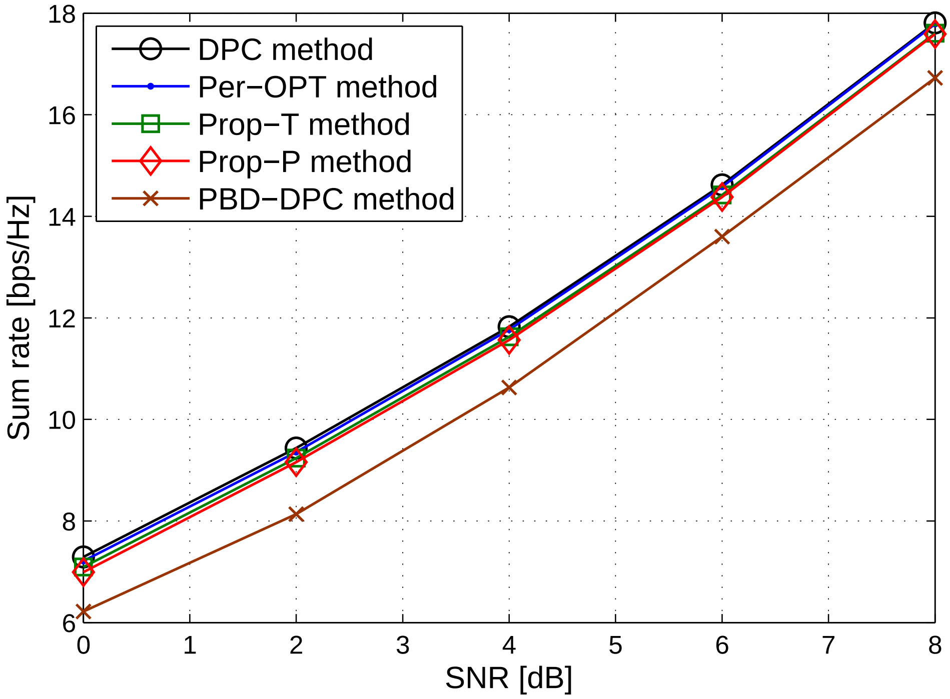

Figure 4.3: Sum rate comparison of DPC method, Per-OPT method, PBD-DPC method, Prop-T method and Prop-P method, $N_t = 8$, $N_{r,k} = 2$, $\forall k$, $K = 4$, and $P_n = \frac{P_T}{N_t}$.

PBD-DPC method, this comes from the fact that PBD-DPC method allocates a number of data streams to one user without considering each individual subchannel gain. But in Prop-P method, only one data stream is allocated to the user who can contribute the largest increase of the throughout. Moreover, it can also be observed in Figure 4.3 that Prop-P method provides almost the same sum rate as Per-OPT method, while the computational complexity, which will be shown later, is lower.

In Figure 4.4, the required iteration number is presented with a random realization, the configuration is the same as Figure 4.3, and the tolerant error in Per-OPT method is preset as 10^{-3}. Note that the total needed iterations over the whole data stream allocation procedure is recorded in Prop-P method. As shown in Figure 4.4, Per-OPT method requires more iterations to converge compared with Prop-P method. This is due to that the subgradient in Per-OPT method needs not be a descent function, an iteration can even decrease the objective function [26]. Moreover, the convergence rate in Per-OPT method relies strongly on the problem scale. As seen in Figure 4.5, the entire processing time of a random realization by Matlab is recorded. Not only the required number of iterations in Per-OPT method increases, but also the processing time

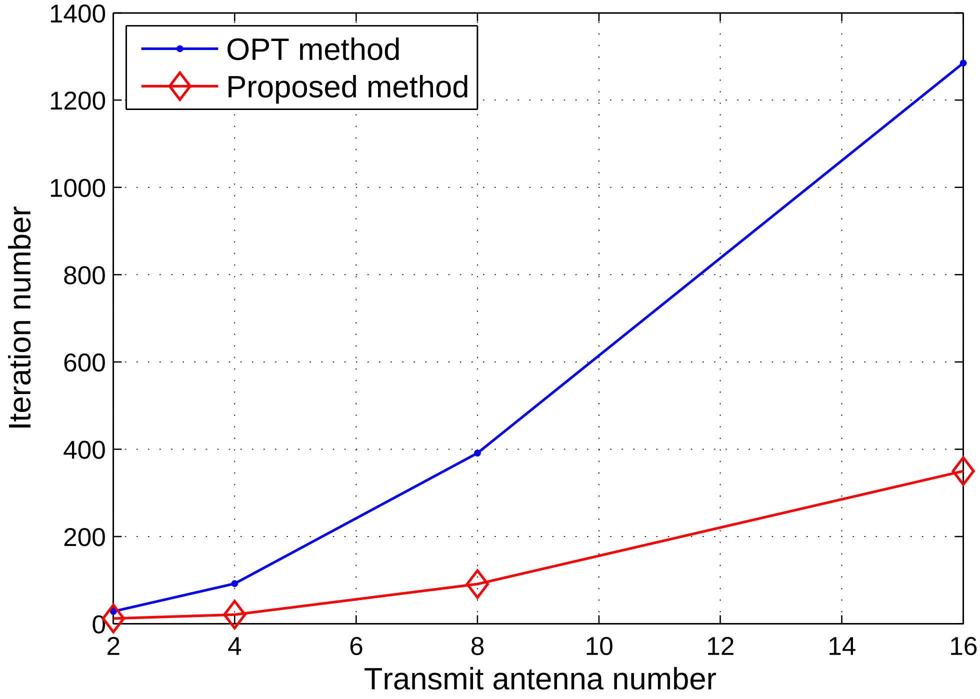

Figure 4.4: Iteration number comparison of Per-OPT method and Prop-P method, $N_t = 8$, $N_{r,k} = 2,\ \forall k$, and $K = 4$.

needed for each iteration increases rapidly with the number of transmit antenna growing. Therefore, the proposed method is more promising for practical implementation compared with Per-OPT method.

Second, We consider MU-MIMO broadcast channels with $N_t = 4$ transmit antennas at the base station, and $K = 2$ users under a natural order, each user is equipped with two receive antennas (i.e., $N_{r,k} = 2, \forall k$). The power of each transmit antenna is limited to $P_n = P_T/N_t$. The threshold ϵ is preset as 10^{-3}. Figure 4.6 illustrates the sum capacity provided by DPC method, the average sum rates provided by Prop-T method, Prop-P method and PBD-DPC method [26]. Similar observations can be found as that in Figure 4.3. In Figure 4.7, the power of the nth transmit antenna is limited to $P_n = \frac{P_T}{\sum_{l=1}^{N_t} l} n$ similarly to [25], we can see that Prop-P method also has a significant global throughput improvement compared with PBD-DPC method.

Next, we consider another MU-MIMO broadcast channels configuration to evaluate the performance of the related methods. The base station with $N_t = 4$ transmit antennas serves $K = 4$ users, each user is equipped with $N_{r,k} = 4$ receive antennas. The power of each transmit antenna is limited to $P_n = P_T/N_t$. Since the number of total receive antennas is greater than that of transmit antennas, PBD-DPC does not

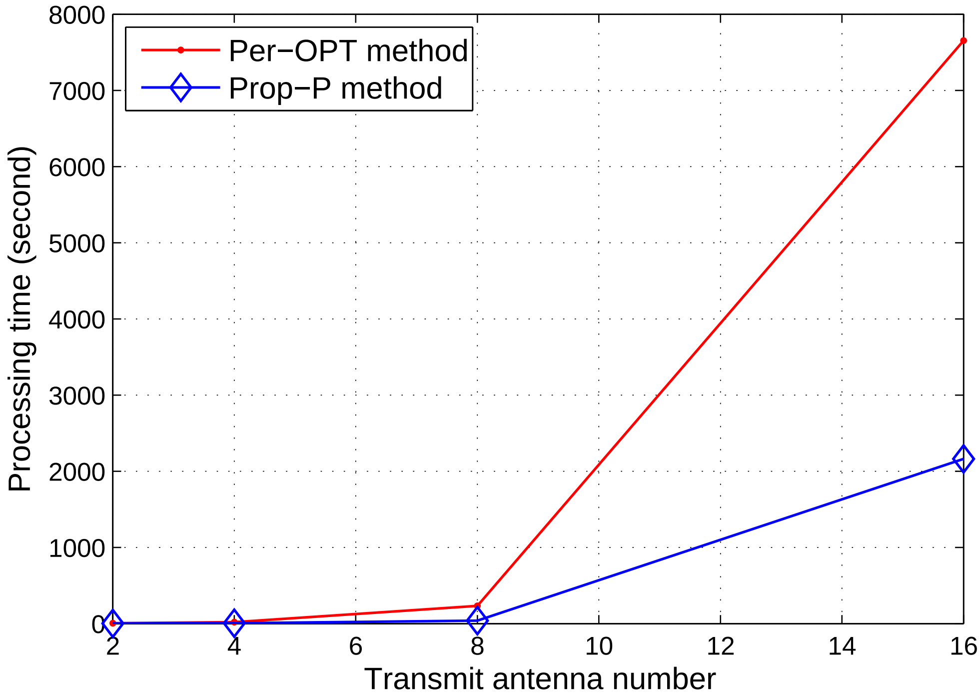

Figure 4.5: Processing time comparison of Per-OPT method and Prop-P method, $N_t = 8$, $N_{r,k} = 2,\ \forall k$, and $K = 4$.

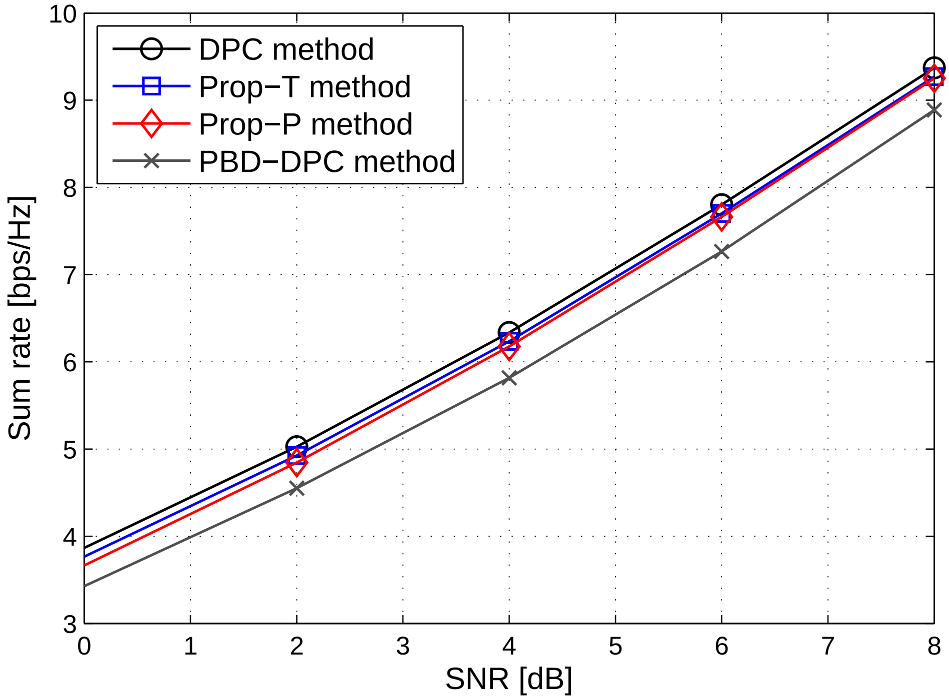

Figure 4.6: Sum rate comparison of DPC method, PBD-DPC method, Prop-T method, and Prop-P method, $N_t = 4$, $N_{r,k} = 2$, $\forall k$, $K = 2$, and $P_n = \frac{P_T}{N_t}$.

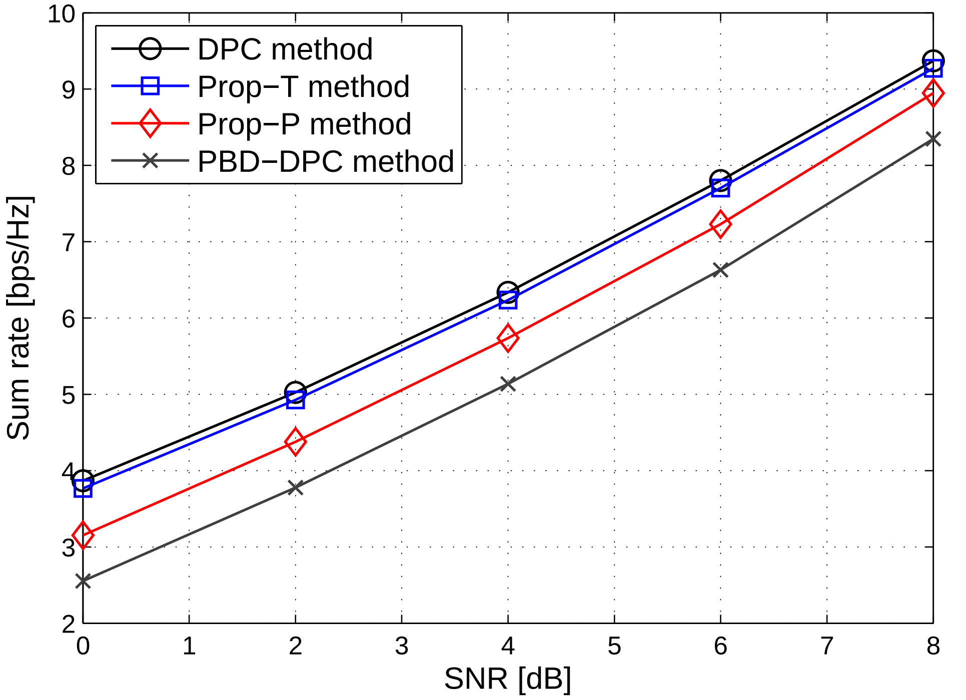

Figure 4.7: Sum rate comparison of DPC method, PBD-DPC method, Prop-T method, and Prop-P method, $N_t = 4$, $N_{r,k} = 2$, $\forall k$, $K = 2$, and $P_n = \frac{P_T}{\sum_{l=1}^{N_t} l} n$.

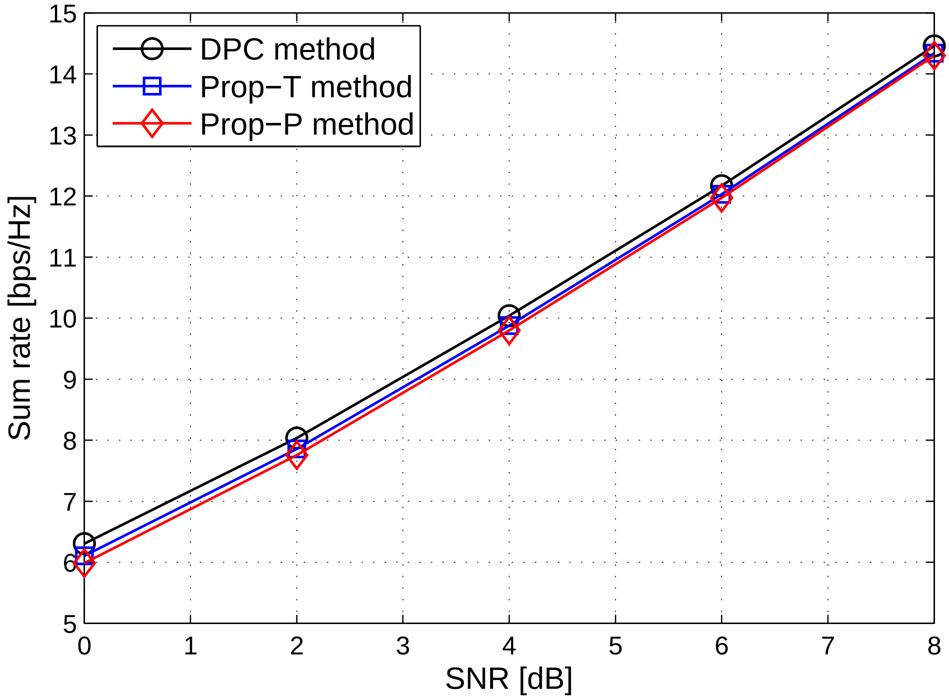

Figure 4.8: Sum rate comparison of DPC method, Prop-T method, and Prop-P method, $N_t = 4$, $N_{r,k} = 4$, $\forall k$, $K = 4$, and $P_n = \frac{P_T}{N_t}$.

work in this case. In Figure 4.8, we show the sum capacity provided by DPC method, the average sum rates provided by Prop-T and Prop-P methods. It can be observed in this figure that Prop-P method provides almost the same sum rate as that with Prop-T method, especially as SNR increases, since the equal power allocation is optimal at high SNR region. In Figure 4.9, we illustrate the power allocation over the transmit antennas by a random channel realization. The system is configured as that in the previous simulation. Moreover, SNR is zero, the total transmit power P_T is one, and the power of each transmit antenna is limited to $P_n = \frac{P_T}{N_t}$. We can observe from Figure 4.9 that the power of each transmit antenna varies widely under the total power constraint, while under the per-antenna power constraint, the power is the same for each transmit antenna, which shows that Prop-P method is attractive in practical implementation.

Considering the computational complexity of PBD-DPC and Prop-T methods, the average processing time of Matlab program of these two methods is calculated over 100 realizations, which is given in Table 4.1. We assume that every user has the same number of receive antennas in each configuration, and four different system configurations are simulated. Generally, we can see that the complexity of Prop-P method is

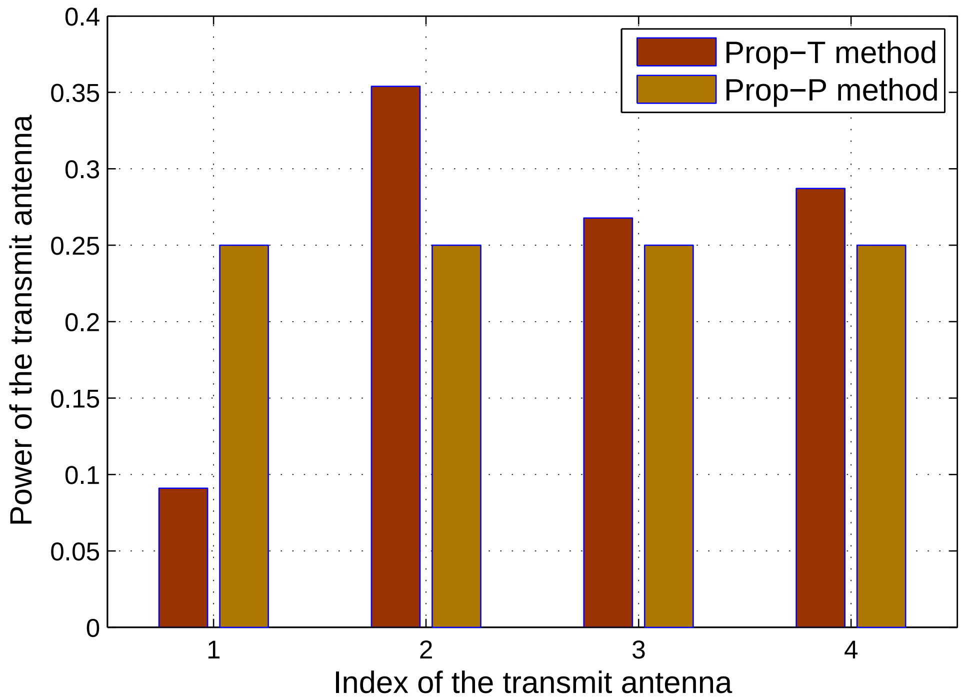

Figure 4.9: Illustration of the power allocation over the transmit antennas under the total power constraint and the per-antenna power constraint, $N_t = 4$, $N_{r,k} = 4$, $\forall k$, $K = 4$, $P_T = 1$, and $P_n = 0.25$, $\forall n$.

higher than that of PBD-DPC method, and the complexity ratio between PBD-DPC and Prop-P methods becomes larger with the increase of the number of transmit antennas.

Table 4.1: Average processing time of Prop-P and PBD-DPC methods (second).

N_t, N_r, K	2, 1, 2	4, 2, 2	6, 3, 2	8, 4, 2
PBD-DPC method	0.1656	0.2481	0.2950	0.4126
Prop-P method	0.3402	1.0913	2.0825	3.2930

4.4 Conclusion

In this chapter, a new beamforming design method is proposed under the total power constraint. Compared with SA-DPC method, the same transmit beamforming and receive combining vectors are obtained, but the proposed method can be easily adapted to the more realistic per-antenna power constraint when practical issues are considered. Then, a new data allocation and transmit beamforming design method under the per-antenna power constraint is proposed. Since the computational complexity of the optimal solution in the literature is high, we transform the original problem into two classical subproblems, which can be efficiently solved with current available techniques. The transmit beamforming vectors are designed to remove one part of interference, and the residual interference which is non-causally known at the base station is pre-subtracted through DPC technique. Compared with PBD-DPC method in the literature, significant improvement of sum rate is achieved. The number of total receive antennas may be larger than that of transmit antennas. Simulation results validate the benefits of the two proposed methods.

5

Proposed power allocation method

As described in Section 3.4, interference is negligible compared with the additive Gaussian noise at low SNR region, it is not crucial to suppress the interference thoroughly. So, zero interference condition can be relaxed, but the beamforming optimization and the power allocation problems become non-convex problems in this case [21]. In the literature, several suboptimal approaches that jointly optimize transmit beamforming vectors and power allocation are developed. In [28], authors analyze the situation where each user has a single receive antenna. This scheme is generalized in [21] where each user is equipped with multiple receive antennas and multiple data streams. However, the power allocation method in [21] requires formulating and solving a geometric programming (GP) iteratively, which exhibits a pretty high computational complexity. In [29] the average signal-to-interference-plus-noise ratio (SINR) of each user is adopted as the performance measurement instead of the SINR of every single data stream. This criterion is not applicable in some cases. For instance, the degradation of bit error rate (BER) performance is caused mainly by the lower subchannel SINR value, even in the case of a higher average SINR. In [30], the connection between the individual SINR and the weighted sum rate is established under the assumption of a single receive antenna per user. The weighted sum rate optimization in multicell MIMO scenarios is investigated. In [31] and [32], it is shown that the power allocation problem is NP hard, and the optimal result is found at the expense of the exponential computational complexity. In [33], [34] and [35], local optimum solutions are studied with appropriate complexity.

Motivated by [21], a new power allocation method in MU-MIMO broadcast interference channels is proposed. We adopt the MMSE beamforming technique in [21], which is an efficient strategy for such optimization problem [21], [29] and [36]. In addition, unlike GP power allocation in [21], the proposed power allocation method allocates the total transmit power iteratively through the very simple water-filling principle, by which more power is allocated to the subchannels with large channel gains. This strategy substantially reduces the computational complexity compared with GP method. Also, since the proposed method allocates the total transmit power iteratively taking the interference from the previous allocated power into account, it also results in a near-capacity sum rate. Therefore, the proposed algorithm is attractive for practical implementation. Additionally, notice that the beamforming vectors are determined separately for each data stream, the number of receive antennas can be larger than that of transmit antennas.

5.1 Proposed power allocation

Consider the downlink MU-MIMO broadcast channels with K users, where a base station is equipped with N_t transmit antennas and transmits $\sum_k L_k = L$ data streams to the users, each user has $N_{r,k}$ receive antennas and receives L_k data streams. The channel state information (CSI) is supposed to be perfectly known at the base station.

In Section 3.4.5, we know that the transmit and receive beamforming vectors can be obtained by MMSE receive beamforming method when the zero forcing condition is relaxed. In this section, we focus on investigating the power allocation optimization problem in (3.59) that is described as

$$\begin{aligned} &\max_{\{p_l\}} \sum_{l=1}^{L} \log_2\left(1 + \frac{p_l G_{ll}}{\sum_{j=1, j \neq l}^{L} p_j G_{lj} + \sigma^2}\right) \\ &\text{subject to } \sum_{l=1}^{L} p_l \leq P_T,\ p_l \geq 0, 1 \leq l \leq L \end{aligned} \tag{5.1}$$

where G_{lj} and G_{ll} represent the coefficient of the interference to the lth subchannel caused by the jth subchannel and the channel gain of the lth subchannel, respectively, which are determined by MMSE receiver, they are supposed to be known in this section.

Since this problem is non-convex, the methods proposed to find optimal solutions in [32], [83] and [84] have exponential computational complexity, which are hard to implement in practice. In [85], a low complexity suboptimal method is proposed, which approximates the original non-convex problem as a geometric programming (GP) optimization one in high SNR region. However, high SNR region can not be always guaranteed in practice, and the method in [85] results in serious distortion in low SNR region. In [21], authors try to solve this problem for general SNR region using GP iteratively, but the computational complexity is still significantly high.

In parallel and interference free subchannels (PIFS), it is well-known that the computationally efficient water-filling power allocation algorithm achieves the optimal solution [86]. Motivated by this, we propose an effective iterative power allocation method (Algorithm 14) which not only reduces substantially the computational complexity but provides high global throughput as well. The proposed method works in both low and high SNR region.

It is known that the power allocation problem in PIFS can be optimally solved through the water-filling principle, which allocates more power to the subchannels who have larger SNR. It also means that the subchannels with weaker channel gain or stronger AWGN will be excluded by this principle. As it will be shown later in this thesis, this principle is also useful to fight against the interference. Particularly, in low SNR region, for the lth ($1 \leq l \leq L$) subchannel in (5.1), the objective function shows that AWGN plays a dominant role compared with interference (i.e., $\sum_{j\neq l}^{L} p_j G_{lj} << \sigma^2$). Thus, in this case, we can approximately consider the subchannels as PIFS-like ones by neglecting the interference. In other words, the PIFS-like hypothesis can be used to approximate the subchannels if the transmit power is small. Inspired by these observations, we propose the following iterative method. The total transmit power constraint P_T is divided into N equal portions (i.e., $\Delta P = \frac{P_T}{N}$, with $\Delta P << \sigma^2$). Then at each iteration n ($1 \leq n \leq N$), we only distribute one small power portion ΔP over the approximated PIFS-like subchannels. Notice that even the interference of the current iteration is omitted, the interference that comes from the previous iterations must be taken into account. For the lth ($1 \leq l \leq L$) subchannel, we denote both the interference coming from the previous iterations and AWGN as $\beta_l^{(n-1)}$, that is

$$\beta_l^{(n-1)} = \sum_{j\neq l}^{L} \Delta p_j^{(n-1)} G_{lj} + \sigma^2 \tag{5.2}$$

where $\Delta p_j^{(n-1)}$ is the already allocated power to the jth subchannel at the previous

Algorithm 14 The proposed iterative power allocation algorithm

$\Delta P = \frac{P_T}{N}$, $\Delta p_l^{(0)} = 0,\ 1 \leq l \leq L$.
for iteration $n = 1 : N$
  for the lth subchannel
    Calculate the term $\beta_l^{(n-1)}$ (5.2).
  end
  Allocate power ΔP by the water-filling algorithm (5.3).
  Calculate $\Delta p_l^{(n)} = \Delta p_l^{(n-1)} + \Delta p_l, \forall l$.
end

$n-1$ iterations. Then the approximated power allocation problem is

$$\begin{aligned} &\max_{\{\Delta p_l\}} \sum_{l=1}^{L} \log_2(1 + \frac{\Delta p_l G_{ll}}{\beta_l^{(n-1)}}) \\ &\text{subject to} \sum_{l=1}^{L} \Delta p_l = \Delta P,\ \Delta p_l \geq 0,\ 1 \leq l \leq L \end{aligned} \tag{5.3}$$

This PIFS-like approximation is tight since the interference has minor effect compared with AWGN, it is quite close to PIFS, especially in low SNR region. In high SNR region, it is also approximately true at the beginning of this approach. With the iteration moving on, the performance degrades a little since $\Delta p_j^{(n-1)} G_{lj}$ plays a more important role in (5.2). It is easy to show that (5.3) is a convex optimization problem, which can be optimally solved by the water-filling algorithm. The total transmit power is allocated after N iterations.

In the proposed method, according to (5.2), since we consider the interference and AWGN together as $\beta_l^{(n-1)}$ at each iteration, the water-filling principle will exclude the subchannel if $\beta_l^{(n-1)}$ is too strong, which is equivalent to omit the subchannels who have strong interference. So, this approach suppresses the strong interference as well. Therefore, the more portions the total transmit power is divided into, the larger global throughput we can expect. These observations are validated by the simulation results in the next section.

Algorithm 14 summarizes the above proposed iterative water-filling power allocation algorithm. This method can also be easily extended to weighted sum rate optimization problems. Similarly to [21], the transmit beamforming and receive combining vectors are optimized alternatively in virtual uplink multiuser access channels and downlink MU-MIMO broadcast channels, respectively.

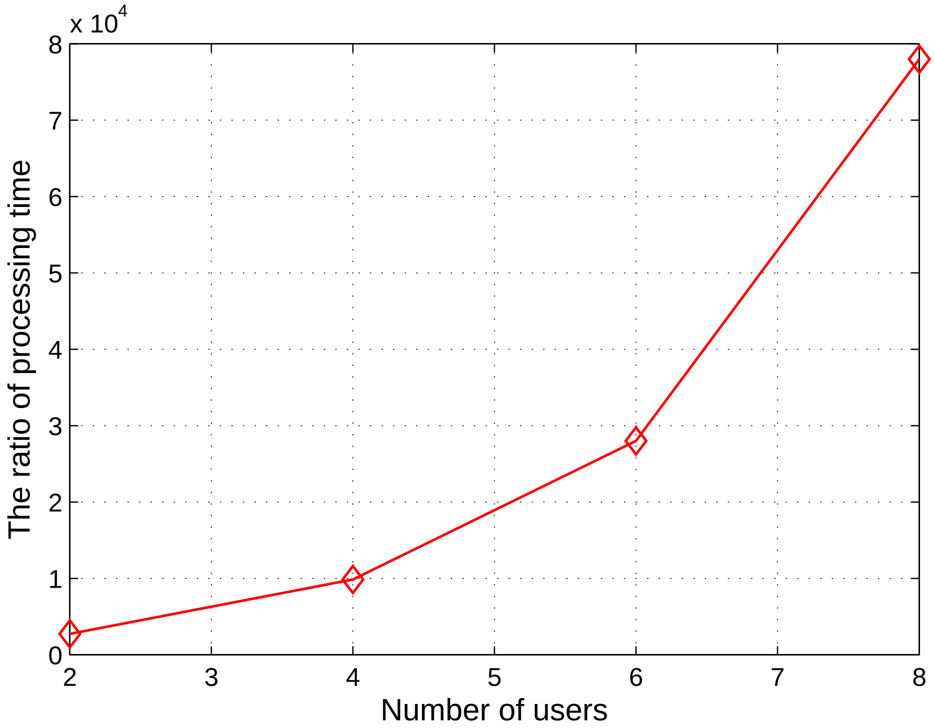

Figure 5.1: The ratio of the processing time between GP method over the proposed method.

5.2 Simulation results

This section presents the simulation results of the proposed iterative power allocation method. In Figure 5.1 the ratio of processing time of GP method in [21] to the proposed method is presented. The base station has $N_t = 4$ transmit antennas, each user is equipped with $N_r = N_{r,k} = 2$ $(\forall k)$ receive antennas, and the number of users is $K = 2, 4$, 6, or 8. Similarly to [21], the number of served subchannels is $L = KN_r$. The processing time is recorded for 8 iterations for GP method, which is a moderate number of iterations in [21]. In the proposed method, the total transmit power P_T is divided into 5 equal portions (i.e. $N = 5$). Note that the water-filling algorithm is used N times, and the computational complexity of the classical water-filling algorithm is $\mathbb{O}(L\log_2 L)$ [86]. It can be concluded that the computational complexity of the proposed method is $\mathbb{O}(NL\log_2 L)$. In Figure 5.1 we observe that the proposed method performs much faster than GP method, and the ratio increases when the number of users becomes larger.

Figure 5.2 illustrates the sum capacity provided by DPC algorithm [13], the achievable sum rates of GP algorithm and the proposed algorithm, respectively. We can

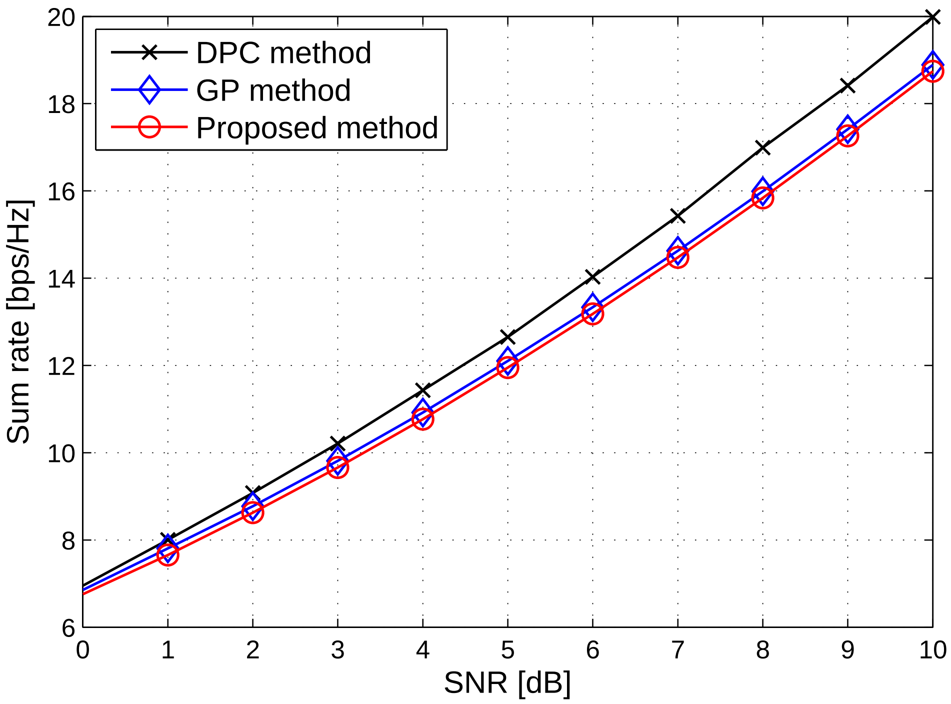

Figure 5.2: Sum rate comparison of DPC method, GP method, and the proposed method. $N_t = 8$, $N_r = 2$, and $K = 4$.

observe that GP algorithm performs only slightly better than the proposed method. Therefore, the proposed method is more practical due to the greatly reduced computational complexity.

To show that the sum rate of the proposed method is close to the sum capacity, in Figure 5.3, we show the simulation results with respect to the power division number N, where we consider a MU-MIMO broadcast channels system with $N_t = 4$, $N_r = 4$, and $K = 4$. Note that $N = 1$ is the classical water-filling algorithm, the subchannels are considered as PIFS, so it does not work here. The proposed method ($N \geq 2$) assigns transmit power iteratively by taking the interference that comes from the previous iterations into account, and then follows the water-filling principle. Figure 5.3 illustrates that the gap between the sum rate provided by the proposed method and sum capacity is reduced when the power division number is increased. Notice that even with a quite small power division number N, most of the throughput can be achieved. Therefore, the proposed technique is promising for practical implementation.

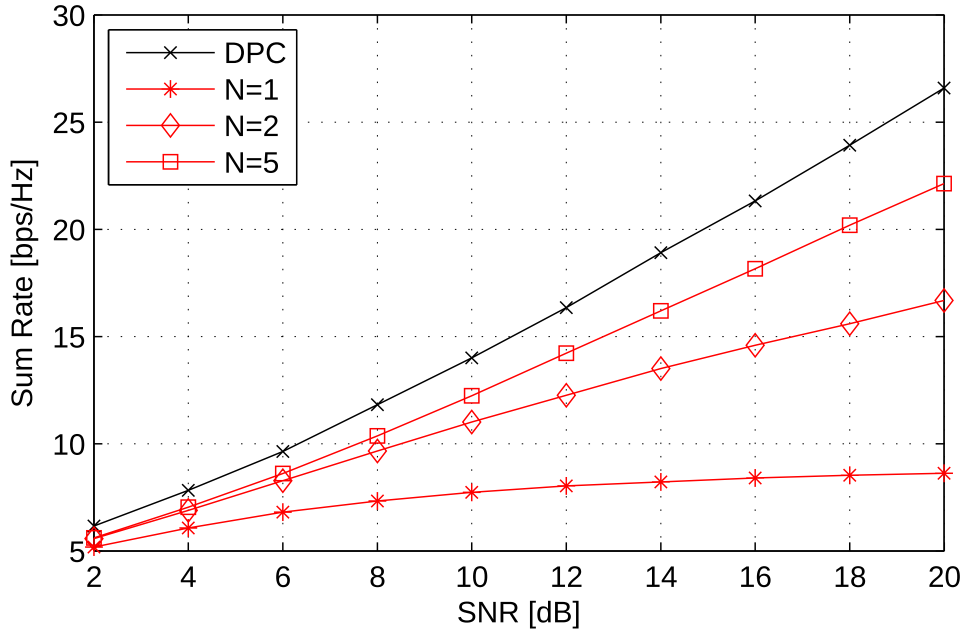

Figure 5.3: Illustration of the sum rate approaching the sum capacity with power division number $N = 1, 2$ and 5, respectively.

5.3 Conclusion

We propose an efficient power allocation method taking advantage of the water-filling principle. In order to minimize the influence of interference, the total transmit power is first divided into small equal portions, then each portion is allocated by approximating the original problem as a convex solvable one and the simple water-filling algorithm is adopted to obtain the quasi-optimal solution. The computational complexity is substantially reduced compared with GP method in the literature, with only a slight performance degradation. This new method is very attractive in practice, some numerical results validate the proposed technique.

6

Conclusion and future work

Generally, for algorithm design, there is always a trade off between the performance and the computational complexity. A better performance often requires a higher computational complexity, whereas a simple implementation usually results in a loss in performance. After investigating these two aspects of state of the art methods in the field of MU-MIMO capacity optimization, this dissertation proposes two compromised and promising algorithms with comparable performance as the optimal one but requiring much less computational complexity. This final chapter summarizes the thesis and a brief discussion on possibilities for future work is provided.

6.1 Conclusion

In this thesis, we explored the transmit beamforming techniques for multiuser MIMO broadcast channels in terms of capacity optimization. Because of the lack of radio spectrum, the benefits of MIMO system make itself a very promising way to improve the spectrum efficiency. We know that the capacity can grow linearly with the number of transmit or receive antennas. An extension of MIMO named MU-MIMO is also very interesting because the spatial signatures induced by the spatially dispersed users can be exploited. However, due to the coupled structure of the transmit signals, the optimization problems in this system are usually non-convex and thus are difficult to solve.

Due to the high computational complexity of the optimal transmit beamforming design in multiuser MIMO broadcast channel, a great effort is made to explore suboptimal methods that have appropriate performance and are easy to implement. Generally, there are two directions for the transmit beamforming design considering the suboptimal methods, we denote them as non-linear methods and linear methods. For the non-linear method, the transmit signals are usually encoded in a successive way, then the interference is separated into two different parts. The part caused by the previously encoded signals is known non-causally, and thus can be removed by DPC techniques. This gives more degrees of freedom for transmit beamforming design, since only the second part of interference caused by the subsequently encoded signals is need to suppress. This strategy usually leads the performance of the non-linear methods to be very close to the sum capacity. But notice that DPC has to be adopted in these methods, the structures of the transmitter and receiver become more complex. Considering the linear methods, most of them aims to suppress the entire interference via transmit beamforming vectors, which means that a more strict constraint on the transmit beamforming design is imposed compared with non-linear method. This strategy results in simple structures of the transmitter and receiver, at the expense of some more signal processing.

In practice, the transmit power of each antenna is usually limited by the power amplifier. So, transmit beamforming techniques are often performed under per-antenna power constraint. In the literature, this kind of optimization problem is usually transformed into a semidefinite programming, then the existing standard optimization packages can be used to solve it. However, the optimal transmit beamforming design has very high computational complexity. Considering the suboptimal method with low computational complexity, BD method and DPC technique are often adopted. It is commonly known that BD strategy has limited performance, since this method calculates a number of transmit beamforming vectors simultaneously, which is obviously suboptimal because the channel gain of each individual data stream may vary widely, and the contribution of those weak subchannels might be negligible but severe constraints are imposed on the subsequent subchannels. Consequently, there is still a large room for improvement in the transmit beamforming design under the per-antenna power constraint.

According to the above issues and motivated by the idea of successive allocation strategy, we propose a new transmit beamforming method under the per-antenna power constraint. At each step, one subchannel that has largest channel gain among the candidates is found for transmitting the signal, since this optimization problem is still difficult to solve, we cope with it in an iterative way. At each iteration, one transmit beamforming (or receive combining) vector is optimized with the receive combining (or transmit beamforming) vector fixed, and it is shown that it converges within a small number of iterations. The computational complexity is significantly reduced compared with the optimal solution and the gap of the performance between them is very small.

When the interference is not removed completely (e.g. MMSE method), power allocation becomes a NP hard problem. The optimal result can be found at the expense of the exponential computational complexity, which prohibits its implementation in practice. Motivated by the classical water-filling algorithm, more power should be assigned to the channels with larger channel gains. We propose an iterative water-filling power allocation method with low complexity. Simulation results show that the performance of this proposed method is close to the optimal value.

6.2 Future work

Base on this work, some suggestions for future works are given as follows:

- Complexity is a very essential aspect to evaluate an algorithm. In this thesis, we give a rough complexity analysis of the proposed methods via simulation results. The analytical theoretical complexity analysis is very important direction in the future.
- In this thesis, the perfect channel state information (CSI) is assumed to be known at the base station. However, perfect CSI is not realistic in practice. How the methods behave when CSI is imperfect is a promising direction. In addition, most of the existing methods only consider frequency flat-fading channels. The extension to frequency selective channels also should be studied.
- In the proposed methods, THP technique is used to pre-subtract the non-causally known interference, and modulo device has to be used for the implementation of THP technique, which increases the complexity of the transmitter and the receiver. In the future work, other techniques that are easier to implement DPC technique should be investigated.
- In this dissertation, only global throughput is optimized, SNR of each subchannel is not considered. However, SNR has a strong effect on the bit error rate (BER), low SNR often leads to a high BER, which is not permitted in practice. In addition, fairness should be also considered in practice. In order to maximize the global throughput, some weak users are neglected in this work. While in practice, the system sometimes has to guarantee that every registered user is

served. In the future work, these parameters should be taken into account for system design.

– Multi-cellular MU-MIMO broadcast channels which consider multiple cellulars simultaneously attract a great attention recently. Not only the interference from one cellular is removed, the interference that comes from other cellulars also should be taken into account. In the future work, a promising direction is to adopt the proposed methods in these more general multi-cellular MU-MIMO broadcast channels.

7

Appendixes

7.1 Appendix A : Proof of the general BC-MAC duality

We prove the general BC-MAC duality by formulating a set of $\{q_{i,j}\}$ for the dual MAC that satisfy the following two constraints : the transmit power constraint, $\sum_{i,j} \sigma^2 q_{i,j} = P_T$, and the SINR constrains, $SINR_{i,j} = SINR^M_{i,j}, \forall i,\ j$.

To meet the SINR constraints, we choose the transmit power for each data stream as

$$q_{i,j} = \frac{SINR_{i,j} \boldsymbol{v}_{i,j}^H \left(\sum_{k=1}^{i-1} \sum_{l=1}^{N} q_{k,l} \boldsymbol{H}_k{}^H \boldsymbol{u}_{k,l} \boldsymbol{u}_{k,l}{}^H \boldsymbol{H}_k + \sum_{l=1}^{j-1} q_{i,l} \boldsymbol{H}_i{}^H \boldsymbol{u}_{i,l} \boldsymbol{u}_{i,l}{}^H \boldsymbol{H}_i + \boldsymbol{I} \right) \boldsymbol{v}_{i,j}}{\left| \boldsymbol{v}_{i,j}^H \boldsymbol{H}_i^H \boldsymbol{u}_{i,j} \right|^2} \tag{7.1}$$

It is evident that the resulting $SINR^M_{i,j}$ satisfies the condition $SINR_{i,j} = SINR^M_{i,j}$, $\forall i,\ j$.

We next show that the selected $\{q_{i,j}\}$ given by (7.1) satisfy the transmit power constraint. According to the definition of $SINR_{i,j}$ and $SINR_{i,j}^{M}$, it follows that

$$\frac{q_{i,j}\left|\boldsymbol{v}_{i,j}^{H}\boldsymbol{H}_i^{H}\boldsymbol{u}_{i,j}\right|^2}{\boldsymbol{v}_{i,j}^{H}\left(\sum_{k=1}^{i-1}\sum_{l=1}^{N} q_{k,l}\boldsymbol{H}_k{}^{H}\boldsymbol{u}_{k,l}\boldsymbol{u}_{k,l}{}^{H}\boldsymbol{H}_k + \sum_{l=1}^{j-1} q_{i,l}\boldsymbol{H}_i{}^{H}\boldsymbol{u}_{i,l}\boldsymbol{u}_{i,l}{}^{H}\boldsymbol{H}_i + \boldsymbol{I}\right)\boldsymbol{v}_{i,j}} = \frac{p_{i,j}\left|\boldsymbol{u}_{i,j}^{H}\boldsymbol{H}_i\boldsymbol{v}_{i,j}\right|^2}{\sum_{k=i+1}^{K}\sum_{l=1}^{N} p_{k,l}\left|\boldsymbol{u}_{i,j}^{H}\boldsymbol{H}_i\boldsymbol{v}_{k,l}\right|^2 + \sum_{l=j+1}^{N} p_{i,l}\left|\boldsymbol{u}_{i,j}^{H}\boldsymbol{H}_i\boldsymbol{v}_{i,l}\right|^2 + \sigma^2} \tag{7.2}$$

By rearranging the terms in (7.2), we can list the equations related to all the data streams as follows :

$$p_{1,1}\boldsymbol{v}_{1,1}^{H}\boldsymbol{v}_{1,1} = q_{1,1}\left(\sum_{k=2}^{K}\sum_{l=1}^{N} p_{k,l}\left|\boldsymbol{u}_{1,1}^{H}\boldsymbol{H}_1\boldsymbol{v}_{k,l}\right|^2 + \sum_{l=2}^{N} p_{1,l}\left|\boldsymbol{u}_{1,1}^{H}\boldsymbol{H}_1\boldsymbol{v}_{1,l}\right|^2 + \sigma^2\right) \tag{7.3}$$

$$p_{1,2}\left(\boldsymbol{v}_{1,2}^{H}(q_{1,1}\boldsymbol{H}_1{}^{H}\boldsymbol{u}_{1,1}\boldsymbol{u}_{1,1}{}^{H}\boldsymbol{H}_1 + \boldsymbol{I})\boldsymbol{v}_{i,2}\right) = q_{1,1}\left(\sum_{k=2}^{K}\sum_{l=1}^{N} p_{k,l}\left|\boldsymbol{u}_{1,2}^{H}\boldsymbol{H}_1\boldsymbol{v}_{k,l}\right|^2 + \sum_{l=3}^{N} p_{1,l}\left|\boldsymbol{u}_{1,2}^{H}\boldsymbol{H}_1\boldsymbol{v}_{1,l}\right|^2 + \sigma^2\right) \tag{7.4}$$

$$p_{K,N}\left(\boldsymbol{v}_{K,N}^{H}(\sum_{k=1}^{K-1}\sum_{l=1}^{N} q_{k,l}\boldsymbol{H}_k{}^{H}\boldsymbol{u}_{k,l}\boldsymbol{u}_{k,l}{}^{H}\boldsymbol{H}_k + \sum_{l=1}^{N-1} q_{K,l}\boldsymbol{H}_K{}^{H}\boldsymbol{u}_{K,l}\boldsymbol{u}_{K,l}{}^{H}\boldsymbol{H}_K + \boldsymbol{I})\boldsymbol{v}_{K,N}\right) = q_{K,N}\sigma^2 \tag{7.5}$$

By adding the above equations together, we obtain

$$\sum_{i=1}^{K}\sum_{j=1}^{N} p_{i,j}\boldsymbol{v}_{i,j}^{H}\boldsymbol{v}_{i,j} = \sum_{i=1}^{K}\sum_{j=1}^{N}\sigma^2 q_{i,j} \tag{7.6}$$

therefore, we have

$$\text{trace}(\sum_{i=1}^{K}\sum_{j=1}^{N} p_{i,j}\boldsymbol{v}_{i,j}\boldsymbol{v}_{i,j}^{H}) = \sum_{i=1}^{K}\sum_{j=1}^{N}\sigma^2 q_{i,j} = P_T \tag{7.7}$$

7.2 Appendix B : Use Lagrange duality method to solve (4.11)

The Lagrangian of (4.11) can be written as [24]

$$\begin{aligned}
&L(\{\boldsymbol{F}_j\},\{\mu_n\}) = \\
&\sum_{j=1}^{l}\log_2(1+\frac{\boldsymbol{h}_j^H\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H\boldsymbol{h}_j}{\sigma^2}) - \sum_{n=1}^{N_t}\mu_n\left(\text{trace}(\sum_{j=1}^{l}\boldsymbol{B}_n\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H) - P_n\right) \\
&= \sum_{j=1}^{l}\log_2(1+\frac{\boldsymbol{h}_j^H\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H\boldsymbol{h}_j}{\sigma^2}) - \sum_{j=1}^{l}\text{trace}(\sum_{n=1}^{N_t}\mu_n\boldsymbol{B}_n\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H) + \sum_{n=1}^{N_t}\mu_n P_n \\
&= \sum_{j=1}^{l}\left(\log_2(1+\frac{\boldsymbol{h}_j^H\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H\boldsymbol{h}_j}{\sigma^2}) - \text{trace}(\sum_{n=1}^{N_t}\mu_n\boldsymbol{B}_n\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H)\right) + \sum_{n=1}^{N_t}\mu_n P_n
\end{aligned} \tag{7.8}$$

where $\boldsymbol{B}_{n(N_t\times N_t)} = \text{diag}(0,\cdots 0,1,0,\cdots,0)$ is a diagonal matrix with only the nth diagonal element being 1. Then the Lagrange dual function can be written as

$$g(\{\mu_n\}) = \max_{\boldsymbol{F}_j\succeq\boldsymbol{0}} L(\{\boldsymbol{F}_j\},\{\mu_n\}) \tag{7.9}$$

and the Lagrange dual problem is [82]

$$\min_{\mu_n\geq 0,\forall n} g(\{\mu_n\}) \tag{7.10}$$

Since the problem (4.11) is convex and satisfies the Slater's condition, the duality gap between (4.11) and (7.10) is zero. Thus, (4.11) can be solved equivalently by solving (7.10). Moreover, (7.10) is convex and can be solved by the ellipsoid method in [76] when the optimal $\{\boldsymbol{F}_j^\star\}$ in (7.9) is found. Specifically, $\boldsymbol{g}$ is defined as $\boldsymbol{g} = [g_1,\cdots,g_{N_t}]^T$ with

$$g_n = \frac{\partial L(\{\boldsymbol{F}_j\},\{\mu_n\})}{\partial\mu_n} = P_n - \sum_{j=1}^{l}\text{trace}(\boldsymbol{B}_n\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H), \quad 1\leq n\leq N_t \tag{7.11}$$

and $\boldsymbol{\mu}$ is defined as $\boldsymbol{\mu} = [\mu_1,\cdots,\mu_{N_t}]^T$ with random initialized value. If $\boldsymbol{\mu}^+$ and $\boldsymbol{M}^+$ are denoted as the upgraded $\boldsymbol{\mu}$ and the upgraded $\boldsymbol{M}$ after each iteration, respectively, and $\boldsymbol{M}$ is initialized as $\boldsymbol{M} = \boldsymbol{I}_{\boldsymbol{N_t}}$, then according to the ellipsoid method, $\boldsymbol{\mu}$ is upgraded as

$$\boldsymbol{\mu}^{+} = \boldsymbol{\mu} - \frac{1}{(N_t+1)\sqrt{\boldsymbol{g}^T\boldsymbol{M}\boldsymbol{g}}}\boldsymbol{M}\boldsymbol{g} \tag{7.12}$$

and $\boldsymbol{M}$ is upgraded as

$$\boldsymbol{M}^{+} = \frac{N_t^2}{N_t^2-1}\left(\boldsymbol{M} - \frac{2}{(n+1)\boldsymbol{g}^T\boldsymbol{M}\boldsymbol{g}}\boldsymbol{M}\boldsymbol{g}\boldsymbol{g}^T\boldsymbol{M}\right) \tag{7.13}$$

We alternatively calculate (7.12) and solve problem (7.9) until all $\mu_n' s$ converge to a preset threshold. Here, we focus on solving (7.9) for $\{\boldsymbol{F}_j^\star\}$ with fixed $\{\mu_n\}$. Notice that the maximization problem in (7.8) can be separated into l independent subproblems each involving only one $\boldsymbol{F}_j$. By discarding the irrelevant terms, the corresponding subproblem for a given j can be expressed as

$$\max_{\boldsymbol{F}_j>0} \log_2(1+\frac{\boldsymbol{h}_j^H\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H\boldsymbol{h}_j}{\sigma^2}) - \text{trace}(\sum_{n=1}^{N_t}\mu_n\boldsymbol{B}_n\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H) \tag{7.14}$$

If $\boldsymbol{B}_\mu = \sum_{n=1}^{N_t}\mu_n\boldsymbol{B}_n$ is defined, since trace$(\boldsymbol{X}\boldsymbol{Y})$ = trace$(\boldsymbol{Y}\boldsymbol{X})$, and $\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j$ of the dimension $(N_t-j+1)\times(N_t-j+1)$ is an invertible matrix, then

$$\text{trace}(\boldsymbol{B}_\mu\boldsymbol{A}_j\boldsymbol{F}_j\boldsymbol{A}_j^H) = \text{trace}((\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{\frac{1}{2}}\boldsymbol{F}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{\frac{1}{2}}) \tag{7.15}$$

We define

$$\hat{\boldsymbol{F}}_j = (\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{\frac{1}{2}}\boldsymbol{F}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{\frac{1}{2}} \tag{7.16}$$

therefore

$$\boldsymbol{F}_j = (\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{-\frac{1}{2}}\hat{\boldsymbol{F}}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{-\frac{1}{2}} \tag{7.17}$$

(7.14) can be reformulated to

$$\max_{\hat{\boldsymbol{F}}_j>0} \log_2(1+\frac{\boldsymbol{h}_j^H\boldsymbol{A}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{-\frac{1}{2}}\hat{\boldsymbol{F}}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{-\frac{1}{2}}\boldsymbol{A}_j^H\boldsymbol{h}_j}{\sigma^2}) - \text{trace}(\hat{\boldsymbol{F}}_j) \tag{7.18}$$

Note that the SVD of $\boldsymbol{h}_j^H\boldsymbol{A}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{-\frac{1}{2}}$ is

$$\boldsymbol{h}_j^H\boldsymbol{A}_j(\boldsymbol{A}_j^H\boldsymbol{B}_\mu\boldsymbol{A}_j)^{-\frac{1}{2}} = \delta_j\hat{\boldsymbol{v}}_j^H \tag{7.19}$$

therefore, (7.18) is maximized by choosing $\hat{\boldsymbol{F}}_j = \lambda_j\hat{\boldsymbol{v}}_j\hat{\boldsymbol{v}}_j^H$ and λ_j is decided by water-filling algorithm, i.e.,

$$\lambda_j = \left[\eta - \frac{1}{\delta_j}\right]^+ \tag{7.20}$$

where $[x]^+ = \max(0, x)$, and η is the water level. Finally, we find $\boldsymbol{F}_j$ according to (7.17)

$$\begin{aligned}\boldsymbol{F}_j &= \lambda_j (\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-\frac{1}{2}} \hat{\boldsymbol{v}}_j \hat{\boldsymbol{v}}_j^H (\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-\frac{1}{2}} \\ &= \lambda_j / \delta_j^2 (\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-1} \boldsymbol{A}_j^H \boldsymbol{h}_j \boldsymbol{h}_j^H \boldsymbol{A}_j (\boldsymbol{A}_j^H \boldsymbol{B}_\mu \boldsymbol{A}_j)^{-1}\end{aligned} \tag{7.21}$$

(7.21) and (7.12) are alternatively calculated until $\{\mu_n\}$ converges. Note that the closed-form of $\{\boldsymbol{F}_j^\star\}$ is found by (7.21) when $\{\mu_n\}$ converges.

Bibliography

[1] INSTITUE OF ELECTRICAL AND ELECTRONICS ENGINEERS (IEEE). *Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: enhancements for very high throughput for operation in bands below 6GHz,*. Tech. Rep., IEEEP802.11ac/D1.0 Standard, 2012. 5, 19

[2] 3GPP. *TR 36.913 Requirements for further advancements for E-UTRA (LTE-Advanced),*. Tech. Rep., 3GPP, 2011. 5, 19

[3] J.H. WINTERS. **On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment**. *Selected Areas in Communications, IEEE Journal on*, **5**(5):871–878, Jun. 1987. 5, 19

[4] G. J. FOSCHINI. **Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas**. *Bell Labs Technical Journal*, **1**(2):41–59, Autumn 1996. 5, 6, 19, 21, 23

[5] EMRE TELATAR. **Capacity of Multi-antenna Gaussian Channels**. *European Transactions on Telecommunications*, **10**(6):585–595, 1999. 5, 19, 30, 43

[6] N. JINDAL, S. VISHWANATH, AND A. GOLDSMITH. **On the duality of Gaussian multiple-access and broadcast channels**. *Information Theory, IEEE Transactions on*, **50**(5):768–783, May 2004. 5, 19, 36, 40

[7] A. GOLDSMITH, S.A. JAFAR, N. JINDAL, AND S. VISHWANATH. **Capacity limits of MIMO channels**. *Selected Areas in Communications, IEEE Journal on*, **21**(5):684–702, Jun. 2003. 5, 19

[8] B. SKLAR. *Digintal communications: fundamentals and applications,*. Prentice Hall PTR, 2nd edition, 2002. 5, 21, 22

[9] A. KURVE. **Multi-user MIMO systems: the future in the making**. *Potentials, IEEE*, **28**(6):37–42, Nov. 2009. 5, 22

[10] AROGYASWAMI PAULRAJ, ROHIT NABAR, AND DHANANJAY GORE. *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003. 5, 21, 22, 28, 29, 30, 33

[11] J. G. PROAKIS AND M. SALEHI. *Digintal communications,*. McGraw-Hill, 5th edition, 2007. 5, 21, 22

[12] LINGJIA LIU, RUNHUA CHEN, S. GEIRHOFER, K. SAYANA, ZHIHUA SHI, AND YONGXING ZHOU. **Downlink MIMO in LTE-advanced: SU-MIMO vs. MU-MIMO**. *Communications Magazine, IEEE*, **50**(2):140–147, Feb 2012. 5, 22

[13] N. JINDAL, W. RHEE, S. VISHWANATH, S. A. JAFAR, AND A. GOLDSMITH. **Sum power iterative water-filling for multi-antenna Gaussian broadcast channels**. *Information Theory, IEEE Transactions on*, **51**(4):1570–1580, Apr. 2005. 6, 23, 40, 105

[14] G. CAIRE AND S. SHAMAI. **On the achievable throughput of a multi-antenna Gaussian broadcast channel**. *Information Theory, IEEE Transactions on*, **49**(7):1691–1706, Jul. 2003. 6, 23, 35, 47, 48, 84

[15] P. TEJERA, W. UTSCHICK, G. BAUCH, AND J.A. NOSSEK. **Subchannel Allocation in Multiuser Multiple-Input-Multiple-Output Systems**. *Information Theory, IEEE Transactions on*, **52**(10):4721–4733, Oct. 2006. 6, 7, 23, 24, 52

[16] A. D. DABBAGH AND D. J. LOVE. **Precoding for Multiple Antenna Gaussian Broadcast Channels With Successive Zero-Forcing**. *Signal Processing, IEEE Transactions on*, **55**(7):3837–3850, Jul. 2007. 6, 23, 48, 50

[17] C. B. PEEL, B. M. HOCHWALD, AND A. L. SWINDLEHURST. **A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization**. *Communications, IEEE Transactions on*, **53**(1):195–202, Jan. 2005. 6, 23, 55

[18] Q. H. SPENCER, A. L. SWINDLEHURST, AND M. HAARDT. **Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels**. *Signal Processing, IEEE Transactions on*, **52**(2):461–471, Feb. 2004. 6, 23, 56

[19] C. B. CHAE, D. MAZZARESE, T. INOUE, AND R. W. HEATH. **Coordinated Beamforming for the Multiuser MIMO Broadcast Channel With Limited Feedforward**. *Signal Processing, IEEE Transactions on*, **56**(12):6044–6056, Dec. 2008. 6, 23, 59

[20] C. GUTHY, W. UTSCHICK, R. HUNGER, AND M. JOHAM. **Efficient Weighted Sum Rate Maximization With Linear Precoding**. *Signal Processing, IEEE Transactions on*, **58**(4):2284–2297, Apr. 2010. 6, 23, 63, 84

[21] M. CODREANU, A. TOLLI, M. JUNTTI, AND M. LATVA-AHO. **Joint Design of Tx-Rx Beamformers in MIMO Downlink Channel**. *Signal Processing, IEEE Transactions on*, **55**(9):4639–4655, Sep. 2007. 7, 8, 23, 24, 65, 67, 68, 69, 101, 102, 103, 104, 105

[22] X. ZHENG, Y. XIE, J. LI, AND P STOICA. **MIMO Transmit Beamforming Under Uniform Elemental Power Constraint**. *Signal Processing, IEEE Transactions on*, **55**(11):5395–5406, Nov. 2007. 7, 23

[23] A. WIESEL, Y.C. ELDAR, AND S. SHAMAI. **Zero-Forcing Precoding and Generalized Inverses**. *Signal Processing, IEEE Transactions on*, **56**(9):4409–4418, Sep. 2008. 7, 23

[24] R. ZHANG. **Cooperative Multi-Cell Block Diagonalization with Per-Base-Station Power Constraints**. *Selected Areas in Communications, IEEE Journal on*, **28**(9):1435–1445, Dec. 2010. 7, 23, 90, 115

[25] L.-N. TRAN, M. JUNTTI, M. BENGTSSON, AND B. OTTERSTEN. **Beamformer Designs for MISO Broadcast Channels with Zero-Forcing Dirty Paper Coding**. *Wireless Communications, IEEE Transactions on*, **12**(3):1173–1185, Mar. 2013. 7, 23, 24, 89, 93

[26] L.-N. TRAN, M. JUNTTI, M. BENGTSSON, AND B. OTTERSTEN. **Weighted Sum Rate Maximization for MIMO Broadcast Channels Using Dirty Paper Coding and Zero-forcing Methods**. *Communications, IEEE Transactions on*, **61**(6):2362–2373, Jun. 2013. 7, 23, 24, 25, 78, 80, 81, 91, 92, 93

[27] L. ZHANG, R. ZHANG, Y.-C. LIANG, Y. XIN, AND H.V. POOR. **On Gaussian MIMO BC-MAC Duality With Multiple Transmit Covariance Constraints**. *Information Theory, IEEE Transactions on*, **58**(4):2064–2078, Apr. 2012. 7, 23, 24, 36, 37, 41, 42, 43, 44, 76, 77, 91

[28] M. SCHUBERT AND H. BOCHE. **Solution of the multiuser downlink beamforming problem with individual SINR constraints**. *Vehicular Technology, IEEE Transactions on*, **53**(1):18–28, Jan. 2004. 8, 40, 101

[29] Y. H. YANG, S. C. LIN, AND H. J. SU. **Multiuser MIMO Downlink Beamforming Design Based on Group Maximum SINR Filtering**. *Signal Processing, IEEE Transactions on*, **59**(4):1746–1758, Apr. 2011. 8, 40, 65, 84, 101, 102

[30] C. W. TAN, M. CHIANG, AND R. SRIKANT. **Maximizing Sum Rate and Minimizing MSE on Multiuser Downlink: Optimality, Fast Algorithms and

Equivalence via Max-min SINR. *Signal Processing, IEEE Transactions on*, **59**(12):6127–6143, Dec. 2011. 8, 67, 68, 101

[31] E. Bjornson, G. Zheng, M. Bengtsson, and B. Ottersten. **Robust Monotonic Optimization Framework for Multicell MISO Systems**. *Signal Processing, IEEE Transactions on*, **60**(5):2508–2523, May 2012. 8, 101

[32] W. Utschick and J. Brehmer. **Monotonic Optimization Framework for Coordinated Beamforming in Multicell Networks**. *Signal Processing, IEEE Transactions on*, **60**(4):1899–1909, Apr. 2012. 8, 101, 103

[33] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo. **Coordinated Beamforming for MISO Interference Channel: Complexity Analysis and Efficient Algorithms**. *Signal Processing, IEEE Transactions on*, **59**(3):1142–1157, Mar. 2011. 8, 101

[34] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He. **An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel**. *Signal Processing, IEEE Transactions on*, **59**(9):4331–4340, Sep. 2011. 8, 101

[35] E. Bjornson, N. Jalden, M. Bengtsson, and B. Ottersten. **Optimality Properties, Distributed Strategies, and Measurement-Based Evaluation of Coordinated Multicell OFDMA Transmission**. *Signal Processing, IEEE Transactions on*, **59**(12):6086–6101, Dec. 2011. 8, 101

[36] A. Khabbazibasmenj, F. Roemer, S.A. Vorobyov, and M. Haardt. **Sum-Rate Maximization in Two-Way AF MIMO Relaying: Polynomial Time Solutions to a Class of DC Programming Problems**. *Signal Processing, IEEE Transactions on*, **60**(10):5478–5493, Oct. 2012. 8, 102

[37] J. Hoydis, M. Kobayashi, and M. Debbah. **Green Small-Cell Networks**. *Vehicular Technology Magazine, IEEE*, **6**(1):37–43, March 2011. 19

[38] IEEE Standards Association. *IEEE 802.11n-2009 Amendment 5: Enhancements for Higher Throughput,*. Tech. Rep., IEEE, 2011. 20

[39] Working Group on Broadband Wireless Access Standards. *Air Interface for Fixed and Mobile Broadband Wireless Access System,*. Tech. Rep., IEEE, 2011. 20

[40] S. Alamouti. **A simple transmit diversity technique for wireless communications**. *Selected Areas in Communications, IEEE Journal on*, **16**(8):1451–1458, Oct. 1998. 20

[41] VAHID TAROKH, HAMID JAFARKHANI, AND A.R. CALDERBANK. **Space-time block coding for wireless communications: performance results**. *Selected Areas in Communications, IEEE Journal on*, **17**(3):451–460, Mar. 1999. 20

[42] LIZHONG ZHENG AND D.N.C. TSE. **Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels**. *Information Theory, IEEE Transactions on*, **49**(5):1073–1096, May 2003. 20

[43] G. DIVID. *Introduction to electrodynamics,*. Prentice Hall, third edition, 1999. 21

[44] W. JAKES. **Diversity Techniques,**. *Microwave Mobile Communications*, 1974. 21

[45] P.W. WOLNIANSKY, G.J. FOSCHINI, G.D. GOLDEN, AND R. VALENZUELA. **V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel**. In *Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on*, pages 295–300, Sep. 1998. 21

[46] C. WINDPASSINGER. *Detection and precoding for multiple input multiple output channels,*. PhD thesis, University Erlangen-Nurnberg, Germany, 2004. 22

[47] CHI-YUK CHIU, JIE-BANG YAN, AND R.D. MURCH. **24-Port and 36-Port Antenna Cubes Suitable for MIMO Wireless Communications**. *Antennas and Propagation, IEEE Transactions on*, **56**(4):1170–1176, Apr. 2008. 22

[48] F. RUSEK, D. PERSSON, BUON KIONG LAU, E.G. LARSSON, T.L. MARZETTA, O. EDFORS, AND F. TUFVESSON. **Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays**. *Signal Processing Magazine, IEEE*, **30**(1):40–60, Jan. 2013. 22

[49] S. VISHWANATH, N. JINDAL, AND A. GOLDSMITH. **Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels**. *Information Theory, IEEE Transactions on*, **49**(10):2658–2668, Oct. 2003. 23, 36, 40

[50] WILLIAM C. JAKES AND DONALD C. COX, editors. *Microwave Mobile Communications*. Wiley-IEEE Press, 1994. 27

[51] WILLIAM C. Y. LEE, editor. *Mobile Communications Engineering: Theory and Applicationss*. McGraw-Hill, New York, second version edition, 1982. 27

[52] J. D. PARSONS, editor. *The Mobile Radio Propagation Channel*. Wiley, New York, second version edition, 1992. 27

[53] THEODORE RAPPAPORT, editor. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, second version edition, 2001. 27

[54] H.L. BERTONI. *Radio propagation for modern wireless systems*. Prentice Hall PTR, 2000. 27

[55] F.D. NEESER AND J.L. MASSEY. **Proper complex random processes with applications to information theory**. *Information Theory, IEEE Transactions on*, **39**(4):1293–1302, Jul. 1993. 29

[56] E. T. JAYNES. **Information Theory and Statistical Mechanics**. *Phys. Rev.*, **106**:620–630, May 1957. 32

[57] M. COSTA. **Writing on dirty paper**. *Information Theory, IEEE Transactions on*, **29**(3):439–441, May 1983. 34, 35

[58] P. VISWANATH AND D.N.C. TSE. **Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality**. *Information Theory, IEEE Transactions on*, **49**(8):1912–1921, Aug 2003. 36, 40

[59] W. YU. **Sum-capacity computation for the Gaussian vector broadcast channel via dual decomposition**. *Information Theory, IEEE Transactions on*, **52**(2):754–759, Feb. 2006. 36, 40, 41

[60] WEI YU AND J.M. CIOFFI. **Sum capacity of Gaussian vector broadcast channels**. *Information Theory, IEEE Transactions on*, **50**(9):1875–1892, Sep. 2004. 36, 40, 41

[61] H. WEINGARTEN, Y. STEINBERG, AND S. SHAMAI. **The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel**. *Information Theory, IEEE Transactions on*, **52**(9):3936–3964, Sep. 2006. 36

[62] F. RASHID-FARROKHI, K.J.R. LIU, AND L. TASSIULAS. **Transmit beamforming and power control for cellular wireless systems**. *Selected Areas in Communications, IEEE Journal on*, **16**(8):1437–1450, Oct 1998. 40

[63] D.N.C. TSE AND P. VISWANATH. **Downlink-uplink duality and effective bandwidths**. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, Jul. 2002. 40

[64] W. YU AND T. LAN. **Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints**. *Signal Processing, IEEE Transactions on*, **55**(6):2646–2660, Jun. 2007. 40, 41, 63, 67, 84

[65] D. TSE AND P. VISWANATH. *Fundamentals of Wireless COmmunications*. Cambridge, U.K.: Cambridge University Press, 2005. 41

[66] RUI ZHANG, YING-CHANG LIANG, AND SHUGUANG CUI. **Dynamic Resource Allocation in Cognitive Radio Networks**. *Signal Processing Magazine, IEEE*, **27**(3):102–114, May 2010. 43

[67] M. TOMLINSON. **New automatic equaliser employing modulo arithmetic**. *Electronics Letters*, **7**(5):138–139, Mar. 1971. 44

[68] H. HARASHIMA AND H. MIYAKAWA. **Matched-Transmission Technique for Channels With Intersymbol Interference**. *Communications, IEEE Transactions on*, **20**(4):774–780, Aug. 1972. 44

[69] C. WINDPASSINGER, R. F. H. FISCHER, T. VENCEL, AND J. B. HUBER. **Precoding in multiantenna and multiuser communications**. *Wireless Communications, IEEE Transactions on*, **3**(4):1305–1316, Jul. 2004. 45, 47

[70] M. SCHUBERT AND S. SHI. **MMSE transmit optimization with interference pre-compensation**. In *proc. IEEE 61st Vehicular Technology Conference*, **2**, pages 845–849, May 2005. 45

[71] M.B. SHENOUDA AND T.N. DAVIDSON. **Tomlinson-Harashima Precoding for Broadcast Channels with Uncertainty**. *Selected Areas in Communications, IEEE Journal on*, **25**(7):1380–1389, Sep. 2007. 45

[72] L.-N. TRAN AND E.-K. HONG. **Multiuser Diversity for Successive Zero-Forcing Dirty Paper Coding: Greedy Scheduling Algorithms and Asymptotic Performance Analysis**. *Signal Processing, IEEE Transactions on*, **58**(6):3411–3416, Jun. 2010. 50

[73] C. SWANNACK, E. UYSAL-BIYIKOGLU, AND G. WORNELL. **Low complexity multiuser scheduling for maximizing throughput in the MIMO broadcast channel**. In *Proc. 42nd Ann. Allerton Conf. Commun., Control, Comput.*, pages 440–449, 2004. 63

[74] JIANQI WANG, D.J. LOVE, AND M.D. ZOLTOWSKI. **User Selection With Zero-Forcing Beamforming Achieves the Asymptotically Optimal Sum Rate**. *Signal Processing, IEEE Transactions on*, **56**(8):3713–3726, Aug. 2008. 64

[75] Z.-Q. LUO AND W. YU. **An introduction to convex optimization for communications and signal processing**. *Selected Areas in Communications, IEEE Journal on*, **24**(8):1426–1438, Aug. 2006. 69

[76] R. G. BLAND, D. GOLDFARB, AND M. J. TODD. **The ellipsoid method: a survey**. *Operatioins Research*, **29**(6):1039–1091, 1981. 78, 115

[77] Z. Shen, R. Chen, J.G. Andrews, R.W. Heath, and B.L. Evans. **Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization**. *Signal Processing, IEEE Transactions on*, **54**(9):3658–3663, Sep. 2006. 86

[78] L.-N. Tran, M. Bengtsson, and B. Ottersten. **Iterative Precoder Design and User Scheduling for Block-Diagonalized Systems**. *Signal Processing, IEEE Transactions on*, **60**(7):3726–3739, Jul. 2012. 86

[79] G.H. Golub and C.F. Van Loan. *Matrix computations*. Baltimore, MD: The John Hopkins Univ. Press, 3rd edition, 1996. 86

[80] M. Grant and S. Boyd. **CVX: Matlab software for disciplined convex programming, version 1.21.**, Apr. 2011. available: http://cvxr.com/cvx. 89

[81] J. Löfberg. **YALMIP : A Toolbox for Modeling and Optimization in MATLAB**. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. 89

[82] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. 89, 115

[83] P.C. Weeraddana, M. Codreanu, M. Latva-aho, and A. Ephremides. **Weighted Sum-Rate Maximization for a Set of Interfering Links via Branch and Bound**. *Signal Processing, IEEE Transactions on*, **59**(8):3977–3996, Aug. 2011. 103

[84] L. Qian, Y.J.A. Zhang, and J Huang. **MAPEL: Achieving global optimality for a non-convex wireless power control problem**. *Wireless Communications, IEEE Transactions on*, **8**(3):1553–1563, Mar. 2009. 103

[85] D. Julian, M. Chiang, D.O. Neill, and S. Boyd. **QoS and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks**. In *Proc, IEEEINFOCOM, 2002.*, **2**, pages 477–486, Ju. 2002. 103

[86] D.P. Palomar and J.R. Fonollosa. **Practical algorithms for a family of waterfilling solutions**. *Signal Processing, IEEE Transactions on*, Feb. 2005. 103, 105

Thèse de Doctorat

Lei ZHAO

TECHNIQUES DE FORMATION DE VOIES ET ALLOCATION DE PUISSANCE POUR CANAUX MULTI-UTILISATEURS MIMO

BEAMFORMING AND POWER ALLOCATION TECHNIQUES FOR MULTIUSER MIMO BROADCAST CHANNELS

Résumé

La maximisation du débit des transmissions en canaux de diffusion multi-utilisateurs MIMO est étudiée dans cette thèse. Tout d'abord, nous donnons un aperçu de l'état de l'art des techniques de la littérature, et proposons une nouvelle méthode de formation de voies sous contrainte de puissance totale. Ensuite, considérant que la puissance de chaque antenne à l'émission est limitée par la linéarité de son amplificateur de puissance, nous imposons une contrainte plus réaliste et formulons une optimisation du débit avec une contrainte de puissance par antenne. Dans cette thèse, ce problème initial difficile à résoudre est alors divisé en deux sous-problèmes classiques, dont l'un est un problème d'optimisation SDP (semidefinite programming), et l'autre se résout par la technique MRC (maximal-ratio-combining). Par comparaison avec les méthodes de la littérature, la technique proposée permet une augmentation du débit global. De plus, la méthode proposée s'applique aussi dans le cas où le nombre total d'antennes de réception est plus grand que le nombre d'antennes à l'émission. Enfin, si l'interférence n'est pas complètement supprimée entre les canaux de diffusion multi-utilisateurs MIMO, l'allocation de la puissance parmi les canaux devient un problème NP difficile. Dans ce contexte, une nouvelle technique d'allocation de puissance efficace en temps de calcul pour l'optimisation du débit global est proposée. En comparaison avec les méthodes de la littérature, la complexité de calcul est significativement réduite, et le débit obtenu reste proche de la valeur optimale.

Mots clés

MU-MIMO, Optimisation de débit, Allocation de puissance, Formation de voies.

Abstract

The sum rate maximization in multiuser MIMO broadcast channels is investigated in this thesis. Firstly, we give an overview of the state of the art of the different beamforming techniques in the literature, and a new beamforming method under total power constraint is proposed. Secondly, considering that the power of each antenna is limited individually by the linearity of its power amplifier in practice, we impose a more realistic per-antenna power constraint to optimize the sum rate. It can be seen that the original problem is difficult to solve. Then in this thesis, this problem is divided into two classical subproblems, one of which becomes a semidefinite programming (SDP) problem, and the other one can be solved by the maximal-ratio combining (MRC) technique. Compared with the methods in the literature, a better sum rate performance is achieved. Moreover, the proposed method works even if the number of total receive antennas is larger than that of the transmit antennas. Thirdly, if the interference is not removed completely in multiuser MIMO broadcast channels, power allocation becomes a NP hard problem. In this thesis, an efficient suboptimal power allocation method is proposed to maximize the sum rate. Compared with the methods in the literature, the computational complexity is substantially reduced, and the performance is close to the optimal value.

Key Words

MU-MIMO, Sum rate optimization, Power allocation, Beamforming.