1 Introduction to Interference Management

In this chapter, we provide a high-level introduction to interference management in wireless networks, including a historical perspective on wireless cellular networks, and an overview of the remaining chapters in the book. We also summarize the notation used in the book.

1.1 Interference Management in Cellular Networks: A Historical Perspective

Managing interference from other users sharing the same frequency bands has been the key driver for mobile wireless communications. The first wireless phone systems served as extensions to the wired public switched telephone network [1]. These systems were "single cell" systems in the sense that mobile terminals could be connected to only one basestation during a call, with the call being lost when out of range of the basestation, akin to losing an FM radio signal while driving out of range of the station. Interference in these networks could be managed by simply orthogonalizing the users in the time–frequency plane, i.e., through the use of time-division multiple-access (TDMA) or frequency-division multiple-access (FDMA), or some combination of the two. Interference between basestations operating in the same frequency band was managed by ensuring that they are geographically far apart, again akin to the way in which radio stations operating in the same frequency band are placed.

1.1.1 Cellular Concept

A major breakthrough toward improving both the capacity and the mobility in wireless phone systems came with the introduction of the cellular concept [2]. In the cellular system design, a given geographical region is split into contiguous regions called "cells," without any gaps in coverage. The system is designed so that cells that use the same frequency band are far enough from each other to cause little interference to each other. The number of different frequency bands is called the *reuse factor* of the system. The reuse factor is a measure of spectral efficiency in the system, with a larger reuse factor corresponding to a smaller efficiency. A key innovation in the cellular concept is the introduction of *handoff* between neighboring cells operating in different frequency bands, which allows a mobile user to maintain a continuous connection while moving through the geographical region. Interference management within each of the cells is achieved by orthogonalizing the users in the time–frequency plane.

Early cellular communication systems, both analog (e.g., AMPS [2]) and digital (e.g., GSM [3]), adopted narrowband communication links within each cell. For example, in GSM the available bandwidth is divided into 200 kHz channels, with each channel serving eight users through TDMA. Since the users within a cell occupy orthogonal time–frequency resources, there is no in-cell interference. However, users in neighboring cells communicating in the same time/frequency slot cause *co-channel* interference, which is controlled through the reuse factor.

1.1.2 Code-Division Multiple-Access

An alternative way to separate the users within a cell is through code-division multiple-access (CDMA) [4], where each user's signal occupies the entire time–frequency plane, and the users are separated through the use of different code sequences. Interference from in-cell users can be controlled by using orthogonal codes, which can be implemented in the downlink since the downlink transmissions from a basestation can easily be synchronized. On the uplink, the tight time synchrony required for orthogonalization is difficult to implement, and user separation can be accomplished through simple matched filtering or through the use of more sophisticated successive interference cancellation [5]. A major advantage of CDMA cellular systems over TDMA/FDMA systems comes from the fact that it is possible through the use of pseudorandom overlay codes to randomize the interference across cells in the network so that this interference simply adds to the noise floor for a given user's communication channel. This allows for universal reuse of spectrum (i.e., a reuse factor of one), although there is a loss in spectral efficiency due to the out-of-cell interference effectively raising the noise floor within each cell. This loss in spectral efficiency is generally limited to a factor of two due to the power-law decay of transmitted power with distance [4]. Some other advantages of CDMA systems from an interference management viewpoint include: (i) no frequency planning is needed since the reuse factor is one; (ii) there is a graceful degradation of performance with the number of users in a cell; and (iii) any technique that reduces the power of interferers (e.g., soft handoff, voice activity detection, power control, etc.) increases the capacity. There are some disadvantages that offset these advantages to some extent, including the fact that in-cell interference cannot be eliminated completely and hence reduces capacity, and that tight power control is needed to manage both in-cell and out-of-cell interference, and may be difficult to implement, especially for data applications, which have a low duty cycle.

1.1.3 Orthogonal Frequency-Division Multiplexing

A question that cellular communication system designers asked in the late 1990's when wireless data applications started to grow rapidly was: Can we simultaneously have universal reuse and keep the in-cell users orthogonal? The answer came in the form of multiple access based on orthogonal frequency-division multiplexing (OFDM) [6].

The idea is to split up the bandwidth into narrowband *subchannels*, with every user having access to all the subchannels. The basic unit of resource is a *virtual channel* (i.e., a hopping sequence in time across the subchannels). A given user may be assigned one or more virtual channels for communication. The virtual channels for all users within a given cell are designed to be orthogonal in the time–frequency plane, akin to orthogonal CDMA. Due to the narrowband nature of the subchannels, the orthogonality across users can be maintained almost as easily in the uplink as it is in the downlink. Furthermore, the hopping patterns in adjacent cells are chosen so that there is minimal overlap between any pair of virtual channels across cells, thus averaging the out-of-cell interference to appear as white noise, as in CDMA, as opposed to being localized, as in FDMA/TDMA.

1.2 Additional Resources for Interference Management

In addition to time–frequency separation and geographical separation, there are a number of other resources that can be exploited to manage interference in wireless networks.

1.2.1 Multiple Antennas

The use of multiple antennas at either end of a wireless link provides resilience to fading due to the diversity in the fading seen by the different antennas. Having multiple antennas at both ends of the wireless link forms a multiple-input multiple-output (MIMO) channel, which can be exploited to create multiple parallel streams for communication [7]. This leads to multiplexing or degrees of freedom (DoF) gains. From the viewpoint of interference management, multiple antennas can be used for beamforming toward desired receivers, while minimizing the interference to other receivers. This allows for a more flexible design of interference management schemes, as we will see in Chapters 2, 4, 5, and 6.

1.2.2 Cooperation and Relaying

Cooperation among basestations equipped with multiple antennas can be used for coordinated beamforming across cells so as to maximize the signal-to-interference-plus-noise ratio (SINR) at the receivers [8]. Moreover, such cooperation can also be used for coordinated multi-point (CoMP) transmission and reception by the basestations, which can greatly enhance the DoF achievable in cellular wireless networks, a topic of particular emphasis in this book from Chapter 5 onwards.

Cooperation among mobiles in the networks can also be exploited for interference management, with one mobile relaying the information to or from another mobile [9]. This way potential interferers can become helpers in terms of relaying information to the receiver. Such relaying is particularly useful in distributed interference management in *ad hoc* and *mesh* networks [10].

1.2.3 Cognitive Radio and Dynamic Spectrum Access

Another way to manage interference in wireless networks is through active interference avoidance, as in *cognitive radio* [11]. The key idea is to use sensing to determine the times when a specific licensed band is not used at a particular place and use this band for unlicensed transmissions, without causing interference to the licensed user (referred to as the "primary user"). An important part of designing such systems is the development of a dynamic spectrum access scheme for channel selection. The cognitive radio (also called the "secondary user") needs to adopt the best strategy for selecting channels for sensing and access. The sensing and access policies should jointly ensure that the probability of interfering with the primary user's transmission meets a pre-specified constraint.

1.3 Motivation for This Book

The last few years have seen an exponential growth in data traffic over wireless networks. Wireless service providers are having to accommodate this exponential growth without any significant new useful spectrum. Spectral efficiency gains from improvements in the physical layer are quite limited, with error control coding and decoding being performed near Shannon limits. One way to accommodate the increasing demand for wireless data services is through the addition of basestations in the networks in a hierarchical manner, going from macro to micro to pico to femto basestations, but this comes with a significant cost and makes the interference management problem more difficult through the traditional means described above. Infrastructure enhancements such as cooperative transmission and reception have the potential for increasing the spectral efficiency at low cost, through efficient interference management, but new techniques for interference management are needed. The main motivation for this book is to develop a deep understanding of the fundamental limits of interference management in wireless networks with cooperative transmission and reception, and to use this understanding to develop practical schemes for interference management that approach these limits.

1.4 Overview of This Book

In Chapter 2, we introduce a mathematical model for a K-user interference channel, and use this model to develop some basic information-theoretic bounds on the rates for communication on the channel; particular cases where the sum capacity of the channel can be analyzed exactly are also discussed in this chapter. In Chapter 3, we take an alternative approach to characterizing the rate of communication on an interference channel, based on a *degrees of freedom* analysis, which will be followed in the remainder of the book. In particular, we describe the important technique of interference alignment (IA) in Chapter 3, which is justified through a DoF analysis.

In Chapter 4, we study iterative algorithms for approaching the interference alignment solutions for interference management.

In Chapter 5, we start discussing the value of cooperative communication in large interference networks by studying the DoF of fully connected interference networks when each message can be available at more than one transmitter. In Chapter 6, we extend this setting by studying locally connected networks where each of the transmitters is only connected to a set of neighboring receivers. In Chapter 7, we consider an average backhaul load constraint, where the average number of transmitters per message cannot exceed a set value. We then study cooperative reception schemes for cellular uplink in Chapter 8. In Chapter 9, we study dynamic interference networks, where we alter our interference network model to take into account the deep fading conditions that can result in random link erasures. In Chapter 10, we discuss some recent advances and open problems.

1.5 Notation

We use lower-case and upper-case letters for scalars, lower-case letters in bold font for vectors, and upper-case letters in bold font for matrices. For example, we use h, x, and K to denote scalars, $\boldsymbol{h}$ and $\boldsymbol{x}$ to denote vectors, and $\boldsymbol{H}$, $\boldsymbol{A}$ to denote matrices. Superscripts denote sequences of variables in time. For example, we use x^n and $\boldsymbol{x}^n$ to denote sequences of length n of scalars and vectors, respectively.

We use the notation $\boldsymbol{A}(d)$ for the dth column of the matrix $\boldsymbol{A}$. When we use this notation in general we will refer to a collection of matrices $\boldsymbol{A}_i$, and therefore in our notation $\boldsymbol{A}_i(d)$ is the dth column of the ith matrix $\boldsymbol{A}_i$. Also, $\boldsymbol{x}(\ell)$ is sometimes used to denote the ℓth element of the vector $\boldsymbol{x}$. The matrix $\boldsymbol{I}$ denotes the identity matrix, $\boldsymbol{A}^\dagger$ is the conjugate transpose of $\boldsymbol{A}$, and $\mathrm{diag}(x_1,\ldots,x_N)$ is an $N \times N$ diagonal matrix with $x_1,\ldots,x_N$ on the diagonal.

We use $\boldsymbol{\Sigma}_{\boldsymbol{x}}$ and $\mathrm{Cov}\,(\boldsymbol{x})$ to denote the covariance matrix of a random vector $\boldsymbol{x}$. We use $\boldsymbol{\Sigma}_{\boldsymbol{y}|\boldsymbol{x}}$ and $\mathrm{Cov}\,(\boldsymbol{y}|\boldsymbol{x})$ to denote the covariance matrix of the minimum mean square estimation error in estimating the random vector $\boldsymbol{y}$ from the random vector $\boldsymbol{x}$, with similar notation for random scalars. We use $\mathcal{CN}(0,\boldsymbol{\Sigma})$ to denote the circularly symmetric complex Gaussian vector distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}$, with similar notation for random scalars. We use $\mathsf{H}\,(.)$ to denote the entropy of a discrete random variable, $\mathsf{h}\,(.)$ to denote the differential entropy of a continuous random variable or vector, and $\mathsf{I}\,(.;.)$ to denote the mutual information. Finally, we use $[K]$ to denote the set $\{1,2,\ldots,K\}$, where the number K will be obvious from the context.

Cambridge University Press
978-1-107-16500-7 — Interference Management in Wireless Networks
Venugopal V. Veeravalli , Aly El Gamal
Excerpt
More Information

2 System Model and Sum Capacity Characterization

In this chapter, we introduce a mathematical model for the K-user interference channel, and use this model to develop some basic information-theoretic bounds on the rates for communication on the channel. The focus will be on characterizing bounds on the sum-rate (throughput) of the channel. Particular cases where the sum capacity of the channel can be analyzed exactly will be discussed.

2.1 System and Channel Model

The K-user (fully connected) Gaussian interference channel, illustrated in Figure 2.1, consists of K transmitter–receiver pairs, where every transmitter is heard by every receiver. The signal $\boldsymbol{y}_k \in \mathbb{C}^{N_\mathrm{r}}$ received by receiver k is given by

$$\boldsymbol{y}_k = \sum_{j=1}^{K} \boldsymbol{H}_{kj}\boldsymbol{x}_j + \boldsymbol{z}_k, \ \forall k \in [K], \tag{2.1}$$

where $\boldsymbol{x}_j \in \mathbb{C}^{N_\mathrm{t}\times 1}$ denotes the signal of transmitter j, $\boldsymbol{z}_k \in \mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{I}_{N_\mathrm{r}}\right)$ denotes the additive white Gaussian noise at receiver k, and $\boldsymbol{H}_{kj} \in \mathbb{C}^{N_\mathrm{r}\times N_\mathrm{t}}$ denotes the channel transfer matrix from transmitter j to receiver k. Each transmitter is assumed to have N_t transmit antennas, and each receiver is assumed to have N_r receive antennas.[1] An interference channel with N_t and N_r taking arbitrary values is referred to as a multiple-input multiple-output (MIMO) interference channel. The special cases with $N_\mathrm{t} = 1$ or $N_\mathrm{r} = 1$ or $N_\mathrm{t} = N_\mathrm{r} = 1$ are referred to as the single-input multiple-output (SIMO), multiple-input single-output (MISO), and single-input single-output (SISO) interference channels, respectively. The transmitters are assumed to operate under average power constraints $\{P_k\}$; i.e., for each $k \in [K]$, the power consumed by transmitter k is not allowed to exceed P_k on average.

2.1.1 Achievable Schemes

Consider the problem of communicating K messages over the interference channel (2.1). For each $k \in [K]$, the message W_k is available at transmitter k, and is desired

[1] More generally, the number of antennas could be different at each transmitter and each receiver, as we consider in Section 2.2.

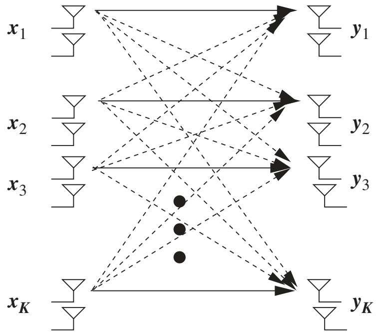

Fig. 2.1 The K-user MIMO Gaussian interference channel.

by receiver k. A communication scheme consists of an encoder–decoder pair for each message. The encoder at transmitter k maps the message W_k onto the physical signal $\boldsymbol{x}_k$ that is transmitted on the channel. The decoder at receiver k reconstructs the message W_k from the received signal $\boldsymbol{y}_k$. The communication scheme is said to be reliable if all the messages can be reconstructed at their respective receivers with high probability. For the single-user channel, Shannon [12] established that the key to reliable communication over noisy channels is coding over multiple symbols. We consider the same *block coding* framework, where the communication scheme operates over n symbols at a time. For a fixed rate tuple $(R_1, R_2, \ldots, R_K) \in \mathbb{R}_+^K$ and a block length $n \geq 1$, the message W_k takes values from the set $\mathcal{W}_k = \left\{1, 2, \ldots, \lceil 2^{nR_k} \rceil\right\}$. The block code consists of the encoders

$$\boldsymbol{x}_k^n : \mathcal{W}_k \to \mathbb{C}^{N_t \times n}, \ \forall k \in [K],$$

and the decoders

$$\hat{W}_k : \mathbb{C}^{N_r \times n} \to \mathcal{W}_k, \ \forall k \in [K].$$

Assuming that the message W_k is a uniform random variable taking values in the set $\mathcal{W}_k$, the probability of a decoding error is defined as

$$e_n = \max_{k \in [K]} \mathbb{P}\left(\hat{W}_k\left(\boldsymbol{y}_k^n\right) \neq W_k\right).$$

We say that the rate tuple $(R_1, R_2, \ldots, R_K)$ is achievable if and only if there exists a sequence of block codes satisfying the average power constraints

$$\mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n} ||\boldsymbol{x}_k(t)||^2\right] \leq P_k, \ \forall k \in [K]$$

such that the probability of error $e_n \to 0$ as $n \to \infty$. The *capacity region* $\mathcal{C}$ is defined as the closure of the set of achievable rate tuples. Except in some special cases, determining the exact capacity region of the Gaussian interference channel remains an open problem.

The *sum capacity* is defined as:

$$\mathcal{C}_{\text{sum}} = \max_{(R_1, R_2, \ldots, R_k) \in \mathcal{C}} R_1 + R_2 + \cdots + R_K. \tag{2.2}$$

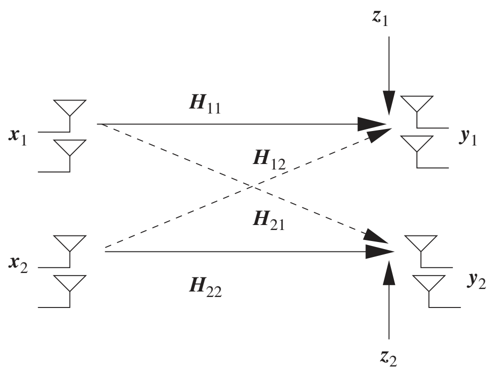

Fig. 2.2 Two-user interference channel. ©[2017] IEEE. Reprinted, with permission, from [13].

2.1.2 Channel Knowledge

We assume that the various channel coefficients are known at all the transmitters and at all the receivers at which they are required for a given achievable scheme. In practice, the channel knowledge is obtained by transmitting known signals, called pilots, at regular intervals and estimating the channel coefficients at the receivers. The estimated (local) channel coefficients are then distributed to other transmitters and receivers. Although the processes of channel estimation and distribution can incur significant overhead, it is difficult to accommodate this overhead in information-theoretic capacity analyses. The common practice, which is also followed in this book, is to perform the capacity analysis assuming channel knowledge where needed, and account for the overhead when designing practical achievable schemes.

2.2 Two-User Interference Channel

In this section, we consider the Gaussian interference channel (2.1) in the two-user case, assuming that multiple antennas are available at the transmitter and receiver:

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{H}_{11}\boldsymbol{x}_1 + \boldsymbol{H}_{12}\boldsymbol{x}_2 + \boldsymbol{z}_1, \\ \boldsymbol{y}_2 &= \boldsymbol{H}_{21}\boldsymbol{x}_1 + \boldsymbol{H}_{22}\boldsymbol{x}_2 + \boldsymbol{z}_2, \end{aligned} \tag{2.3}$$

where $z_i \in \mathcal{CN}(0, \boldsymbol{I})$, and the average power constraints at transmitters 1 and 2 are denoted by P_1 and P_2, respectively. This is depicted in Figure 2.2. Let $N_{1\mathrm{t}}, N_{2\mathrm{t}}$ denote the number of transmit antennas at transmitters 1 and 2, respectively, and $N_{1\mathrm{r}}, N_{2\mathrm{r}}$ denote the number of receive antennas at receivers 1 and 2, respectively. The dimensions of the channel matrices, the signal vectors, and the noise vectors are defined appropriately. We are interested in determining the best sum-rate achievable by using Gaussian inputs and treating interference as noise, and also the sum capacity (maximum throughput) of the two-user MIMO Gaussian interference channel.

We start by studying the problem of determining the best achievable sum-rate. The use of multiple antennas at the transmitters and the receivers provides spatial dimensions to suppress the interference and improve the achievable sum-rate. While it is easy to express the achievable sum-rate as a function of the spatial beams at the transmitter and receiver, the design of beams that maximize the achievable sum-rate is known to be a difficult problem. The main difficulty stems from the fact that the sum-rate optimization problem cannot be posed as a convex (concave) optimization problem, which makes the optimization problem difficult to solve analytically or even numerically. We introduce a technique based on convex approximation and optimization to solve this nonconvex optimization problem. Specifically, we upper-bound the achievable sum-rate with a concave function, and use this to obtain an upper bound to the original sum-rate optimization problem. We show that if the channel parameters satisfy certain conditions, then the bounds coincide, leading to an exact characterization of the best achievable sum-rate by using Gaussian inputs and treating interference as noise.

The problem of determining the best achievable sum-rate by treating interference as noise is important from a practical perspective, because coding schemes that approach the rates promised by the information-theoretic analysis can be designed in the same way as schemes for point-to-point Gaussian channels, a topic that is well understood [14]. Therefore, it is also important to understand the gap between the sum capacity and the sum-rate achievable by treating interference as noise. If the lower and upper bounds on the achievable sum-rate coincide, then the best achievable sum-rate is indeed equal to the sum capacity. Using the Karush–Kuhn–Tucker (KKT) conditions [15], we obtain necessary and sufficient conditions for the bounds to coincide, leading to an exact characterization of the sum capacity. We observe that the conditions are satisfied in a *low interference regime* where the interfering signal levels are small compared to the desired signal levels. We end the section by providing some nontrivial examples of the two-user Gaussian interference channel in the low interference regime. In particular, we consider the special cases of symmetric MISO and SIMO interference channels, and derive a simple closed-form condition on the channel parameters for the channels to be in the low interference regime. We also specialize the results to SISO interference channels.

2.2.1 Standard Form

The following assumptions can be made about the two-user MIMO Gaussian interference channel (2.3) without any loss of generality, as we establish below:

- The direct channel matrices $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{22}$ have unit (Frobenius) norm.
- The cross channel matrices $\boldsymbol{H}_{12}$ and $\boldsymbol{H}_{21}$ are diagonal with real and nonnegative entries.
- The numbers of transmit and receive antennas $(N_{1t}, N_{2t}, N_{1r}, N_{2r})$ satisfy

$$N_{1t} \leq \text{rank}\left\{\begin{bmatrix} \boldsymbol{H}_{11} \\ \boldsymbol{H}_{21} \end{bmatrix}\right\}, N_{2t} \leq \text{rank}\left\{\begin{bmatrix} \boldsymbol{H}_{12} \\ \boldsymbol{H}_{22} \end{bmatrix}\right\},$$

and

$$N_{1r} \leq \text{rank}\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\},\ N_{2r} \leq \text{rank}\{[\boldsymbol{H}_{21}\ \boldsymbol{H}_{22}]\}.$$

The second assumption implies that the cross channel matrices can be expressed as

$$\boldsymbol{H}_{12} = \begin{bmatrix} \tilde{\boldsymbol{H}}_{12} \\ \boldsymbol{0} \end{bmatrix},\ \boldsymbol{H}_{21} = \begin{bmatrix} \tilde{\boldsymbol{H}}_{21} \\ \boldsymbol{0} \end{bmatrix},$$

where $\tilde{\boldsymbol{H}}_{12}$ and $\tilde{\boldsymbol{H}}_{21}$ are diagonal matrices with full row rank. This is the only assumption we use in the development of the outer bound techniques presented in this chapter. The other two assumptions are used in Section 2.2.13 to simplify the presentation.

The first assumption can easily be justified by scaling the transmit power constraints P_1 and P_2 appropriately. We now justify the other two assumptions. First, consider the singular value decomposition (SVD) of $\boldsymbol{H}_{12}$ and $\boldsymbol{H}_{21}$:

$$\boldsymbol{H}_{12} = \boldsymbol{U}_1 \boldsymbol{\Lambda}_{12} \boldsymbol{V}_2^{\dagger},$$
$$\boldsymbol{H}_{21} = \boldsymbol{U}_2 \boldsymbol{\Lambda}_{21} \boldsymbol{V}_1^{\dagger},$$

where $\boldsymbol{\Lambda}_{12}, \boldsymbol{\Lambda}_{21}$ are diagonal matrices with real and nonnegative entries, and $\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{U}_1, \boldsymbol{U}_2$ are unitary matrices. We obtain an equivalent Gaussian interference channel, satisfying the second assumption, by projecting the received signals along $\boldsymbol{U}_1, \boldsymbol{U}_2$, and the transmitted signals along $\boldsymbol{V}_1, \boldsymbol{V}_2$, i.e., by making the following substitutions:

$$\boldsymbol{x}_j \leftarrow \boldsymbol{V}_j^{\dagger} \boldsymbol{x}_j,$$
$$\boldsymbol{y}_i \leftarrow \boldsymbol{U}_i^{\dagger} \boldsymbol{y}_i,$$
$$\boldsymbol{z}_i \leftarrow \boldsymbol{U}_i^{\dagger} \boldsymbol{z}_i,$$
$$\boldsymbol{H}_{ij} \leftarrow \boldsymbol{U}_i^{\dagger} \boldsymbol{H}_{ij} \boldsymbol{V}_j.$$

Observe that the average transmit power constraint and the distribution of the receive noise terms remain unchanged because $\boldsymbol{U}_1, \boldsymbol{U}_2, \boldsymbol{V}_1, \boldsymbol{V}_2$ are unitary matrices.

The third assumption can be justified by appropriately choosing the unitary matrices. For example, suppose $N_{1r} > \text{rank}\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\}$. Consider the SVD of $\boldsymbol{H}_{12} = \boldsymbol{U}_1 \Lambda_{12} \boldsymbol{V}_2^{\dagger}$. Observe that the span of the first rank$\{\boldsymbol{H}_{12}\}$ columns of $\boldsymbol{U}_1$ is equal to the column space of $\boldsymbol{H}_{12}$, and we have flexibility in choosing the remaining $N_{1r} - \text{rank}\{\boldsymbol{H}_{12}\}$ columns. We may choose those columns such that the span of the first rank$\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\}$ columns of $\boldsymbol{U}_1$ is equal to the column space of $[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]$, so that the last $N_{1r} - \text{rank}\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\}$ columns of $\boldsymbol{U}_1$ are orthogonal to the columns of $[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]$. Therefore, the last $N_{1r} - \text{rank}\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\}$ rows of the channel matrices $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{12}$ in the new channel are equal to zero, i.e., receiver 1 sees only Gaussian noise from the last $N_{1r} - \text{rank}\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\}$ antennas. Hence, we can ignore these antennas and assume that $N_{1r} = \text{rank}\{[\boldsymbol{H}_{11}\ \boldsymbol{H}_{12}]\}$. We can repeat the same argument at receiver 2, and also at transmitters 1 and 2, to justify the other inequalities in the third assumption.