Chapter 2

Preliminaries and System Overview

2.1 Overview of MIMO

MIMO system is a key technology which is responsible for enhancing data rate in wireless communications [57]. In the MIMO system, transmit and/or receive antenna elements are kept together while keeping some distance from each other, which is an order of carrier wavelength, in a form to implement antenna diversity as shown in Fig. 2.1, which is same as time diversity, or frequency diversity [57]. Antenna diversity is also known as space diversity or spatial diversity [57]. Further, these antenna elements are separated by a particular distance and can be deployed in linear or planner arrays. It exploits more degrees of freedom in the spatial dimension [57] which can not be achieved in a single antenna communication system. In fact, spatial degree of freedom can be used to significantly enhance the spectral efficiency. On the other hand, MIMO systems also encounters fading in wireless communication channel, and suppress interference [57]. It has been seen that MIMO system improve the system performance without additional power and bandwidth requirements. Due to this characteristic, MIMO has played vital role in many standards of wireless communication such as 3GPP (3rd Generation Partnership Project) [58], WiMAX [59], WLAN IEEE 802.11 [60], LTE [59], etc.

2.1.1 MIMO System

Consider a downlink MIMO system having a transmitter and a receiver. Transmitter is equipped with N antenna elements, and receiver is equipped with K antenna elements. Further, antenna elements in linear array are kept $\frac{\lambda}{2}$ apart from each other at both ends to avoid antenna correlation. Thus, flat fading MIMO system model

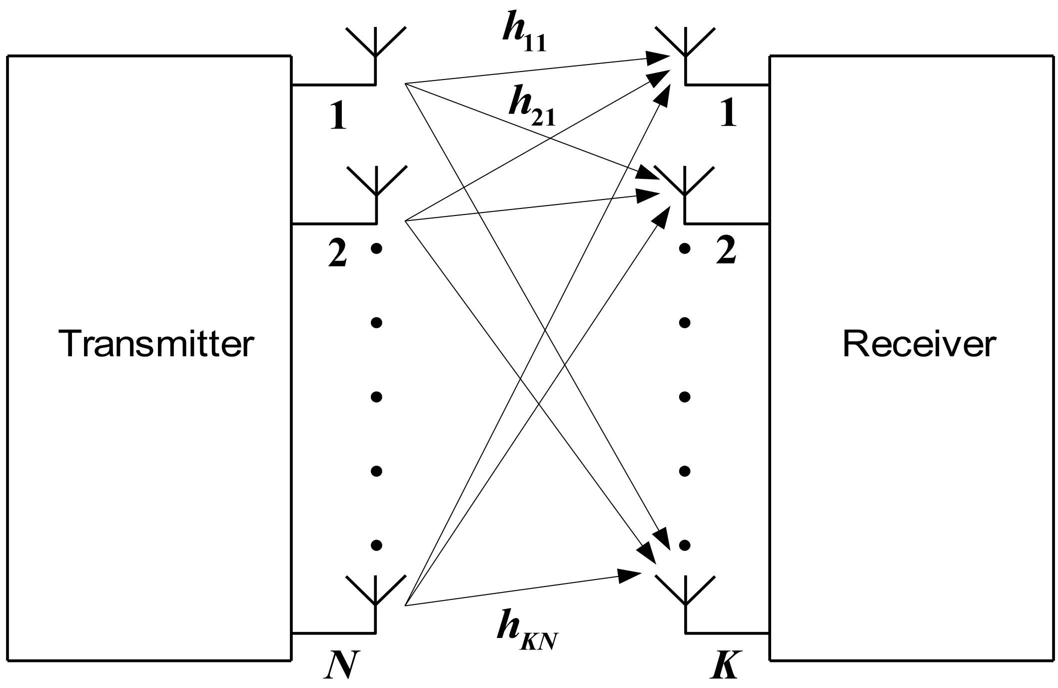


Figure 2.1 MIMO System

for downlink is described by the input-output relationship as,

$$\mathbf{y} = \mathbf{H}^H \mathbf{x} + \mathbf{w}, \tag{2.1}$$

where $\mathbf{H}^H \in \mathbb{C}^{K \times N}$ is the MIMO channel matrix whose matrix elements are h_{kn}, $k = \{1, \ldots, K\}, n = \{1, \ldots, N\}$ and H is a conjugate transpose or Hermitian operator. $\mathbf{x} \in \mathbb{C}^{N \times 1}$ is the transmitted symbol vector, and $\mathbf{y} \in \mathbb{C}^{K \times 1}$ is the received information vector. Further, the total average power constraint is satisfied as $\mathbb{E}[\| \mathbf{x} \|^2] \leq \rho$, ρ is the total transmitted power. $\mathbf{w}$ is an additive white Gaussian noise (AWGN) vector with $\mathbf{w} \sim \mathbb{C}\mathcal{N}(0, N_0 \mathbf{I}_k)$.

2.1.2 MIMO System Capacity

Claude Shannon had discussed the capacity of a channel as the mutual information between the channel input and output or maximum rate of reliable communication which is denoted as C. According to Shannon's definition, the channel capacity can be written as [40],

$$C = \mathrm{B} \log_2(1 + \mathrm{SNR}) \text{ b/s} \tag{2.2}$$

where B and SNR denote the channel bandwidth and received signal-to-noise-ratio, respectively. When there is a perfect channel state information at the receiver and at the transmitter, the capacity of MIMO channel is defined as [57],

$$C = \max_{\mathrm{Tr}(\Lambda)\leq\rho} \log_2 \det\left(\mathbf{I} + \frac{\mathbf{H}^H\Lambda\mathbf{H}}{N_0}\right) \text{ b/s/Hz} \tag{2.3}$$

where N_0 is the noise power, $\mathbf{I}$ is the identity matrix, and Λ denotes the covariance matrix of the transmitted symbol vector such that

$$\Lambda = \mathbb{E}[\mathbf{x}\mathbf{x}^H]. \tag{2.4}$$

2.1.3 **MIMO Water-Filling Power Allocation Algorithm**

The capacity of a MIMO system can be further increased, if we allocate power according to the water-filling algorithm to the parallel channels [57]. But, one must know the CSI both at the transmitter and receiver. Further, MIMO channel $\mathbf{H}^H$ can be converted into parallel channels, and these parallel channels are free of interference and each channel is considered as single-input-single-output (SISO). The process of singular value decomposition (SVD) of matrix $\mathbf{H}^H$ converts these into interference-free parallel channels. Thus, SVD is defined as [57],

$$\mathbf{H}^H = \mathbf{U}\Sigma\mathbf{V}^H, \tag{2.5}$$

where $\mathbf{U} \in \mathbb{C}^{K\times K}$ and $\mathbf{V} \in \mathbb{C}^{N\times N}$ are a unitary matrices, and $\Sigma \in \mathbb{R}^{K\times N}$ is a diagonal matrix with non-negative entries. The diagonal elements of Σ are known as the singular values of $\mathbf{H}^H$ and denoted by σ_i, which are arranged in descending order as

$$\sigma_1 \geq \sigma_2 \cdots \geq \sigma_{\min(K,N)}. \tag{2.6}$$

The precoding and postcoding process of the single user MIMO system can be simply represented as in Fig. 2.2, where the information vector is multiplied by matrix $\mathbf{V}$ before transmission, i.e.

$$\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}. \tag{2.7}$$

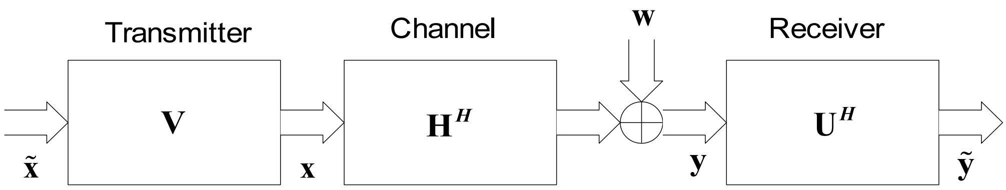


Figure 2.2 Precoding and Postcoding with SVD decomposition

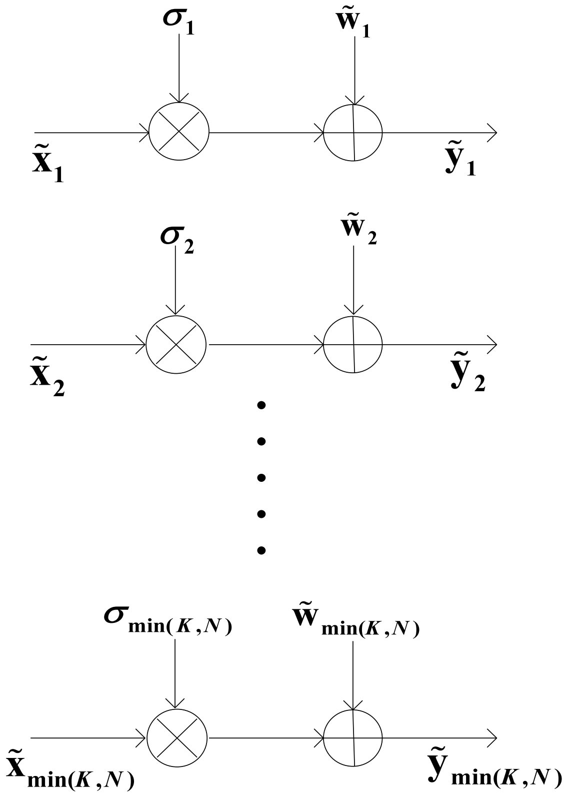


Figure 2.3 MIMO SVD parallel channels

At the receiver, received signal is multiplied by $\mathbf{U}^H$. Hence, using (2.1), the output is given as

$$\tilde{\mathbf{y}} = \mathbf{U}^H\mathbf{y}, \tag{2.8}$$

$$\tilde{\mathbf{y}} = \mathbf{U}^H(\mathbf{H}^H\mathbf{x} + \mathbf{w}). \tag{2.9}$$

Substituting (2.5) and (2.7) into (2.9), one can obtain

$$\tilde{\mathbf{y}} = \mathbf{U}^H(\mathbf{U}\Sigma\mathbf{V}^H\mathbf{V}\tilde{\mathbf{x}} + \mathbf{w}), \tag{2.10}$$

$$\tilde{\mathbf{y}} = \Sigma\tilde{\mathbf{x}} + \tilde{\mathbf{w}}, \tag{2.11}$$

where $\tilde{\mathbf{y}} = \mathbf{V}\mathbf{y}$, $\tilde{\mathbf{w}} = \mathbf{U}^H\mathbf{w}$, and Σ is a diagonal matrix. Further, one can obtain

$$\tilde{y}_i = \sigma_i \tilde{x}_i + \tilde{w}_i, \quad i = 1, 2, \ldots, \min(K, N), \tag{2.12}$$

where σ_i is the Eigen value. Thus, the transmitted signal is given by

$$s_i = \sigma_i \tilde{x}_i, \quad i = 1, 2, \ldots, \min(K, N). \tag{2.13}$$

Hence, $\mathbb{E}[\| s_i \|^2]$ is the signal power to the i^{th} channel and it is defined as

$$\mathbb{E}[\| s_i \|^2] = \sigma_i^2 p_i, \tag{2.14}$$

where p_i denotes power allocated to i^{th} channel. The received SNR for i^{th} channel can be written as

$$\text{SNR}_i = \frac{\sigma_i^2 p_i}{N_0}. \tag{2.15}$$

Thus, non-interfering parallel channels are obtained as shown in Fig. 2.3, and the MIMO capacity can be obtained as

$$C = \sum_{i=1}^{\min(K,N)} \log_2 \left(1 + \frac{\sigma_i^2 p_i}{N_0} \right). \tag{2.16}$$

The non-interfering parallel channels comprise of different Eigen values due to considerable difference in singular values. This permits the use of water-filling algorithm, hence power is optimally allocated to the non-interfering parallel channels by using the following equation [57],

$$p_i = \left(\frac{1}{\lambda} - \frac{N_0}{\sigma_i^2} \right)^+, \tag{2.17}$$

where $(x)^+ = \max\{0, x\}$, here λ is a Lagrange multiplier, and $\frac{1}{\lambda}$ is the water level chosen to satisfy the total power constraint with $\sum_{i=1}^{\min(K,N)} p_i \leq \rho$. This power allocation scheme is coined as water-filling strategy and it optimally allocates the power to the non-interfering parallel channels. Suppose there are $\min(K, N)$ vessels, as shown in Fig. 2.4, and the height of i^{th} vessel is $\frac{N_0}{\sigma_i^2}$. Further, water is poured in each vessel to the level $\frac{1}{\lambda}$, then the level of water in the i^{th} vessel is $\left(\frac{1}{\lambda} - \frac{N_0}{\sigma_i^2} \right)^+$. It is observed that the water or power allocated to the i^{th} vessel is proportional to the

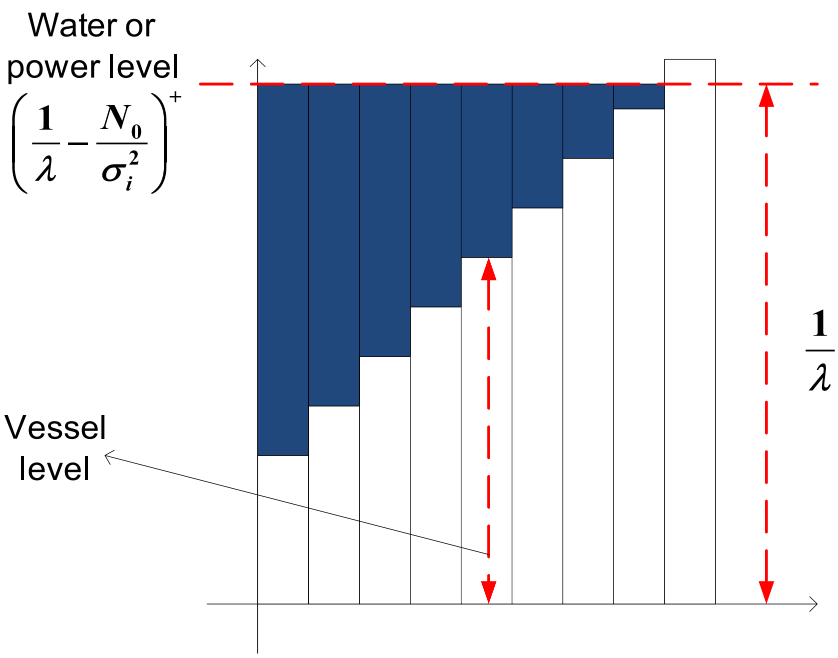


Figure 2.4 Water filling analogy with water

Eigen value. It means, if σ_i is the largest Eigen value, the poured water in this vessel is the highest. Further, it is observed that this (2.17) is a non-linear equation which can be solved iteratively as follows. Let's assume $\min(K, N) = N$, and solving below expression

$$\sum_{i=1}^{N}\left(\frac{1}{\lambda}-\frac{N_0}{\sigma_i^2}\right)^{+} = \rho, \tag{2.18}$$

one can obtain the λ. Next, check whether $\left(\frac{1}{\lambda}-\frac{N_0}{\sigma_i^2}\right) \geq 0$, $i \in \{1, \ldots, N\}$ or not. If $\left(\frac{1}{\lambda}-\frac{N_0}{\sigma_N^2}\right) < 0$, then set $N = N-1$, and again repeat the process.

It has been observed that at low SNR, the water-filling algorithm allocates less power to those channels whose Eigen values is less, and allocates high power to those channels whose Eigen value is high. More specifically, the channels with good condition are allocated more power than those the worst channel conditions. At high SNR, water-filling algorithm allocates power approximately equal across all non-interfering parallel channels.

2.2 Overview of MU-MIMO

In the MU-MIMO system, multiple users interact with a transmitter simultaneously. There are two communication links: a uplink communication systems, where multiple users transmit their information to the same transmitter, and a downlink communication systems, where the transmitter attempts to transmit signals to multiple users.

In the downlink communication systems, i.e., the downlink channel, the transmitter is simultaneously transmitting information to multiple users, as shown in the Fig. 2.5. In such situation, transmitter attempts to transmit information over different channels to multiple users, but there is some MUI [61]. However, one can mitigate MUI at transmitter by designing transmitted informations intelligently, if the CSI is available at transmitter. So, the transmitter is aware of what interference would be experienced by each user. This pre-known MUI can be mitigated by cleverly designing beamforming vector for each user [62], and it has been explained in section 2.3. Further, beamforming vector of each user is contained by a precoding matrix. Finally, the challenge in the downlink communication systems is primarily in the precoding done by the transmitter to separate the users' informations. Similarly, in the uplink communication systems, users transmit their information to the transmitter over different channels. In this context, the primary challenge for the transmitter is to separate the information transmitted by the users. However, users can optimize their signal with respect to each other before transmitting it to the transmitter pertaining to some conditions, i.e., some channel feedback is possible from the users to the transmitter along with some interaction is possible among users. Further, it is required that each user should know all the other users' channel vectors including its own. In contrast, if users can't talk to each other, then the challenge in the uplink communication system is primarily in the postcoding done by the transmitter to separate the users' informations.

2.2.1 **MU-MIMO System**

In this work, we are considering a downlink MU-MIMO system with a transmitter and K users. Transmitter is equipped with N antenna elements, and each user is equipped with a single receive antenna, and $K < N$. Further, at transmitter, antenna elements in linear array are kept $\frac{\lambda}{2}$ apart from each other to avoid antenna correlation. λ is the wavelength of carrier frequency used. Thus, the flat fading MU-

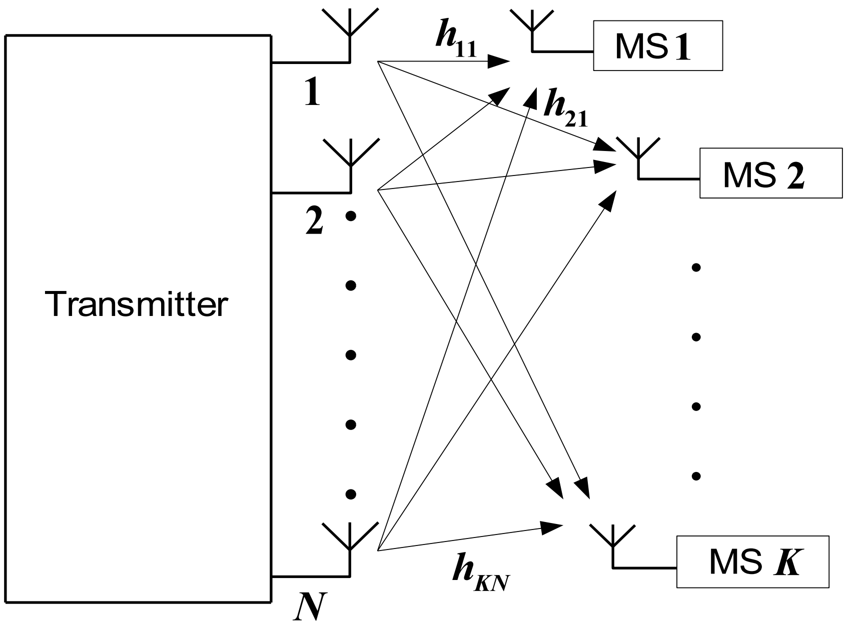


Figure 2.5 MU-MIMO System Model

MIMO system model for downlink is described by the input-output relationship as

$$\mathbf{y} = \mathbf{H}^H \mathbf{P} \mathbf{x} + \mathbf{w}, \tag{2.19}$$

where $\mathbf{H}^H \in \mathbb{C}^{K \times N} = [\mathbf{h}_1^H, \ldots, \mathbf{h}_K^H]^H$ is the MU-MIMO channel matrix with elements as $h_{kn}, k = \{1, \ldots, K\}, n = \{1, \ldots, N\}$, and $\mathbf{h}_k$ is the channel vector for user k. $\mathbf{x} \in \mathbb{C}^{K \times 1}$ is the transmitted symbol vector, and $\mathbf{y} \in \mathbb{C}^{K \times 1}$ is the received information vector. $\mathbf{w}$ is an AWGN vector with $\mathbf{w} \sim \mathbb{C}\mathcal{N}(0, N_0 \mathbf{I}_k)$. $\mathbf{P} \in \mathbb{C}^{N \times K}$ is the precoding matrix which comprises of precoding vectors corresponding to each user, respectively

$$\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_K], \tag{2.20}$$

where $\mathbf{p}_K$ is the precoding vector for user K. Further, one need to satisfy a power constraint as

$$\mathbb{E}[\| \mathbf{P}\mathbf{x} \|^2] \le \rho, \tag{2.21}$$

where ρ is the total transmitted power.

2.2.2 MU-MIMO System Capacity

The main advantage of precoding in downlink MU-MIMO systems is to pre-cancel the MUI at the transmitter. This is achieved by multiplying the transmitted information vector with a precoding matrix $\mathbf{P}$ before transmission. But, knowledge of CSI must be available at the transmitter in order to design the precoder. The transmitted signal is given as

$$\tilde{\mathbf{x}} = \sum_{k=1}^{K} \mathbf{p}_k x_k, \tag{2.22}$$

where x_k is the symbol of user k. The received signal at user k can be written as

$$y_k = \mathbf{h}_k \mathbf{p}_k x_k + \sum_{j \neq k} \mathbf{h}_k \mathbf{p}_j x_j + w_k. \tag{2.23}$$

The SINR which can be obtained at user k is

$$\text{SINR}_k = \frac{p_k |\mathbf{h}_k \mathbf{p}_k|^2}{\sum\limits_{j \neq k} p_j |\mathbf{h}_k \mathbf{p}_j|^2 + N_0}, \tag{2.24}$$

where p_k is the power allocated to user k, and $p_k |\mathbf{h}_k \mathbf{p}_k|^2$ is the received signal power at user k. $\sum\limits_{j \neq k} p_j |\mathbf{h}_k \mathbf{p}_j|^2$ is the interference power generated due to interference of information signal of user k with the remaining users' information signal, and N_0 is the noise power. Further, MUI is nullified for each user k through precoding as

$$\sum_{j \neq k} p_j |\mathbf{h}_k \mathbf{p}_j|^2 = 0. \tag{2.25}$$

One can redefine the SINR obtained at user k as

$$\text{SINR}_k = \frac{p_k |\mathbf{h}_k \mathbf{p}_k|^2}{N_0}. \tag{2.26}$$

Thus, capacity of MU-MIMO system is defined as

$$C = \sum_{k=1}^{K} \log_2 \left(1 + \frac{p_k |\mathbf{h}_k \mathbf{p}_k|^2}{N_0} \right) \text{ b/s/Hz}. \tag{2.27}$$

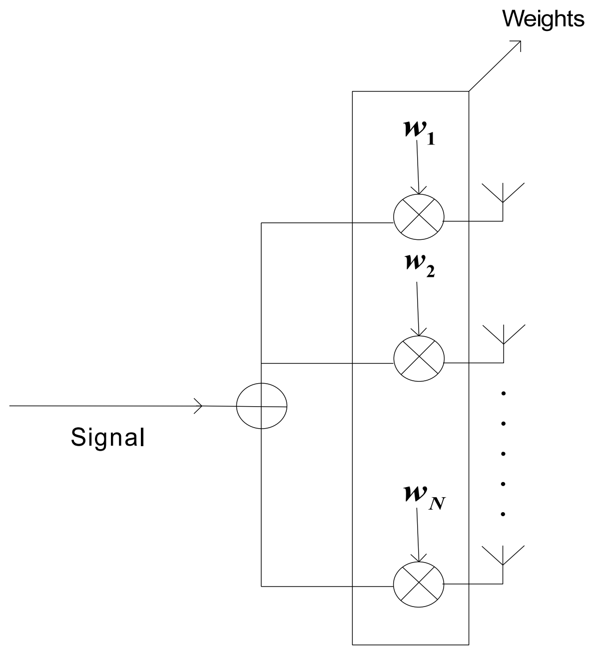


Figure 2.6 MIMO Beamforming

2.3 Overview of MIMO Beamforming

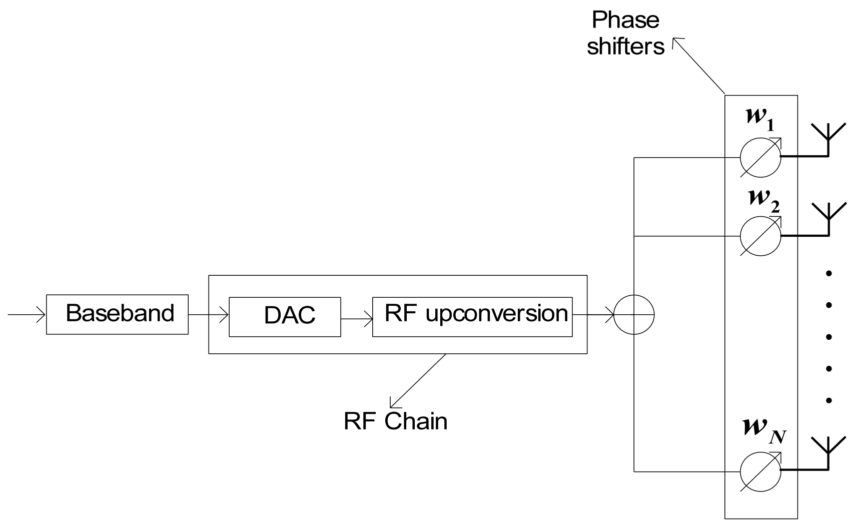
Beamforming is essentially a spatial filtering operation typically using an array of antenna elements or radiators to capture and/or radiate energy from/in a specific direction over its aperture [57], as shown in the Fig. 2.6. Thus, the improvement achieved over omnidirectional transmission/reception is the transmit/receive gain. Modern communication systems deploy smart antenna systems which can combine array gain along with interference mitigation to further increase capacity of the communication link. This is achieved by electronic beam steering using a phased array, which is a multi-antenna radiation device with a specific configuration array [57], [63], i.e., linear phased array or planner phased array.

Using phased antenna array, it is possible to control the shape and direction of the signal beam from multiple antenna elements, and antenna elements are kept at the specific spacing in the array. At each antenna element, a signal is multiplied with weight (a complex number) before transmitting, as shown in Fig. 2.6. In turn, it generates a beam pattern in the specific direction. In other words, the creation of the beam using the technique of constructive interference is called beamforming. Further, the spatial power distribution, termed as the antenna array radiation pattern, can be determined by the vector sum of the fields radiated by individual antenna element [57], [63]. It can be expressed in terms of the array factor, which is a function of the antenna array geometry and amplitude/phase shifts applied to individual antenna elements [57], [63].

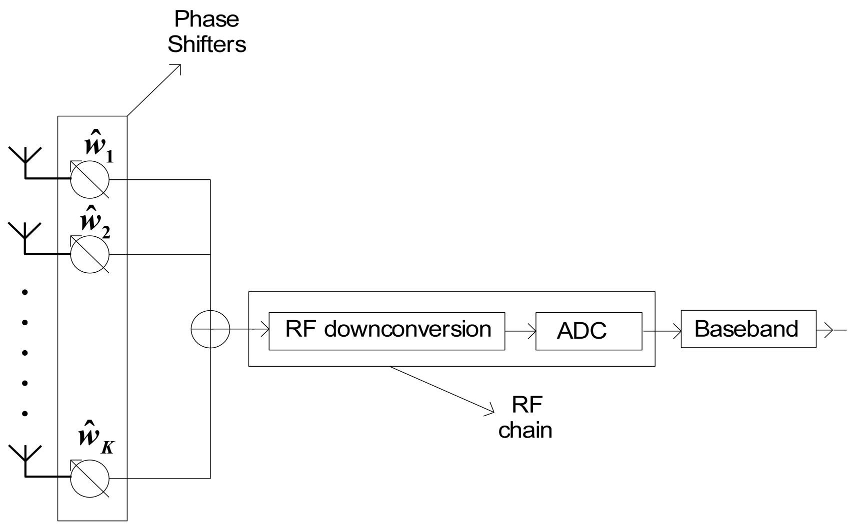
2.3.1 **Analog Beamforming**

Analog beamforming is a technique to transmit and/or receive a signal in/from a specific direction by combining the linearly weighted signal in the analog domain. These weights comprise of amplitude and phase-shifters and are applied to each antenna element, as shown in Fig. 2.7. The antenna array coherently adds up the signal at a particular angle and destructively cancels out other signals. One can achieve a signal with high SINR as the analog beamforming is surpassing the interference [64]. Further, the transmitted/received signal is fed to digital to analog (DAC)/ analog to digital (ADC) converter as a input after up-conversion/down-conversion. So, a signal with high SINR is obtained and it allows the use of low-resolution ADCs and DACs. The ADC/DAC power consumption is $f_s \times 2^{2R}$, where f_s is the sampling frequency, R is the ADC/DAC resolution in bits [64]. Therefore, the signal obtained with high SINR through analog beamforming can be converted to digital/analog using low-resolution ADC/DAC, without loss of any information.

A low-resolution ADC/DAC based single-user analog beamforming structure was proposed in [64] which combines the weighted signals in the baseband domain. The antenna weights are adjusted by minimizing the mean squared error of the desired signal [64]. Combining the weights in the baseband domain require a number of ADCs/DACs and RF chains, which is equal to a number of antenna elements. To alleviate the number of ADCs, a single ADC and RF based mmWave analog beamforming structure for a single user was proposed in [65], as shown in Fig. 2.7. It combines the weighted signal in the RF domain, and the linearly added weighted signals are given to ADC as input after down conversion and the antenna weights or beamforming vectors are adjusted through a gradient descent method. However, fine adjustment of beamforming vectors lead to hardware complexity and high power consumption. To address these issues, the weights are adjusted through



(a) Transmitter



(b) Receiver

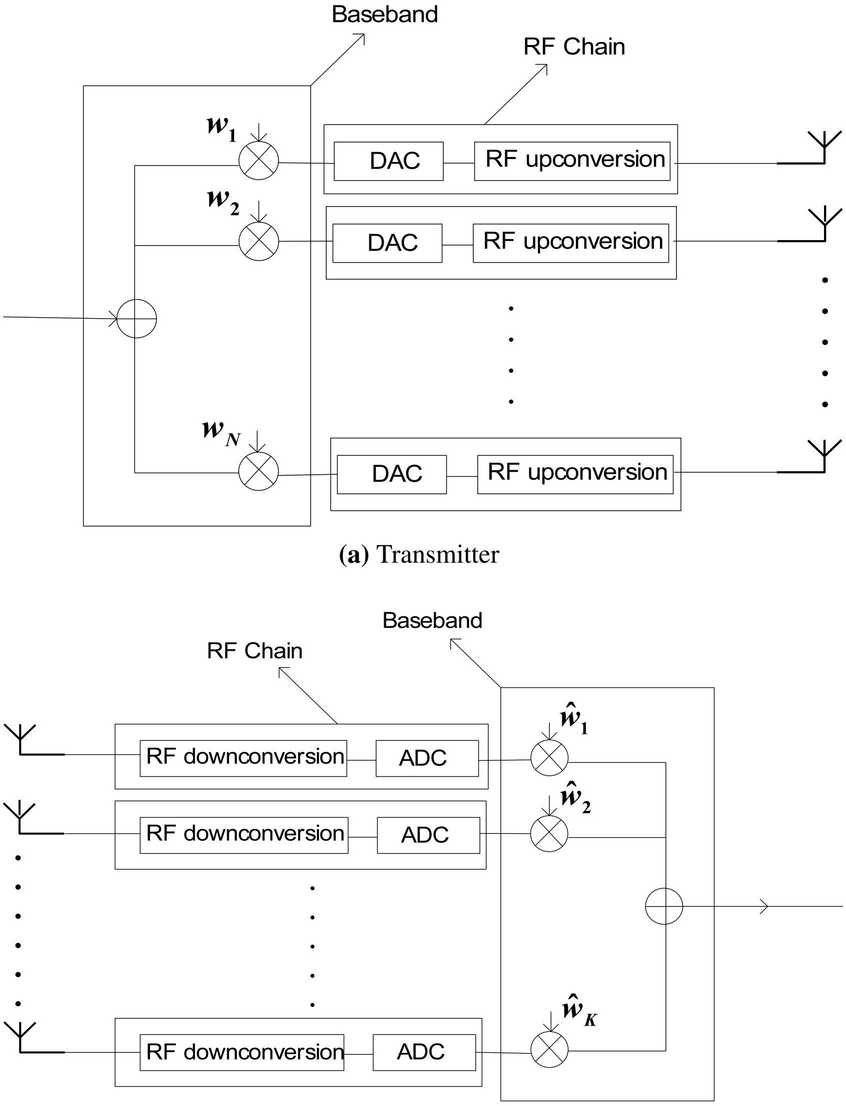
Figure 2.7 Analog Beamforming

a codebook, which comprises of beamforming vectors made of low-resolution RF-phase shifters in the predefined directions. Next, the receiver feedback an index to the transmitter indicating the best beamforming vector to be used at the transmitter. An exhaustive search algorithm is used to find the best beamforming vector. This algorithm sequentially tests all the beamforming vectors and finds the best beamforming vector. However, the overall search time is prohibitive because the number of beamforming vectors are usually large for mmWave communication. Further, to improve the search efficiency, a hierarchy of codebook is used to search the best beamforming vector [66, 67]. A hierarchical codebook consisting of a small number of low-resolution RF-phase shifters covering wide angle at the top level of the codebook and a large number of high-resolution RF-phase shifters offering high directional beamforming gain at the bottom level.

2.3.2 **Digital Beamforming**

Unlike analog beamforming, digital beamforming is a technique to transmit and/or receive a signal in/from a specific direction by combining the linearly weighted signal in the baseband or digital domain. Most of the analog beamforming structure adjust their weights using gradient descent algorithm [65]. But, the convergence rate of the gradient descent algorithm is sometimes unacceptable [68]. An alternative algorithm, Gram Schmidt orthogonalization is much faster [68]. But, it is computational prohibitive and require high precision which is possible in digital domain. Further, one must aim to estimate weights of high-resolution which means it requires a non-linear processing. Thus, the requirement of faster convergence and high-resolution weights make a prerequisite for an advent of digital beamforming. In this context, a structure of digital beamforming comprising of a number ADCs/DACs equal to number of antenna elements was proposed in [68] , as shown in Fig. 2.8. Still there are some hardware complexities with regard to dynamic range of ADC/DAC, self calibration require accurate amplitude and phase reference at each element. Evidently, one can utilize the maximum number of degrees of freedom in the antenna array with digital beamforming [57], as it requires up/down converters, ADCs/DACs, at each antenna element.

Further, an architecture of digital beamforming was proposed in [69], which require less hardware. It is cost effective and easy to implement. The system comprises of the RF down converter, ADCs, and an adaptive algorithm. The signal received from each antenna element, i.e., both real and imaginary signals are combined separately using a multiplexer (MUX) into vector signals. These real and imaginary vector signals are then digitized separately using two ADCs for both real and imaginary vector signals. The digitized real and imaginary vector signal is down converted using a digital down converter, the real and imaginary vectored signals are then separately de-interleaved using a demultiplexer (DEMUX) into two linear mean square (LMS) algorithm, one corresponding to the real part and the other cor-



(a) Transmitter

(b) Receiver

Figure 2.8 Digital Beamforming

responding to the imaginary part. This improved structure has reduced hardware computation as compared to the conventional digital beamforming structure.

For fifth-generation mmWave communications, a MIMO transceiver with fully digital beamforming of 64 channels, deployed as a 2-D array with 16 columns and 4 rows for a better beamforming resolution, is tested in [70]. It operates at 28-GHz band with a 500-MHz signal bandwidth. Further, the system performance is tested to verify the feasibility of the digital beamforming based massive MIMO transceiver for mmWave communications, and achieve a steady 5.3-Gb/s throughput for a single user in fast mobile environment using the beam-tracking technique and two streams of 64-QAM signals. Thus, the digital beamforming based mmWave

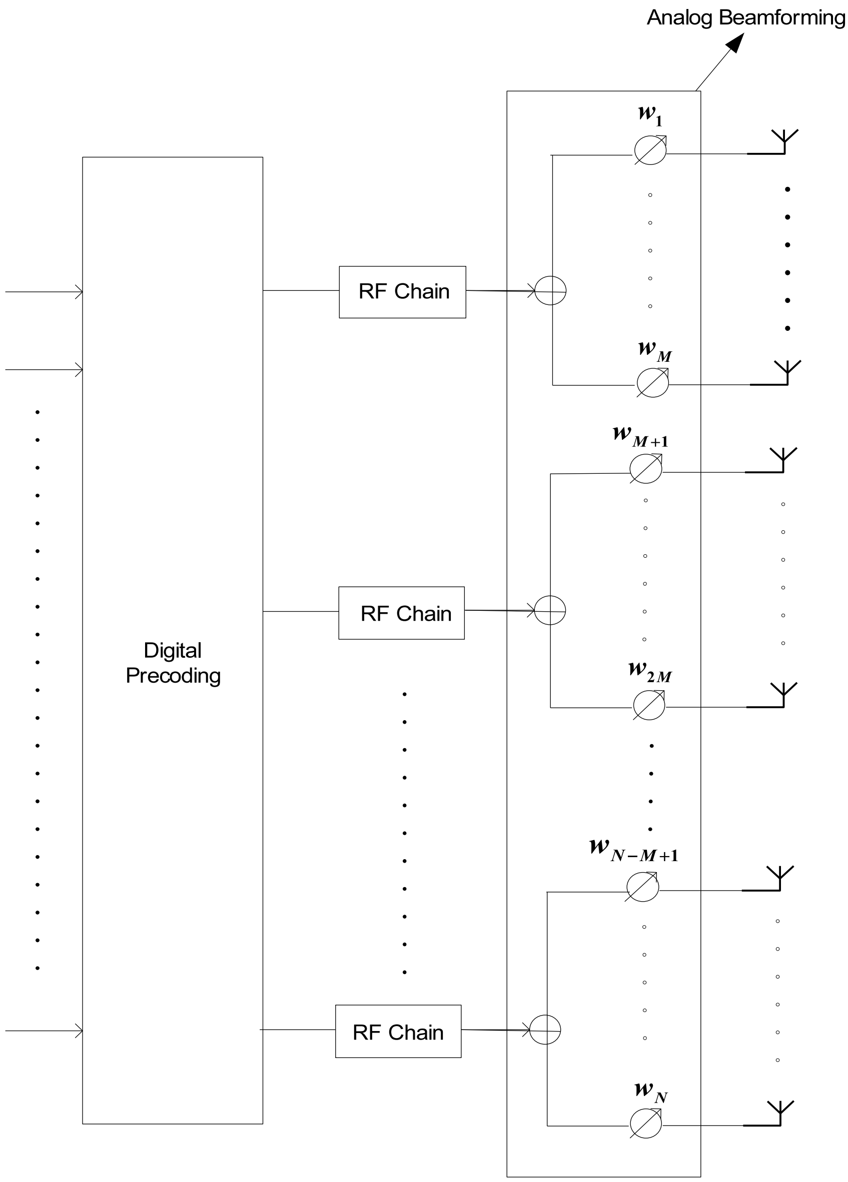


Figure 2.9 Partially Connected Hybrid Beamforming

MIMO transceiver is a hopeful choice for future 5G communications.

2.3.3 **Hybrid Beamforming**

A large number of antenna elements are employed to overcome the severe path loss, absorption loss in the mmWave spectrum [71]. However, the high cost and power consumption of mixed-signal components prevent a separate RF chain for each antenna from using MIMO baseband beamforming/precoding schemes, which supports multi-stream for a single user and multi-user system. To overcome these hardware limitations, the experts suggested splitting the baseband beamforming/precoding processing between analog and digital domains by designing hybrid analog-digital beamforming/precoding schemes [71]. Note that beamforming with multiple data streams, is known as precoding, can achieve high-performance. But, the hardware limitation employ constraints while designing hybrid, precoder or, beamformer. Many research papers [72–79] have been investigating the design of low-complexity hybrid analog-digital beamformer for single-user and multi-user mmWave communication systems.

Two types of architectures of hybrid beamforming [80, 81] have been discussed in the literature, i.e., partially-connected and fully-connected hybrid beamforming as shown in the Fig. 2.9 and 2.10. Let's assume a transmitter equipped with N antenna elements. Each antenna element is placed in a single row with a critical distance, i.e., $\frac{\lambda}{2}$, where λ is the carrier frequency in the mmWave spectrum. Note that the antenna elements are kept apart $\frac{\lambda}{2}$ distance to avoid any correlation among antenna elements at transmitter and receiver end. The partially-connected hybrid beamfroming uses a separate antenna array, also called a subarray, for the analog beamforming. Each sub-array is connected to a individual RF chain. While in the fully-connected hybrid beamforming, each RF chain is connected to all the antenna elements. Theoretically, the performance with the fully-connected ~~hybird~~ beamforming architecture is supposed to be better than that of the partially-connected hybrid beamfroming [82]. But, additional components are required to combine the RF signals from different RF chains make RF circuits design challenging.

Compared to the fully connected architecture, each RF chain in the partially connected architecture has access to less number of antenna elements such that its analog beamforming will have a wider beam width [80, 81]. Thus, the less directive signal will have a stronger interference from other analog beamforming. In spite of these disadvantages, the increased MUI among different sub-arrays can be effectively mitigated with MIMO baseband precoding and combining at transmitter and receiver ends [80–82]. Considering the circuit designs challenges and the performance losses, partially-connected hybrid beamforming is preferred in practice for mmWave communications [82]. Hybrid beamforming has been a prerequisite to enable mmWave communication for the indoor as well as outdoor environment.

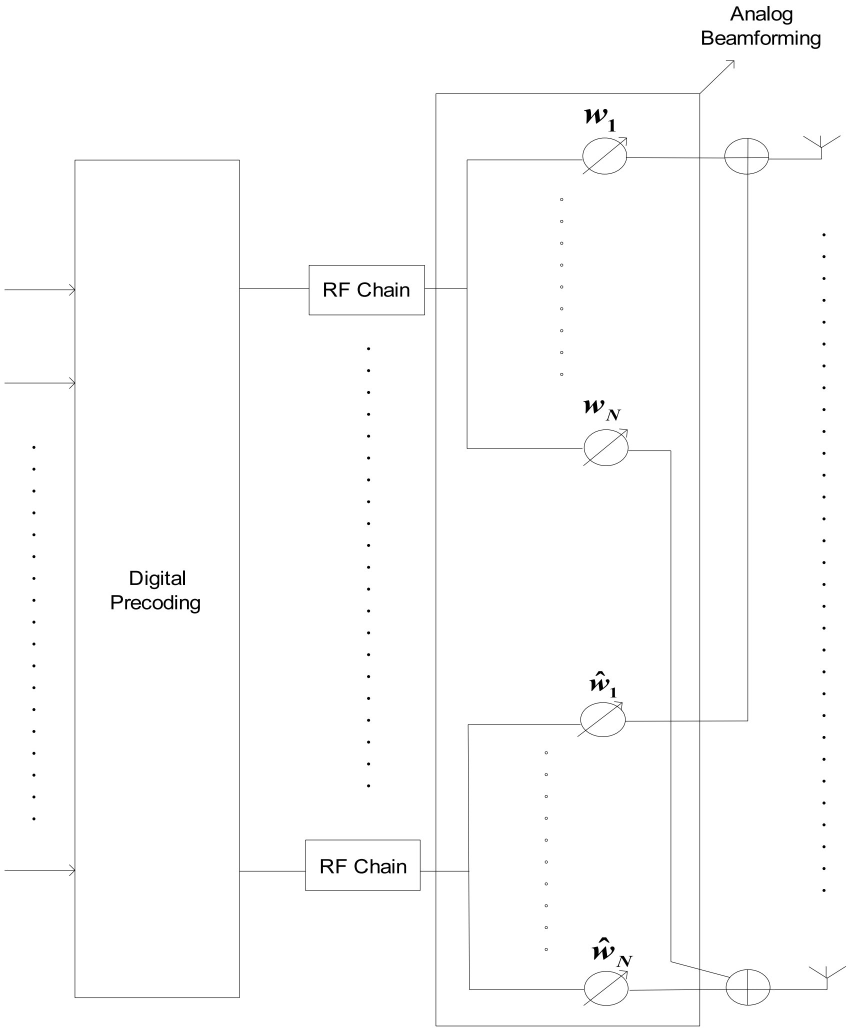


Figure 2.10 Fully Connected Hybrid Beamforming

But, designing of hybrid precoder has been a bottleneck for such single user and multi-user mmWave communication systems.

A single user multi-stream mmWave system is designed with hybrid beamforming at both ends, namely, transmit, and receive beamforming. Further, to design a hybrid precoder and combiner for such a system is cumbersome. However, different methods have been applied to find an optimal precoder through sparse recovery methods, i.e., basis pursuit [72], orthogonal matching pursuit [73], compressive sensing [74]. Such methods exploit sparse nature of the mmWave channels. It is not possible to acquire full channel information at both ends due to hardware limitations and noise. Hence, a novel hybrid precoder and combiner is discussed in [74], using the principle of matching pursuit while assuming partial channel information, in the form angle of arrival/departure knowledge, at transmitter and receiver. Further, to acquire partial channel information at both ends is possible through estimating the mmWave channel. Therefore, an adaptive channel estimation algorithm is proposed in [75], for single LoS path with practical assumptions on the analog beamforming vectors. The mmWave channel estimation is not restrained only for LoS path [75]. Hence, Zhenyu Xiao *et al.*, [76] developed a channel estimation algorithm for multipath (LoS and NLoS) path. Here, a hierarchical multi-beam search scheme is used to improve search accuracy for analog precoder. It uses a pre-designed analog hierarchical codebook and it achieves the performance, i.e., spectral efficiency, close to [75]. Different frequency components of signal experience different fading for wide signal bandwidth, i.e, frequency selective fading. Therefore, orthogonal OFDM-based precoder and combiner for frequency selective channels is proposed in [77–79] which achieves the performance close to digital precoder and combiner.

Digital precoding in hybrid beamformer for multi-user systems is equally important compared to analog beamforming. It is implemented to reduce the MUI. Hence, developing hybrid beamforming vectors for multi-user mmWave systems is also of special interest. The hybrid beamforming at both ends, transmit and receive beamforming, for MU-MIMO mmWave systems is discussed in [77, 78, 83–95]. Design of hybrid beamforming comprising of analog beamforming and digital precoding together is not a simple task. However, separately designed analog beamforming and digital precoder are concatenated to obtain a beamforming. This methodology reduces hardware complexity at both ends. The vectors or weights for analog beamforming for each user are adjusted through low-resolution phase shifters in RF domain at both ends. In contrast, digital precoding is designed in baseband domain to mitigate the effect of MUI. Digital precoding such as ZF and minimum mean sqaure (MMSE) are commonly used to mitigate the effect of MUI. Note that, one must know full CSI to design a digital precoder. However, S. Park *et al.* have developed [77] a digital precoder with partial channel information.

Evidently, a transmitter in the cellular system deployed at base station could have unlimited resources such as power, bandwidth, etc. But, the mobile station or users will have limited resources because of their compact size. Therefore, the mobile station is considered to be equipped with a single receive antenna. Hence, beamforming is not possible at the mobile station. A downlink mmWave MU-MIMO system proposed in [96] has a transmitter, and K users with single receive antenna, where transmitter is equipped with N number of antenna elements, but driven by a small number of RF chains. Such hybrid beamforming is deployed at the transmitter to reduce hardware complexity, which has jointly optimized analog beamforming and digital precoding to maximized the sum rate [96]. Similarly, a robust and low complexity hybrid beamforming is proposed [97] for uplink mmWave MU-MIMO system, here analog beamforming is obtained using the low complexity Gram-Schmidt method, and the digital precoding matrix is obtained using MMSE with the low dimensional effective channel. Due to practical constraints on RF pahse shifters, i.e, low-resolution RF phase shifters, insignificant loss in the system performance is observed in the analog beamforming.

2.4 Overview of mmWave MU-MIMO

In section 2.2, we had discussed a MU-MIMO system equipped with conventional or digital beamforming. Further, it is not possible to employ digital beamforming/precoding at mmWave frequencies because of hardware limitations and the requirement of a large antenna array. Digital beamforming using such a large antenna array require large number of RF chains and ADCs/DACs. Hence, hybrid beamforming is better approach to curtail the required number of RF chains and ADCs/DACs. mmWave MU-MIMO system with hybrid beamforming is detailed in this section.

2.4.1 mmWave MU-MIMO System

We are considering a downlink mmWave MU-MIMO system equipped with a transmitter, and K mobile stations (MS) or users, as shown in Fig. 2.11. Transmitter is equipped with a uniform linear array (ULA), and each user is having a single receive antenna. ULA consists of N antenna elements, and here $K << N$. Further, antenna elements are kept $\frac{\lambda}{2}$ apart from each other. Thus, the flat fading mmWave MU-MIMO system model for downlink is described by the input-output relationship as

$$\mathbf{y} = \mathbf{H}^H \mathbf{P}\mathbf{x} + \mathbf{w}, \tag{2.28}$$

where $\mathbf{H}^H \in \mathbb{C}^{K\times N} = [\mathbf{h}_1^H, \ldots, \mathbf{h}_K^H]^H$ is a mmWave MU-MIMO channel matrix, and $\mathbf{h}_k$ is channel vector to user k, $k = \{1, \ldots, K\}$, $n = \{1, \ldots, N\}$, . $\mathbf{x} \in \mathbb{C}^{K\times 1}$ is the transmitted symbol vector, and $\mathbf{y} \in \mathbb{C}^{K\times 1}$ is the received information vector. $\mathbf{w}$ is an AWGN vector with $\mathbf{w} \sim \mathbb{C}\mathcal{N}(0, N_0\mathbf{I}_k)$. $\mathbf{P} \in \mathbb{C}^{N\times K}$ is a precoding matrix, which is comprising of analog beamforming and digital precoding matrix as $\mathbf{P} = \mathbf{A}\mathbf{D}$, where $\mathbf{A}$ is a analog beamforming matrix and $\mathbf{D}$ is a digital precoding matrix. Thus, precoding matrix is consisting precoding vectors corresponding to each user, respectively

$$\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_K], \tag{2.29}$$

where $\mathbf{p}_K$ is the precoding vector for user K. Further, we need to satisfy a power constraint as

$$\mathbb{E}[\| \mathbf{P}\mathbf{x} \|^2] \leq \rho, \tag{2.30}$$

where ρ is the total transmitted power.

2.4.2 mmWave MU-MIMO Channel Model

It has been observed through different channel measurements and found that mmWave communications is dominated by LoS paths and few dominant NLoS paths, which is a kind of quasi optical nature of propagation, resulting in sparse channel characteristics [98]. Thus, the channel model is developed for sparse MIMO channels, appropriate for mmWave frequencies, and propose an approach, i.e., user's localization based MIMO channel matrix as a function angle of arrival (AOA) of the signal. Most existing techniques focus on the information provided by the LoS path. However, LoS propagation is not always guaranteed in a real world environment, e.g. urban or indoor sights. Thus, the channel model for mmWave communication system [99] for user k is given as

$$\mathbf{h}_k = \sqrt{\frac{N}{L+1}} \sum_{\ell=0}^{L} \beta_k^{(\ell)} \mathbf{a}\left(\theta_k^{(\ell)}\right), \tag{2.31}$$

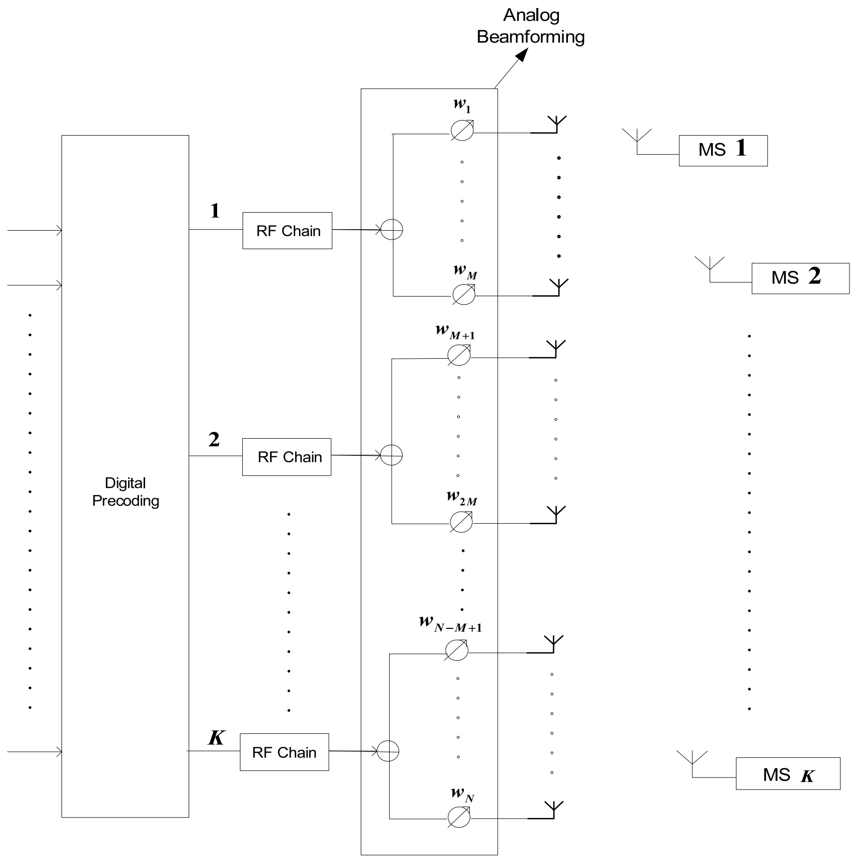


Figure 2.11 mmWave MU-MIMO System

where $\beta_k^{(0)}$ denotes the complex-valued path gain of the LoS and $\beta_k^{(\ell)}, \ell = 1, \ldots, L$ denotes the complex-valued path gain of ℓ^{th} NLoS, respectively. Further, the array steering vector corresponding to the ℓ^{th} path for user k is given by,

$$\mathbf{a}\left(\theta_k^{(\ell)}\right) = \frac{1}{\sqrt{N}}\left[\exp(-\jmath 2\pi\theta_k^{(\ell)} i)\right]_{i\in\mathcal{I}_N}, \tag{2.32}$$

where $\theta_k^{(\ell)}$ is the spatial frequency evaluated by the angle of arrival (AOA) $\phi_k^{(\ell)}$ corresponding to the l^{th} path for user k as

$$\theta_k^{(\ell)} = \frac{d}{\lambda}\sin\phi_k^{(\ell)}, \tag{2.33}$$

where $\phi_k^{(\ell)}$ is uniformly distributed in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and $\mathcal{I}_N = \{i - (N - 1)/2, i = 0, 1, \ldots, N - 1\}$ is a symmetric set of index which is centered around 0.

2.4.3 **mmWave MU-MIMO Channel Sparsity**

Recent works in wireless communications consider rich multipaths. Thus, the research on MIMO systems was developed by results based on an identical independent distribution (i.i.d.) channel model, which represents a rich multipath environment. However, there is an experimental evidence that mmWave wireless channels exhibit a sparse multipath structure accompanied with large bandwidths [100–105]. Enough experiments were conducted at mmWave frequencies to estimate the channel [106], [107] which show some interesting results; the LoS path domionates over NLoS paths, hence NLoS paths 10-20 dB are weaker than LoS path [98, 108, 109]. Therefore, mmWave communications is considered as a directional communication, and it is treated as a LoS communication. In turn, mmWave communication propagation exhibits channel sparsity.

In section 2.4.2, we had discussed the conventional or spatial channel model for mmWave MU-MIMO system. Due to directional nature of mmWave communication, transmitter requires accurate information of Angle of Arrival (AOA) for each user. However, it is cumbersome job. Further, the channel model can be transformed from the spatial domain, for exploiting channel sparsity, to the beamspace domain by employing discrete lens array (DLA) rather employing ULA at transmitter [98, 108, 109]. Functionality of DLA is the same as ULA except transforming the channel into beamspace domain from spatial domain. Hence, beamspace domain allows us to work on a few predefined directions only, and these predefined directions cover entire angular region. However, DLA is replacing the phase shift networks, as shown in Fig. 2.12. Hence, it is possible to generate multiple beams in the predefined directions using DLA. This reduces the hardware complexity of the MIMO hybrid architecture.

2.4.4 **Beamspace Representation**

The above discussion presents the channel model in conventional spatial domain. Due to the highly directional nature of propagation at mmWave communications, LoS component dominates over NLoS components and the mmWave channel is

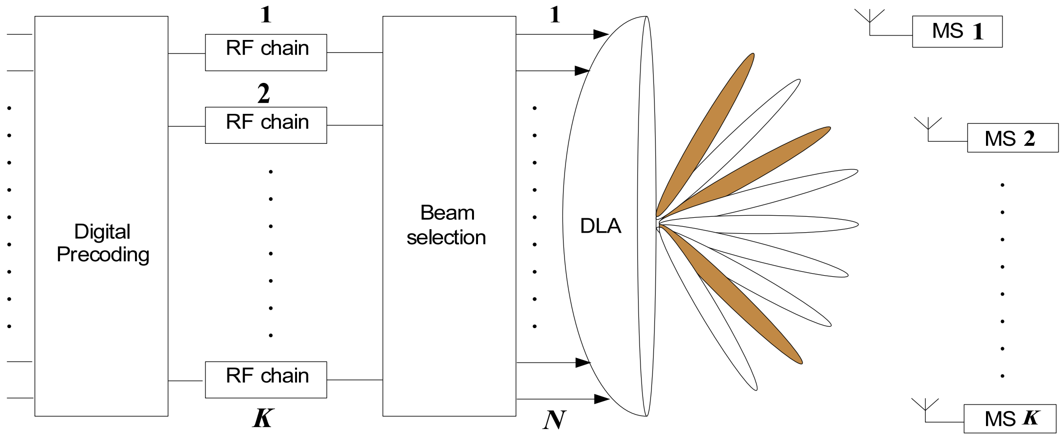


Figure 2.12 mmWave Beamspace MU-MIMO System

sparse. Beamspace domain, i.e., angular domain, representation enables us to exploit the inherent sparsity in such channels. The conventional spatial channel can be transformed into the beamspace domain by employing DLA at the transmitter. DLA plays the role of a *spatial discrete Fourier transform* (DFT), which can be represented by the matrix $\mathbf{U} \in \mathbb{C}^{N \times N}$. The columns of $\mathbf{U}$ are array response vectors corresponding to N fixed spatial frequencies or N orthogonal predefined directions given by

$$\theta_i = \frac{i}{N}, i \in \mathcal{I}(N). \tag{2.34}$$

These predefined directions are covering the entire angular space, and the beamforming matrix $\mathbf{U}$ can be expressed as

$$\mathbf{U} = \left[\mathbf{a} \left(\theta_i = \frac{i}{N} \right) \right]_{i \in \mathcal{I}(N)}, \tag{2.35}$$

where the columns of beamforming matrix correspond to a pre-defined spatial direction. Thus, the columns of matrix $\mathbf{U}$ play a role of spatial filtering in the predefined directions, and it is analogous to analog beamforming. Further, the column of matrix $\mathbf{U}$ are DFT vectors, and exhibits a property, i.e., $\mathbf{U}^H \mathbf{U} = \mathbf{I}$. The beamspace representation of mmWave beamspace MU-MIMO system is given by

$$\mathbf{y}_b = \mathbf{H}_b^H \mathbf{P}_b \mathbf{x} + \mathbf{w}_b, \tag{2.36}$$

where $\mathbf{H}_b = \mathbf{U}^H \mathbf{H} \in \mathbb{C}^{N \times K}$ is the beamspace channel matrix, $\mathbf{P}_b \in \mathbb{C}^{K \times N}$ is a digital precoding matrix. Each $\mathbf{h}_{b,k} = \mathbf{U}^H \mathbf{h}_k \in \mathbb{C}^{N \times 1}$, $k = 1, \ldots, K$, will have few dominant entries (significantly less than N) around the LoS direction $\theta_k^{(0)}$ and thus, $\mathbf{H}_b$ captures the inherent sparsity in the mmWave channel.

2.4.5 **Problem Definition: Beam Selection**

Problem definition is to select only K beams out of N without incurring considerable loss in the sum rate $R = \sum_{k=1}^{K} R_k$, where R_k is the data rate achieved by user k after beam selection. Such a reduced-dimensional system requires only K RF chains, rather than N RF chains required by a full-dimensional system. The K-dimensional system, after beam selection, can be expressed as

$$\tilde{\mathbf{y}}_b = \tilde{\mathbf{H}}_b^H \tilde{\mathbf{P}}_b \mathbf{x} + \tilde{\mathbf{w}}_b, \tag{2.37}$$

where $\tilde{\mathbf{H}}_b^H = [\tilde{\mathbf{h}}_{b,1}^H, \ldots, \tilde{\mathbf{h}}_{b,K}^H]^H \in \mathbb{C}^{K\times K}$ is the beamspace channel matrix corresponding to the K selected beams and $\tilde{\mathbf{P}}_b \in \mathbb{C}^{K\times K}$ is a (reduced-dimensional) digital precoding matrix. $\tilde{\mathbf{w}}_b$ is AWGN noise with $\tilde{\mathbf{w}}_b \sim \mathbb{C}\mathcal{N}(0, N_0\mathbf{I})$.

One can obtain $\tilde{\mathbf{H}}_b^H$, by appropriately selecting K columns from $\mathbf{H}_b^H$, and understand the existing beam selection algorithm to obtain $\tilde{\mathbf{H}}_b^H$ from $\mathbf{H}_b^H$. The primary beam selection algorithm was maximum magnitude (MM) [108], which maximizes magnitude of the beam or channel gain corresponding to each user. However, multiple users can select the same beam. Hence, "MM" beam selection is not a practically viable algorithm until there is user-wise beam selection along with a suitable users' topoloy [109]. Further, several beam selection algorithms were proposed which select distinct beams for each user and outperform "MM" beam selection algorithm. Pierluigi V. Amadori *et al.* proposed two beam selection algorithms [109]; maximization of the signal-to-interference-plus-noise-ratio (M-SINR), and maximization of the capacity (MC). "MC" performs the beam selection with two different approaches, i.e., decremental and incremental. Decremental "MC" selects one-by-one beam which are not to be used and incremental "MC" selects one-by-one beam which is to be used. However, incremental and decremental "MC" having different computational complexity perform equal for each SNR. Next, "M-SINR" and "MC" perform approximately equal. "MC" performs inferior to "M-SINR" at low

SNR, but performs equal at high SNR. But, “M-SINR” and “MC” beam selection algorithms are computationally much complex than “MM.” Therefore, Xinyu Gao *et al.* proposed a near optimal beam selection algorithm named interference-aware (IA) beam selection algorithm [110]. “IA” beam selection is using the concept of “MM” and “M-SINR.” In the first step, “IA” selects beam for non-interference-users (NIUs) and in the second step, “IA” selects beam for interference-users (IUs). For NIUs, the beams with larger magnitude are selected while for IUs, the beams are selected by a low-complexity incremental algorithm based on the criterion of maximization of the “M-SINR.” Finally, “IA” beam selection achieves near optimal performance to “M-SINR” with less complexity.

The existing beam selection algorithms are discussed so far. Although these algorithms proposed in [108–110] achieve satisfactory performance, but it is not considerably close to full dimensional system performance with all beams are active. In order to achieve considerable close to full dimensional system performance, we are trying to develop beam selection algorithms while taking account of computation complexity.

2.4.6 **Zeroforcing Precoding/Beamforming Scheme**

In digital precoding, the symbol vector corresponding to all users are passed through a digital precoder before being transmitted over different phase shifters or analog beamformer, and then the information vector pass over different antennas, as shown in Fig. 2.11 and 2.12. In fact, digital precoder corresponding to each user is a weight vector and it is designed with a specific method in order to cancel the interference caused by the others [108, 109]. More specifically, digital precoder will act constructively in the desired directions and destructively in the undesired directions. In turn, this will enhance the received SNR at the users and cancel the MUI.

Considering a K-dimensional system model as explained in section 2.4.5 where $\tilde{\mathbf{P}}_b \in K \times K$ is a digital precoding matrix, and the transmitted signal has average power constraint, while perfect channel state information is assumed at transmitter.

$$\mathbb{E}[\| \tilde{\mathbf{P}}_b \mathbf{x} \|^2] \leq \rho. \tag{2.38}$$

In digital precoding, direction of transmitted information for each user is different. The transmitted signal vector is given as

$$\tilde{\mathbf{x}} = \sum_{k=1}^{K} \tilde{\mathbf{p}}_{b,k} x_k \tag{2.39}$$

where x_k, $\tilde{\mathbf{p}}_{b,k}$ are the symbol and the digital precoding vector, respectively, and it is defined as

$$\tilde{\mathbf{P}}_b = [\tilde{\mathbf{p}}_{b,1}, \ldots, \tilde{\mathbf{p}}_{b,k}]. \tag{2.40}$$

The received signal at user k can be written as

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{h}}_{b,k}\tilde{\mathbf{p}}_{b,k}x_k + \sum_{j\neq k} \tilde{\mathbf{h}}_{b,k}\tilde{\mathbf{p}}_{b,j}x_j + \tilde{\mathbf{w}}_{b,k}. \tag{2.41}$$

In ZF precoding, the precoders has to satisfy a condition which is given as

$$\tilde{\mathbf{h}}_{b,k}\tilde{\mathbf{p}}_{b,j} = 0, j \neq k, \tag{2.42}$$

then the matrix $\tilde{\mathbf{P}}_b$ can be selected to be the inverse of the channel matrix $\tilde{\mathbf{H}}_b$ as

$$\tilde{\mathbf{P}}_b = (\tilde{\mathbf{H}}_b)^{-1}. \tag{2.43}$$

Now, power constraint must satisfy the condition given in (2.38). Thus, $\mathbf{T}$ is a matrix, and can be redefine as

$$\mathbf{T} = \alpha\tilde{\mathbf{P}}_b, \tag{2.44}$$

where α is a power scaling coefficient that guarantees

$$\mathbb{E}[\| \mathbf{T}\mathbf{x} \|^2] \leq \rho, \tag{2.45}$$

$$\mathbb{E}[\| \alpha\tilde{\mathbf{P}}_b\mathbf{x} \|^2] \leq \rho, \tag{2.46}$$

and α can be evaluated as

$$\alpha \leq \sqrt{\frac{\rho}{\mathrm{Tr}(\tilde{\mathbf{P}}_b\mathbf{\Lambda}\tilde{\mathbf{P}}_b^H)}}, \tag{2.47}$$

where $\mathbf{\Lambda} = \mathbb{E}[\mathbf{x}\mathbf{x}^H]$ is the input covariance matrix. Generally, $\mathbf{\Lambda}$ has to be taken an identity matrix $\mathbf{I}$. Thus, α can be redefined as

$$\alpha \leq \sqrt{\frac{\rho}{\mathrm{Tr}(\tilde{\mathbf{P}}_b \tilde{\mathbf{P}}_b^H)}} \tag{2.48}$$

2.4.7 **mmWave beamspace MU-MIMO System Capacity**

A digital precoder, ZF is to precancel the MUI at the transmitter for mmWave beamspace MU-MIMO system. This is achieved by pre-multiplying ZF precoding matrix $\tilde{\mathbf{P}}_b$ with transmitted symbol vector before transmission. But, it is necessary to have knowledge of channel state information at the transmitter in order to design the ZF precoders. Thus, we obtain the SINR for user k as

$$\mathrm{SINR}_k = \frac{\alpha^2 p_k |\tilde{\mathbf{h}}_{b,k} \tilde{\mathbf{p}}_{b,k}|^2}{\sum\limits_{j=1, j\neq k,}^{K} \alpha^2 p_j |\tilde{\mathbf{h}}_{b,k} \tilde{\mathbf{p}}_{b,j}|^2 + N_0}, \tag{2.49}$$

where p_k is the power allocated to user k, and $p_k|\tilde{\mathbf{h}}_{b,k} \tilde{\mathbf{p}}_{b,k}|^2$ is the received signal power at user k. $\sum\limits_{j=1, j\neq k,}^{K} p_j |\tilde{\mathbf{h}}_{b,k} \tilde{\mathbf{p}}_{b,j}|^2$ is the interference power generated due to interference of user k signal with the signal of remaining users, and N_0 is the noise power. Further, after nullifying the MUI,

$$|\tilde{\mathbf{h}}_{b,k} \tilde{\mathbf{p}}_{b,k}|^2 = 1, \tag{2.50}$$

$$|\tilde{\mathbf{h}}_{b,k} \tilde{\mathbf{p}}_{b,j}|^2 = 0, \quad j \neq k. \tag{2.51}$$

We can redefine SINR obtained for user k as

$$\mathrm{SINR}_k = \frac{\alpha^2 p_k}{N_0}. \tag{2.52}$$

After ZF precoding, the channel of each user experiences the same fading. Hence, we consider equal power allocation $p_k = \frac{\rho}{K}, k \in \{1, \ldots, K\}$, and can redefine SINR_k as

$$\mathrm{SINR}_k = \frac{\alpha^2 \rho}{K N_0}. \tag{2.53}$$

Thus, capacity of the system is defined as

$$C = \sum_{k=1}^{K} \log_2 \left(1 + \text{SINR}_k\right) \text{ b/s/Hz} \tag{2.54}$$

$$C = K \log_2 \left(1 + \frac{\alpha^2 \rho}{K N_0}\right) \text{ b/s/Hz} \tag{2.55}$$

2.5 Concluding Remarks

A brief introduction of MIMO downlink system and its capacity is discussed in this chapter. Further, one can use an optimal power allocation algorithm to improve the MIMO capacity. In this context, an optimal power allocation algorithm, i.e., water filling algorithm is discussed. Next, a MU-MIMO downlink system, and its capacity is discussed including a method to nullify the MUI at user's end. Further, to mitigate the MUI at user's end different beamforming — analog, digital, and hybrid techniques — are discussed. In the next section, mmWave MU-MIMO system, its channel model, and sparsity is discussed. To exploit the channel sparsity, one can transform the channel into beamspace domain from spatial domain, hence the beamspace representation of mmWave MU-MIMO system is discussed. After explaining mmWave beamspace MU-MIMO system, problem definition of such a system as well as previous research work is discussed. Next, mmwave MU-MIMO capacity is discussed while nullifying MUI using ZF precoder. Finally, it has been observed that mmWave communication is a directional communication, and useful for various applications.