CHAPTER 2

MIMO SYSTEM AND LINEAR PRECODING
TECHNIQUES

2.1 INTRODUCTION

By employing multiple antennas at both the transmitter and the receiver in wireless communication systems, popularly known as MIMO technology, a very high data rate can be achieved under the condition of deployment in the rich scattering propagation medium. This interesting property of MIMO system suggests their use in high rate and high quality wireless communication system. Also, several precoding criteria for SU-MIMO and MU-MIMO technologies are discussed. Finally, the basic principles of linear precoding concepts and their developments are discussed.

2.2 MIMO IN WIRELESS NETWORKS

Wireless networks may be broadly classified as cellular or ad hoc networks (Biglieri et al. 2007). A cellular network is a radio network distributed over land areas called cells and is characterized by centralized communication. Each cell is served by at least one fixed transceiver known as Base Station (BS). The multiple users within a cell communicate with a BS that controls all transmission/reception and forwards data to the users. On the contrary, an ad hoc network is a decentralized type of wireless network and it does not rely on a pre-existing infrastructure i.e., any terminal can act as a sender or receiver of data or as a relay for other transmissions. This subsection briefly reviews the use of MIMO technology in each of these
networks and also discusses the new form of technology known as distributed MIMO.

**MIMO cellular networks:** In a cellular wireless communication network, each cell uses a different set of frequencies from neighbouring cells to avoid interference. The network capacity and coverage can be increased by the reuse of frequency resources, provided the same frequencies are not reused in adjacent neighbouring cells as that would cause co-channel interference.

Multiple users may be separated in time (time-division) or frequency (frequency-division) or code (code-division). The spatial dimension in MIMO channels provides an extra dimension to separate users, allowing more aggressive reuse of time and frequency resources, thereby increasing the network capacity.

**MIMO Ad Hoc Networks:** At a given instant of time, some of the terminals will act as sources of data and another subset of terminals will act as the intended destinations. The terminals in the network that are neither sources nor destinations may act as relays to assist data transmission in the network (Tse & Viswanath 2005). The terminals which would act as source / destination will be determined dynamically on the basis of network connectivity. Thus, the number of operating modes in an ad hoc network is very large and will, in general, comprise of combinations of multiple-access, broadcast, relay and interference channels. An ad hoc network typically refers to any set of networks where all devices have equal status on a network and are free to associate with any other ad hoc device in link range.

**Distributed MIMO:** While MIMO technology provides substantial performance gains, the cost of deploying multiple antennas at terminals in a network can be prohibitive, at least in the immediate future. Distributed
MIMO is a means of realizing the gains of MIMO with single-antenna terminals in a network, allowing a gradual migration to a true MIMO network.

This approach requires some level of cooperation between network terminals. This can be Source Inactive Destination. With the advent of Internet and rapid proliferation of computational and communication devices, the demand for higher data rates is ever growing. In many circumstances, the wireless medium is an effective means of delivering a high data rate at a cost lower than that of wire line techniques (such as DSL or cable modem). Limited bandwidth and power make MIMO technology indispensable in meeting the increasing demand for data. MIMO technology is now at the core of many existing and emerging wireless standards such as IEEE 802.11 (for wireless local area networks or WLAN), IEEE 802.16 (for wireless metropolitan area networks or WMAN) and IEEE 802.20 (for mobile broadband wireless access or MBWA). While MIMO is compatible with any modulation scheme, to tackle the increased bandwidth delay spread product, the preferred choice for next-generation networks is orthogonal frequency-division multiplexing (OFDM). OFDM significantly reduces the computational complexity of equalization at the receiver by dividing an otherwise frequency-selective channel into narrower frequency-flat fading sub-channels called tones. The associated multi-user format is orthogonal frequency division multiple access (OFDMA).

2.3 MULTIPATH FADING

The wireless communication systems design has focused on improving link reliability and increasing spectral efficiency. The major problems faced by wireless communication systems are scarce wireless spectrum resource and the complicated wireless propagation environment. The ever growing development of industry and business requires the
increased radio spectrum which leads to the scarce wireless channel band width. In land mobile communication systems, the radio waves may encounter trees, buildings and mountains etc, and the transmitted signal will be reflected, scattered and diffracted. Hence, the wireless channel is a multipath time-varying channel. The multiple paths arise from signals reflecting off multiple random scatterers in the propagation environment; it is intensified in mobile communications with moving transmitter or receiver (or both). These paths combine sometimes constructively and sometimes destructively, creating a channel with multi-tap impulse response, in which each tap has a random phase and a time-varying amplitude. The wireless channel is therefore often characterized statistically.

Hence, the received signal in wireless communication system is the sum of multi path arrival of transmitted signal through differing angles and/or differing time delay and/or differing frequency shifts due to scattering of electromagnetic waves in the propagation environment. Multipath propagation results in spreading of signal over angle of arrival, time and frequency. The random fluctuation in signal level known as fading is caused by the destructive addition of multipath components. Correspondingly, angle spread, delay spread and Doppler spread are obtained. Moreover, fading, delay spread, Doppler spread, and angle spread are the main channel effects.

A wireless channel generally obeys large, medium and small scale fading. Large scale fading is a deterministic effect attributed to path loss. Medium scale effect is a random effect observed in spatial dimension when moved over several tens of wavelengths. Small scale fading is the random effect observed in temporal and spatial dimension. This can be categorized into slow and fast fading and flat and frequency selective fading.
2.4 DIVERSITY AND MULTIPLEXING

The quality and reliability of the wireless communication system is hence affected by multi path fading, noise and interference between the users. In order to overcome the effect of noise, matched filtering method to maximize the Signal to Noise Ratio (SNR) can be employed. To combat fading and interference, thereby to improve reliability, the diversity plays a major role. Diversity is the process of transmitting several replicas of information bearing signal simultaneously over independent fading channel. Thus, there is a good likelihood, that atleast one of the received signals will not be severely degraded by channel fading. There are several forms of diversity such as time diversity, frequency diversity and spatial diversity. In time diversity, the same information bearing signal is transmitted at different time slots with interval between successive time slots being equal to or greater than coherence time of the channel. In frequency diversity, the information bearing signal is transmitted by means of several carriers that are sufficiently Frequency spacing is chosen to be equal to greater than the coherence bandwidth of channel. In space diversity, multiple transmit / receive antennas or both are used with the spacing between adjacent antennas being chosen so as to ensure independence of possible fading events occurring in the channel. Recently, the use of spatial diversity through multiple transmit antennas and receive antennas becomes popular, due to its increased spectral efficiency. Hence, the use of multiple antennas at transmitter and receiver in wireless system, commonly known as MIMO technology has rapidly gained popularity over the past decades, because of its powerful performance enhancing capabilities and its tremendous potential in addressing the limited spectrum resource and the system performance problems.

Under suitable fading conditions, MIMO channel provides an additional spatial dimension for communication called spatial multiplexing.
The concept of spatial multiplexing is that the data source at transmitter is divided into several parallel sub data streams, where each sub stream can be modulated and coded independently, and transmitted simultaneously onto the MIMO channel. Thus, spatial multiplexing leads to an increase in capacity. The spatial multiplexing transmission offers a linear increase in transmission rate for the same bandwidth and with no additional power expenditure (Paulraj et al. 2003). The capacity of such MIMO channel with N transmit and receive antennas is proportional to N. The spatial multiplexing configuration can also be applied in a multiuser format. This allows a capacity increase proportional to the number of transmit antennas and the number of users.

Other major benefits of MIMO technology are related to achieving the performance gains such as spatial diversity gain, spatial multiplexing gain, interference reduction and array gain.

- Diversity gain mitigates random fluctuation (or fading) of radio signals, and can be made available at the receiver, the transmitter or both (Alamouti 1998 and Tarokh et al 1998, 1999 & Lozano & Zindal 2010);

- Spatial multiplexing offers a linear increase in capacity at no additional expenditure of power or bandwidth, under conducive channel conditions, such as rich scattering (Foschini &Gans1998, Telatar 1999 & Legnain 2013).

- Interference reduction algorithms cancel or reduce the co-channel interference (CCI) by differentiating between spatial signatures of the desired and CCI signals (Paulraj & Papadias 1997).
• Array gain results from coherent combining, increases average receive SNR, and can be achieved through processing at the transmit and receive side (Andersen 2000).

It is not possible to employ all the above mentioned benefits simultaneously. However, using some combination of benefits will result in improved reliability and capacity. The basic idea is to supply the receiver with multiple independently faded replicas of the same information symbol, so that the probability that all the signal components fade simultaneously is reduced.

2.5 MIMO SYSTEM MODEL

For a general MIMO system with $N_r$ receive antennas and $N_t$ transmit antennas, the maximum diversity order that can be achieved is $D = N_t \times N_r$ where the channel between each transmit-receive antenna pair is assumed to fade independently. An MIMO system is viewed as a single user point to point communication system. The special case with $N_r = N_t = 1$ is called a Single Input Single Output (SISO) system. When $N_t = 1 \& N_r \geq 2$, the system is called as Single Input and Multi Output (SIMO) system. Finally if $N_r = 1 \& N_t \geq 2$, then the system is called as Multi Input and Single Output (MISO).

![Figure 2.1 MIMO System Model](image)
2.5.1 Single user – MIMO

Wireless communication system with $N_t$ transmit antennas at the transmitter and $N_r$ receive antennas at the receiver sides with single user in the system is considered. In a flat fading channel, each signal path is represented by random complex fading coefficient. Hence, the MIMO channel is described by the matrix $H$ and its $(i,j)^{th}$ element $h_{ij}$ denotes the channel gain from the transmit antenna $i$ to receive antenna $j$ where, $i = 1, 2, \ldots N_t, j = 1, 2, \ldots N_r$.

$$H = \begin{bmatrix}
    h_{11} & \cdots & h_{1N_t} \\
    \vdots & \ddots & \vdots \\
    h_{N_t1} & \cdots & h_{N_tN_t}
\end{bmatrix} \quad (2.1)$$

The random channel gains are modelled by circularly symmetric complex Gaussian random variables as mentioned in the books (Proakis 2001, and Rappaport 1996), denoted as $h_{ij} \sim CN(m, 1)$. If the mean of channel gain $m$ is non-zero, the channel is said to undergo Rician fading. If $m = 0$, the channel undergoes Rayleigh fading. Throughout this thesis, the channel is assumed to be Rayleigh fading. These spatial elements can have different statistical parameters. Their statistics characterize antenna correlation, channel mean and temporal auto-correlations.

In 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

$$y = Hs + n \quad (2.2)$$

where $s$ is the transmitted signal vector of size $N_t \times 1$ and $n$ is the Gaussian noise vector with i.i.d entries of zero mean variance $\sigma^2$. 
2.5.1.1 Spatial correlation

Spatial channel information, either instantaneous or statistical, can bring significant improvement in system performance, such as increasing the transmission rate and enhancing reliability.

A Rayleigh fading MIMO channel is said to be spatially white if 
\[ E[h_i h_m^H] = 0, \quad i, m = 1, 2, \ldots, N_t, \quad j, n = 1, 2, \ldots, N_r, \quad i \neq m \text{ and } j \neq n, \]
and is denoted by \( H_w \). In many practical applications, the transmit and/or the receive antennas can be correlated. For example, in cellular wireless communication environment, base station antennas are often situated at high elevation. So, they are typically unobstructed and have no local scatterers. Thus, channels arising from the transmit antennas during downlink are correlated. As a result, the MIMO channel entries do not fade independently. The spatial diversity gain, one of the MIMO channel parameters, depends on the antenna correlation. If the antennas are highly correlated, spatial diversity gain is small and vice-versa if the antennas are uncorrelated.

The channel correlation matrix captures the spatial correlation among all transmit and receive antennas. It is defined among all \( N_t \times N_r \) channel elements as a \( N_t \times N_r \) matrix \( R_0 = E[hh^H] \), where \( h = \text{vec}(H) \) and \( (\cdot)^H \) is the conjugate transpose. \( R_0 \) is the positive semi definite Hermitian matrix. The correlation matrix is assumed to have a simplified kronecker structure explained in the literature (Shiu 2000). The kronecker model assumes that the correlation of the scalar channels seen from \( N_t \) transmit antennas to a single receive antenna i.e corresponding to single row of \( H \), is the same for any receive antenna, i.e., any row of \( H \) and equals to \( R_t \) \( (N_t \times N_t) \). Let \( h_i^H \) be row ‘i’ of the channel matrix \( H \), then \( R_t = E[h_i h_i^H] \) for any \( i \). If \( \phi_{ij} \) is the correlation coefficient between \( i^{th} \) and \( j^{th} \) transmit antennas, then the transmit correlation matrix \( R_t \) is written as
\[
R_t = \begin{pmatrix}
1 & \phi_{11} & \cdots & \phi_{1N_t} \\
\phi_{11} & 1 & \cdots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0_{N_t1} & \cdots & 0_{N_tN_t}
\end{pmatrix}
\] (2.3)

Also, covariance of the scalar channels seen from the single transmit antenna to all receive antennas i.e corresponding to single column of \( \mathbf{H} \) is assumed to be same for any transmit antenna, i.e., any column of \( \mathbf{H} \) and equals to \( R_r (N_r \times N_r) \). Let \( h_j^H \) be column \( j \)' of the channel matrix \( \mathbf{H} \), then \( R_r = \mathbb{E}[h_j h_j^H] \) for any \( j \). Receive covariance matrix \( R_r \) is also defined similarly with receive correlation coefficients.

However, for an urban transmission environment, the exponential model has been used in the literature (Loyka 2001, Chizhik et al. 2003 and Choi &Murch2004) for transmit and receive correlation, in which \( (j,i)^{th} \) element of \( R_t \) are given by

\[
R_t(j, i) = \phi_t^{||j-i||}
\] (2.4)

for \( i, j = 1, 2, \ldots N_t \), where \( \phi_t \) is the correlation coefficient between two neighbouring antennas. The receive correlation matrix is similarly defined with \( \phi_t \) replaced with \( \phi_r \) and with indices ranging from \( i, j = 1, 2, \ldots N_r \). The exponential model is very simple because the correlation matrix is controlled by one parameter. The exponential model is useful for uniform planar arrays scenarios. In this work, this exponential correlation model is used. The covariance matrices \( R_t \) and \( R_r \) are complex positive semi definite Hermitian matrices. In the exponential model, the channel correlation matrix \( R_o \) is now decomposed as

\[
R_o = R_t^\dagger \otimes R_r
\] (2.5)

In the literature (Yu et al. 2001, Kermoal et al. 2002 and Bliss et al. 2004), the kronecker model has been experimentally verified with $3 \times 3$ antenna configuration in the indoor environment and with $8 \times 8$ antenna configuration in the outdoor environment respectively. The channel in Equation (2.1) is now written as

$$H = R_{r}^{1/2} H_{w} R_{t}^{1/2}$$ (2.6)

### 2.5.1.2 Temporal correlation

In addition to the spatial correlation, MIMO channel also exhibits temporal correlation. The channel is assumed to be Wide Sense Stationary (WSS), i.e., the temporal channel auto correlation depends only on time difference. Variation of the channel can be modelled by the correlation coefficient $\rho$ between two channel coefficients at two different time periods and is given by,

$$\rho = E[h_{ij}(n)h_{ij}(n + 1)]$$ (2.7)

where $h_{ij}(n)$ is channel coefficient between the transmit antenna $i$ to the receive antenna $j$ at the time instant $n$. It is also assumed that the fading corresponding to different obstacles is uncorrelated and this type of channel is described as wide sense stationary uncorrelated scattering (WSSUS). The temporal correlation model is based on sum of sinusoids of Jakes model and thus the theoretical power spectral density of the received signal has a well known $U$ shaped band limited spectrum (Jakes 1994)

$$S(F) = \frac{1}{2\pi f_D \left|1 - \left(\frac{F}{f_D}\right)^2\right|}$$ (2.8)
where $f_D$ is the maximum Doppler frequency in Hz. Using the Jakes model, the temporal magnitude correlation of two channel realization separated $\Delta$ time apart is given by $J_0(2\pi f_D \Delta)$, where $J_0(.)$ is the zeroth order Bessel function of first kind. The channel estimates available at the base station of a cellular system are not only usually noisy but may also be outdated due to feedback delay, velocity of the mobile user or duplex time. Since the channel gain $h_{ij}(n)$ and $h_{ij}(n + 1)$ are zero mean circularly symmetric complex joint Gaussian random variables, the relationship between $h_{ij}(n + 1)$ and $h_{ij}(n)$ is modeled as

$$h_{ij}(n + 1) = \rho h_{ij}(n) + \sqrt{1 - \rho^2} w$$  \hspace{1cm} (2.9)

where $w$ is auxiliary zero-mean unit-variance circularly symmetric complex joint Gaussian random variable independent of $h_{ij}(n)$ and $\rho$ is the correlation coefficient. Moreover, the correlation between two subsequent random variables is equal to $\rho$.

Proof: $E[h_{ij}(n)h_{ij}(n + 1)^H] = E[h_{ij}(n)\rho h_{ij}(n)^H + h_{ij}(n)\sqrt{1 - \rho^2}w^H]$  

$$= \rho E[h_{ij}(n)h_{ij}(n)^H] + \sqrt{1 - \rho^2}E[h_{ij}(n)w^H] = \rho $$  \hspace{1cm} (2.10)

Furthermore, $h_{ij}(n + 1)$ is a random variable with the same distribution as $h_{ij}(n)$ and $w$.

### 2.5.2 Multiuser MIMO (MU-MIMO)

MU-MIMO systems have drawn wide attention in recent times because of the potential to achieve higher multiplexing gains. MU-MIMO receivers are significantly more complex than SU-MIMO systems, since the signals from all users must be detected simultaneously. In multiuser cellular
communication systems, there are two scenarios. The first case considers multiuser communication, in particular on the downlink of a cellular communication system, in which the base station with multiple antennas \( \left( N_t \right) \) serves the \( K \) users (mobile stations) in the corresponding cell and sends information to number of mobile terminals in that area. Each mobile terminal is equipped with a single antenna.

![Figure 2.2 MU- MIMO System Model](image)

The MIMO channel matrix of size \( K \times N_t \) is now described as

\[
H = \begin{bmatrix}
h_{11} & \cdots & h_{1K} \\
\vdots & \ddots & \vdots \\
h_{K1} & \cdots & h_{KN_t}
\end{bmatrix}
\]  

(2.11)

The second scenario considers the downlink cellular communication channel with \( N_t \) transmit antennas at the base station. There are \( K \) users in the system with each user equipped with \( N_i \) antennas, \( i = 1, 2, \ldots K \). The total number of receive antennas is \( N_r = \sum_{i=1}^{K} N_i \). From the system model, the combined channel matrix \( H \) of size \( (N_r \times N_t) \) is given by

\[
H = [H_1^H, H_2^H, \ldots, H_K^H]^H
\]  

(2.12)
where $H_i$ ($N_t \times N_t$) is the $i^{th}$ mobile user’s channel matrix. For both the cases, the spatial correlation matrices are described as in section 2.5.1.1

2.6 CHANNEL KNOWLEDGE AT THE TRANSMITTER

Generally, the knowledge about the radio channel, also called as channel state information, can be used at the transmitter side to adapt the transmitting signal to the radio channel and at the receiver side to decode the received signal. The channel knowledge at the receiver is crucial for the whole transmission as this knowledge is essential for proper detection of data symbols. Because of noise and channel variation in time, the receiver does not know the radio channel perfectly and hence reduces the channel capacity. However, if the channel matrix $H$ is known only at the receiver, the transmitter will treat all the transmitted signals in the same way and will allocate equal power to all of them. However, this strategy will be ineffective for many cases because some data signals are heavily attenuated during transmission. In space time coded or spatially multiplexed system, no CSI is needed at the transmitter. Hence, those schemes are not optimized over instantaneous channel realizations. They assume an i.i.d matrix channel for good system performance. But in practice, the line of sight paths from the transmitter to the receiver and transmit and receive antenna correlation gives a MIMO channel matrix which is not i.i.d.

2.6.1 System Capacity and Duality

The basic measure of performance of a wireless communication system is the capacity of the channel. The channel capacity was pioneered by Claude Shannon (1948, 1949), using a mathematical theory of communication. The capacity of a channel, denoted by $C$, is the maximum rate at which the small probability of error can be achieved without any constraints on transmitter and receiver complexity.
The channel capacity of a time invariant Additive White Gaussian Noise (AWGN) channel with bandwidth $B$ and SNR $\left(\frac{P}{N}\right)$, where $P$ is the signal power and $N$ is the noise power, is commonly defined as

$$C = B \log_2 (1 + \frac{P}{N}) \text{bps}. \quad (2.13)$$

For time varying channels, when the instantaneous channel gain, (also known as channel state information (CSI)), are known perfectly at the transmitter and the receiver, the transmitter can adapt its transmission strategy relative to the instantaneous channel state; and Shannon capacity for this case is the maximum mutual information averaged over all channel states. When the channel state is varying quickly, the performance metric is Ergodic capacity. With CSI at the transmitter (CSIT), Ergodic capacity can be achieved using an adaptive transmission policy where the power and data rate vary relative to the channel state variations as referred in the literature (Goldsmith & Varaiya 1997). An alternate capacity definition for time varying channels with perfect CSI at transmitter and receiver is outage capacity. This outage capacity is the appropriate capacity metric in slowly varying channels.

The channel capacity of a point-to-point MIMO channel is a real number which is the fundamental limit on reliable communication: any rate strictly smaller than the capacity is achievable, while all rates strictly larger than the capacity are not achievable. For multi-user channels, the channel capacity has a similar definition, but the capacity is a region (i.e. a set in $K$-dimensional space) instead of a single number because different rates are associated with the multiple users. The channel capacity for single user MIMO system with the received signal obtained in the Equation (2.2) is expressed as
\[ C = \max_{\text{tr}(R_{ss})= \rho_t} \log_2 \det \left( I + \frac{E_s}{N_0} H R_{ss} H^H \right) \]  

(2.14)

where \( R_{ss} = E[ss^H] \), the autocorrelation of the transmitted signal \( s \), \( E_s \) is the total average transmitted energy and \( N_0 \) is the noise power. The capacity expressed in (2.14) is dependent on CSI at the transmitter side. If CSI is not known at the transmit side and if only the statistics of \( H \) are known, the transmit energy is equally distributed to all the transmit antennas, that is \( R_{ss} = I_{N_t} \) and the channel capacity for this case is given as

\[ C = \log_2 \det \left( I + \frac{E_s}{N_0} H H^H \right) \]  

(2.15)

The capacity for the MU-MIMO channels will be defined separately for Downlink (DL) / Broadcast channel (BC) and Uplink (UL) / Multiple Access Channel (MAC). In MAC, each user is assumed to have an independent message for the base-station, and thus a different rate is associated with each user. In BC, the transmitter is assumed to have a different (and independent) message for each of the users and similarly a different rate is associated with each transmission. The capacity region is therefore defined as the set of rates that can simultaneously be achieved with an arbitrarily small probability of error.

The capacity region of two user single antenna AWGN MAC, with \( N_t = N_r = 1 \) and \( K > 1 \), is given by

\[ R_1 \leq \log(1 + |h_1|^2 P_1) \]
\[ R_2 \leq \log(1 + |h_2|^2 P_2) \]
\[ R_1 + R_2 \leq \log(1 + |h_1|^2 P_1 + |h_2|^2 P_2) \]  

(2.16)

where \( P_1, P_2 \) and \( R_1, R_2 \) are the transmitting powers and sumrate of user1 and user 2 respectively. The capacity region clearly corresponds to a pentagonal
region by considering $R_1$ and $R_2$ as corner points. For the case, $N_t > 1, N_t = 1 \& K > 1$, the sum rate capacity is proportional to $\min (K, N_t)$, and is similar to SU-MIMO systems.

$$C_{MAC} = \log_2 \det \left( I + \frac{P}{K} H H^H \right)$$

(2.17)

where each antenna transmits with power $\frac{P}{K}$. For the case, $N_t > 1, N_t > 1 \& K > 1$, suppose a user $k$ transmits with a fixed covariance $Q_k$, then

$$Q_k = \mathbb{E} [x_k x_k^H]$$

(2.18)

The sum rate of $K$ users is given in the literature (Huang et al 2012),

$$C_{MAC} = \log_2 \det \left( I + \sum_{j=1}^{K} \frac{1}{N_0} H_j Q_j H_j^H \right)$$

(2.19)

For maximization of the sum rate in Equation (2.19), each user needs to determine its optimum covariance by using convex optimization techniques, which requires the information of all other users’ channel as referred in the literature (Viswanathan et al. 2003). Unlike multiple-access channel, a general expression for capacity region of broadcast channel is unknown. Yet, by using DPC strategy proposed by Costa (1983), a duality between BC and MAC has been established, and has been proved in the literature (Caire & Shamai 2003, Vishwanath & Tse 2003). Vishwanath et al. (2003) have shown that the dirty paper region is exactly equal to the capacity region of dual MIMO MAC, with the transmitters having same sum power constraint as MIMO BC, which is given by

$$C_{BC} = \mathbb{U}_{\sum_{i=1}^{K} P_i = P} G_{MAC}(\{H_i\}|\{P_i\})$$

(2.20)
Thus, CSIT helps to improve the transmission rate, to enhance coverage and to reduce the receiver complexity in MIMO wireless systems.

## 2.6.2 Channel Estimation in FDD and TDD Systems

CSI available at the transmitter side is known as CSIT and at the receiver side is known as CSIR. At the transmitter side, since the signal to be transmitted enters the channel after leaving from the transmitter, the CSIT is acquired indirectly. The receiver, however, can estimate the channel directly from channel modified received signal. Pilot signals are usually inserted into transmitted signal to facilitate channel estimation at the receiver. The CSIT can indirectly be obtained either by using the reciprocity principle or by using the feedback from receiver.

**Figure 2.3 Channel Estimation TDD Systems**

The reciprocity principle of wireless communication is defined as, the channel from antenna 1 to antenna 2 is identical to the transpose of channel from 2 to 1. This reciprocity will be in force if both the forward and the reverse links occur at same frequency, same time and at the same locations. In practical systems, the reciprocity principle will hold approximately in some situations. This is possible if the difference in any of these dimensions (time, frequency, space) is relatively smaller compared to the channel variation across the referred dimension.
In TDD systems, since the forward and reverse links are having identical frequency band and antennas, reciprocity principle is used for practical channel acquisition. The time lag between the forward and the reverse link of the TDD system should be smaller as compared to the channel coherence time. Further, due to necessity of good calibration for measurement of the channel, reciprocity is very difficult to accomplish.

The reciprocity principle is not applicable in Frequency Division Duplexing (FDD) system, where separate frequency band is allocated to each user for forward and reverse links, because the frequency separation between these links is much larger than coherence bandwidth. Hence, a CSIT can be obtained using feedback from receiver of the forward link. The channel is measured in receiver at antenna 2 during the forward link (1 to 2) transmission, and the information is sent to transmitter at antenna 1 on the reverse-link. Feedback is not limited by reciprocity requirements. Same principles are applicable in multi user communication system also. However, system overhead due to the feedback channel will be a major drawback in this system.

2.7 LINEAR PRECODER.

The advantage of MIMO system gets enhanced, if the precoding techniques, which reduce performance loss, caused by interference and fading, are applied at data symbols before they are transmitted. Precoding is a pre-processing technique at the base station which exploits CSIT to match
transmission to the instantaneous channel conditions. It is well known that BS is readily available with powerful computing ability and power supply is not a major issue at the transmit side. Hence, the receiver based interference cancellation techniques and the detection techniques can be transferred to the base station. Thus the structure of mobile unit is simplified and significant amount of power consumption is obtained which is useful during the mobility of mobile unit. The precoding techniques can either be linear or non-linear.

Popular non-linear precoding scheme namely Tomlinson Harashima precoding (THP), originally proposed by Tomlinson (1971) and Harashima & Miyakawa (1972) for dispersive SISO systems, can also be applied to MIMO systems in the literature (Ginis & Cioffi 2000 and Fischer et al 2002). The famous THP is the non-linear precoding based on DPC theory. Other TH precoding techniques have been developed in the literature (Stankovic & Haardt 2005, Miao et al. 2008 and Takeda et al. 2008) to combat multiuser co-channel interference. THP is similar to decision feedback equalization (DFE), but contrary to DFE, THP feeds back the already transmitted symbols to reduce the interference caused by these symbols at the receivers. Although THP is based on the application of nonlinear modulo operators at the receivers and the transmitter, similar optimizations as for linear transmit filters can be used to find the THP filters. THP, with partial channel knowledge, channel estimation error and delay is considered in the literature (Athanasios 2005). The statistical model for channel time variation and estimation errors are used to design the robust TH precoder with new cost function by adopting a Bayesian approach. At low SNR and slow channel time variation, robust TH precoder outperforms the MMSE-DFE; however at high SNR MMSE-DFE outperforms the robust THP.

Generally, linear precoding schemes require low implementation schemes at the expense of performance as compared to non-linear schemes.
Linear precoding is a simple and efficient method to utilize CSIT. The linear precoding strategies are used to simplify the MIMO receiver. Linear precoders include zero-forcing (ZF), matched filtering (MF), Wiener filtering, and Regularized zero-forcing (RZF). Some other design parameters considered for the design of linear precoders are sum rate capacity, Pairwise Error Probability (PEP) and Minimum Bit Error Rate (MBER).

2.7.1 Zero-Forcing Linear Precoding

The most common linear precoding scheme is zero-forcing scheme. ZF is a simple method which decouples the multiuser channel into multiple independent sub channels and reduces the design to a power allocation problem. ZF precoding often involves channel inversion analyzed by Haustein et al. (2002) using pseudo-inverse of the channel or other generalized inverses mentioned in the literature (Joham et al. 2005). The ZF precoder completely eliminates the interference at the receiver. There are several studies on ZF precoding in the literature (Vojcic& Jang 1998, Dimic & Sidiropoulos 2005, Yoo & Goldsmith 2005, Boccardi et al. 2006 and Bartolome & Perez Neira 2006) to focus on different design criteria. The two common design criteria are throughput and fairness under a total (or per antenna) power constraint (Wiesel et al. 2008).

To design the zero forcing precoder the following optimization problem is solved

\[ T = \arg\min_T E[|\|Tx\|^2|] \]  

subject to \( HT = I \)

The resulting ZF transmit filter is given by

\[ T_{ZF} = \mathbf{H}\mathbf{H}^{-1} \]
It is shown in the literature (Hochwald et al. 2005) that the plain, direct channel inversion performs poorly due to singular value spread of the channel matrix at all SNRs and for all number of users.

### 2.7.2 Regularized Zero-Forcing Precoding

In this technique, the channel inversion is regularized by adding scaled identity matrix before the inverse is taken. The Regularized Zero-forcing filter is now given as

\[ T_{RFZ} = H H^H (\beta I + H H^H)^{-1} \]  \hspace{1cm} (2.23)

where \( \beta \) is the regularization factor which is obtained to maximize SINR at each receiver. Hence, the performance is improved at low SNRs.

### 2.7.3 Matched Filter Precoding

Transmit matched filter (TxMF) was introduced by Esmailzadeh et al. (1993) by moving channel matched filter from the receiver to the transmitter, but they gave only an intuitive explanation. Transmit matched filter maximizes Signal to Interference Ratio (SIR) at the receiver and is optimum for high SNR region. TxMF is derived in the literature (Joham et al. 2005) by maximizing the ratio between the power of the desired signal portion in the received signal and the signal power under the transmit power constraint.

The transmit matched filter is given as

\[ T_{MF} = \beta H^H \]  \hspace{1cm} (2.24)

with \( \beta = \frac{1}{\sqrt{\text{tr}(H^H H)}} \)
2.7.4 Wiener Filter Precoding

It is shown in the literature (Joham et al. 2005) that the performance of matched filter precoding is worse than the zero-forcing precoding for high SNR, but outperforms $T_{ZF}$ for low SNR. Transmit Wiener filter is designed to minimize weighted MSE function and it is shown that the transmit Wiener filter finds optimum trade off between signal maximization of the matched filter precoding and the interference elimination of zero forcing precoding.

$$T_{WF} = \beta F^{-1}H^H$$

(2.25)

with $F = (HH^H + \frac{N}{\rho} I)$ and $\beta = \frac{1}{\sqrt{\text{tr}(\beta^2HH^H)}}$ where $\beta$ can be interpreted as optimum gain for the combined precoder and channel.

Block Diagonalization (BD) based precoding techniques, which are generalization of the zero forcing precoding, have been proposed in the literature (Spencer et al. 2004 and Choi & Murch 2004), for MU-MIMO systems. BD methods decompose a MU-MIMO channel into multiple SU-MIMO channels in parallel to completely eliminate Multi User Interference (MUI) with noise by making use of null space.

When BD is employed, precoding matrices $T_i \forall i$ are chosen such that $H_k T_i = 0 \forall k \neq i$, thus eliminating the MUI. This requires to determine left null space of the matrix formed by stacking all $H_i \forall i \neq k$, matrices together. Defining $\overline{H}_k$ as

$$\overline{H}_k = [H_j^T \ldots H_{k-1}^T H_{k+1}^T H_k^T]^T$$

(2.26)
then any suitable \( T_k \) lies in the null space of \( \overline{H}_k \). With BD, each users precoding matrix lie in the null space of all other users channel. Let the Singular Value Decomposition (SVD) of \( \overline{H}_k \) be

\[
\overline{H}_k = \overline{U}_k \overline{D}_k \begin{bmatrix} \overline{V}_k^{(1)} & \overline{V}_k^{(0)} \end{bmatrix}^H
\]

(2.27)

where \( \overline{U}_k \) and \( \overline{D}_k \) are left singular vector matrix and matrix containing the singular values of \( \overline{H}_k \) respectively, and \( \overline{V}_k^{(1)} \) and \( \overline{V}_k^{(0)} \) denote the right singular matrices, each corresponding to non-singular values and zero singular values respectively. Thus, \( \overline{V}_k^{(0)} \) forms the orthogonal basis for null space of \( \overline{H}_k \) and hence the MUI is completely eliminated in BD. Hence, it performs better than other channel inversion techniques. However, BD based precoding algorithm considers only the elimination of MUI, and suffers a performance loss at low SNR, where the noise is more dominant. Therefore, the null space vectors are normalized using regularization factor to avoid noise amplification problem effectively.

### 2.7.5 Precoder for Sum Rate Maximization

It is quite apparent that mutual information between the channel input and channel output can be obtained as

\[
I(X; Y) = \log_2 \det \left( I + \frac{E_s}{N_0}(TH)R_{ss}(TH)^H \right)
\]

(2.28)

where \( T \) is the precoding matrix. Ergodic capacity and optimal signaling are obtained by maximizing this mutual information subject to the transmit power constraint and is given as

\[
\max_{\text{tr}(TT^H)=1} \log_2 \det \left( I + \frac{E_s}{N_0}(TH)R_{ss}(TH)^H \right)
\]

(2.29)
Optimum precoding matrix $\mathbf{T}$ is designed to maximize the ergodic capacity. For MU-MIMO case, it is already described in the literature (Vishwanah et al. 2003 and Vishwanath & Tse 2003) that, information theoretic approach characterizes the downlink sum capacity by solving the sum capacity of the equivalent uplink MAC and applying the duality result. However, under single user precoding, each user treats the interference as a noise, and user $k$ can achieve the rate

$$R_k = \log\frac{\det\left(\sum_{j=1}^{K} \mathbf{H}_k^H \mathbf{Q}_j \mathbf{H}_k + \mathbf{N}_0 I\right)}{\det\left(\sum_{j=k}^{K} \mathbf{H}_k^H \mathbf{Q}_j \mathbf{H}_k + \mathbf{N}_0 I\right)}$$

(2.30)

where $\mathbf{Q}_j$ is the transmit covariance matrix $\mathbb{E}\left[\mathbf{T}s_j(\mathbf{T}s_j)^H\right]$ and $\mathbf{s}_j$ is the data vector intended for user $j$.

It is practically observed that the demand of downlink data transfer is much higher than that of uplink. However, it is often harder to find the optimum transmission strategy for the downlink. It has been shown that the capacity region of Gaussian BC is equal to the capacity region of dual Gaussian MAC subject to the same sum power constraint. In the literature (Vishwanath et al. 2003 and Weingarten et al. 2006), it has been proved that the capacity region of Gaussian MIMO MAC and BC is identical under the same sum power constraint. This concept of having similarities between MAC and BC as well as their capacities and optimal transmission strategies, is known as duality. Duality greatly simplifies numerical computation of the multiple broadcast channel, sum rate capacity and achievable regions.

### 2.7.6 Pairwise Error Probability (PEP) Precoder

The probability of error, averaged over the channel fading is a common system performance measure. This type of precoding design minimizes target error probability and also depends on the modulation
detection rule. PEP is normally obtained using per-distance criterion proposed in the literature (Tarokh et al. 1998) and average PEP criterion proposed in the literature (Zheng & Tse 2002). If $S(H_k)$ denotes the transmit symbol vector corresponding to the hypothesis $H_k$, the PEP is defined as error probability of choosing in favor of $S(H_l)$ instead of $S(H_k)$. Assuming ideal CSI, PEP is well approximated by

$$P(S(H_k) \rightarrow S(H_l)) \leq e^{-\frac{-d_{\text{min}}^2}{2}}$$ (2.31)

where $d_{\text{min}}^2 = \min_{k,l} \left| [S(H_k) - S(H_l)]^H (TH)^H R_{nn}^{-1} (TH) [S(H_k) - S(H_l)] \right|$, where $R_{nn}$ is the noise covariance, $T$ and $H$ are the precoding matrix and channel matrix respectively. The minimum distance is lower bounded as mentioned in the book (Lutkepohl 1996)

$$d_{\text{min}}^2 = \lambda_{\text{min}} \left( (TH)^H R_{nn}^{-1} (TH) \right) \min_{k,l} \| S(H_k) - S(H_l) \|^2$$ (2.32)

On substituting the Equation (2.32) into (2.31), PEP is now written as

$$P(S(H_k) \rightarrow S(H_l)) \leq e^{-\lambda_{\text{min}} \left( (TH)^H R_{nn}^{-1} (TH) \right) \min_{k,l} \| S(H_k) - S(H_l) \|^2}$$ (2.33)

The upper bound on PEP can be minimized if the term

$$\lambda_{\text{min}} \left( (TH)^H R_{nn}^{-1} (TH) \right)$$

is maximized.

In average PEP criterion, the optimization problem is obtained by averaging the PEP over the coding distribution and the fading statistics. Here, the average PEP is independent of the minimum distance ($d_{\text{min}}^2$)
2.7.7 Minimum Bit Error Rate (MBER) Precoder

The design of linear precoder based on minimum MSE criteria is simple and mathematically easily tractable. However, for a communication system, the system’s BER is required. The MMSE criterion is used to minimize the MSE value. Although, MMSE precoders perform reasonably well in practice, MMSE criterion does not guarantee minimum BER (MBER). Hence, for MIMO systems, the MBER precoder for ZF equalization of block transmission, proposed by Ding et al. (2003), and block transceivers with MMSE decision feedback equalization proposed by Xu et al. (2005) are designed to minimize BER. The block transceiver design for MMSE-DFE in the literature (Xu et al. 2005) provides a closed-form solution to approximate MBER (AMBER) precoding for successive interference cancellation (SIC) receivers. In the literature (Wang & Blostein 2007), MBER is employed as an optimization criterion, and an Approximate MBER (AMBER) power allocation algorithm is proposed for a MIMO wireless communication system. In this precoder design, the authors analyze the special cases of noisy CSI and power feedback.

2.8 PRECODER DESIGN FOR IMPERFECT CSI

The linear precoding design criteria discussed in section 2.7 assumes that CSI available at the transmitter side is perfect. However, it is well known that the perfect CSI is impractical due to often inaccurate channel estimation and CSI feedback errors. Consequently, designing the linear precoders without taking into account CSI imperfection will result in performance degradation. Further, for the time varying channel, the channel knowledge at the transmitter side used for precoding is often outdated and should be updated whenever the channel changes. Designing optimal transmitters based on statistical information about underlying stationary random channels, is thus well motivated.
Consequently, it is necessary to design a system robust to imperfect CSIR. Practical situations indicate that some forms of partial CSIT can be available. For example, partial CSIT can be acquired by transferring the CSIR to the transmitter via a feedback link. CSIT obtained using feedback is always imperfect due to channel estimation errors, erroneous CSIR and limitation of the feedback link. The Precoders with Statistical information of CSI and partial CSI are discussed in the following subsection.

2.8.1 Mean Feedback and Covariance Feedback Precoders

Consequent upon using multiple antennas in MIMO wireless communication system, there will be a large number of channel state parameters. Hence, the feedback overhead could be large. As a result, CSI is being quantized and sent to the transmitter over the limited rate feedback channel. Otherwise, the channel statistics knowledge, namely, channel mean or covariance, which varies very less frequently, is fed back to the transmitter. The channel statistics knowledge can be acquired through field measurements, ray tracing simulation or using physical channel models (Shiu et al. 2000). By employing long term averaging of channel realizations, the channel correlation can be estimated and the same is fed back to the transmitter through a low rate feedback channel and is referred to as covariance feedback as referred in the literature (Jafar & Goldsmith 2001, Jafar et al. 2001 and Visotsky & Madhow 2001). Given CSIT and channel statistics, conditional mean of the actual channel gain matrix can be calculated and used for mean-feedback precoding.

On the other hand, since the channel covariance matrix is primarily determined by antenna correlation, which can readily be evaluated at the transmitter, the feedback requirement for covariance precoding is much smaller than for mean-feedback precoding. The precoder design using either channel mean or covariance is studied in the literature (Narula et al. 1998,
Jafar & Goldsmith 2001, and Visotsky & Madhow 2001) from channel capacity point of view.

In the literature (Jafar & Goldsmith 2001, Jafar et al. 2001 and Visotsky & Madhow 2001) the covariance based optimal precoder has been designed based on capacity criterion, which specifies the theoretical maximum rate of reliable communication achievable in the absence of delay and processing constraints. The spatial channel $\mathbf{h}$ is generally modeled as complex Gaussian random vector with non-zero mean $\bar{\mathbf{h}}$, and non-white covariance matrix $\mathbf{\Sigma}_h$; the mean feedback is achieved in slowly time varying channel, in which the knowledge of the channel mean is assumed to be known and the covariance matrix is modelled as white with $\mathbf{\Sigma}_h$ proportional to an identity matrix. The covariance feedback is motivated when the channel $\mathbf{h}$ varies too rapidly for the transmitter to track its mean. In this case, the channel mean is set to zero and non-white covariance matrix $\mathbf{\Sigma}_h$ is assumed while requirement for covariance precoding is much smaller than for mean-feedback precoding. Naturally, the quality of mean-feedback precoding will be degraded due to estimation errors, and, in general, it is more sensitive to channel time variations and feedback delay than covariance precoding. In contrast, covariance precoding may become less effective when the mean feedback is accurate.

For a CDMA system with partial CSI, Montalbano & Slock (1999) developed a covariance-based transmit filter aiming at the suppression of intra-cell interference. Forster et al. (2000) developed a similar transmit filter for Global System for Mobile Communications (GSM) in the frequency domain. By maximizing the minimum signal-to-interference-plus-noise ratio (SINR), Schubert & Boche (2004) proposed an optimal scheme which based on covariance knowledge as well. In the literature (BennoZerlin et al. 2006), the transmit matching filter, the transmit Wiener filter and the transmit zero
forcing filter are designed for DS-CDMA system, based on the knowledge of channel covariance.

SER bound optimal multi antenna transmit precoder has been designed based on channel correlation and channel mean feedback in the literature (Zhou & Giannakis 2002 and Zhou & Giannakis 2003). The spatial correlation in multiple antenna transmission schemes always degrades the BER / SER performance when the CSI is not available at the BS. There have been several studies on the impact of spatial correlation on Ergodic capacity with perfect CSI (Björnson & Ottersten 2010) and with various types of CSI in the literature (Narula et al. 1999, Visotsky & Madhow 2001, Ivrlac et al. 2003, Jorswieck & Boche 2004 and Björnson and Ottersten 2009). The effect of spatial correlation on channel capacity was derived and evaluated analytically in the literature (Ivrlac et al. 2003, Jorswieck & Boche 2004).

In the literature (Björnson & Ottersten 2010), orthogonal space time block code transmission and linear precoder to minimize SER, based on the statistical knowledge of the channel for the spatially correlated system is analyzed. Yu Fu et al. (2009) designed the linear and non-linear precoders to minimize BER for OSTBC MIMO OFDM downlink when the conditional mean of the channel gain matrix and the channel gain covariance matrix are available at the transmitter. These precoders are designed for the systems in which the receiver imperfectly estimates CSI and feeds back the inaccurate estimates of the CSI to the transmitter through the feedback channel, which causes delay. Hence, the authors proposed the precoders, which take into account the estimation errors, channel time variation due to feedback delay and offer optimal power allocation. Further, dual mode precoding is proposed to switch between the mean feedback precoding and the covariance feedback precoding in accordance with the channel conditions. Covariance based linear and non linear precoding techniques for MIMO OFDM system had been
proposed by Yu Fu et al. (2010) to mitigate effect of transmit antenna and path correlations.

2.8.2 Precoder with Cut-Off Rate Criterion

In numerous communication systems, Packet Error Rate (PER) is an indicator of the quality of system. For example, system with an ARQ protocol requires accurate and real throughput for the retransmission of erroneous packets. Hence, minimization of PER is directly related with the maximization of throughput. It is difficult to design a linear precoder to minimize PER directly. Hence, the precoder to maximize channel cut-off rate has been proposed by Rey (2010), which minimizes code word error probability over the ensemble of binary channel codes. The cut off rate has been first proposed by Wozencraft (1966) and Massey (1974) as a design parameter for the design of coded modulation with block and convolutional codes. Then, it has been used in the design of space time codes under the assumption of no CSI neither at the transmitter nor at the receiver. In the literature (Rey 2010), the linear precoding design has been proposed for MIMO OFDM coded system that is robust to channel estimation errors at the transmitter with cut-off rate maximization as an optimization criterion.

2.9 SUMMARY

This chapter describes the basic concept of MIMO wireless communication, multipath fading and the precoding concepts. The linear precoding concepts with perfect CSI and imperfect CSI are also described in this chapter. The spatial correlation and temporal correlation addressed in this chapter is incorporated in the imperfect CSI model used in chapter 3 and 4. The sum rate problem with imperfect CSI is addressed in chapter 5.