ON Semiconductor

Is Now

To learn more about onsemi™, please visit our website at www.onsemi.com

ON Semiconductor®

Introduction to MIMO

Originally published in Dec. 2013 under the company name Quantenna® Communications, Inc.

ON Semiconductor®

Introduction to MIMO

Abstract

This paper provides an introduction to MIMO as used in modern wireless communication systems. We introduce the various MIMO concepts and give a brief theoretical analysis of the performance benefits. We then evaluate how these benefits are achieved in real−world wireless LAN systems like 802.11n and 802.11ac.

Introduction

Traditional wireless systems use a single antenna at the transmitter and a single antenna at the receiver to communicate. Multiple−Input Multiple−Output (MIMO) on the other hand refers to any communication system where both the transmitter and the receiver have multiple antennas. Strictly speaking, MIMO refers to the channel that exists between such a transmitter and such a receiver, not to the transmitter or receiver itself. The channel has multiple "inputs" in the form of signals that are sent "into" the channel by the antennas at the transmitter and multiple "outputs" in the form of signals that are observed at the antennas of the receiver. Using the same logic, a system with a single transmit and a single receive antenna is also referred to as SISO (single−input, single−output).

With single−antenna systems, throughput can be increased by either using wider bandwidth for the transmission or by increasing the transmit power used for the transmission. Both bandwidth and power are precious and strictly regulated resources, so one is typically not free to increase them at will. The breakthrough in MIMO is the fact that it provides another "dimension" that can be used for increasing throughput. The concepts of MIMO as applied to wireless systems were introduced in [1], which may be one of the most significant developments in modern communications. What's more, the improvement offered by using MIMO concepts is achieved without an increase in transmit power or bandwidth. The origin of the improvement can be intuitively understood by imagining that the use of multiple antennas at transmitter and receiver creates multiple independent paths between transmitter and receiver that can be used for data transmission. Whether this happens or not depends not only on the presence of multiple antennas, but also on the topology of the environment. Typically, wireless environments with significant scattering create multi−path environments that meet the conditions needed to exploit the presence of multiple paths between transmitter and receiver. The power of MIMO is that it exploits the multi−path environments rather than trying to mitigate it as SISO systems do. As a side effect, MIMO performs better outside over−controlled environments like anechoic chambers and it is particularly well suited to indoor communications.

The use of MIMO systems can provide benefits in a number of different ways. When transmitter and/or receiver have multiple antennas, the following techniques can be used to improve either the quality or the throughput of the connection:

1. Diversity
   When using diversity, the same data is sent through multiple independent paths between transmitter and receiver. This improves the reliability of the communication, but does not increase the data rate.
2. Spatial Multiplexing
   When employing spatial multiplexing, the transmitter sends independent data over each of the available paths. This results in an increase in data rate without an increase in either bandwidth or transmit power. In ideal circumstances, the data rate will increase with a factor proportional to the number of transmit or receiver antennas.
3. Transmit Beamforming (Tx BF)
   Transmit beamforming assumes the transmitter has some knowledge of the channel. This knowledge can be obtained by the transmitter itself or through feedback provided by the receiver. Suitable allocation of transmit signals to the various transmit antennas based on channel information allows the transmitter to steer energy more accurately in the direction of the intended receiver. This process is called Transmit Beamforming.
4. Multi−User MIMO (MU−MIMO)
   When doing MU−MIMO, the independent spatial paths are also used to send independent data, but the data is destined for different users. This allows a transmitter to simultaneously send data to more than one user and thus it is particularly effective when the channel to a single receiver does not allow the transmitter to exploit the maximum number of streams that it may use. Transmit signals are precoded to eliminate interference between these concurrently transmitted signals. Like Transmit Beamforming, channel information at the transmitter is required for MU−MIMO.

The advantages offered by MIMO are so compelling that they have become an integral part of various wireless standards, such as IEEE 802.11 (Wi−Fi®) and IEEE 802.16 (LTE).

MIMO Theory and Concepts

Modeling a MIMO Channel

Figure 1 illustrates a transmitter with four transmit antennas communicating with a receiver with four receive antennas. A signal sent on any of the transmit antennas will propagate through space and a version of that signal will be received by all receive antennas. In total, this creates $N_{RX} \times N_{TX}$ channels between the transmit antennas and the receive antennas. The transmit antennas are collocated at a single devices and are typically spaced about half a wavelength apart. The same applies at the receiver.

When talking about MIMO systems, the convention is to refer to these systems as "$N_{TX} \times N_{RX}$ systems", where the first number denotes the number of transmit antennas and the second number the number of receive antennas. For instance a 2x4 system means the transmitter has two antennas and the receiver has four antennas.

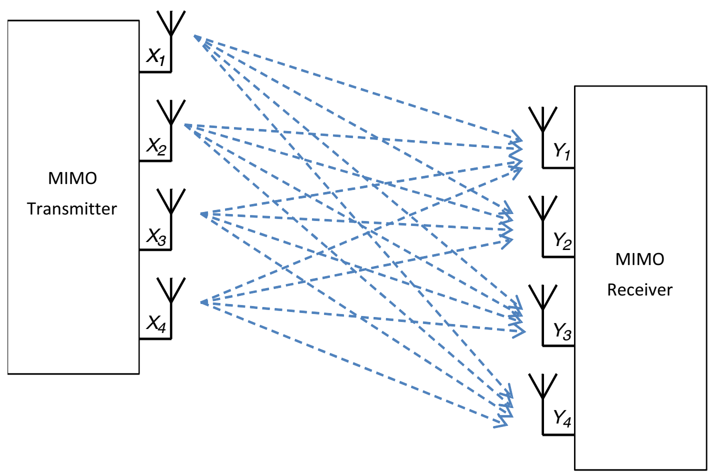

Figure 1. Illustration of MIMO System

As shown in Figure 1, a signal X_i is transmitted from transmit antenna *i*. The signal received at receive antenna *j* is denoted as Y_j. There exist a total of $N_{RX} \times N_{TX}$ paths (or channels) between the N_{TX} transmit antennas and the N_{RX} receive antennas. Each of these paths can be characterized by a value h_{ji}, which denotes the gain or loss that is experienced over the path between transmit antenna *i* and receive antenna *j*.

Any transmit signal X_i causes some signal to be received at receive antenna *j*. This received signal at receive antenna *j* will be proportional to X_i and can be written as $h_{ji}\,X_i$. The total received signal Y_j at antenna *j* will be the sum of the signals originating from all transmit antennas. This can be written as:

$$Y_j = \sum_{i=1}^{N_{TX}} h_{ji} X_i, \; j = 1, \ldots, N_{RX} \qquad \text{(eq. 1)}$$

The $N_{RX} \times N_{TX}$ channel coefficients h_{ji} can be conveniently grouped into a $N_{RX} \times N_{TX}$ matrix:

$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_{TX}} \\ h_{21} & \ddots & & h_{2N_{TX}} \\ \vdots & & \ddots & \vdots \\ h_{N_{RX}1} & h_{N_{RX}2} & \cdots & h_{N_{RX}N_{TX}} \end{bmatrix} \qquad \text{(eq. 2)}$$

With (eq. 2), expression (eq. 1) can be rewritten in matrix form as:

$$Y = HX \qquad \text{(eq. 3)}$$

Both the transmitted signal and the received signal are now denoted as vectors.

To complete the description, we add a term *N* that captures the amount of noise that is added at the receiver. Together, knowledge of the channel and the noise will allow us to make predictions about the achievable rates of a MIMO system. The MIMO transmission model in the presence of noise is given by:

$$Y = HX + N \qquad \text{(eq. 4)}$$

This is a convenient compact notation that allows further mathematical analysis of the properties of MIMO. In some ways, the expression is “too simple” because it treats each of the paths h_{ji} as a single value, which corresponds to a flat−fading channel with a single attenuation over the full bandwidth of the signal. Typically, the channel between the transmitter and the receiver will vary with both time and frequency, which would make the expressions (eq. 3) and (eq. 4) a function of both time and frequency.

However, the frequency dependence can be eliminated or ignored if we assume that the transmission bandwidth is sufficiently narrow. In that case, the channel can be approximated by a flat attenuation over the bandwidth and the assumption of a frequency−independent H_{ji} holds. For wider channel bandwidths with significant variation over the width of the channel, OFDM can be used to get around the frequency dependence. OFDM divides the full bandwidth up in a large number of independent narrow−band channels, which can be treated independently for modeling purposes. This is illustrated in Figure 2. Even though the assumption of flat fading clearly does not hold true over the entire bandwidth of the channel shown in Figure 2, it is approximately valid for each of the sub−channels of the OFDM transmission (illustrated by the gray bars in Figure 2). When using OFDM, an independent channel matrix *H* and noise vector *N* (see (eq. 4)) will be used to model each of the subchannels.

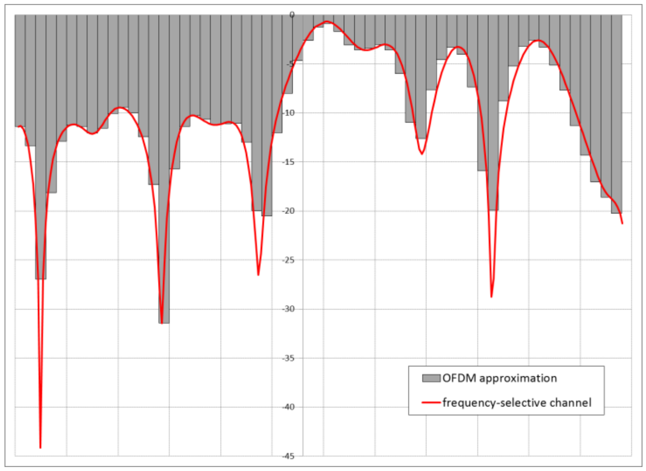

Figure 2. Frequency−Selective Channel Can Be Treated as a Series of Narrow−Band Flat−Fading Channels

Time dependence can be ignored if the channel is semi−stationary and the duration of any transmission is within the coherence time of the channel. In that case, the channel matrix may vary from one transmission to the next, but it can be assumed to be constant over the duration of a single transmission. Wireless protocols are designed to transmit short “bursts” of signal to stay within the coherence time of the channel.

This means that the expressions (eq. 3) and (eq. 4) can be used as−is, including for time and frequency−varying channels, provided that it is understood that they apply to narrow−band channels (which could be part of an OFDM signal) and that the channel does not vary appreciably over the duration of the transmission. The use of matrix notation provides a powerful tool for further analysis and many of the properties of MIMO transmission can be now explained in terms of matrix terminology.

Now that we have established an expression for MIMO transmission, we need a metric to quantify the advantages that MIMO brings. To this end, we’ll use the concept of channel capacity.

Channel Capacity

Channel capacity is a number associated with a communication system that predicts the maximum error−free rate that can be sent by that system over a given channel. Any transmission at or below the channel capacity can theoretically be sent at arbitrarily small bit error. However, it is not possible to send at a rate higher than the channel capacity without introducing bit errors. This result was established in 1948 by Claude Shannon [2]. In honor of Shannon, channel capacity is also referred to as Shannon capacity.

Shannon established that the channel capacity (*C*) is a function of SNR (Signal to Noise Ratio). He obtained the following mathematical expression for *C*:

$$C = \log_2(1 + SNR) \qquad \text{(eq. 5)}$$

In (eq. 5), C is expressed in bits/second/Hz. For a signal with bandwidth W, the total capacity in bits/second can be obtained by integrating (eq. 5) over the bandwidth of the signal.

For fading channels, a slightly modified form of (5) is used:

$$C = \log_2(1 + \rho\|h\|^2) \qquad \text{(eq. 6)}$$

In (eq. 6), ρ is the *average* SNR and $||h||$ h is the channel fading coefficient for a particular channel realization of a random channel. $\rho||h||^2$ is also called the instantaneous SNR. If the channel *h* is a fading channel that frequently changes over time, the average channel capacity can be obtained by calculating (eq. 6) for a large number of channel realizations and taking the average over that ensemble. This corresponds to modeling a system where the channel can be assumed to be constant for the duration of a single transmission, but can change between one transmission and the next. This formulation (eq. 6) is more appropriate for the modeling of wireless channels and is the expression for channel capacity that will be used in the remainder of this paper.

Figure 3 shows the shape of the channel capacity (eq. 6) as a function of ρ. This result is obtained for random complex−valued Gaussian channels. From this graph, one can see for instance that at 10 dB average SNR, a theoretical throughput of about 2.92 bits/sec/Hz is possible. To double the capacity to 5.84 bits/sec/Hz, an increase in average SNR of about 10 dB would be needed (from 10 to 20 dB). This is equivalent to increasing the transmit power with a factor of 10.

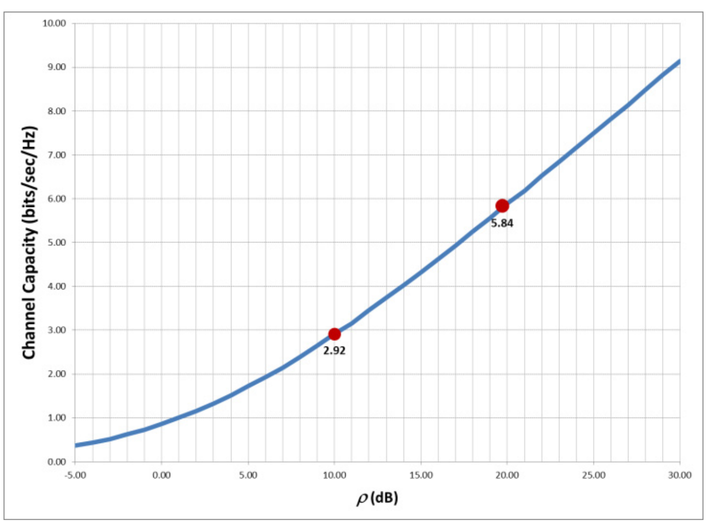

Figure 3. Channel Capacity as Function of Average SNR (see (6))

Shannon's channel capacity theorem does not say which coding or modulation should be used to achieve the predicted theoretical rate. It only establishes that it should be possible to design a system that can come arbitrarily close to achieving this capacity, but that it is impossible to exceed it, no matter the amount of processing and resources used by the system. In practice, systems using multi−level modulation and advanced coding such as LDPC can achieve the channel capacity to within a couple of dB. Even if real systems don't actually achieve the channel capacity, we can still use it as a way to compare the optimal performance that could be achieved by various MIMO and SISO configurations and to assess the relative gains that can be realized by various MIMO concepts.

The important thing to realize is that the formula (eq. 6) applies to SISO systems. For MIMO system, the existence of multiple paths will lead to important modifications that would seem the significantly violate the channel capacity predicted by (eq. 6) and Figure 3.

Advantages of MIMO: Theoretical Analysis

As mentioned earlier, MIMO can be used in a number of different ways. Specifically, each of the following approaches will benefit from MIMO concepts:

1. Diversity
2. Spatial multiplexing
3. Transmit Beamforming
4. Multi−user MIMO

In the following sub−sections, we'll illustrate the improvements that MIMO brings in each of these cases. We'll use channel capacity as a metric to gauge the level of improvement that is available in theory. The actual performance gain from using MIMO concepts depends to a large extent on the properties of the wireless channel. Optimal gains are achieved if the channel has a maximum number of truly independent paths (which explains why indoor environments with a lot of scattering and multipaths are well suited for MIMO, while in pure line−of−sight setups, MIMO show little benefit). In the following analysis, we will model the elements of the matrix channels as independent random complex variables following a Gaussian distribution. In a way, this gives us the best−case scenario for the illustration of the advantages of MIMO.

In section Performance of 802.11 MIMO systems , we'll also evaluate the improvements in a modeled wireless channel, which will be more representative of real−world performance.

Diversity

Diversity refers to situations where the same data is transmitted through multiple paths. This generally improves the quality of the reception, which leads to performance gain.

Under the general heading of “diversity”, we’ll distinguish between “receive diversity” and “transmit diversity”.

Let’s first consider the case of receiver diversity, which refers to the effect of additional receive antennas on system performance. To isolate the effects of receiver diversity, we consider a transmission system with a transmitter having a single antenna and a receiver having multiple antennas. Such a system is also referred to as “SIMO” (Single Input, Multiple Output). Note that “Input” and “Output” refer to the channel, not the equipment. A single signal is sent “into” the channel and N signals come “out of” the channel.

In SIMO, each of the receive antennas receives an independent copy of the transmit channel. At the receiver, all these signals can be combined into a stronger signal. As a result, the channel capacity of a SIMO system is given by:

$$C = \log_2\left(1 + \rho\left(\|h_1\|^2 + \ldots + \|h_{N_{RX}}\|^2\right)\right) \qquad \text{(eq. 7)}$$

Since the powers of uncorrelated signals can be added to find the total power, the average SNR at the receiver is simply the sum of SNRs of the individual paths. Assuming that all paths have roughly the same attenuation, the average SNR at the multi−antenna receiver will be approximated by $N_{RX}SNR$, with SNR the Signal−to−Noise Ratio of a single channel. This leads to the following approximation for the channel capacity for a system using receiver diversity:

$$C = \log_2\left(1 + N_{RX}\,\rho\|h_{avg}\|^2\right) \qquad \text{(eq. 8)}$$

Figure 4 shows the channel capacity for various SIMO configurations, with the number of receive antennas equal to 1, 2, 3 and 4.

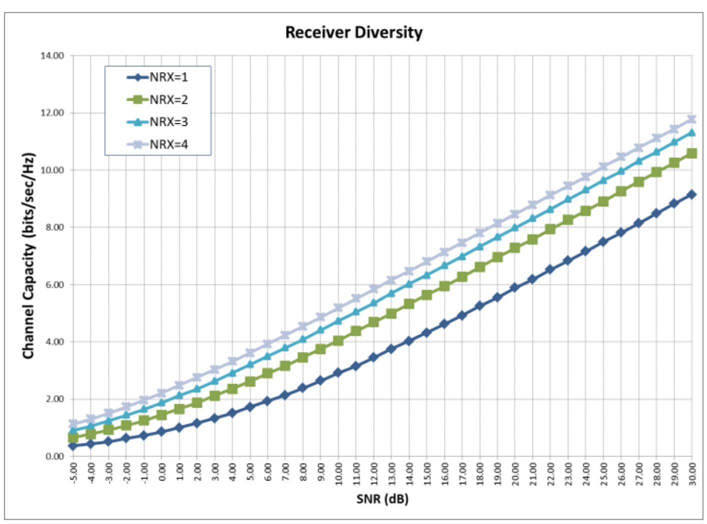

Figure 4. Channel Capacity for Various SIMO Configurations

As can be seen from Figure 4, receiver diversity provides a significant boost in capacity. Where the 1x1 system provided a theoretical capacity of 2.92 bits/sec/Hz at 10 dB SNR, this number improves to 4.05 bits/sec/Hz for a 1x2 system, 4.73 bits/sec/Hz for a 1x3 system and 5.19 bits/sec/Hz for a 1x4 system. Note that this improvement is achieved without any increase in power or bandwidth, but instead relies entirely on the extra degrees of freedom that are available in a MIMO channel. This is an extremely powerful result and illustrates the amazing potential of MIMO systems. To get a similar improvement from 2.92 bits/sec/Hz to 5.19 bits/sec/Hz using a 1x1 system would require a power increase of about 8 dB or a 70% increase in bandwidth. In power−limited and bandwidth−limited systems, one does not have the freedom to raise power or increase bandwidth, so MIMO is the only available avenue for increasing throughput.

The results in Figure 4 only apply to receiver diversity. A similar analysis can be done for transmit diversity. To isolate the effect of transmit diversity, we consider a case of a transmitter with N_{TX} antennas and a receiver with a single antenna. Such a system is also referred to as a MISO system (Multiple Input, Single Output). For now, we assume that the transmitter has no knowledge of the channel. For such a case, the optimum power allocation is an equal distribution of power among the N_{TX} antennas. In this case, the powers of the various signals can again be added at the receiver. The important difference is that each of those signals now only carries $1/N_{TX}$ of the power. As a result, the capacity is given by:

$$C = \log_2\left(1 + \rho\left(\frac{\|h_1\|^2}{N_{TX}} + \ldots + \frac{\|h_{N_{TX}}\|^2}{N_{TX}}\right)\right) \qquad \text{(eq. 9)}$$

This can be approximated as:

$$C = \log_2\left(1 + \rho\|h_{avg}\|^2\right) \qquad \text{(eq. 10)}$$

This means that transmit diversity does not provide the kind of gain that can be obtained from receiver diversity, at least not when no channel information is available at the transmitter. Even though the approximate expression (10) does not show the improvement, transmit diversity does provide a beneficial effect, albeit of a smaller order than receive diversity. Since the same data now reached the receiver through up to N_{TX} independent paths, the effects of fading are averaged out to some extent. The probability of losing a signal drops exponentially with the number of uncorrelated transmission paths. As a result, the received signal will be less vulnerable to the fading on any one of the multiple paths, compared to a SISO system. Simulations results do show this effect, as can be seen in Figure 5, where we consider a MISO system with the number of transmit antennas equal to 1, 2, 3 and 4.

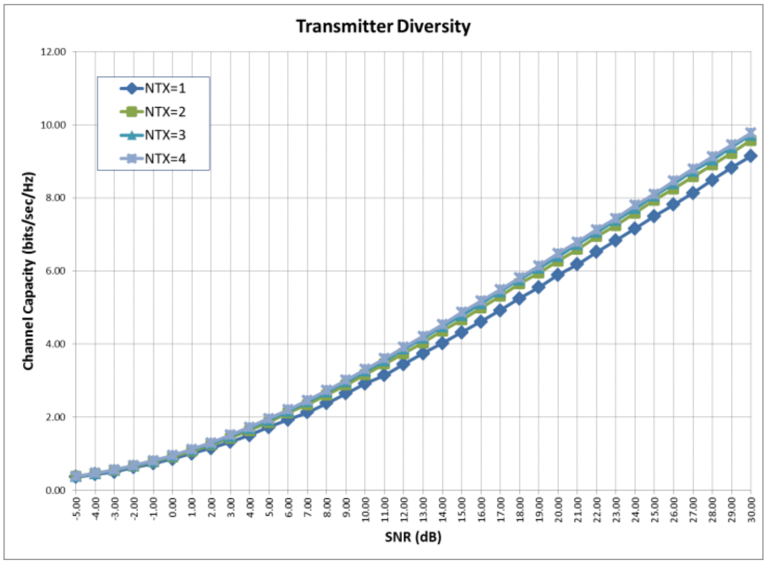

Figure 5. Channel Capacity for Various MISO Configurations

Overall, we can see from Figure 5 that the effect of transmit diversity is smaller than that of receiver diversity. This is only the case when no channel information is present at the transmitter. With channel information, the transmitter can enable transmit beamforming, which will lead to much larger gains. Transmit beamforming will be discussed in more detail in section Transmit Beamforming.

Spatial Multiplexing

Spatial multiplexing involves sending independent data streams through different independent paths. This stands in contrast to the diversity that was discussed above, where the same data is sent over the independent channels. To evaluate the performance of MIMO when spatial multiplexing is used, the Channel Capacity theorem has to be extended to the case of MIMO channels.

Although we won't go into the details of how to derive this expression, it can be shown that the channel capacity of a MIMO channel characterized by channel matrix H can be expressed as:

$$C = \log_2\left(\det\left(I_{N_{RX}} + \frac{\rho}{N_{TX}} HH^{\dagger}\right)\right) \quad \text{(eq. 11)}$$

Where:

- N_{RX} is number of receive antennas
- N_{TX} is number of transmit antennas
- I_{NRX} is the N_{RX} x N_{RX} identity matrix
- ρ is the average Signal−to−Noise ratio
- H is the N_{RX} x N_{TX} channel matrix
- det( ) is the determinant of the matrix

We assume that H is a complex−valued random matrix of independently distributed Gaussian variables with standard deviation equal to 1.

Note that one can show that (eq. 11) reduces to (eq. 8) and (eq. 10) for the cases of SIMO and MISO respectively. For SISO, (eq. 11) reduces to the traditional expression (eq. 6).

Expression (eq. 11) can be used to predict the channel capacity for various MIMO configurations. Results are shown in Figure 6.

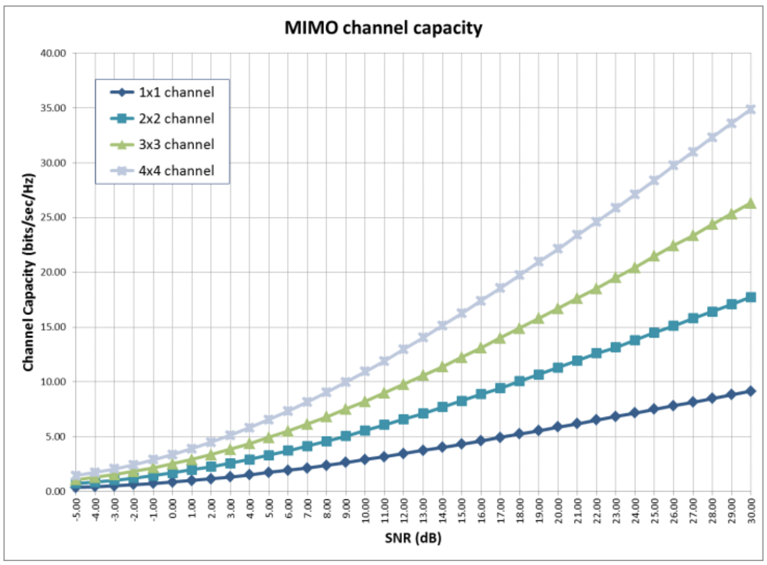

Figure 6. Channel Capacity for Various MIMO Configurations

Figure 6 clearly shows the tremendous potential of using spatial multiplexing. While a 1x1 system achieves a channel capacity of 2.92 bits/sec/Hz at 10 dB SNR, a 2x2 achieves 5.55 bits/sec/Hz without any increase in transmit power or signal bandwidth. A 3x3 and a 4x4 system achieve 8.22 bits/sec/Hz and 10.94 bits/sec/Hz respectively.

What’s more, the increase in capacity for a given SNR value grows linearly as a function of the number of transmission modes. To be exact, the capacity scales linearly with min(N_{RX},N_{TX}). This is illustrated in Figure 7 for various values of the average SNR. For comparison, the channel capacity of a SISO system grows only logarithmically as a function of average SNR.

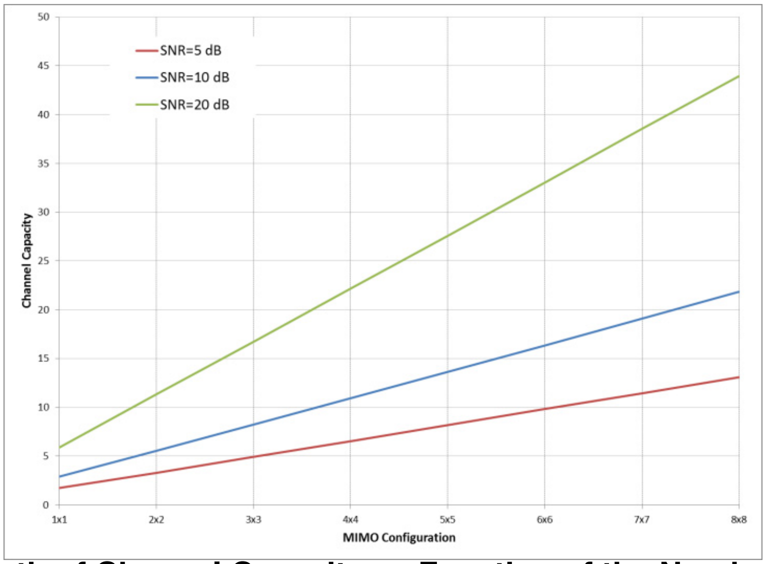

Figure 7. Growth of Channel Capacity as Function of the Number of Antennas (for SNR = 5, 10 and 20 dB)

Transmit Beamforming

Transmit beamforming is a form of transmit precoding that allows the transmitter to direct the signal more efficiently to the intended receiver. It assumes that channel state information is available at the transmitter. This information can either be supplied by the receiver through a feedback channel or can be autonomously obtained by the transmitter from the reverse channel.

To understand the principles behind beamforming, we have to make more concrete the notion of "independent paths" we have been using somewhat intuitively so far.

Mathematically, it can be shown that an N_{RX} x N_{TX} matrix has at most min(N_{RX},N_{TX}) independent modes that can be used for transmission. The actual number of independent modes is equal to the rank of the matrix H. Additionally, the matrix will have $|N_{RX}-N_{TX}|$ or more zero modes. Zero modes can be thought of as transmission paths that produce no signal at the receiver, regardless of the signal that is applied at the transmitter. Zero modes effectively cancel the signal at the receiver.

All this is best understood by performing the so−called SVD decomposition of the channel matrix. It can be shown that any matrix can be written as:

$$H = U \sum V^{+} \quad \text{(eq. 12)}$$

Where:

- U is an N_{RX} x N_{RX} unitary matrix
- $\sum$ is and N_{RX} x N_{TX} matrix with all off−diagonal elements equal to zero
- V is an N_{TX} x N_{TX} unitary matrix

The diagonal elements of the matrix Σ are called the singular values of the channel matrix H. We'll denote these as $\sigma_i, i=1,\ldots,min(N_{RX}, N_{TX})$. These values are real, non−negative and correspond to the independent modes of the channel. The numerical value of the singular values determines the quality of the respective modes.

Using the SVD form (eq. 12) of the matrix, we can rewrite the MIMO channel capacity (eq. 11) as:

$$C = \log_2\left(\det\left(I_{N_{RX}} + \frac{\rho}{N_{TX}}\Sigma\Sigma^{\dagger}\right)\right) = \sum_{i=1}^{\min(N_{RX},N_{TX})} \log_2\left(1 + \frac{\rho}{N_{TX}}\sigma_i^2\right) \qquad \text{(eq. 13)}$$

Expression (eq. 13) shows how the capacity of the matrix channel can be expressed as the aggregate capacity of multiple independent SISO channels (compare with e.g. (eq. 6)). This provides us with a firmer interpretation of the terms "independent paths" we have been using so far. Each of the singular values can be identified with such a path. It also gives us a metric to judge the reliability of these independent paths. The smaller the singular value, the higher the path loss over this particular path and the less reliable it will be for transmission. Some of the singular values may be zero, meaning that they can not be used for data transmission at all (their channel capacity is zero). Any signal sent over such a path is irrevocably lost.

Without beamforming, transmit signal is allocated in equal amounts to the different transmit antennas and some of the energy ends up in weak or useless transmission paths. Transmit beamforming on the other hand allows the transmitter to precode the signal in such a way that the transmission will use only the most reliable available mode or modes for transmission of the signal. As a result, the same channel available to a MIMO system that does not use TxBF, can be exploited by TxBF in the most efficient way, thus improving the channel capacity, for the same exact conditions (environment, number of antennas, etc).

Simulation results illustrating the benefits of beamforming are shown in Figure 8.

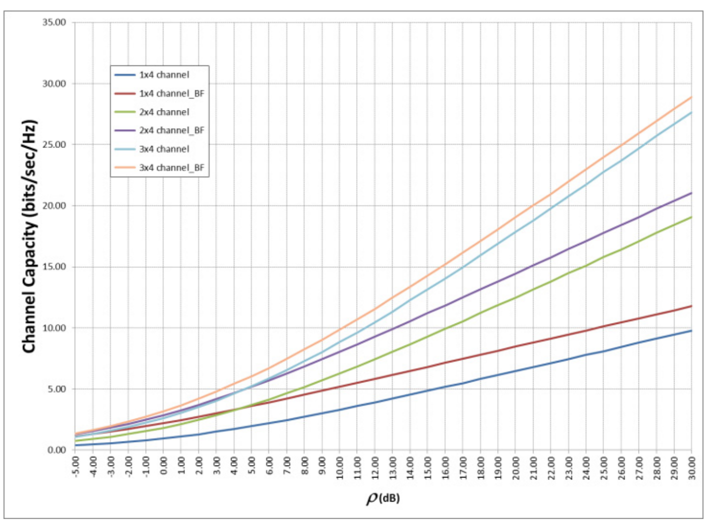

Figure 8. Channel Capacity with and without Transmit Beamforming, using 4 Transmit Antennas

Multi–User MIMO (MU–MIMO)

MU–MIMO uses the spatial diversity that is present in a matrix channel in a somewhat different way. When we use transmit diversity as described in the section Spatial Multiplexing, we send independent data streams using the multiple independent paths. Even so, all data is sent from a single transmitter to a single receiver. In MU–MIMO, some of the data streams are sent to one receiver, while other streams are sent to a different receiver. The transmit precoding that is applied attempts to eliminate the mutual interference that would otherwise result from sending two simultaneous transmissions.

MU–MIMO covers both scenarios with a single transmitter addressing multiple receivers and scenarios where a single receiver receives data from multiple transmitters. The former is called Down–Link MU–MIMO (DL MU–MIMO) and the latter Up–Link MU–MIMO (UL MU–MIMO). In the remainder of the discussion, we limit ourselves to DL MU–MIMO since this is the one that is most relevant to wireless LAN networks.

Because there is now a single transmitter and multiple receivers, there will be more than one matrix channel involved in modeling this scenario.

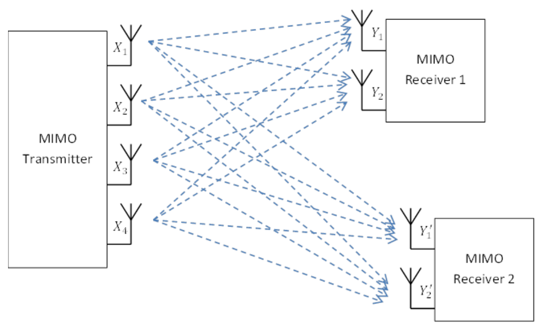

Figure 9. Illustration on MU–MIMO Consisting of One Transmitter and Two Receivers

In transmit beamforming, the precoding of the transmit signals was designed to send the transmit signals into the most reliable modes of the MIMO channel, avoiding weak modes or zero modes. In MU–MIMO, a different criterion is used. The precoding at the transmitter now has to meet following criteria:

1. The precoded transmit signal intended for the first receiver should result in zero signal power at the second receiver since this signal represents interference to the second receiver
2. Likewise, the precoded transmit signal intended for the second receiver should result in zero signal power at the first receiver
3. Within the previous constraints, signals should attempt to use the most reliable modes of the respective channels

There are alternative criteria that don’t attempt to completely cancel the interfering signals, but we will stick with the scheme given above for simplicity of explanation.

Therefore, successful use of MU−MIMO relies on finding transmission paths that provide good−quality reception at one of the receivers, while simultaneously being a zero−mode towards the other receiver. Whether such a solution exists depends on the channels from the transmitter to the various receivers, so the gains of MU−MIMO will vary. For instance, if two receivers are in the same or approximately same location, it is obviously not possible for the transmitter to simultaneously optimize and cancel signal in that direction. Based on the channel information available at the transmitter, the transmitter has to make an informed decision on when to use MU−MIMO and when not to use it.

The example below illustrates the operation of MU−MIMO. Consider a transmitter with four transmit antennas ($N_{TX} = 4$) sending to three receivers, each with a single receive antenna ($N_{RX1} = N_{RX2} = N_{RX3} = 1$). Each of the channels is a 1x4 vector. For instance:

$$\begin{aligned} H_1 &= \begin{bmatrix} -1.506 & -0.4446 & -0.1559 & 0.2761 \end{bmatrix} \\ H_2 &= \begin{bmatrix} -0.2612 & 0.4434 & 0.3919 & -1.2507 \end{bmatrix} \\ H_3 &= \begin{bmatrix} -0.9480 & -0.7411 & 0.5078 & -0.3206 \end{bmatrix} \end{aligned} \qquad \text{(eq. 14)}$$

The transmitter can perform a MU transmission by sending independent data to each of the receivers. If we denote the data for the various receivers as X_1, X_2 and X_3, the combined transmission is simply a suitably chosen linear combination of the individual transmissions, i.e:

$$T_X = Q_1X_1 + Q_2X_2 + Q_3X_3 \qquad \text{(eq. 15)}$$

Where the Q_i are (4 x 1) precoding coefficients. A possible choice for the values of Q_i is given in formula (eq. 16) below.

$$Q_1 = \begin{bmatrix} -0.0430 \\ -0.5478 \\ 0.8321 \\ 0.0755 \end{bmatrix},\ Q_2 = \begin{bmatrix} 0.0262 \\ -0.5120 \\ 0.8320 \\ -0.2117 \end{bmatrix},\ Q_3 = \begin{bmatrix} 0.0206 \\ -0.2812 \\ 0.9403 \\ 0.1906 \end{bmatrix} \qquad \text{(eq. 16)}$$

One can verify that:

$$\begin{aligned} H_1Q_2 &= H_1Q_3 = 0 \\ H_2Q_1 &= H_2Q_3 = 0 \\ H_3Q_1 &= H_3Q_2 = 0 \end{aligned} \qquad \text{(eq. 17)}$$

This leads to the following signals at the respective receivers:

$$\begin{aligned} R_{x_1} &= H_1Q_1X_1 \\ R_{x_2} &= H_2Q_2X_2 \\ R_{x_3} &= H_3Q_3X_3 \end{aligned} \qquad \text{(eq. 18)}$$

This means that when the precoded signal (eq. 15) travels to first receiver over channel H_1, only the signal proportional to X_1 will remain, where X_1 is exactly the data that was intended for this receiver. The signals for the other two users, which would otherwise show up as interference at the first receiver have been entirely removed through the precoding (eq. 15), (eq. 16). Likewise, one can see that the signals at the second and third receiver are received free of interference.

The challenge for MU−MIMO is to find a set of precoding coefficients that will achieve this result for the given number of users and channel conditions.

Capacity for MU−MIMO systems can not be expressed as a single number. The SNR for each of the users could be different, and even for equal SNR settings there are many possible ways to allocate resources (say, transmit power) between the users. There are a number of ways in which MU capacity is evaluated. One of them is the concept of “capacity region”, which shows the ensemble of rates that can be achieved simultaneously by the users. An alternative way is the concept of “sum capacity”, which calculates the total aggregate capacity that can be achieved by all users in the MU−MIMO system. Note that sum capacity does not necessarily guarantee fairness, since the arrangement that optimizes the sum capacity may favor one user over the others.

In Figure 10, we illustrate the channel capacity of a 2−user MU−MIMO system, with each user receiving a single spatial stream. In this example, the channel capacity corresponds to the sum capacity of the two users. It is also assumed that both users have been allocated the same transmit power and that the SNR on both links is the same.

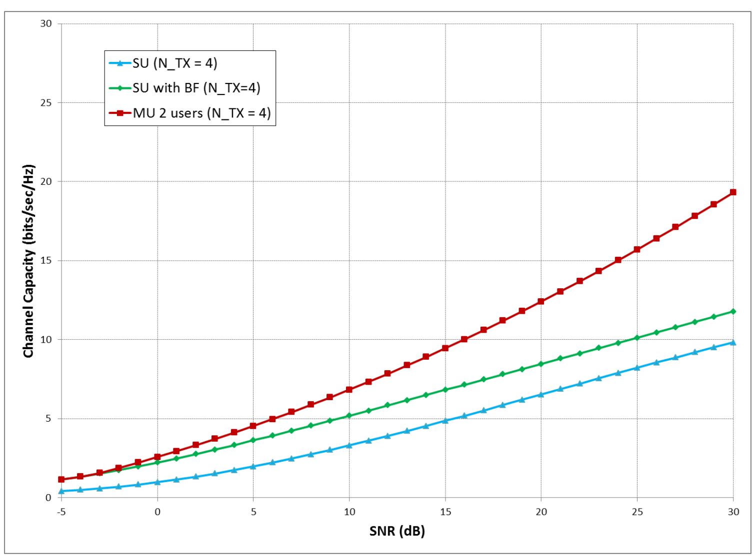

Figure 10. Sum Channel capacity for 2−user MU−MIMO scenario, compared with SU

Obviously, the total power of the transmit signal can not exceed the total power used for SU transmissions. In Figure 10, we also compare the MU sum capacity with the capacity that can be achieved with an SU transmission using the same transmit power. We consider both the case with and without transmission beamforming.

An interesting observation is that for low SNR, the MU sum channel capacity converges to the channel capacity of a single user with transmission beamforming. This can be understood by observing that in low SNR, more interference can be tolerated without impacting the performance. Therefore, most of the precoding could still be used to achieve beamforming gain. In addition, in the low SNR region, the capacity is approximately a linear function of SNR, so two users that each get half the total transmit power should be able to achieve the same capacity as a single user using the total transmit power. The curves in Figure 10 confirm this.

Also note that, outside of the low SNR region, the MU channel capacity always exceeds the channel capacity of a single user with beamforming.

Putting it all together

The preceding sections (Diversity – Multi−User MIMO (MU−MIMO)) summarized the various benefits that can be achieved from using MIMO concepts. The question in practice is whether one should try to achieve diversity gain or spatial multiplexing, or some combination of both. The optimal strategy depends on the properties of the matrix channel. As we've seen in the section Transmit Beamforming, the matrix channel can be thought of as a number of parallel SISO channels, each with their own capacity (which could be zero for some channels). The rank of the matrix determines the number of independent channels that is available. At the same time, the SVD decomposition tells us something about the reliability of each of those channels.

A practical system should include the intelligence to decide how many spatial streams to use for transmission. If the number of spatial streams is less than $\min(N_{RX},N_{TX})$, the remaining MIMO degrees of freedom will automatically be used for diversity gain. How to decide the number of spatial streams is part of the rate adaptation that is implemented at transmitter and receiver.

Likewise, a system should intelligently decide on when to use MU−MIMO vs. SU transmission. The channel feedback that is provided by the receivers allows the transmitter to assess the feasibility of multi−user operation. If the channel conditions are such that MU−MIMO does not provide the optimal aggregate throughput (e.g. when channels to different users are highly correlated), the transmitter should stay with SU transmission instead. This too should be part of the rate adaptation implemented by the transmitter and receiver(s).

Diversity, Spatial Multiplexing, Beamforming and MU−MIMO are all tools that are available to the transmitter to optimize performance in every situation. Some of the tools will be more effective in certain situations, but together they guarantee the best performance across the range of scenarios that will be encountered in a wireless environment.

MIMO in action: 802.11n and 802.11ac

MIMO concepts were introduced to wireless LAN technology with the approval of the 802.11n amendment to the IEEE 802.11 standard, popularly known as “Wi−Fi”. The recent 802.11ac amendment further extends the number of MIMO options available to wireless stations.

IEEE 802.11n

The 802.11n amendment to the 802.11 standard was approved in 2009. It provides support for systems with up to four transmit and four receive antennas. It allows full spatial multiplexing of up to four data streams. Support of two spatial streams is mandatory at the AP. STAs only need to support a single stream. Support of higher number of spatial streams is optional.

802.11n transmits data in short bursts (or frames). Each frame is preceded by a preamble that allows the receiver to estimate the MIMO channel and learn the parameters of the transmission (e.g. the number of spatial streams). To enable beamforming, 802.11n defines a channel sounding and feedback protocol that informs the transmitter of the MIMO channel measured at the receiver. The transmitter uses this information to calculate the transmit precoding.

The maximum PHY rate possible with 802.11n is 600 Mbps, which is achieved when four spatial streams are used with the maximum modulation.

802.11n also includes Space−Time Block Coding (STBC), which is another way to exploit diversity in a MIMO environment.

IEEE 802.11ac

The 802.11ac amendment to the 802.11 standard is scheduled for approval in February 2014. It further increases the number of spatial streams that can optionally be supported by APs and STAs to a maximum of eight. 802.11ac is also the first technology to enable MU−MIMO. Only down−link (DL) MU−MIMO is supported, i.e.: only transmissions from AP to STAs can benefit from the use of MU−MIMO. When using MU−MIMO, an AP can simultaneously transmit to up to four STAs, with a maximum of four spatial streams allocated to any of the users. The total number of spatial streams in a MU−MIMO transmission can not exceed eight.

The frame structure and preamble for 802.11ac are similar to the one used in 802.11n.

Like 802.11n, 802.11ac also supports exchange of Channel State Information (CSI) to enable Transmit Beamforming. The same CSI is also used to determine precoding matrices for MU−MIMO. The main difference between 802.11n and 802.11ac is that 802.11ac has reduced the three possible formats for CSI exchange that were defined in 802.11n to a single format.

Performance of IEEE 802.11 MIMO Systems

In section Advantages of MIMO: theoretical analysis, we analyzed the maximum potential improvement that can be gained from MIMO technologies from a theoretical point of view. The results discussed there correspond to the optimal case using the assumptions that channels are fully random and uncorrelated complex Gaussian channels. For such case, the advantages of MIMO can be exploited to the fullest. Moreover, the metric used to compare different approaches is the channel capacity, which does not take into account the practical implementation of modulation and coding. Channel capacity provides the performance of an ideal system without any implementation restrictions.

In reality, the characteristics of wireless channels will vary wildly between deployments. As such, while the illustrations in section Advantages of MIMO: theoretical analysis provided a theoretical understanding of the potential of MIMO, they may not reflect the actual performance of a 2x2, 3x3 or 4x4 system in a real−world wireless environment, but they do however provide a maximum achievable limit.

In the following sections, we'll investigate the performance of various MIMO configurations in channels that more closely model the actual wireless channel. The channel models were developed by IEEE as part of an effort to allow simulation of performance of 802.11 systems that is representative of expected performance in real−world scenarios. The channel model used in the simulation below is known as "Channel D" and was designed to represent a typical office environment. Figure 11 shows an example of the frequency variation of a Channel D realization.

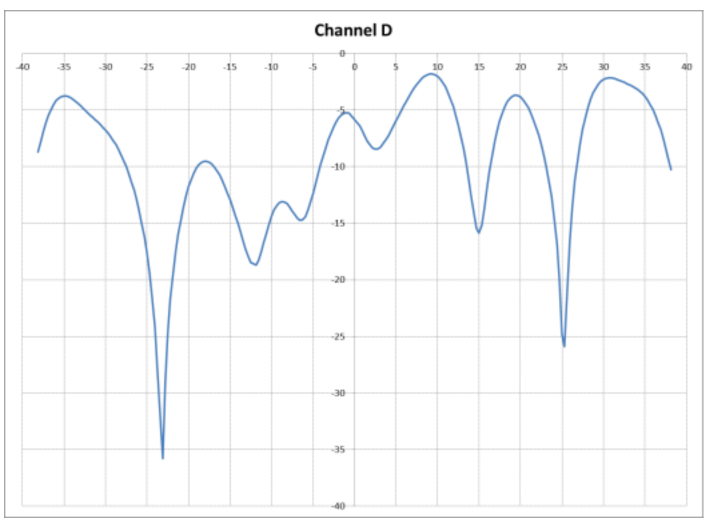

Figure 11. Example of Channel D Realization for 80 MHz Bandwidth

The channel in Figure 11 shows an obvious fading over the 80 MHz channel bandwidth. Both 802.11n and 802.11ac use OFDM modulation to deal with the fading and multiple paths that occur in the wireless environment.

Instead of using channel capacity to evaluate the performance of a system, we'll instead express performance as a rate−reach curve, which shows the throughput that can be achieved at a given path loss (or, equivalently, SNR) of the wireless channel. This approach follows more closely the way wireless systems are evaluated in practice.

We will also restrict the modulation and coding to the options that are allowed in 802.11. In order to isolate the effects of MIMO, we assume perfect rate adaptation algorithm, i.e. at each attenuation, the system uses the optimal number of spatial streams and the optimal modulation of the spatial streams. Simulations below do not assume perfect channel information at the receiver. Instead, the receiver has to estimate the channel in accordance with the training opportunities that are provided by the 802.11 frame format.

Diversity

Diversity involves sending the same signal through multiple independent paths, to improve reliability of the reception (see the section Diversity). We consider a case with a single transmit antenna and multiple receive antennas to clearly isolate the effect of receive diversity.

Figure 12 shows how receive diversity impacts performance for a practical MIMO system, in this case an 802.11ac system using 80 MHz bandwidth. The benefit of multiple antennas at the receiver can be clearly seen, the gain is not negligible, even in front on a client that has a single antenna like in the case study.

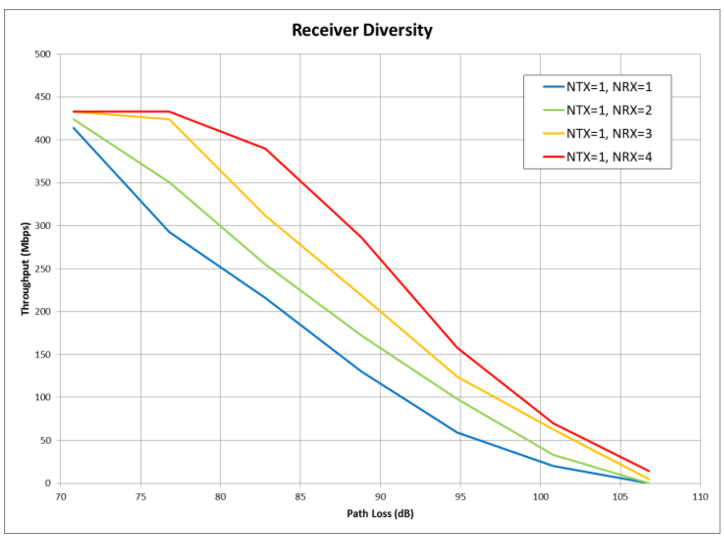

Figure 12. Effect of Multiple Receive Antennas on System Throughput

Takeaways:

There is a non−negligible throughput gain in receiving traffic on N antennas from systems with fewer antennas.

For medium−range attenuations (−80 to −99), for a given number of TX antennas, adding one extra RX antenna does increase throughput considerably (up to more than 50 Mbps on a 11ac 80 MHz system)

Spatial Multiplexing

Spatial multiplexing (see the section Spatial Multiplexing) involves the transmission of independent data streams over the different independent transmission modes that exists between systems with multiple transmit and receive antennas. Simulation results for an 80 MHz 802.11ac system are shown in Figure 13. These results assume that no channel information is present at the transmitter. Consequently, power is allocated equally across all transmit antennas.

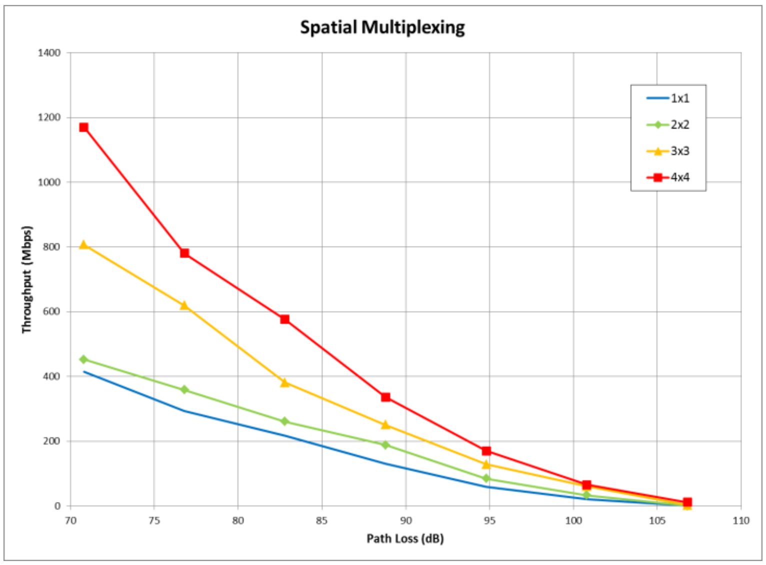

Figure 13. Effect of Spatial Multiplexing on System Throughput

Figure 13 shows the ever−increasing throughput as one goes from 1x1 to 2x2, 3x3 and 4x4. The use of multiple spatial streams allows for significant performance gains. The number of spatial streams that is used at each path loss is chosen automatically to optimize performance. The channel may not allow the maximum number of spatial streams for each configuration (e.g. 4 spatial streams for a 4x4 system). In that case, the transmitter will autonomously determine the right balance between spatial multiplexing and diversity.

Takeaways:

The use of spatial multiplexing has non−negligible throughput gain. For example on a 11ac 80 MHz system, the use of 4 spatial streams results in 50% throughput increase wrt 3 spatial streams.

Transmit Beamforming

Transmit beamforming (see the section Transmit Beamforming) means that the transmitter allocates transmit power optimally to the different modes of the channel so to use the best transmission modes available in the current channel. To do this, channel state information (CSI) is required at the transmitter. Simulation results for an 80 MHz 802.11ac system are shown in Figure 14.

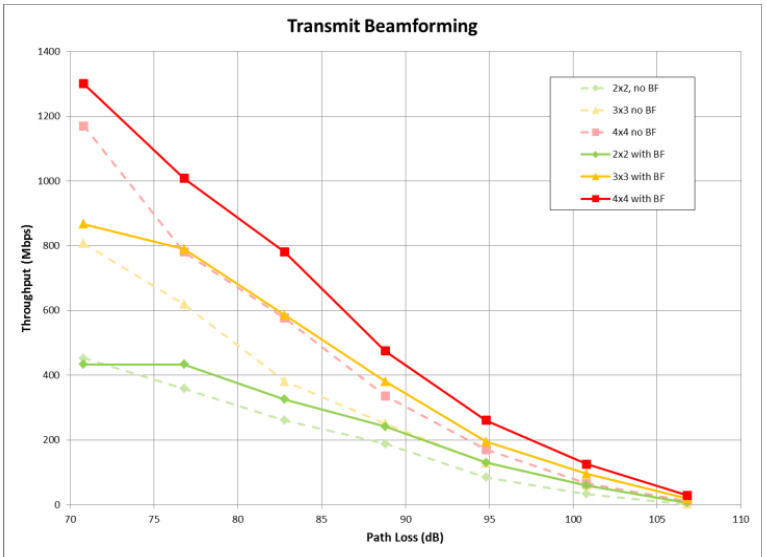

Figure 14. Effect of Transmit Beamforming on System Throughput

The performance improvement resulting from transmit beamforming cam be clearly seen from Figure 14. The simulations assume the channel feedback mechanism that is defined in 802.11. This means that feedback is based on the receiver's estimate of the channel, including any errors in the estimation due to noise and channel conditions.

To obtain the optimal throughput, the transmitter autonomously determines the optimal spatial stream allocation and the modulation for the spatial streams. In case the channel is not full rank (i.e. does not support the maximum number of spatial streams), the system will automatically pick the configuration that optimizes performance. Beamforming weights are determined by the transmitter based on the feedback reported in CSI by the receiver.

Takeaways:

The use of TX beamforming results in non−negligible throughput gain. For example on an 11ac 80 MHz system, at medium range attenuations, the use of TX beamforming on 4 spatial streams results in 25% throughput increase wrt to the same 4 spatial streams with no TX beamforming; moreover, the use of TX beamforming on 4 spatial streams results in 67% of throughput increase wrt to a 3 spatial stream system with no TX beamforming.

MU−MIMO

The gains of MU−MIMO depend on the channels to the different users. The more independent those channels are, the better the performance can be with MU−MIMO precoding. In order to get some idea of the expected performance with MU−MIMO, it is possible to establish a simple bound on the achievable throughput. Assuming that the channels to the respective users and the precoding are such that interference can be completely eliminated without loss in received signal power, the MU capacity would reduce to a sum of SU capacities. One only has to consider the fact that the sum of the transmit powers for all users should not exceed the transmit power used for a single user.

Two examples are shown in Figure 15 and Figure 16. In these figures, we show the sum capacity of MU−MIMO for the specific case of two users with the same SNR.

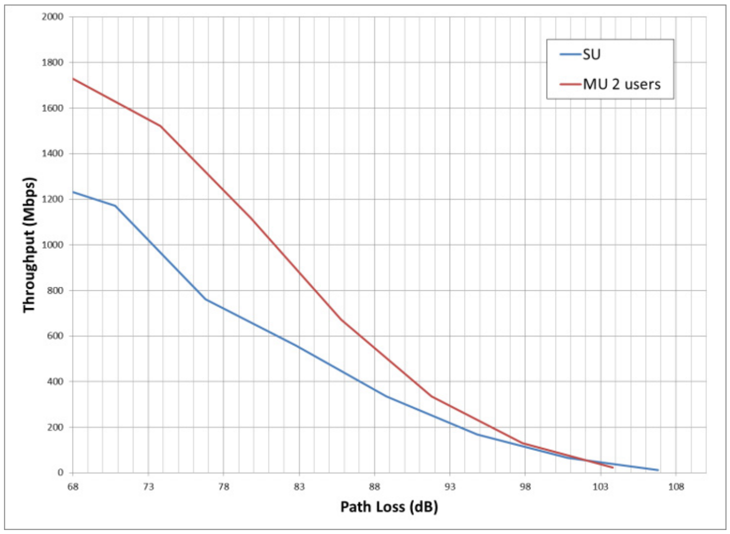

Figure 15. Extrapolated MU−MIMO Aggregate Performance

Figure 15 compares the performance of an SU system with the aggregate performance of an MU system serving two users. In this case, it was assumed that both users are allocated half of the total power and that they both experience the same SNR. The total number of spatial streams supported by the AP is limited to four. For this case, MU−MIMO offers the most benefit for lower path loss values, with gains of about 400−500 Mbps.

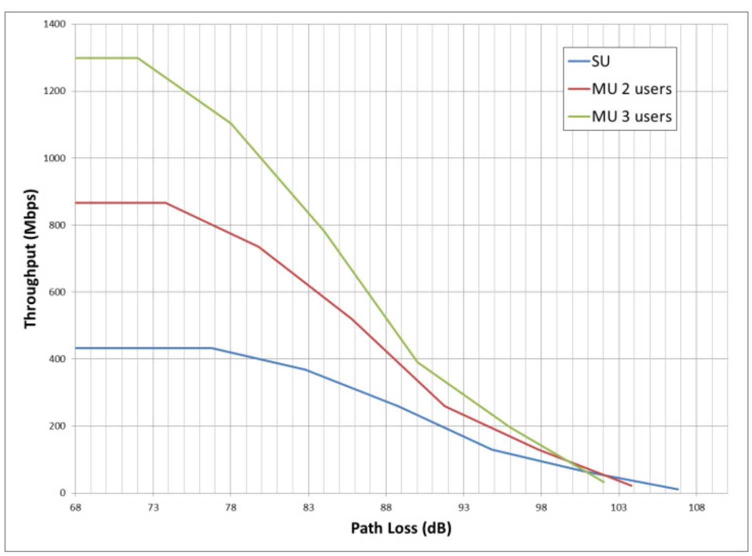

Figure 16. Extrapolated MU−MIMO Aggregate Performance When Receivers Can Only Process One Spatial Stream

Figure 16 considers a slightly different case, where the STAs are one−stream devices. This would be representative of e.g. Wi−Fi enabled phones or other hand−held devices. In this case, the total aggregate throughput is significantly higher than what would be transmitted to an SU device.

Takeaways:

Even with only 2 concurrent users, the use MU−MIMO increases the bandwidth available to a wifi network by up to 40%.

In presence of single stream devices, the use of MU−MIMO increases the available bandwidth linearly with the number of users that can be served simultaneously.

Conclusion

We have presented an overview of MIMO techniques and advantages, specifically in the context of wireless communications.

First, we evaluated the different MIMO concepts in terms of their theoretical channel capacity to highlight the maximum theoretical gains that can be achieved by using MIMO techniques. We then studied the rate versus range gains introduced by these techniques when applied to 802.11, showing the performance increases that are achieved for real−world wireless systems.

References

[1] G. J. Foschini and M.. J. Gans (January 1998),"On limits of wireless communications in a fading environment when using multiple antennas", Wireless Personal Communications 6 (3): 311–335

[2] A mathematical theory of communications, C. Shannon, 1948, Bell Systems Technical Journal, vol 27, pp 347−423

[3] G. J. Foschini, "Layered space−time architecture for wireless communications in a fading environment when using multi−element antennas", Bell Labs Technical Journal 1 (2): 41–59

[4] E. Telatar (1999), "Capacity of Multi−antenna Gaussian Channels". European Transactions on Telecommunications 10 (6): 585–595

PUBLICATION ORDERING INFORMATION

LITERATURE FULFILLMENT:
Email Requests to: orderlit@onsemi.com

ON Semiconductor Website: www.onsemi.com

TECHNICAL SUPPORT
North American Technical Support:
Voice Mail: 1 800−282−9855 Toll Free USA/Canada
Phone: 011 421 33 790 2910

Europe, Middle East and Africa Technical Support:
Phone: 00421 33 790 2910
For additional information, please contact your local Sales Representative