Chapter 2
Printed Antennas for High-Speed Communication Systems: Preliminaries and Review

2.1 Ultra-Wideband Communication Systems

2.1.1 Brief History of UWB Technology

UWB technology has its roots that date back to more than a century. UWB technology came to limelight in early 1960s from work on electromagnetic wave propagation study [13]. Historically, UWB systems used the concept of impulse radio where transmission of high data rates is achieved by transmitting a few nano second duration pulses [14]. On 12th December, 1901, spark gap radio transmitter was used by Guglielmo Marconi to transmit Morse code across Atlantic Ocean although he designed it in the late 1890s [4]. Very wide bandwidth pulses were induced by this radio transmitter [15]. The effective recovery of wideband energy and also discrimination from many other wideband signals were not possible at that time. Ultimately wideband communication was abandoned. The communication world turned to narrowband radio transmitter which had the traits of easy regulation and coordination. Efforts were made to lower the interference and to increase the reliability during 1942–1945. This led to filing of several patents on this aspect. Most of the patents were frozen because they were suitable for military applications by the U.S. government [16].

In 1960s, two major substantial breakthroughs in UWB technology took place. An impulse measurement technique to calculate transient behavior of microwave networks was developed by Gerald F. Ross. In 1962, Hewlett-packard developed sampling oscilloscope. This equipment was used to observe, analyze and measure the impulse response of microwave networks and catalyzed methods for generation of sub-nanosecond pulses.

In 1960s, antenna designers developed antennas whose impedance and pattern are almost frequency independent up to certain frequency. Rumsey and Dyson developed logarithmic spiral antennas [17, 18]. The wideband radiating antenna design was carried out by Ross using impulse measurement techniques [19]. This boosted short pulse radar communications. In April, 1973, the short-pulse receiver is the first accepted patent in UWB communications which was filed by Ross [20].

J. Malik et al., *Compact Antennas for High Data Rate Communication*,
Springer Topics in Signal Processing 14, DOI 10.1007/978-3-319-63175-2_2

The unlicensed wideband communication in the bands allocated for Industrial Scientific and Medicine (ISM) was allowed in the mid 1980s. Since then, WLAN and Wi-Fi have grown tremendously. Subsequently the communication industry began to study the merits and implications of wider bandwidth communications.

In 1990, the term ‘UWB’ was first used in a radar study by Defense Advanced Research Agency [21]. In early days, UWB was considered as ‘carrier-free’ or ‘baseband’ signals having extremely fast rise time. Pulse or step signals were used as wideband signal which excites a wide-band antenna. Advancements in semiconductor technology encouraged UWB for commercial applications.

On 14th February, 2002, FCC amended rules that allowed the unlicensed use of 7.5 GHz bandwidth from 3.1 to 10.6 GHz [3]. This allocation led to the substantial growth in research and development of RF circuits and antennas for UWB communications. Industrial investment have increased in this revolutionary technology to provide wide range of services for customers. UWB technology is an emerging technology which is considered as a promising solution to achieve high data rate communication.

2.1.2 General Concepts of UWB Technology

The UWB systems are different from conventional narrow band systems in terms of transmission of information. In conventional narrow band systems, the baseband signal is modulated and transmitted using a carrier frequency. In UWB systems, the transmission and reception of information is done using low duty cycle, short duration (picosecond-nanosecond) pulses. The low power spectral density signals are transmitted over a huge frequency band. Techniques for generation of short pulses that are used in UWB impulse radio communications are discussed in [44–46]. The general representation of narrowband and wideband signals in time and frequency domain is shown in Fig. 2.1.

Even though the power level of the transmitted pulse is very low, concerns regarding the potential interference to other narrow band systems was taken seriously. Therefore, an emission mask which defines the maximum allowable radiated power for UWB devices was specified by FCC. Hand held communication devices, imaging systems and vehicular radar systems are defined as UWB systems by FCC in its first report and order. The spectral mask for these systems (indoor and outdoor) is illustrated in Fig. 2.2. The emission mask for outdoor and indoor applications in the frequency range 3.1–10.6 GHz is −41.3 dBm/MHz. For frequencies higher than 10.6 GHz, the emission mask for outdoor applications is 10 dB lower than the indoor mask [3].

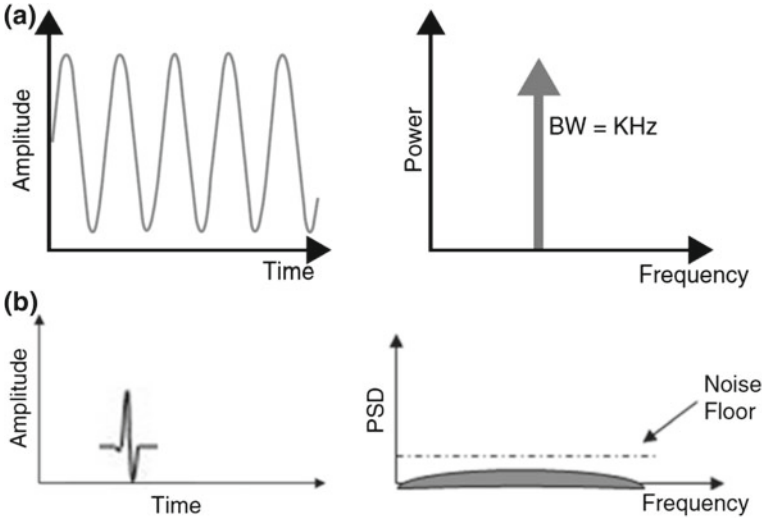

Fig. 2.1 Time-domain and Frequency-domain representation of **a** narrow band signal **b** UWB signal [22] [From: F. Nekoogar and Farid Dowla, Ultra-Wideband Radio Frequency Identification Systems, Springer New York - Heidelberg (ISBN 978-1-4419-9700-5) © Springer Science+Business Media 2011. Reproduced courtesy of the Springer.]

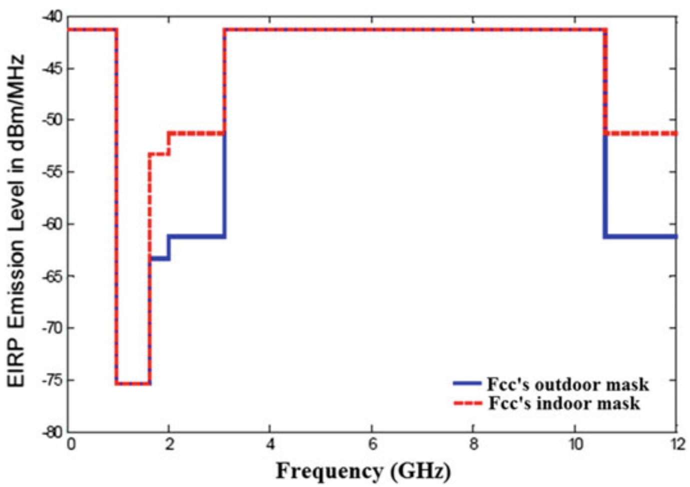

Fig. 2.2 Indoor and outdoor power spectral density emission masks specified by FCC [14] [From: J. Liang, "Antenna study and design for ultra-wideband communication," Ph.D. Thesis, Queen Mary, University of London, United Kingdom, July 2006. Reproduced courtesy of Queen Mary, Uni. of London.]

2.1.3 Advantages of UWB Technology

Ultra-wideband is considered better solution for radar and communication applications because of its unique merits over other technologies. Some of the key advantages of UWB technology are listed below [4, 22].

- Low cost.
- Large channel capacity.
- Low interference to legacy systems.
- Highly immune to multipath interference.
- Less probability of interception and detection.
- Higher accuracy in range measurement and range resolution.

Channel capacity varies linearly with the signal bandwidth at higher frequencies. It is possible to achieve gigabits per second (Gbps) data rate at gigahertz of bandwidth. However, the power limitation on UWB transmission restricts this speed to short ranges up to 10 m [23]. Hence it makes the system more suitable for small range, high data rate wireless applications.

The maximum radiation limit as −41.3 dBm/MHz or 75 nW/MHz makes UWB systems categorized under unintentional radiators. These systems reside well below the noise floor for any conventional narrow-band receiver. Figure 2.3 shows the general idea of coexistence of UWB with narrowband and wideband technologies.

The transmitted signal undergoes multiple reflections from various surfaces and may combine constructively or destructively leading to the distortion of the signal. This mechanism is called as multi-path interference. As the duration of UWB pulse is very short, the chance of collision of reflected pulse with the line of sight (LOS) pulse is very less. Increased immunity to multi-path interference is obtained using UWB signals.

UWB pulses have low average transmission power, hence it appears as noise-like to narrow band communication system. It is very difficult to intercept UWB signal

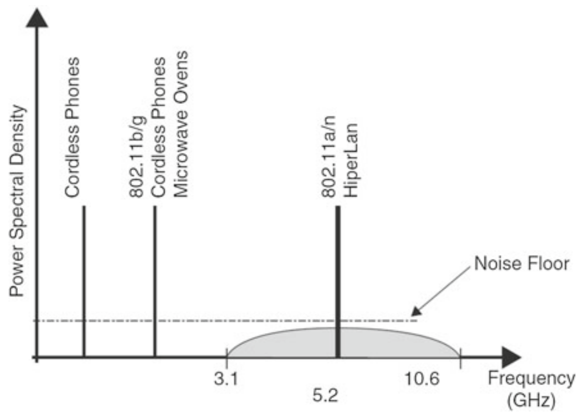

Fig. 2.3 Coexistence of UWB signals with narrowband and wideband signals [22] [From: F. Nekoogar and Farid Dowla, Ultra-Wideband Radio Frequency Identification Systems, Springer New York—Heidelberg (ISBN 978-1-4419-9700-5) © Springer Science+Business Media 2011. Reproduced courtesy of the Springer.]

without a prior knowledge. To secure UWB pulses, unique codes are used to time modulate UWB pulses which are unique to a transmitter/receiver pair [23]. It is impossible to guess a picosecond duration pulse without any prior information. In UWB transmission, modulation of data with a specific carrier frequency is not done, that eliminates the need for RF mixer unlike conventional narrowband technologies [24]. Therefore, fewer RF components are required for carrierless transmission as compared to carrier-based transmission. This decreases the cost and complexity of UWB transceiver.

Higher accuracy in range measurement and range resolution is obtained using a short pulse UWB radar system as compared to a conventional radar system.

2.1.4 Various Applications of UWB Technology

The high data rate characteristic of UWB makes it well suitable to be used in Wireless Personal Area Networks (WPAN) applications. The data is transmitted within a distance of 10 m or less, at rates up to 100–500 Mbps. The wireless connectivity of personal computers to storage devices, printers, scanners, video cameras, using high speed wireless universal serial bus (WUSB) is possible. Download/upload high quality video from mobile devices, audio streaming on MP3 players, wireless high definition digital video receiver, and wireless connectivity of speakers are enabled using UWB in WPANs [25]. The need of cables for interconnection between the mentioned devices is eliminated.

The UWB technology also finds its application in sensor networks. Sensor networks consist of group of sensors distributed spatially in a region for monitoring. The sensor nodes are either static or mobile. The sensor nodes are mobile when equipped on automobiles, firemen, robots, soldiers and emergency response situations [5]. Robustness, multi-functionality, low cost, and low power are the desired properties of sensor networks which can be attained by using UWB technology. An enormous quantity of sensory data can be gathered and disseminated in a timely manner with the help of UWB communication systems. The noticeable drop in installation and maintenance costs is possible without usage of wires.

UWB is an excellent candidate for indoor positioning and tracking applications due to higher data rate at smaller range. The advanced tracking mechanism can be used to track movement of objects in an indoor environment with centimeters of accuracy [15]. They can be used to locate objects or missing people in situations such a burning building, causalities in a remote area and so on [14].

Reflections of UWB pulses from the target causes time shift and amplitude distortion. Scattering affects UWB waveforms significantly as compared to conventional radar systems. Taking advantage of this property, UWB can be used to replace X-ray systems in medical diagnostics, through wall imaging under grounding imaging using ground-penetrating radar for rescue operations.

Resolution of conventional proximity and motion sensors can be improved using UWB-based sensors. The feature of high range accuracy combined with the ability to differentiate targets can be used to build intelligent collision-avoidance systems [5]. Improvement in deployment of airbag and adaptation of breaking/suspension systems with conditions of road are some key applications of vehicular UWB radar systems.

2.1.5 Antenna Requirements for UWB Technology

The design process of UWB antenna is more complex than narrowband antennas as higher number of performance parameters are to be considered in design. The huge bandwidth requirement of UWB antennas makes it different from other antennas. As per FCC regulations, a suitable UWB antenna should achieve at least 500 MHz of absolute bandwidth or at least 20% fractional bandwidth. Generally, the UWB antenna needs to be functional over the entire 3.1–10.6 GHz band.

A consistent performance of UWB antenna over the whole 3.1–10.6 GHz band is required. Ideally, stable radiation pattern, gain and good impedance matching over the entire band is desired. For some applications, in order to make the coexistence possible with other conventional systems, band-rejection characteristic in UWB antenna is required [37–43]. Depending upon the application, an omni-directional antenna or a directional antenna can be used. From application to application, the required radiation property of the antenna also differs. For mobile and hand-held communication systems, an omnidirectional antenna is preferred so that the signal can be received from any direction. The gain of such UWB antenna is of the order of 5 dBi over the entire band. For applications where high gain is required such as radar systems, a directional antenna is used.

The limitation imposed on power spectral density of transmitted pulse is very low in UWB systems. The performance of the system may degrade if losses increase. So, the losses should be as minimum as possible to achieve maximum radiation efficiency. Use of low loss dielectric materials leads to higher radiation efficiency, if a proper impedance matching is maintained at the input port. A greater emphasis is to be laid on size reduction techniques while designing. In order to be compatible with the present day mobile and portable systems, the UWB antenna needs to compact and planar. The antenna should also have the features of low profile and should be compatible with the printed-circuit-board (PCB) technology.

Apart from the above discussed frequency domain performance parameters, a good time-domain characteristics are demanded for UWB antennas. The efficacy of a narrow band antenna is same in its whole bandwidth, while the same cannot be assumed for a UWB antenna because of its huge operational bandwidth. The UWB antenna exhibits significant impact on the transmitted signal. The group delay should be nearly constant over the whole bandwidth, which implies that the phase response should be linear. This minimizes the pulse shape distortion. The ringing duration needs to negligibly small since it lead to inter symbol interference (ISI). The general requirements for a UWB antenna for high-speed communication is given in Table 2.1.

Table 2.1 Antenna parameter considerations for impulse radio communications

Parameters	Requirement
Impedance bandwidth	3.1–10.6 GHz
Radiation pattern	Omni-directional
Phase	Nearly linear
Physical profile	Compact and planar

2.1.6 Characterization of UWB Antennas in Time-Domain

Unlike narrowband communication systems, UWB antennas are used to transmit and receive short pulses (order of pico-seconds) in time domain. Since the antenna radiates over a wide spectral bandwidth, it's necessary to investigate the impulse response of these antennas for successful integration to impulse-radio systems. The parameters mentioned in the previous section are used to characterize UWB antennas in frequency-domain. The time-domain performance of UWB antennas is evaluated based on a variety of parameters. To characterize wideband antennas, some extra parameters are formulated and utilized. These parameters include impulse response, system fidelity factor (SFF) and group delay.

UWB antenna system can be modeled as a two port network as shown in Fig. 2.4 [26]. The antenna system measurement is usually carried out outside anechoic chamber considering a real world situations Fig. 2.5. The antenna system consists of two identical UWB antennas connected to two ports of a vector-network-analyzer (VNA). Keeping in mind of the maximum power output of VNA, the antenna are usually separated by a small distance. To investigation of the system performance, two different spatial orientation between the transmitter and receiver are considered. In one case the antennas remain oriented face-to-face to each other, whereas for second case it is oriented side-by-side to each other. In face-to-face orientation, the two front side of the antennas face each other and in side by side orientation, the front side of the antennas point to the same direction.

The magnitude of system transfer function ($|S_{21}|$) and group delay can be used to analyze the antenna dispersion characteristics. The transmitted pulse will undergo distortion if the phase of S_{21} exhibits nonlinearity at the magnitude drop portion in

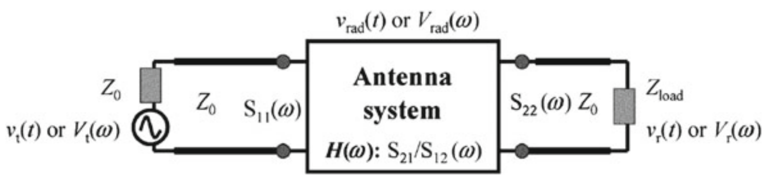

Fig. 2.4 UWB antenna system as a 2-port network [26] [From: Z. N. Chen, X. H. Wu, H. F. Li, N. Yang, and M. Y. Chia, "Considerations for source pulses and antennas in UWB radio systems," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 7, pp. 1739-1748, July 2004. © IEEE 2004. Reproduced courtesy of IEEE, USA.]

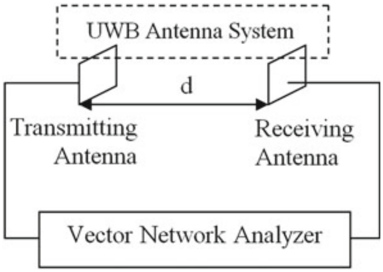

Fig. 2.5 UWB antenna measurement setup using a VNA

the S_{21} plot. But the amount by which the pulse distortion takes place cannot be said with system transfer function alone.

Group Delay parameter is generally encountered when characterizing two port devices. Group delay measures the total phase distortion between the input signal and the output signal. The transfer function $H(\omega)$ of the device is expressed in (2.1) and the group delay is defined as the derivative of the phase response ($\angle H(\omega)$) versus frequency is given in (2.3) as follows:

$$H(\omega) = |H(\omega)|e^{j\angle H(\omega)} \tag{2.1}$$

$$\tau = -\frac{\partial}{\partial \omega}\angle H(\omega) = -\frac{1}{360^0}\frac{\partial}{\partial f}\angle H(f) \tag{2.2}$$

The dispersion produced by the antenna would be less when the transfer function magnitude varies less and the group delay is nearly constant over the desired frequency range. A fairly linear phase of S_{21} and small the group delay variation (<1 ns) over the entire UWB band ensures pulse transmission with minimal distortion.

The impulse reponse of the antenna system can be obtained using the method of Hermitian processing as described in [27]. First, the pass-band signal can be obtained with zero padding from the lowest frequency measured in VNA down to DC. The complex conjugate of the signal can be obtained and when reflected to the negative frequencies gives a double sided spectral response. This is done to obey the property of time domain real signals that the Fast Fourier Transform (FFT) of real signals is conjugate symmetric in nature. Hence the Inverse Fast Fourier Transform (IFFT) of the complex conjugate signal gives real valued signal. The resultant double-sided spectrum after reflection corresponds to a real signal spectrum. Then IFFT of the double sided spectrum is performed to obtain the impulse response in time domain. The algorithm for obtaining impulse response is shown in Fig. 2.6. The received signal can be obtained by simply doing convolution of impulse response with the transmitted signal.

The transmitted signal is subjected to dispersion by the antenna as well as the channel. The fidelity factor quantifies the amount of resemblance between the input driving voltage and the radiated waveform of a transmitting antenna [28]. The correlation between the normalized input signal and the normalized radiated field gives

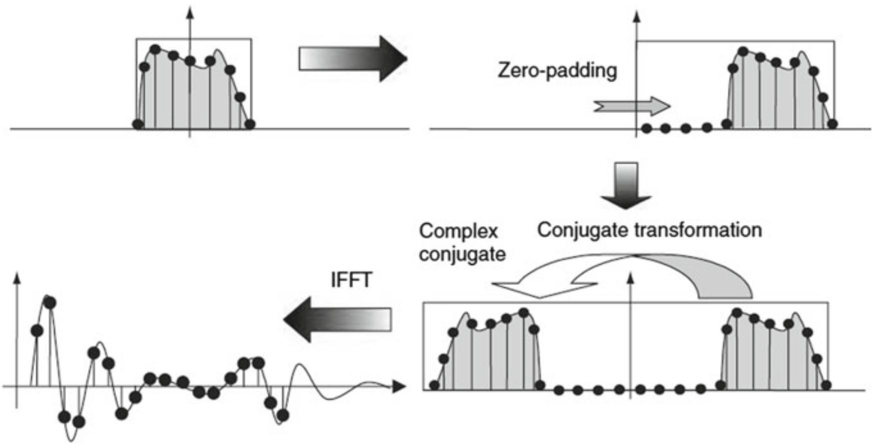

Fig. 2.6 Zero padding, conjugate reflection and resulting impulse response [27] [From: I. Oppermann, M. Hamalainen, and J. Iinatti, UWB Theory and Applications, John Wiley & Sons Ltd., 2004. Reproduced courtesy of John Wiley, USA.]

fidelity factor [29]. In practical scenario, it is difficult to measure radiated field. Alternately SFF is defined as correlation between the transmitted signal and the received signal. From the S_{21} parameter of the antenna system (transmitting antenna, channel, receiving antenna), SFF quantifying the extent of resemblance between the transmitted signal and the received signal. Fidelity factor considers the distortion yielded by transmit antenna only while the system fidelity factor considers the distortion yielded by two antennas (transmitting antenna and receiving antenna) [30]. The correlation between the normalized transmitted and the normalized received pulse gives SFF as:

$$SFF = max_{\tau} \frac{\int_{-\infty}^{\infty} T_s(t).R_s(t+\tau)dt}{\left[\int_{-\infty}^{\infty} |T_s(t)|^2\, dt\right]^{1/2} \left[\int_{-\infty}^{\infty} |R_s(t)|^2\, dt\right]^{1/2}} \tag{2.3}$$

Normalization is performed so that only the shape of the signals is taken into account without considering magnitudes, since the received signal magnitude is always less than transmitted signal [79]. The value of SFF lies between 0 and 1. The received signal is entirely different from transmitted signal when the SFF value is 0, and is identical when the SFF value is 1. The signal is unrecognizable when the distortion is greater than 50% or SFF is <0.5.

Ringing is defined as undesired oscillation of in signal. Ringing effect in antennas arise due to the stored energy or multiple reflections between the input port and the antenna [31]. The time taken by the envelope of a signal h(t) to reduce from peak value to a certain lower value (5% of the peak) is called the ringing duration.

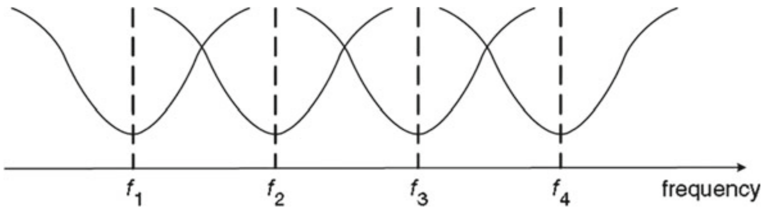

Fig. 2.7 Overlapping of multiple resonance modes [14] [From: J. Liang, "Antenna study and design for ultra-wideband communication," Ph.D. Thesis, Queen Mary University of London, United Kingdom, July 2006. Reproduced courtesy of Queen Mary Uni. of London.]

2.1.7 Revisiting Antennas for UWB Communications

The huge impedance bandwidth of planar monopole antennas is achieved by overlapping of closely spaced resonance modes. General illustration of this phenomenon is shown in Fig. 2.7.

An extensive study on the performance characterization of a CPW fed circular disc monopole antenna has been conducted in [127]. The critical parameters that define the performance of the antenna are the feed gap, width of ground plane and radius of monopole disc. Since the first two parameters depend on the ground plane, it plays a significant role that affect the operational bandwidth and input impedance matching. It acts as an impedance matching circuit. The UWB characteristic is achieved by the overlapping of multiple resonances and harmonics.

A CPW fed planar monopole antenna which is in the shape of L-I is presented in [128]. The wideband operation is achieved using the concept of resonance overlapping. The I-shaped open stub and an L-shaped monopole are connected to the center metal of the CPW feed line. The L-shaped monopole antenna excites resonant frequencies at 4.1 and 9.2 GHz while the I-stub excites resonant frequency at 5.4 GHz. The three resonant frequencies overlap to attaining a 3–11 GHz impedance bandwidth. In [84], a UWB antenna with huge bandwidth from 2–20 GHz is designed using three pairs of multi-resonant split ring loops. The resonances of these split ring loops overlap leading to the huge bandwidth.

To improve impedance matching between the feeding structure and radiator, an impedance matching network can be employed. It can be done in two ways. Introducing an on-patch matching network by cutting out the lower part of the patch near the feed in various shapes. An impedance matching network can be introduced in the feed line without modifying the radiator. The on-patch matching network transfers the high impedance at the radiating edge of the patch to the 50 Ω characteristic impedance of the feed. Suitable step or linear tapering on the lower part of the patch near the feeding structure leads to the gradual variation of impedance instead of sudden variation. The impedance matching is considerably increased at higher frequencies because of the enhancement in traveling wave modes. Stepped trimming at the lower side of the patch is done forming two steps [129], and four stepped for wide bandwidth [130]. In [72], lower side of the patch is trimmed linearly to achieve wide band operation. Patch in the shape of a bell is designed for impedance matching over a wide bandwidth [131].

Stepped impedance transformer acting as an impedance matching network in the feeding structure to achieve wideband impedance bandwidth is presented in [132, 133]. In [134], a linearly tapered impedance transformer is used in a coplanar-fed monopole antenna for impedance matching.

The antennas impedance bandwidth depends on the ground plane because of the coupling effect between the ground and lower side of the radiating patch. The ground plane is also a part of impedance matching network and it also contributes to radiation. The ground plane edge lying below the patch can be step tapered, linearly tapered or exponentially tapering leading to the reduction of capacitance between the bottom of the patch and the ground. This improves the impedance bandwidth of the antenna.

In [135], the outer part of the ground plane is trimmed that resulted in an inverted stair-style ground. In [136], the ground plane is linearly tapered and the matching at higher frequencies is considerably increased with no significant effect at the lower frequencies. After linearly tapering the outer part of the ground plane near the lower side of the radiating patch, antennas impedance bandwidth is increased by more than two times. Exponentially tapered ground plane is discussed in [138]. Because of tapered ground, multi resonance characteristics are obtained at around 5, 10 and 16 GHz. In [137], a combination of above mentioned techniques have been used and a compact CPW fed UWB antenna having an impedance bandwidth 8.4 GHz from 3–11.4 GHz was designed.

A trident shape feeding strip to achieve linear surface current on patch is presented in [139]. This results a wider bandwidth compared to a simple feeding line. Impedance bandwidth have been increased by more than two folds. When the planar monopole antenna is fed by multiple lines, vertical current distribution increases and the distribution of current in horizontal direction is decreased, leading to the improvement in achieved impedance bandwidth [136].

A full disc UWB monopole antenna was designed in [140]. Then the antenna has been cut across the symmetrical plane and the feed line has been tapered to achieve huge bandwidth from 2.42–13.62 GHz exhibiting stable gain and omni-directional radiation patterns. 50% reduction in size was achieved due to symmetric cut.

Dielectric resonator antennas (DRA) are also investigated for wideband applications [56–58, 61]. These antennas are compact in nature and use a high dielectric material as the radiating element with consistent radiation pattern, good radiation efficiency and gain. Patch antennas with fractal geometries are also investigated for multi-band and wideband applications [68–70].

An UWB antenna can be made to reject or suppress certain bands by etching slots on the patch and/or on the ground plane. Destructive interference takes place when the total slot length is adjusted to be approximately $\lambda_g/2$ at the notch frequency desired, hence the antenna is non-radiating at that frequency. The currents on either side of the gap formed are out of phase and hence destructive interference takes place [141]. The effective length of the slot can be calculated by summing the lengths of all the segments that form the slot. Therefore, the notch frequency is dependent on the length of the slot. The gap of the slot can be altered to change the rejected bandwidth. The location of slot on the patch, the length and gap of the slot are the critical parameters. Multiple slots on the patch produce multiple notch bands. An E-shaped slot in [142], C-shaped slot in [143] were used for producing band-notch. A modified π-shaped slot and a half arc shaped slot have been etched on the same patch for producing two notch bands [144].

A slot in the shape of 'L' has been etched in the ground plane of a CPW fed antenna to create notch at 5.5 GHz [145]. The length od the slot is approximately $\lambda_g/4$ at the notch frequency. The stop-band bandwidth changes by varying the location and width of the slot. At the notch frequency, it was observed that the surface current is concentrated only around the slot. In [146], quarter wavelength slots have been etched on the patch to obtain band-notched characteristics at 3.3–3.7 GHz, 5.15–5.825 GHz and 7.25–7.75 GHz. In the equivalent transmission model, each slot can be considered as a shorted stub connected in series with the primary transmission line. At the notch frequency, as its length becomes quarter wavelength, infinite input impedance appears in series between the input transmission line and the antenna. So the radiating patch and the feed line becomes impedance mismatched [147].

In [73], the elliptical radiator has been connected with two resonating inverted L- shaped stubs to produce band rejection characteristic. The L- shaped stub length is approximately $\lambda_g/4$ at notch frequency. An open stub can be connected to the primary transmission line as equivalent for L-shaped stub. At the notch frequency, the stub acts as short on transmission line with no power reaching the antenna, hence no radiation from the antenna.

The quarter wavelength stub can be of any shape. In [74], a meander shaped quarter wavelength stub is used to produce band-notch characteristic. In [72], a parasitic loop is placed beneath radiating patch. The parasitic loop length is $\lambda_g/2$ at notch frequency. The currents on the radiating patch and parasitic loop are opposite to each other at the notch frequency [43], hence band-notch characteristic is obtained. In [148], SRR array is placed on the back side of the UWB antenna to achieve band-notch characteristics. It is observed that the SRR array position affects the higher frequency resonance characteristics. The SRR array acts as cascaded parallel LC circuit. The input signal sees high input impedance at SRRs resonant frequency. So the radiating patch and the feed-line are impedance mismatched [148].

In [75], an UWB antenna with triple band-notch characteristics is designed. A radial slot in ground plane, a pentagonal slot etched on the patch and a pair of stubs connected to the feed-line below the radiating patch are used to obtain band-notch performance. The stub length is quarter wavelength at notch frequency and in the equivalent circuit model it acts as a short circuit. Hence no power reaches the antenna at that frequency.

Varactors have been used to achieve the tunability in notch frequency. In [94], a quarter wavelength open stub resonator is placed under the transmission line and is connected through a 'via' hole to the transmission line. Stubs are loaded with varactors. The capacitance of Varactor changes with the applied bias voltage, hence the electrical length of the stub changes. Consequently the notch frequency is changed. In [149], the slot etched on the UWB antenna, and loaded with a varactor is used to achieve the tunability of the notch frequency in the band 5–6 GHz. The bias circuit is designed using distributed elements. In [150], open loop resonators are placed on either side of the microstrip feed. A varactor is placed in the center of the resonator to tune effective electrical length of the loop to achieve tunability in the notch frequency.

2.2 Multiple-Input-Multiple-Output Communication Systems

2.2.1 Brief Description of MIMO Communication Systems

In a conventional radio system, there is single antenna used at transmitter and at receiver. This system is called a single-input-single-output (SISO) System. The capacity of such systems is limited by Shannon-Nyquist criterion. In order to increase channel capacity of the SISO systems to meet increasing demands for high-speed communications, the bandwidth and transmission power have to be increased significantly. However it is not a viable solution as discussed earlier. Recent developments have shown that using MIMO systems could substantially increase the capacity in wireless communication without increasing the transmission power and bandwidth [9, 10]. Rayleigh fading caused by multipath has been a source of problem for conventional wireless systems. However, to increase the channel capacity, MIMO systems exploit multipath instead of mitigating it. MIMO wireless systems have demonstrated the potential of increased capacity in rich multipath environments. Such systems operate by exploiting the spatial properties of the multipath channel, thereby offering a new dimension which can be used to enhance performance. A typical MIMO system is shown in Fig. 2.8 with multi channel concept. The performance improvements by using MIMO systems are due to array gain, diversity gain, spatial multiplexing gain, and interference reduction [8]. A brief description of these is given below:

1. *Array Gain*: Array gain is the enhancement of signal quality at the receiver through signal processing. The SNR at the receiver can be improved by combining coherently the signals received at each antenna [8].
2. *Diversity Gain*: Diversity is a powerful technique to mitigate fading in wireless links. Diversity techniques rely on transmitting the signal over multiple independently fading paths (in time/frequency/space) [8]. The received signals can be combined together in some optimal way to yield a signal with better SNR.
3. *Spatial Multiplexing Gain*: MIMO channels offer an increase in capacity for no additional power or bandwidth expenditure. This gain, referred to as spatial multiplexing gain, is realized by transmitting independent data signals from the individual antennas. Under favorable channel conditions, such as rich scattering, the receiver can separate the different streams, yielding an increase in the capacity [8].

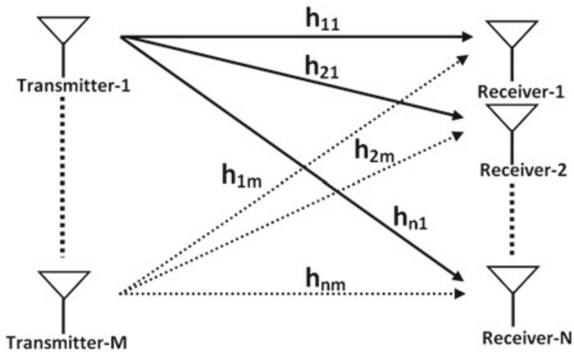

Fig. 2.8 A MIMO system comprising M-transmit antennas and N-receive antennas

4. *Interference Reduction*: When multiple antennas are used, the differentiation between the spatial signatures of the desired signal and interference signals can be exploited to reduce interference. Interference reduction requires knowledge of the desired signals channel. Exact knowledge of the interferers channel may not be necessary [8].

2.2.2 Channel Capacity in MIMO Systems

The channel capacity of a MIMO system with 'N' transmitting and 'N' receiving antenna is discussed here. Figure 2.8 shows a MIMO system with 'M'—transmitting antennas and 'N'—receiving antennas. For generalization, consider M=N for the present case.

For a narrowband channel, the complex transmission coefficients between element $k \in [1, \ldots, N_T]$ at the transmitter and element $j \in [1, \ldots, N_R]$ at the receiver at time 't' is represented by $h_{jk}(t)$. A matrix containing all channel coefficients (channel coefficient matrix, H(t)) can be shown as:

$$H(t) = \begin{pmatrix} h_{1,1}(t) & h_{1,2}(t) & \ldots & h_{1,N_T}(t) \\ h_{2,1}(t) & h_{2,2}(t) & \ldots & h_{2,N_T}(t) \\ \ldots & \ldots & \ldots & \ldots \\ h_{N_R,1}(t) & h_{N_R,2}(t) & \ldots & h_{N_R,N_T}(t) \end{pmatrix} \tag{2.4}$$

Hence, a system transmitting the signal vector $x(t) = [x_1(t), x_2(t), \ldots, x_{N_T}(t)]^T$, where $x_k(t)$ is the signal transmitted from the kth antenna would result in the signal vector $y(t) = [y_1(t), y_2(t), \ldots, y_{N_R}(t)]^T$ being received, where $y_j(t)$ is the signal received by the jth antenna, and

$$y(t) = H(t)x(t) + n(t) \tag{2.5}$$

where n(t) is the noise vector. MIMO systems can achieve orthogonal sub channels between the transmitters and receivers through a rich scattering environment and consequently increase the offered capacity. Mathematically, the number of independent sub channels can be estimated by using the singular value decomposition (SVD) of the channel coefficient matrix H as described in [12] as:

$$H = USV^H \tag{2.6}$$

where $(.)^H$ denotes Hermitian operator. Let the unencoded $N_T \times 1$ transmit vector in the waveform domain be denoted as x'. We encode the transmit vector as $x = Vx'$. Since each element of x' multiplies the corresponding column of V, this operation suggests that each column of V represents array weights for each signal stream. The receiver performs the operation $y' = U^H y$. Indicating that each row of U^H represents the receive array weights for each stream. Because U and V are unitary,

$$y' = U^H y = Sx' + \eta' \tag{2.7}$$

where $\eta' = U^H \eta$. Since the matrix S of singular values is diagonal, (2.7) indicates that y' is a scaled version of the transmit vector corrupted by additive noise. Therefore, this processing has created independent (spatially orthogonal) parallel communication channels in the multipath environment. The number of independent parallel communication channels is equal to rank of H matrix. Therefore in rank deficient channels, the number Q of independent streams should correspond to the number of singular values that are above the noise floor. In this case, only the first Q columns of U, S and V should be used in the processing and analysis.

To see how parallel spatial channels can increase capacity, consider the simple case of Q uncoupled transmission lines. If only one transmission line is used to send data, the Shannon channel capacity will be

$$C = log_2(1 + \rho) \tag{2.8}$$

where ρ is the receiver SNR. If the transmit power is divided equally among the lines, the capacity becomes [12]:

$$C_Q = \sum_{q=1}^{Q} log_2(1 + \rho/Q) = Qlog_2(1 + \rho/Q) \tag{2.9}$$

where equal receiver noise is assumed. For a transmit vector whose elements are complex Gaussian-distributed random variables, the expression for channel capacity is [12]:

$$C = MAX_{R_x:Tr(R_x)\leq P_t} log_2 \left| I + \frac{HR_xH^H}{N_T\sigma^2} \right| \tag{2.10}$$

where $R_x = E(xx^H)$ is the transmit covariance matrix. The diagonal elements of R_x represent transmit power from each antenna and hence the constraint $Tr(R_x) \leq P_t$, where P_t is the total transmit power. Determining the capacity involves identifying the covariance matrix R_x that maximizes (2.10). However, when the transmitter does not know H, it can divide power equally among the transmit antennas to form N_T independent streams or $R_x = (P_T/N_T)I$. Thus the uninformed transmit capacity is given by [12],

$$C = MAX_{R_x:Tr(R_x)\leq P_t} log_2 \left| I + \frac{HR_xH^H}{N_T\sigma^2} \right| \tag{2.11}$$

$$C_{UT} = log_2 \left| I + \frac{P_T HH^H}{N_T\sigma^2} \right| \tag{2.12}$$

which may be decomposed as [8]:

$$C_{UT} = \sum_{i=1}^{r} log_2 \left(1 + \frac{P_T \lambda i}{N_T\sigma^2} \right) \tag{2.13}$$

where r is the rank of H and λ_i (i = 1 to r) are the positive eigenvalues of HH^H. It can be seen from (2.13) that multiple scalar spatial data pipes (also known as spatial modes) open up between transmitter and receiver resulting in significant capacity gains over the SISO case. For example, C_{UT} increases by r bit/s/Hz for every 3-dB increase in transmit power (for high transmit power), as opposed to 1 bit/s/Hz in conventional SISO channels. It has been established that at high SNR, the ergodic capacity of fading MIMO channels is [8]:

$$C = min(N_R, N_T) log_2 \rho + O(1) \tag{2.14}$$

This clearly shows that the capacity increases linearly with the minimum of the number among transmit and receive antennas.

2.2.3 Signalling Schemes in MIMO Systems

There are basically two signalling schemes for MIMO systems.

1. *Spatial Multiplexing*:
   In spatial multiplexing (SM) independent data streams are transmitted simultaneously in parallel channels from each element in an array of antennas. The basic principle of SM is illustrated in Fig. 2.9, wherein a system with two elements at the transmitter and two elements at the receiver is considered [32]. Firstly, the bit stream of data to be transmitted is de-multiplexed into two sub streams, then modulated and transmitted simultaneously from each transmit antenna as

shown in Fig. 2.9. Under favorable channel conditions such as rich scattering, these signals arriving at both the receiving antennas are well separated. The the receiver has knowledge of the channel, hence it can differentiate between the co-channel signals and extract both signals. After demodulating the received signals, the original sub streams can be combined to yield the original bit stream of data. Therefore, spatial multiplexing leads to increase in the channel capacity with the number of transmitting-receiving antenna pairs. This concept can be extended to more general MIMO channels.

The maximum number of parallel channels that can be achieved in an ideal MIMO system is min(N_T, N_R). Shannons capacity formula for SISO system indicates that in the high SNR regime, a 3 dB increase in SNR will approximately increase capacity by 1 bit/s/Hz. However, in the MIMO systems, the capacity in the high SNR regime will increase by min(NT ,NR) bits/s/Hz with every increment of 3 dB in SNR. However, SM does not work well in low SNR environments as it is more difficult for the receiver to identify the uncorrelated signal paths [32].

2. *Space-Time Coding*:
   An alternative method of exploiting MIMO channels is known as space-time coding. It aims to improve the systems performance by using multiple antennas for diversity gain rather than for the spatial-multiplexing gain. It increases network throughput by selecting quality signal paths such that higher data rates can be achieved [32].

   In space-time coded transmitter, a single data stream is encoded across both time and space to produce the symbol streams for each transmitter as shown in Fig. 2.10. Appropriate decoding at the receiver allows a diversity gain to be achieved. This

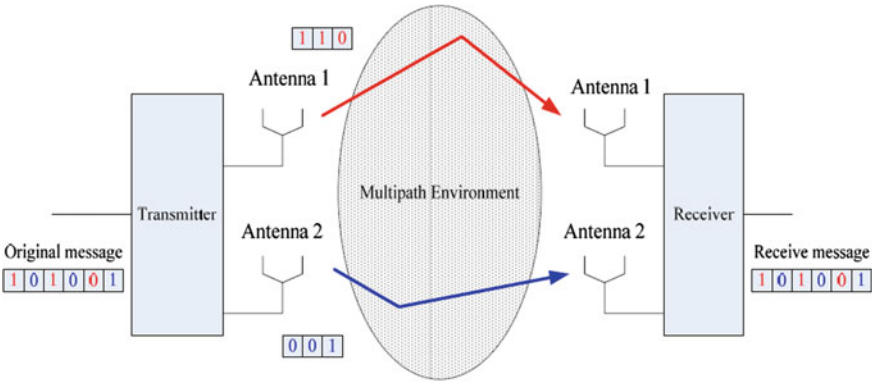

Fig. 2.9 A 2 × 2 MIMO system with a spatial multiplexing scheme [32] [From: C. Chiau, "Study of the diversity antenna array for the MIMO wireless communication systems," Ph.D. Thesis, Department of Electronic Engineering, Queen Mary, University of London, United Kingdom, April 2006. Reproduced courtesy of Queen Mary, Uni. of London.]

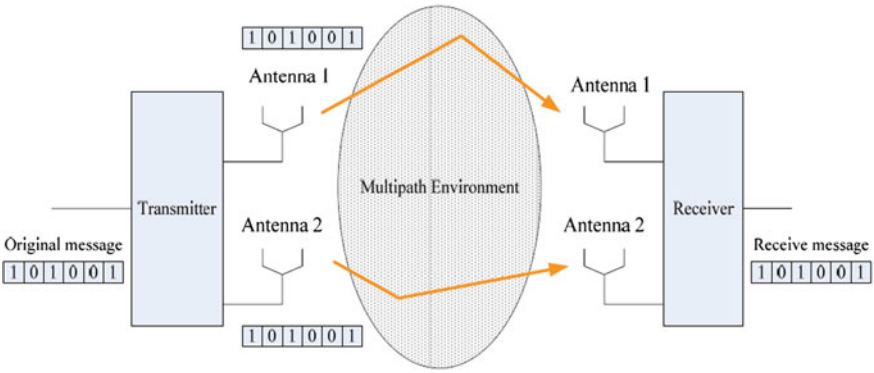

Fig. 2.10 A 2 × 2 MIMO system with a space-time coding scheme [32] [From: C. Chiau, "Study of the diversity antenna array for the MIMO wireless communication systems," Ph.D. Thesis, Department of Electronic Engineering, Queen Mary, University of London, United Kingdom, April 2006. Reproduced courtesy of Queen Mary, Uni. of London.]

method is particularly attractive as it does not require channel knowledge at the transmitter [32]. The resulting diversity gain improves the reliability in a fading wireless link, hence it improves the quality of the transmission. It should be noted that space-time coding scheme does not increase the capacity linearly with the number of transmit/receive elements used. However, it maximizes the wireless range and coverage by improving the quality of the transmission. In order to improve both range and capacity, a MIMO system requires to support both the spatial multiplexing and space-time coding schemes. Examples of diversity coding techniques are the Alamouti coding scheme and delay diversity.

2.2.4 Antenna Diversity in MIMO Systems

Multipath propagation had historically been regarded as an impairment because it causes signal fading. In order to mitigate this problem, diversity techniques were developed. The basic principle of diversity is that the receiver should have more than one version of the transmitted signal available, where each version is received through a different uncorrelated channel. These versions of transmitted signal are combined appropriately to give signal with higher mean SNR at the output compared to a single branch resulting in a diversity gain.

Figure 2.11 shows that a dual-element diversity antenna at a receiver that can receive two different versions of transmitted signals and combine them using a combining circuit. There are three kinds of diversity combining techniques

1. *Selection Diversity*: Chooses the path with the highest SNR, and performs detection based on the signal from the selected path.

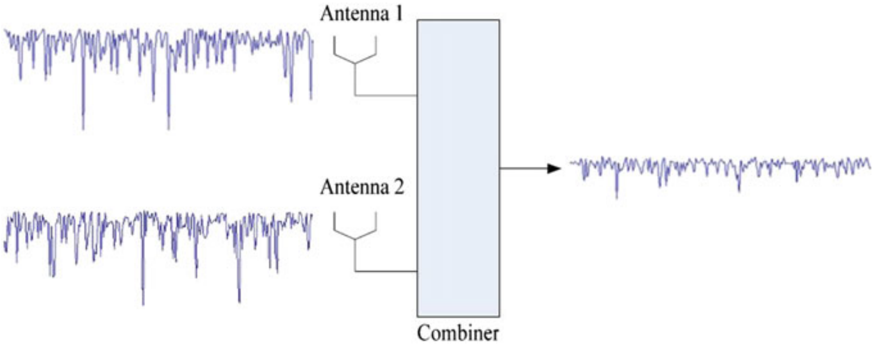

Fig. 2.11 Two signals are combined in a basic diversity receiver [32] [From: C. Chiau, "Study of the diversity antenna array for the MIMO wireless communication systems," Ph.D. Thesis, Department of Electronic Engineering, Queen Mary, University of London, United Kingdom, April 2006. Reproduced courtesy of Queen Mary, Uni. of London.]

2. *Maximal Ratio Combining* (*MRC*): Makes decisions based on an optimal linear combination of the signals.
3. *Equal Gain Combining* (*EGC*): Simply adds the signals after they have been co-phased [33].

MIMO systems exploit the channel spatial degrees of freedom to increase communication performance. In traditional antenna diversity, these resources are used to transmit and/or receive duplicate copies of a single information stream to increase the reliability of detection. In contrast, spatial multiplexing indicates sending distinct information streams over the channels to increase throughput and spectral efficiency.

The combined diversity and spatial multiplexing accomplished with a MIMO system will depend on the required throughput and quality-of-service for an application. This relationship implies that the traditional mechanisms used in diversity systems for reducing antenna correlation generally work to improve MIMO performance as well. It is emphasized, however, that low correlation is a necessary but not sufficient condition for good MIMO performance, since the propagation environment must also possess the appropriate characteristics [34].

It must be noted that if the mean powers in different signal branches are not in same level, then signal in a weaker antenna may not be useful although it is less faded. Therefore to design antenna with a good diversity, the mean power levels in all signal branches should be similar and the correlation between between antennas must be minimum. There are five categories of diversities, i.e. (a) frequency diversity, (b) time diversity, (c) spatial diversity, (d) pattern diversity and (e) polarization diversity. Amongst these five diversities, only the spatial, pattern and polarization diversity techniques are categorized as antenna diversity [32]. These antenna diversity techniques can be applied to obtain uncorrelated signals in MIMO system. The following subsections describe these three types of antennas diversities:

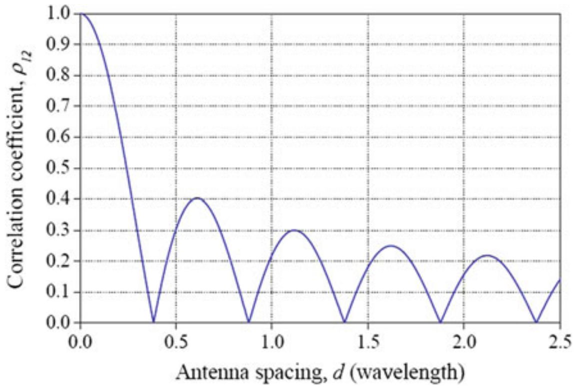

Fig. 2.12 The effect of inter element spacing on correlation coefficient [32] [From: C. Chiau, "Study of the diversity antenna array for the MIMO wireless communication systems," Ph.D. Thesis, Department of Electronic Engineering, Queen Mary, University of London, United Kingdom, April 2006. Reproduced courtesy of Queen Mary, Uni. of London.]

1. *Spatial Diversity*:
   In spatial diversity more than one antenna are employed which are sufficiently separated from each other so that the relative phases of the multipath contributions are significantly different at the two antennas. When large phase differences are present, they give rise to a low correlation between the signals at the antennas. By assuming a two dimensional scenario wherein angular density function is taken to be uniform in azimuth of the mobile environment and no angular density function in elevation, the correlation coefficient for a distance separation 'd' can be obtained from the *zero*th order Bessel function, $J_0(x)$ [32].
   The first null of is $J_0(\beta \mathrm{d})$ is at d = 0.4λ, as shown in Fig. 2.12, where β is the phase constant. As shown graphically in Fig. 2.12, the correlation coefficient has a minima at d = 0.4λ and increases thereafter. However, in suburban areas the measurements show that the first null appears at about d = 0.8λ. This may be due to a lack of uniform angular distribution of wave arrival [32]. Generally, spacing, d of 0.5λ is practically used to obtain two uncorrelated signals at mobile terminals.
2. *Pattern Diversity*:
   Pattern diversity occurs in many instances at the mobile terminals because the antennas will pick up signals coming from different angles. Since the fading signals coming from different directions in a multipath environment are independent, use of pattern diversity can lead to uncorrelated channels. At the mobile terminal, two omni-directional antennas interacting with each other whilst closely spaced can also obtain pattern diversity. Basically, the antennas act as parasitic elements to each other and their patterns change to allow signals to be picked up at different angles. Recent studies conducted on pattern diversity in the MIMO systems have shown that with appropriate dissimilarity in the antenna pattern, the system can achieve large channel capacity [34]. One suggested approach for realizing such a situation involves the use of multimode antennas where the patterns for different modes exhibit high orthogonality (low correlation) [35]. Finding other antenna topologies that offer this orthogonality in a compact form remains an area of active research [59, 60].

3. *Polarization Diversity*:
   Recent work has suggested that in a rich multipath environment, sensing the three Cartesian vector components of the electric and magnetic fields can provide six uncorrelated signals at the receiver. With the use of polarization diversity the size of the antenna structure can be reduced significantly. Polarization diversity can potentially lead to low correlation on at least two branches even when the channel is characterized by little or no multipath scattering. To understand the fundamentals of polarization diversity in MIMO systems, consider two dipoles that are rotated by angle α against one another, are used at the transmitter and at the receiver. Both transmit antennas radiate the same power. For increasing α capacity increases, since the correlation among the channel coefficients of H (the channel matrix) decreases. For $\alpha = 90^0$, the capacity is maximum as two different channels for two different polarizations exist. Those channels are coupled by cross polarization coupling and are not necessarily orthogonal sub channels. The correlation among the two signals is not only influenced by polarization diversity, but also by pattern diversity. Polarization and pattern diversity are always combined and can not be exploited separately [36].

2.2.5 Revisiting Antennas for MIMO Communications

Multiple-input multiple-output (MIMO) systems have emerged as a very interesting strategy to increase the capacity of wireless systems in rich scattering environments [35, 151]. Traditionally, a MIMO system employs several transmit and receive antennas at each end of the radio link, and in order to achieve a high capacity, different signal paths between them should be uncorrelated. The MIMO gains are mainly due to the fact that a rich scattering environment provides independent transmission paths from each transmit antenna to each receive antenna. Antenna diversity is a well-known technique to enhance the performance of wireless communication systems by reducing the multipath fading and co-channel interference [152, 153].

Antenna diversity can be realized in several ways [62–64]. Depending on the environment, space and the expected interference, designers can employ one or more of these methods such as spatial, polarization and pattern diversities to achieve diversity gain. Spatial diversity is employed using spatially separated antenna elements to achieve diversity gain [154–156]. In case of small portable devices, the available space is very limited and it is hard to implement more than one antenna element in such a small space.

However, degradation would result upon placing the antenna elements in close proximity, due to near-field effects, diffraction from finite-ground planes, and strong inductive and capacitive coupling between the elements [120, 152, 157].

The isolation between two antennas or two ports of a single antenna is a critical parameter in diversity antennas and MIMO systems. If the two ports/antennas are strongly coupled then most of the transmitted signal of one port/antenna will not be radiated, but rather received by the second port/antenna which will lead to a reduction of the radiation efficiency due to the dissipation of power in the coupled port/antenna. Therefore, for the two ports/antennas to radiate efficiently they should be sufficiently isolated and the mutual coupling should be small.

In addition, for application in portable devices very small form factors are an important requirement and therefore good isolation between antennas with closely packed antenna elements is necessary. For this reason, polarization and pattern diversity techniques are more suited in the portable devices wherein we can place two antenna elements very close to each other or a single element with more than one feed to achieve diversity gain.

In literature, different techniques have been introduced to increase isolation and reduce the mutual coupling between the two ports/antennas. For example, in [122, 158, 159], modifications on the ground plane have been reported with a view to increase the isolation. An alternate way of reducing the electromagnetic coupling between antenna elements sharing a common ground plane or chassis is to introduce resonant defects or slits in the ground plane. By proper choice of dimensions, the slits resonate and can trap some of the energy between the radiating elements [160, 161]. These resonant slots are designed to be either $\lambda_g/4$ or $\lambda_g/2$ in electrical length depending on open ended or closed type slots. These basically filters/blocks the co-polarized coupling current.

Another strategy to mitigate the coupling between radiating antennas is to use periodic and electromagnetic band-gap (EBG) structures. EBG structures were used extensively in mutual coupling reduction in planar and low-profile antennas [162–166]. These are basically high impedance surfaces which effectively suppress coupling surface currents between closely placed radiators.

Another mechanism proposed previously to isolate highly-coupled monopole antenna elements is by using 180^0 hybrid couplers [167–169]. The method is based on the mode-decomposition network, in which a multi-port network is inserted between the antennas and their driving ports.

To eliminate the use of extra resonating elements for enhancing isolation, use of a neutralization line between antenna elements is another promising technique. A neutralization-line cancels induced current with the direct current, since both currents have opposite phase [170–174]. However, the electrical length of the naturalization line is a critical parameter and frequency dependent (inversely proportional).

Using a parasitic element between two closely placed antennas of a MIMO system, significant enhancement in isolation (low coupling) can be achieved [175–177]. The parasitic resonator basically induces a secondary coupling current of opposite phase that suppresses the direct coupling current between radiating elements. This also decreases the overall radiation efficiency of the system.

Another way to achieve high isolation between antenna elements in a compact MIMO system is to excite different modes having overlapping resonance. The different modes have different surface current distribution that minimize the coupling between the modes [178–180].

http://www.springer.com/978-3-319-63174-5

Compact Antennas for High Data Rate Communication
Ultra-wideband (UWB) and
Multiple-Input-Multiple-Output (MIMO) Technology
Malik, J.; Patnaik, A.; Kartikeyan, M.V.
2018, XV, 103 p. 70 illus., Hardcover
ISBN: 978-3-319-63174-5